

Understanding Biological Behavior 2nd Edition

UNDERSTANDING
BIOLOGICAL BEHAVIOR
2ND EDITION

*Where Biology Meets Behavior - Discover
the Science Behind Human Actions*

EDITED BY EMILY MARLER

University of Missouri-St. Louis Library

St. Louis



Understanding Biological Behavior 2nd Edition Copyright © 2025 by Emily Marler is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/), except where otherwise noted.

CONTENTS

Introduction to Understanding Biological Behavior (Second Edition)	1

PART I. UNIT 1: INTRODUCTION

1. Psychophysiological Methods in Neuroscience	5
Zachary Infantolino and Gregory A. Miller	
2. The Nature-Nurture Question	28
Eric Turkheimer	

PART II. UNIT 2: NERVE CELLS AND HOW THEY COMMUNICATE

3. Neurons	51
Sharon Furtak	

PART III. UNIT 3: THE NERVOUS SYSTEM

4. The Nervous System	83
Aneeq Ahmad	
5. The Brain	112
Diane Beck and Evelina Tapia	

PART IV. UNIT 4: DEVELOPMENT AND PLASTICITY
OF THE BRAIN

6.	Brain Plasticity and Behaviour in the Developing Brain	139
	Bryan Kolb and Robbin Gibb	

PART V. UNIT 5: VISION & PERCEPTION

7.	Vision	143
	Simona Buetti and Alejandro Lleras	

PART VI. UNIT 6: OTHER SENSORY SYSTEMS &
ATTENTION

8.	Hearing	171
	Andrew J. Oxenham	
9.	Taste and Smell	195
	Linda Bartoshuk and Derek Snyder	
10.	Touch and Pain	214
	Guro E. Løseth; Dan-Mikael Ellingson; and Siri Leknes	

PART VII. UNIT 7: THE VESTIBULAR SYSTEM

11.	The Vestibular System	255
	Dora Angelaki and J. David Dickman	

PART VIII. UNIT 8: STATES OF CONSCIOUSNESS

12.	Attention	297
	Frances Friedrich	
13.	States of Consciousness	320
	Robert Biswas-Diener and Jake Teeny	

PART IX. [UNIT 9: REPRODUCTIVE BEHAVIORS](#)

14.	Hormones & Behavior	359
	Randy J. Nelson	
15.	Biochemistry of Love	399
	Sue Carter and Stephen Porges	

PART X. [UNIT 10: SOCIAL NEUROSCIENCE](#)

16.	Social Neuroscience	429
	Tiffany A. Ito and Jennifer T. Kubota	

PART XI. [UNIT 11: THE BIOLOGY OF LEARNING AND MEMORY](#)

17.	Memory (Encoding, Storage, Retrieval)	463
	Kathleen B. McDermott and Henry L. Roediger III	
18.	Conditioning and Learning	502
	Mark E. Bouton	
19.	Forgetting and Amnesia	545
	Nicole Dudukovic and Brice Kuhl	

PART XII. [UNIT 12: PSYCHOLOGICAL DISORDERS](#)

20.	Mood Disorders	567
	Anda Gershon and Renee Thompson	
21.	Schizophrenia Spectrum Disorders	605
	Deanna M. Barch	

Link to the 1st Edition, much thanks to the Noba Project	645

INTRODUCTION TO UNDERSTANDING BIOLOGICAL BEHAVIOR (SECOND EDITION)

“Understanding Biological Behavior” explores the intricate relationship between biology and human behavior. Compiled by Emily Marler with the help of the UMSL Libraries, this second edition offers students a peek into the biological underpinnings of psychological processes and behaviors.

The book is structured into twelve units, each focusing on a crucial aspect of biological psychology:

1. Introduction
2. Nerve Cells and How They Communicate
3. The Nervous System
4. Development and Plasticity of the Brain
5. Vision & Perception
6. Other Sensory Systems & Attention
7. Movement
8. States of Consciousness
9. Reproductive Behaviors
10. Emotional Behaviors
11. The Biology of Learning and Memory
12. Psychological Disorders

The material covers topics ranging from basic neuroanatomy to complex psychological disorders. The textbook aims to provide readers with a solid foundation in psychobiology, integrating knowledge from neuroscience, psychology, and related disciplines.

Key features of this book include:

- In-depth exploration of the brain's structure and function
- Examination of various sensory systems, including vision, hearing, taste, smell, and touch
- Discussion of consciousness, attention, and cognitive processes
- Investigation of the biological basis of emotions, learning, and memory
- Coverage of major psychological disorders from a biological perspective

By bridging the gap between biology and psychology, “Understanding Biological Behavior” equips readers with the knowledge and tools to appreciate the complex interplay between our biology and our actions, thoughts, and experiences.

Acknowledgment: This introduction was created by Claude 3.5 Sonnet, an AI language model, on September 4, 2024 and as AI-generated content requires human review, this has been reviewed by a human.

PART I.

UNIT 1: INTRODUCTION

CHAPTER 1.

PSYCHOPHYSIOLOGICAL METHODS IN NEUROSCIENCE

ZACHARY INFANTOLINO AND GREGORY A. MILLER

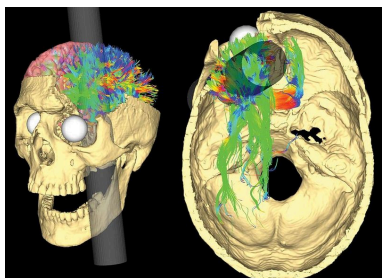
As a generally noninvasive subset of neuroscience methods, psychophysiological methods are used across a variety of disciplines in order to answer diverse questions about psychology, both mental events and behavior. Many different techniques are classified as psychophysiological. Each technique has its strengths and weaknesses, and knowing them allows researchers to decide what each offers for a particular question. Additionally, this knowledge allows research consumers to evaluate the meaning of the results in a particular experiment.

LEARNING OBJECTIVES

- Learn what qualifies as psychophysiology within the broader field of neuroscience.
- Review and compare several examples of psychophysiological methods.
- Understand advantages and disadvantages of different psychophysiological methods.

HISTORY

In the mid-19th century, a railroad worker named Phineas Gage was in charge of setting explosive charges for blasting through rock in order to prepare a path for railroad tracks. He would lay the charge in a hole drilled into the rock, place a fuse and sand on top of the charge, and pack it all down using a tamping iron (a solid iron rod approximately one yard long and a little over an inch in diameter). On a September afternoon when Gage was performing this task, his tamping iron caused a spark that set off the explosive prematurely, sending the tamping iron flying through the air.



Unlike other areas of the body, damage to the brain isn't localized to that specific region; injuries have widespread consequences for other areas, too. [Image: Van Horn JD, Irimia A, Torgerson CM, Chambers MC, Kikinis R, et al., <https://goo.gl/wdhM4o>, CC BY 2.5, <https://goo.gl/OQtWcf>]

Unfortunately for Gage, his head was above the hole and the tamping iron entered the side of his face, passed behind his left eye, and exited out of the top of his head, eventually landing 80 feet away. Gage lost a portion of his left frontal lobe in the accident, but survived and lived for another 12 years. What is most interesting from a psychological perspective

is that Gage's personality changed as a result of this accident. He became more impulsive, he had trouble carrying out plans, and, at times, he engaged in vulgar profanity, which was out of character. This case study leads one to believe that there are specific areas of the

brain that are associated with certain psychological phenomena. When studying psychology, the brain is indeed an interesting source of information. Although it would be impossible to replicate the type of damage done to Gage in the name of research, methods have developed over the years that are able to safely measure different aspects of nervous system activity in order to help researchers better understand psychology as well as the relationship between psychology and biology.

INTRODUCTION

Psychophysiology is defined as any research in which the dependent variable (what the researcher measures) is a physiological measure, and the independent variable (what the researcher manipulates) is behavioral or mental. In most cases the work is done noninvasively with awake human participants. Physiological measures take many forms and range from blood flow or neural activity in the brain to heart rate variability and eye movements. These measures can provide information about processes including emotion, cognition, and the interactions between them. In these ways, physiological measures offer a very flexible set of tools for researchers to answer questions about behavior, cognition, and health.

Psychophysiological methods are a subset of the very large domain of **neuroscience methods**. Many neuroscience methods are **invasive**, such as involving **lesions** of neural tissue, injection of neutrally active chemicals, or manipulation of neural activity via electrical stimulation. The present survey emphasizes **noninvasive** methods widely used with human subjects.

Crucially, in examining the relationship between physiology and overt behavior or mental events, psychophysiology does not attempt to replace the latter with the former. As an example, happiness is a state of pleasurable contentment and is associated with various physiological measures, but one would not say that those physiological measures *are* happiness. We can make inferences about someone's cognitive or emotional state based on his or her self-report, physiology, or overt behavior. Sometimes our interest is primarily in inferences about internal events and sometimes primarily in the physiology itself. Psychophysiology addresses both kinds of goals.

CENTRAL NERVOUS SYSTEM (CNS)

This module provides an overview of several popular psychophysiological methods, though it is far from exhaustive. Each method can draw from a broad range of data-analysis strategies to provide an even more expansive set of tools. The psychophysiological methods discussed below focus on the central



The adult human brain only makes up about 2% (i.e. $\approx$ 3 pounds) of the average adult's weight, but it uses 20% of the body's energy! [Image: _DJ_, <https://goo.gl/eHPH5L>, CC BY-SA 2.0, <https://goo.gl/rxiUsF>]

nervous system. Structural magnetic resonance imaging (sMRI) is a noninvasive technique that allows researchers and clinicians to view anatomical structures within a

human. The participant is placed in a magnetic field that may be 66,000 times greater than the Earth's magnetic field, which causes a small portion of the atoms in his or her body to line up in the same direction. The body is then pulsed with low-energy radio frequencies that are absorbed by the atoms in the body, causing them to tip over. As these atoms return to their aligned state, they give off energy in the form of harmless electromagnetic radiation, which is measured by the machine. The machine then transforms the measured energy into a three-dimensional picture of the tissue within the body. In psychophysiology research, this image may be used to compare the size of structures in different groups of people (e.g., are areas associated with pleasure smaller in individuals with depression?) or to increase the accuracy of spatial locations as measured with functional magnetic resonance imaging (fMRI).

Functional magnetic resonance imaging (fMRI) is a method that is used to assess changes in activity of tissue, such as measuring changes in neural activity in different areas of the brain during thought. This technique builds on the principles of sMRI and also uses the property that, when neurons fire, they use energy, which must be replenished. Glucose and oxygen, two key components for energy production, are supplied to the brain from the blood stream as needed. Oxygen is transported through the blood using **hemoglobin**, which contains binding sites for oxygen. When these sites are saturated with oxygen, it is referred to as **oxygenated hemoglobin**. When the oxygen molecules have all been released from a hemoglobin molecule, it is known as **deoxygenated hemoglobin**. As a set of neurons begin firing, oxygen in the blood surrounding those neurons is consumed,

leading to a reduction in oxygenated hemoglobin. The body then compensates and provides an abundance of oxygenated hemoglobin in the blood surrounding that activated neural tissue. When activity in that neural tissue declines, the level of oxygenated hemoglobin slowly returns to its original level, which typically takes several seconds.

fMRI measures the change in the concentration of oxygenated hemoglobin, which is known as the **blood-oxygen-level-dependent (BOLD)** signal. This leads to two important facts about fMRI. First, fMRI measures blood volume and blood flow, and from this we infer neural activity; fMRI does not measure neural activity directly. Second, fMRI data typically have poor **temporal resolution** (the precision of measurement with respect to time); however, when combined with sMRI, fMRI provides excellent **spatial resolution** (the ability to distinguish one object from another in space). Temporal resolution for fMRI is typically on the order of seconds, whereas its spatial resolution is on the order of millimeters. Under most conditions there is an inverse relationship between temporal and spatial resolution—one can increase temporal resolution at the expense of spatial resolution and vice versa.

This method is valuable for identifying specific areas of the brain that are associated with different physical or psychological tasks. Clinically, fMRI may be used prior to neurosurgery in order to identify areas that are associated with language so that the surgeon can avoid those areas during the operation. fMRI allows researchers to identify differential or convergent patterns of activation associated with tasks. For example, if participants are shown words on a screen and are expected to indicate

the color of the letters, are the same brain areas recruited for this task if the words have emotional content or not? Does this relationship change in psychological disorders such as anxiety or depression? Is there a different pattern of activation even in the absence of overt performance differences? fMRI is an excellent tool for comparing brain activation in different tasks and/or populations. Figure 1 provides an example of results from fMRI analyses overlaid on an sMRI image. The blue and orange shapes represent areas with significant changes in the BOLD signal, thus changes in neural activation.

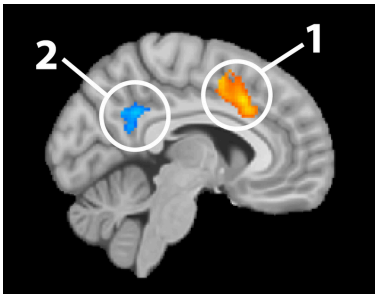


Figure 1. Example of fMRI analyses overlaid on an sMRI image. Area 1 (orange) indicates an increase in the BOLD signal, and Area 2 (blue) demonstrates a decrease in the BOLD signal. We infer that neural activity increased in the Area 1 and decreased in Area 2.

Electroencephalography (EEG) is another technique for studying brain activation. This technique uses at least two and sometimes up to 256 electrodes to measure the difference in electrical charge (the **voltage**) between pairs of points on the head. These electrodes are typically fastened to a flexible cap (similar to a swimming cap) that is placed on the participant's head. From the scalp, the

electrodes measure the electrical activity that is naturally occurring within the brain. They do not introduce any new electrical activity. In contrast to fMRI, EEG measures neural activity directly, rather than a correlate of that activity.

Electrodes used in EEG can also be placed within the

skull, resting directly on the brain itself. This application, called electrocorticography (ECoG), is typically used prior to medical procedures for localizing activity, such as the origin of epileptic seizures. This invasive procedure allows for more precise localization of neural activity, which is essential in medical applications. However, it is generally not justifiable to open a person's skull solely for research purposes, and instead electrodes are placed on the participant's scalp, resulting in a noninvasive technique for measuring neural activity.

Given that this electrical activity must travel through the skull and scalp before reaching the electrodes, localization of activity is less precise when measuring from the scalp, but it can still be within several millimeters when localizing activity that is near the scalp. One major advantage of EEG is its temporal resolution. Data can be recorded thousands of times per second, allowing researchers to document events that happen in less than a millisecond. EEG analyses typically investigate the change in amplitude or frequency components of the recorded EEG on an ongoing basis or averaged over dozens of trials (see Figure 2).

Magnetoencephalography (MEG) is another technique for noninvasively measuring neural activity. The flow of electrical charge (the current) associated with neural activity produces very weak magnetic fields that can be detected by sensors placed near the participant's scalp. The number of sensors used varies from a few to several hundred. Due to the fact that the magnetic fields of interest are so small, special rooms that are shielded from

magnetic fields in the environment are needed in order to avoid contamination of the signal being measured. MEG has the same excellent temporal resolution as EEG. Additionally, MEG is not as susceptible to distortions from the skull and scalp. Magnetic fields are able to pass through the hard and soft tissue relatively unchanged, thus providing better spatial resolution than EEG. MEG analytic strategies are nearly identical to those used in EEG. However, the MEG recording apparatus is much more expensive than EEG, so MEG is much less widely available.

EEG and MEG are both excellent for elucidating the temporal dynamics of neural processes. For example, if someone is reading a sentence that ends with an

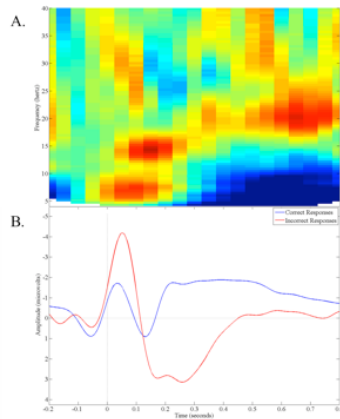


Figure 2. Example of EEG analysis output. Panel A represents changes in the relative strength of different frequencies in the EEG data over time. Panel B represents changes in the amplitude in the instantaneous EEG voltage over time.

unexpected word (e.g., Michelle is going outside to water the book), how long after he or she reads the unexpected word does he or she recognize this as unexpected? In addition to these types of questions, EEG and MEG methods allow researchers to investigate the degree to which different parts of the brain “talk” to each other. This allows for a better understanding of brain networks, such as their role in different tasks and how they may function abnormally in psychopathology.

Positron emission tomography (PET) is a medical imaging technique that is used to measure processes in the body, including the brain. This method relies on a positron-emitting tracer atom that is introduced into the blood stream in a biologically active molecule, such as glucose, water, or ammonia. A positron is a particle much like an electron but with a positive charge. One example of a biologically active molecule is fludeoxyglucose, which acts similarly to glucose in the body. Fludeoxyglucose will concentrate in areas where glucose is needed—commonly areas with higher metabolic needs. Over time, this tracer molecule emits positrons, which are detected by a sensor. The spatial location of the tracer molecule in the brain can be determined based on the emitted positrons. This allows researchers to construct a three-dimensional image of the areas of the brain that have the highest metabolic needs, typically those that are most active. Images resulting from PET usually represent neural activity that has occurred over tens of minutes, which is very poor temporal resolution for some purposes. PET images are often combined with computed tomography (CT) images to improve spatial resolution, as fine as several millimeters. Tracers can also be incorporated into molecules that bind to

neurotransmitter receptors, which allow researchers to answer some unique questions about the action of neurotransmitters. Unfortunately, very few research centers have the equipment required to obtain the images or the special equipment needed to create the positron-emitting tracer molecules, which typically need to be produced on site.

Transcranial magnetic stimulation (TMS) is a noninvasive method that causes **depolarization** or **hyperpolarization** in neurons near the scalp. This method is not considered psychophysiological because the independent variable is physiological, rather than the dependent. However, it does qualify as a neuroscience method because it deals with the function of the nervous system, and it can readily be combined with conventional psychophysiological methods. In TMS, a coil of wire is placed just above the participant's scalp. When electricity flows through the coil, it produces a magnetic field. This magnetic field travels through the skull and scalp and affects neurons near the surface of the brain. When the magnetic field is rapidly turned on and off, a current is induced in the neurons, leading to depolarization or hyperpolarization, depending on the number of magnetic field pulses. Single- or paired-pulse TMS depolarizes site-specific neurons in the cortex, causing them to fire. If this method is used over primary motor cortex, it can produce or block muscle activity, such as inducing a finger twitch or preventing someone from pressing a button. If used over primary visual cortex, it can produce sensations of flashes of light or impair visual processes. This has proved to be a valuable tool in studying the function and timing of specific processes such as the recognition of visual

stimuli. Repetitive TMS produces effects that last longer than the initial stimulation. Depending on the intensity, coil orientation, and frequency, neural activity in the stimulated area may be either attenuated or amplified. Used in this manner, TMS is able to explore **neural plasticity**, which is the ability of connections between neurons to change. This has implications for treating psychological disorders as well as understanding long-term changes in neuronal excitability.

PERIPHERAL NERVOUS SYSTEM

The psychophysiological methods discussed above focus on the central nervous system. Considerable research has also focused on the **peripheral nervous system**. These methods include skin conductance, cardiovascular responses, muscle activity, pupil diameter, eye blinks, and eye movements. Skin conductance, for example, measures the electrical conductance (the inverse of resistance) between two points on the skin, which varies with the level of moisture. Sweat glands are responsible for this moisture and are controlled by the **sympathetic nervous system (SNS)**. Increases in skin conductance can be associated with changes in psychological activity. For example, studying skin conductance allows a researcher to investigate whether psychopaths react to fearful pictures in a normal way. Skin conductance provides relatively poor temporal resolution, with the entire response typically taking several seconds to emerge and resolve. However, it is an easy way to measure SNS response to a variety of stimuli.

Cardiovascular measures include heart rate, heart rate variability, and blood pressure. The heart is innervated

by the **parasympathetic nervous system (PNS)** and SNS. Input from the PNS decreases heart rate and contractile strength, whereas input from the SNS increases heart rate and contractile strength. Heart rate can easily be monitored using a minimum of two electrodes and is measured by counting the number of heartbeats in a given time period, such as one minute, or by assessing the time between successive heartbeats. Psychological activity can prompt increases and decreases in heart rate, often in less than a second, making heart rate a sensitive measure of cognition. Measures of heart rate variability are concerned with consistency in the time interval between heartbeats. Changes in heart rate variability are associated with stress as well as psychiatric conditions. Figure 3 is an example of an electrocardiogram, which is used to measure heart rate and heart rate variability. These cardiovascular measures allow researchers to monitor SNS and PNS reactivity to various stimuli or situations. For example, when an arachnophobe views pictures of spiders, does their heart rate increase more than that of a person not afraid of spiders?

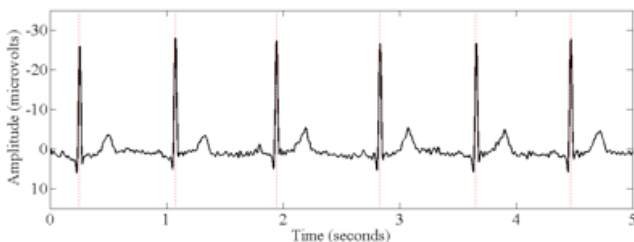


Figure 3. Example of electrocardiogram. The number of strong negative spikes in the output during a given period of time represents the heart rate, whereas the difference in the spacing between those strong negative spikes represents the heart rate variability.

Electromyography (EMG) measures electrical activity produced by skeletal muscles. Similar to EEG, EMG measures the voltage between two points. This technique can be used to determine when a participant first initiates muscle activity to engage in a motor response to a stimulus or the degree to which a participant begins to engage in an incorrect response (such as pressing the wrong button), even if it is never visibly executed. It has also been used in emotion research to identify activity in muscles that are used to produce smiles and frowns. Using EMG, it is possible to detect very small facial movements that are not observable from looking at the face. The temporal resolution of EMG is similar to that of EEG and MEG.

Valuable information can also be gleaned from eye blinks, eye movements, and pupil diameter. Eye blinks are most often assessed using EMG electrodes placed just below the eyelid, but electrical activity associated directly with eye blinks or eye movements can be measured with electrodes placed on the face near the eyes, because there is voltage across the entire eyeball. Another option for the measurement of eye movement is a camera used to record video of an eye. This video method is particularly valuable when determination of absolute direction of gaze (not just change in direction of gaze) is of interest, such as when the eyes scan a picture. With the help of a calibration period in which a participant looks at multiple, known targets, eye position is then extracted from each video frame during the main task and compared with data from the calibration phase, allowing researchers to identify the sequence, direction, and duration of gaze fixations. For example, when viewing pleasant or unpleasant images, people spend different

amounts of time looking at the most arousing parts. This, in turn, can vary as a function of psychopathology. Additionally, the diameter of a participant's pupil can be measured and recorded over time from the video record. As with heart rate, pupil diameter is controlled by competing inputs from the SNS and PNS. Pupil diameter is commonly used as an index of mental effort when performing a task.

WHEN TO USE WHAT

As the reader, you may be wondering, how do I know what tool is right for a given question? Generally, there are no definitive answers. If you wanted to know the temperature in the morning, would you check your phone? Look outside to see how warm it looks? Ask your roommate what he or she is wearing today? Look to see what other people are wearing? There is not a single way to answer the question. The same is true for research questions. However, there are some guidelines that one can consider. For example, if you are interested in what brain structures are associated with cognitive control, you wouldn't use peripheral nervous system measures. A technique such as fMRI or PET might be more appropriate. If you are interested in how cognitive control unfolds over time, EEG or MEG would be a good choice. If you are interested in studying the bodily response to fear in different groups of people, peripheral nervous system measures might be most appropriate. The key to deciding what method is most appropriate is properly defining the question that you are trying to answer. What aspects are most interesting? Do you care about identifying the most relevant brain structures?

Temporal dynamics? Bodily responses? Then, it is important to think about the strengths and weaknesses of the different psychophysiological measures and pick one, or several, whose attributes work best for the question at hand. In fact, it is common to record several at once.

CONCLUSION

The outline of psychophysiological methods above provides a glimpse into the exciting techniques that are available to researchers studying a broad range of topics from clinical to social to cognitive psychology. Some of the most interesting psychophysiological studies use several methods, such as in sleep assessments or multimodal neuroimaging. Psychophysiological methods have applications outside of mainstream psychology in areas where psychological phenomena are central, such as economics, health-related decision making, and brain-computer interfaces. Examples of applications for each method are provided above, but this list is by no means exhaustive. Furthermore, the field is continually evolving, with new methods and new applications being developed. The wide variety of methods and applications provide virtually limitless possibilities for researchers.

OUTSIDE RESOURCES

Book: Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.

Book: Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. New York: Cambridge University Press.

Web: For visualizations on MRI physics (requires a free registration): <https://www.imaio.com/en/e-mri/nmr>.

DISCUSSION QUESTIONS

1. Pick a psychological phenomenon that you would like to know more about. What specific hypothesis would you like to test? What psychophysiological methods might be appropriate for testing this hypothesis and why?
2. What types of questions would require high spatial resolution in measuring brain activity? What types of questions would require high temporal resolution?
3. Take the hypothesis you picked in the first question, and choose what you think would be the best psychophysiological method. What additional information could you obtain using a complementary method? For example, if you want to learn about memory, what two methods could you use that would each provide you with distinct information?
4. The popular press has shown an increasing interest in findings that contain images of brains and neuroscience language. Studies have shown that people often find presentations of results that contain these features more convincing than presentations of results that do not, even if the actual results are the same. Why would images of the brain and neuroscience language be more convincing to people? Given that results with

these features are more convincing, what do you think is the researcher's responsibility in reporting results with brain images and neuroscience language?

5. Many claims in the popular press attempt to reduce complex psychological phenomena to biological events. For example, you may have heard it said that schizophrenia is a brain disorder or that depression is simply a chemical imbalance. However, this type of "reductionism" so far does not appear to be tenable. There has been surprisingly little discussion of possible causal relationships, in either direction, between biological and psychological phenomena. We are aware of no such documented causal mechanisms. Do you think that it will ever be possible to explain how a change in biology can result in a change of a psychological phenomenon, or vice versa?

VOCABULARY

Blood-oxygen-level-dependent (BOLD)

The signal typically measured in fMRI that results from changes in the ratio of oxygenated hemoglobin to deoxygenated hemoglobin in the blood.

Central nervous system

The part of the nervous system that consists of the brain and spinal cord.

Deoxygenated hemoglobin

Hemoglobin not carrying oxygen.

Depolarization

A change in a cell's membrane potential, making the inside of the cell more positive and increasing the chance of an action potential.

Hemoglobin

The oxygen-carrying portion of a red blood cell.

Hyperpolarization

A change in a cell's membrane potential, making the inside of the cell more negative and decreasing the chance of an action potential.

Invasive Procedure

A procedure that involves the skin being broken or an instrument or chemical being introduced into a body cavity.

Lesions

Abnormalities in the tissue of an organism usually caused by disease or trauma.

Neural plasticity

The ability of synapses and neural pathways to change over time and adapt to changes in neural process, behavior, or environment.

Neuroscience methods

A research method that deals with the structure or function of the nervous system and brain.

Noninvasive procedure

A procedure that does not require the insertion of an instrument or chemical through the skin or into a body cavity.

Oxygenated hemoglobin

Hemoglobin carrying oxygen.

Parasympathetic nervous system (PNS)

One of the two major divisions of the autonomic nervous system, responsible for stimulation of “rest and digest” activities.

Peripheral nervous system

The part of the nervous system that is outside the brain and spinal cord.

Positron

A particle having the same mass and numerically equal but positive charge as an electron.

Psychophysiological methods

Any research method in which the dependent variable is a physiological measure and the independent variable is behavioral or mental (such as memory).

Spatial resolution

The degree to which one can separate a single object in space from another.

Sympathetic nervous system (SNS)

One of the two major divisions of the autonomic nervous system, responsible for stimulation of “fight or flight” activities.

Temporal resolution

The degree to which one can separate a single point in time from another.

Voltage

The difference in electric charge between two points.

AUTHORS

Zach Infantolino

Zach Infantolino, Clinical Psychology Graduate Student at the University of Delaware, studies the neural mechanisms associated with motivation and executive functions and how they are impacted by anxiety and depression. He is involved in treating individuals suffering from anxiety and depression, using his research and clinical work to inform one another.



Gregory A. Miller

Gregory A. Miller is Professor and Chair of the Department of Psychology at the University of Delaware and Professor Emeritus at the University of Illinois at Urbana-Champaign. He has served as Director of the Biomedical Imaging Center at Illinois and as President of the Society for Psychophysiological Research.

CREATIVE COMMONS LICENSE



PSYCHOPHYSIOLOGICAL METHODS IN NEUROSCIENCE by [Zachary Infantolino and Gregory A. Miller](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

CHAPTER 2.

THE NATURE-NURTURE QUESTION

ERIC TURKHEIMER

People have a deep intuition about what has been called the “nature–nurture question.” Some aspects of our behavior feel as though they originate in our genetic makeup, while others feel like the result of our upbringing or our own hard work. The scientific field of behavior genetics attempts to study these differences empirically, either by examining similarities among family members with different degrees of genetic relatedness, or, more recently, by studying differences in the DNA of people with different behavioral traits. The scientific methods that have been developed are ingenious, but often inconclusive. Many of the difficulties encountered in the empirical science of behavior genetics turn out to be conceptual, and our intuitions about nature and nurture get more complicated the harder we think about them. In the end, it is an oversimplification to ask how “genetic” some particular behavior is. Genes and environments always combine to produce behavior, and the real science is in the discovery of how they combine for a given behavior.

LEARNING OBJECTIVES

- Understand what the nature–nurture debate is and why the problem fascinates us.
- Understand why nature–nurture questions are difficult to study empirically.
- Know the major research designs that can be used to study nature–nurture questions.
- Appreciate the complexities of nature–nurture and why questions that seem simple turn out not to have simple answers.

INTRODUCTION

There are three related problems at the intersection of philosophy and science that are fundamental to our understanding of our relationship to the natural world: the mind–body problem, the free will problem, and the nature–nurture problem. These great questions have a lot in common. Everyone, even those without much knowledge of science or philosophy, has opinions about the answers to these questions that come simply from observing the world we live in. Our feelings about our relationship with the physical and biological world often seem incomplete. We are in control of our actions in some ways, but at the mercy of our bodies in others; it feels obvious that our consciousness is some kind of creation of our physical brains, at the same time we sense that our awareness must go beyond just the physical. This incomplete knowledge of our relationship with nature leaves us fascinated and a little obsessed, like a cat that climbs into a paper bag and then out again, over and over,

mystified every time by a relationship between inner and outer that it can see but can't quite understand.

It may seem obvious that we are born with certain characteristics while others are acquired, and yet of the three great questions about humans' relationship with the natural world, only nature–nurture gets referred to as a “debate.” In the history of psychology, no other question has caused so much controversy and offense: We are so concerned with nature–nurture because our very sense of moral character seems to depend on it. While we may admire the athletic skills of a great basketball player, we think of his height as simply a gift, a payoff in the “genetic lottery.” For the same reason, no one blames a short person for his height or someone's congenital disability on poor decisions: To state the obvious, it's “not their fault.” But we do praise the concert violinist (and perhaps her parents and teachers as well) for her dedication, just as we condemn cheaters, slackers, and bullies for their bad behavior.

The problem is, most human characteristics aren't usually as clear-cut as height or instrument-mastery, affirming our nature–nurture expectations strongly one way or the other. In fact, even the great violinist might have some inborn qualities—perfect pitch, or long, nimble fingers—that support and reward her hard work. And the basketball player might have eaten a diet while growing up that promoted his genetic tendency for being tall. When we think about our own qualities, they seem under our control in some respects, yet beyond our control in others. And often the traits that don't seem to have an obvious cause are the ones that concern us the most and are far more personally significant. What about how much we drink or worry? What about our honesty,

or religiosity, or sexual orientation? They all come from that uncertain zone, neither fixed by nature nor totally under our own control.



Researchers have learned a great deal about the nature-nurture dynamic by working with animals. But of course many of the techniques used to study animals cannot be applied to people. Separating these two influences in human subjects is a greater research challenge. [Image: Sebastián Dario, <https://goo.gl/OPiIWd>, CC BY-NC 2.0, <https://goo.gl/Filc2e>]

One major problem with answering nature-nurture questions about people is, how do you set up an experiment? In nonhuman animals, there are relatively straightforward experiments for tackling nature–nurture questions. Say, for example, you are interested in aggressiveness in dogs. You want to test for the more important determinant of aggression: being born to aggressive dogs or being raised by them. You could

mate two aggressive dogs—angry Chihuahuas—together, and mate two nonaggressive dogs—happy beagles—together, then switch half the puppies from each litter between the different sets of parents to raise. You would then have puppies born to aggressive parents (the Chihuahuas) but being raised by nonaggressive parents (the Beagles), and vice versa, in litters that mirror each other in puppy distribution. The big questions are: Would the Chihuahua parents raise aggressive beagle puppies? Would the beagle parents raise *nonaggressive* Chihuahua puppies? Would the puppies' *nature* win out, regardless of

who raised them? Or... would the result be a combination of nature *and* nurture? Much of the most significant nature–nurture research has been done in this way ([Scott & Fuller, 1998](#)), and animal breeders have been doing it successfully for thousands of years. In fact, it is fairly easy to breed animals for behavioral traits.

With people, however, we can't assign babies to parents at random, or select parents with certain behavioral characteristics to mate, merely in the interest of science (though history does include horrific examples of such practices, in misguided attempts at “eugenics,” the shaping of human characteristics through intentional breeding). In typical human families, children's biological parents raise them, so it is very difficult to know whether children act like their parents due to genetic (nature) or environmental (nurture) reasons. Nevertheless, despite our restrictions on setting up human-based experiments, we do see real-world examples of nature-nurture at work in the human sphere—though they only provide partial answers to our many questions.

The science of how genes and environments work together to influence behavior is called **behavioral genetics**. The easiest opportunity we have to observe this is the **adoption study**. When children are put up for adoption, the parents who give birth to them are no longer the parents who raise them. This setup isn't quite the same as the experiments with dogs (children aren't assigned to random adoptive parents in order to suit the particular interests of a scientist) but adoption still tells us some interesting things, or at least confirms some basic expectations. For instance, if the biological child of tall parents were adopted into a family of short people, do you suppose the child's growth would be affected? What

about the biological child of a Spanish-speaking family adopted at birth into an English-speaking family? What language would you expect the child to speak? And what might these outcomes tell you about the difference between height and language in terms of nature-nurture?

Another option for observing nature-nurture in humans involves **twin studies**. There are two types of twins: monozygotic (MZ) and dizygotic (DZ). Monozygotic twins, also called “identical” twins, result from a single zygote (fertilized egg) and have the same DNA. They are essentially clones. Dizygotic twins, also known as “fraternal” twins, develop from two



Studies focused on twins have led to important insights about the biological origins of many personality characteristics.

zygotes and share 50% of their DNA. Fraternal twins are ordinary siblings who happen to have been born at the same time. To analyze nature–nurture using twins, we compare the similarity of MZ and DZ pairs. Sticking with the features of height and spoken language, let’s take a look at how nature and nurture apply: Identical twins, unsurprisingly, are almost perfectly similar for height. The heights of fraternal twins, however, are like any other sibling pairs: more similar to each other than to people from other families, but hardly identical. This contrast between twin types gives us a clue about the role genetics plays in determining height. Now consider spoken

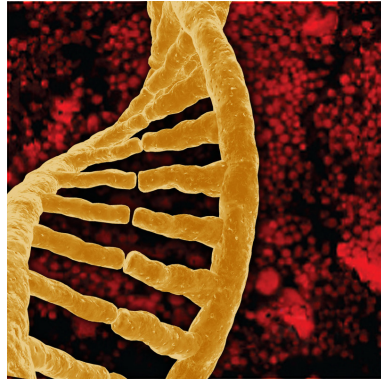
language. If one identical twin speaks Spanish at home, the co-twin with whom she is raised almost certainly does too. But the same would be true for a pair of fraternal twins raised together. In terms of spoken language, fraternal twins are just as similar as identical twins, so it appears that the genetic match of identical twins doesn't make much difference.

Twin and adoption studies are two instances of a much broader class of methods for observing nature-nurture called **quantitative genetics**, the scientific discipline in which similarities among individuals are analyzed based on how biologically related they are. We can do these studies with siblings and half-siblings, cousins, twins who have been separated at birth and raised separately ([Bouchard, Lykken, McGue, & Segal, 1990](#); such twins are very rare and play a smaller role than is commonly believed in the science of nature–nurture), or with entire extended families (see [Plomin, DeFries, Knopik, & Neiderhiser, 2012](#), for a complete introduction to research methods relevant to nature–nurture).

For better or for worse, contentions about nature–nurture have intensified because quantitative genetics produces a number called a **heritability coefficient**, varying from 0 to 1, that is meant to provide a single measure of genetics' influence of a trait. In a general way, a heritability coefficient measures how strongly differences among individuals are related to differences among their genes. But beware: Heritability coefficients, although simple to compute, are deceptively difficult to interpret. Nevertheless, numbers that provide simple answers to complicated questions tend to have a strong influence on the human imagination, and a great deal of time has been spent discussing whether the

heritability of intelligence or personality or depression is equal to one number or another.

One reason nature–nurture continues to fascinate us so much is that we live in an era of great scientific discovery in genetics, comparable to the times of Copernicus, Galileo, and Newton, with regard to astronomy and physics. Every day, it seems, new discoveries are made, new possibilities proposed. When Francis Galton first started thinking about nature–nurture in the late-19th century he was very influenced by his cousin, Charles Darwin, but genetics *per se* was



Quantitative genetics uses statistical methods to study the effects that both heredity and environment have on test subjects. These methods have provided us with the heritability coefficient which measures how strongly differences among individuals for a trait are related to differences among their genes. [Image: EMSL, <https://goo.gl/IRfn9g>, CC BY-NC-SA 2.0, <https://goo.gl/fbv27n>]

unknown. Mendel's famous work with peas, conducted at about the same time, went undiscovered for 20 years; quantitative genetics was developed in the 1920s; DNA was discovered by Watson and Crick in the 1950s; the human genome was completely sequenced at the turn of the 21st century; and we are now on the verge of being able to obtain the specific DNA sequence of anyone at a relatively low cost. No one knows what this new genetic knowledge will mean for the study of nature–nurture, but as we will see in the next section, answers to

nature–nurture questions have turned out to be far more difficult and mysterious than anyone imagined.

WHAT HAVE WE LEARNED ABOUT NATURE–NURTURE?

It would be satisfying to be able to say that nature–nurture studies have given us conclusive and complete evidence about where traits come from, with some traits clearly resulting from genetics and others almost entirely from environmental factors, such as childrearing practices and personal will; but that is not the case. Instead, *everything* has turned out to have some footing in genetics. The more genetically-related people are, the more similar they are—for *everything*: height, weight, intelligence, personality, mental illness, etc. Sure, it seems like common sense that some traits have a genetic bias. For example, adopted children resemble their biological parents even if they have never met them, and identical twins are more similar to each other than are fraternal twins. And while certain psychological traits, such as personality or mental illness (e.g., schizophrenia), seem reasonably influenced by genetics, it turns out that the same is true for political attitudes, how much television people watch ([Plomin, Corley, DeFries, & Fulker, 1990](#)), and whether or not they get divorced ([McGue & Lykken, 1992](#)).



Research over the last half century has revealed how central genetics are to behavior. The more genetically related people are the more similar they are not just physically but also in terms of personality and behavior. [Image: Paul Altobelli, <https://goo.gl/SWLwm2>, CC BY 2.0, <https://goo.gl/9uSnqN>]

It may seem surprising, but genetic influence on behavior is a relatively recent discovery. In the middle of the 20th century, psychology was dominated by the doctrine of behaviorism, which held that behavior could only be explained in terms of environmental factors. Psychiatry concentrated on psychoanalysis, which probed for roots of behavior in individuals' early life-histories. The truth is, neither behaviorism nor

psychoanalysis is incompatible with genetic influences on behavior, and neither Freud nor Skinner was naive about the importance of organic processes in behavior. Nevertheless, in their day it was widely thought that children's personalities were shaped entirely by imitating their parents' behavior, and that schizophrenia was caused by certain kinds of "pathological mothering." Whatever the outcome of our broader discussion of nature–nurture, the basic fact that the best predictors of an adopted child's personality or mental health are found in the biological parents he or she has never met, rather than in the adoptive parents who raised him or her, presents a significant challenge to purely environmental explanations of personality or psychopathology. The message is clear: You can't leave genes out of the equation.

But keep in mind, no behavioral traits are completely inherited, so you can't leave the environment out altogether, either.

Trying to untangle the various ways nature-nurture influences human behavior can be messy, and often common-sense notions can get in the way of good science. One very significant contribution of behavioral genetics that has changed psychology for good can be very helpful to keep in mind: When your subjects are biologically-related, no matter how clearly a situation may seem to point to environmental influence, it is never safe to interpret a behavior as wholly the result of nurture without further evidence. For example, when presented with data showing that children whose mothers read to them often are likely to have better reading scores in third grade, it is tempting to conclude that reading to your kids out loud is important to success in school; this may well be true, but the study as described is inconclusive, because there are genetic *as well as* environmental pathways between the parenting practices of mothers and the abilities of their children. This is a case where "correlation does not imply causation," as they say. To establish that reading aloud causes success, a scientist can either study the problem in adoptive families (in which the genetic pathway is absent) or by finding a way to randomly assign children to oral reading conditions.

The outcomes of nature-nurture studies have fallen short of our expectations (of establishing clear-cut bases for traits) in many ways. The most disappointing outcome has been the inability to organize traits from *more-* to *less-*genetic. As noted earlier, everything has turned out to be at least *somewhat* heritable (passed down), yet nothing has turned out to be *absolutely* heritable, and

there hasn't been much consistency as to which traits are *more* heritable and which are *less* heritable once other considerations (such as how accurately the trait can be measured) are taken into account ([Turkheimer, 2000](#)). The problem is conceptual: The heritability coefficient, and, in fact, the whole quantitative structure that underlies it, does not match up with our nature–nurture intuitions. We want to know how “important” the roles of genes and environment are to the development of a trait, but in focusing on “important” maybe we’re emphasizing the wrong thing. First of all, genes and environment are both crucial to *every* trait; without genes the environment would have nothing to work on, and too, genes cannot develop in a vacuum. Even more important, because nature–nurture questions look at the differences among people, the cause of a given trait depends not only on the trait itself, but also on the differences in that trait between members of the group being studied.

The classic example of the heritability coefficient defying intuition is the trait of having two arms. No one would argue against the development of arms being a biological, genetic process. But fraternal twins are just as similar for “two-armedness” as identical twins, resulting in a heritability coefficient of zero for the trait of having two arms. Normally, according to the heritability model, this result (coefficient of zero) would suggest all nurture, no nature, but we know that’s not the case. The reason this result is not a tip-off that arm development is less genetic than we imagine is because people *do not vary* in the genes related to arm development—which essentially upends the heritability formula. In fact, in this instance, the opposite is likely true: the extent that people differ in arm number is likely the result of accidents and,

therefore, environmental. For reasons like these, we always have to be very careful when asking nature–nurture questions, especially when we try to express the answer in terms of a single number. The heritability of a trait is not simply a property of that trait, but a property of the trait in a particular context of relevant genes and environmental factors.

Another issue with the heritability coefficient is that it divides traits' determinants into two portions—genes and environment—which are then calculated together for the total variability. This is a little like asking how much of the experience of a symphony comes from the horns and how much from the strings; the ways instruments or genes integrate is more complex than that. It turns out to be the case that, for many traits, genetic differences affect behavior under some environmental circumstances but not others—a phenomenon called gene-environment interaction, or $G \times E$. In one well-known example, Caspi et al. (2002) showed that among maltreated children, those who carried a particular allele of the MAOA gene showed a predisposition to violence and antisocial behavior, while those with other alleles did not. Whereas, in children who had not been maltreated, the gene had no effect. Making matters even more complicated are very recent studies of what is known as epigenetics (see module, “Epigenetics” <http://noba.to/37p5cb8v>), a process in which the DNA itself is modified by environmental events, and those genetic changes transmitted to children.

Some common questions about nature–nurture are, how susceptible is a trait to change, how malleable is it, and do we “have a choice” about it? These questions are much more complex than they may seem at first glance. For example, phenylketonuria is an inborn error of metabolism caused by a single gene; it prevents the body from metabolizing phenylalanine. Untreated, it causes intellectual disability and death. But it can be treated effectively by a straightforward environmental



The answer to the nature–nurture question has not turned out to be as straightforward as we would like. The many questions we can ask about the relationships among genes, environments, and human traits may have many different answers, and the answer to one tells us little about the answers to the others. [Image: Sundaram Ramaswamy, <https://goo.gl/Bv8lp6>, CC BY 2.0, <https://goo.gl/9uSnqN>]

intervention: avoiding foods containing phenylalanine. Height seems like a trait firmly rooted in our nature and unchangeable, but the average height of many populations in Asia and Europe has increased significantly in the past 100 years, due to changes in diet and the alleviation of poverty. Even the most modern genetics has not provided definitive answers to nature–nurture questions. When it was first becoming possible to measure the DNA sequences of individual people, it was widely thought that we would quickly progress to finding the specific genes that account for behavioral characteristics, but that hasn’t happened. There are a few rare genes that have been

found to have significant (almost always negative) effects, such as the single gene that causes Huntington's disease, or the Apolipoprotein gene that causes early onset dementia in a small percentage of Alzheimer's cases. Aside from these rare genes of great effect, however, the genetic impact on behavior is broken up over many genes, each with very small effects. For most behavioral traits, the effects are so small and distributed across so many genes that we have not been able to catalog them in a meaningful way. In fact, the same is true of environmental effects. We know that extreme environmental hardship causes catastrophic effects for many behavioral outcomes, but fortunately extreme environmental hardship is very rare. Within the normal range of environmental events, those responsible for differences (e.g., why some children in a suburban third-grade classroom perform better than others) are much more difficult to grasp.

The difficulties with finding clear-cut solutions to nature–nurture problems bring us back to the other great questions about our relationship with the natural world: the mind-body problem and free will. Investigations into what we mean when we say we are aware of something reveal that consciousness is not simply the product of a particular area of the brain, nor does choice turn out to be an orderly activity that we can apply to some behaviors but not others. So it is with nature and nurture: What at first may seem to be a straightforward matter, able to be indexed with a single number, becomes more and more complicated the closer we look. The many questions we can ask about the intersection among genes, environments, and human traits—how sensitive are traits to environmental change, and how common are those

influential environments; are parents or culture more relevant; how sensitive are traits to differences in genes, and how much do the relevant genes vary in a particular population; does the trait involve a single gene or a great many genes; is the trait more easily described in genetic or more-complex behavioral terms?—may have different answers, and the answer to one tells us little about the answers to the others.

It is tempting to predict that the more we understand the wide-ranging effects of genetic differences on all human characteristics—especially behavioral ones—our cultural, ethical, legal, and personal ways of thinking about ourselves will have to undergo profound changes in response. Perhaps criminal proceedings will consider genetic background. Parents, presented with the genetic sequence of their children, will be faced with difficult decisions about reproduction. These hopes or fears are often exaggerated. In some ways, our thinking may need to change—for example, when we consider the meaning behind the fundamental American principle that all men are created equal. Human beings differ, and like all evolved organisms they differ genetically. The Declaration of Independence predates Darwin and Mendel, but it is hard to imagine that Jefferson—whose genius encompassed botany as well as moral philosophy—would have been alarmed to learn about the genetic diversity of organisms. One of the most important things modern genetics has taught us is that almost all human behavior is too complex to be nailed down, even from the most complete genetic information, unless we're looking at identical twins. The science of nature and nurture has demonstrated that genetic differences among people are vital to human moral

equality, freedom, and self-determination, not opposed to them. As Mordecai Kaplan said about the role of the past in Jewish theology, genetics gets a vote, not a veto, in the determination of human behavior. We should indulge our fascination with nature–nurture while resisting the temptation to oversimplify it.

OUTSIDE RESOURCES

Web: Institute for Behavioral Genetics

<http://www.colorado.edu/ibg/>

DISCUSSION QUESTIONS

1. Is your personality more like one of your parents than the other? If you have a sibling, is his or her personality like yours? In your family, how did these similarities and differences develop? What do you think caused them?
2. Can you think of a human characteristic for which genetic differences would play almost no role? Defend your choice.
3. Do you think the time will come when we will be able to predict almost everything about someone by examining their DNA on the day they are born?
4. Identical twins are more similar than fraternal twins for the trait of aggressiveness, as well as for criminal behavior. Do these facts have implications for the courtroom? If it can be shown that a violent criminal had violent parents, should it make a difference in culpability or sentencing?

VOCABULARY

Adoption study

A behavior genetic research method that involves comparison of adopted children to their adoptive and biological parents.

Behavioral genetics

The empirical science of how genes and environments combine to generate behavior.

Heritability coefficient

An easily misinterpreted statistical construct that purports to measure the role of genetics in the explanation of differences among individuals.

Quantitative genetics

Scientific and mathematical methods for inferring genetic and environmental processes based on the degree of genetic and environmental similarity among organisms.

Twin studies

A behavior genetic research method that involves comparison of the similarity of identical (monozygotic; MZ) and fraternal (dizygotic; DZ) twins.

REFERENCES

- Bouchard, T. J., Lykken, D. T., McGue, M., & Segal, N. L. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science*, 250(4978), 223–228.

- Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., Taylor, A. & Poulton, R. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, 297(5582), 851–854.
- McGue, M., & Lykken, D. T. (1992). Genetic influence on risk of divorce. *Psychological Science*, 3(6), 368–373.
- Plomin, R., Corley, R., DeFries, J. C., & Fulker, D. W. (1990). Individual differences in television viewing in early childhood: Nature as well as nurture. *Psychological Science*, 1(6), 371–377.
- Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. M. (2012). *Behavioral genetics*. New York, NY: Worth Publishers.
- Scott, J. P., & Fuller, J. L. (1998). *Genetics and the social behavior of the dog*. Chicago, IL: University of Chicago Press.
- Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*, 9(5), 160–164.

AUTHORS



Eric Turkheimer

Eric Turkheimer is a Professor of Psychology at the University of Virginia. He was President of the Behavior Genetics Association in 2012-2013.

He won the James Shields Award for Twin Research in 2009. His work explores the possibilities and limitations of genetics research as a tool for understanding complex human behavior.

CREATIVE COMMONS LICENSE



THE NATURE-NURTURE QUESTION by [Eric Turkheimer](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

PART II.

**UNIT 2: NERVE CELLS AND
HOW THEY
COMMUNICATE**

CHAPTER 3.

NEURONS

SHARON FURTAK

This module on the biological basis of behavior provides an overview of the basic structure of neurons and their means of communication. Neurons, cells in the central nervous system, receive information from our sensory systems (vision, audition, olfaction, gustation, and somatosensation) about the world around us; in turn, they plan and execute appropriate behavioral responses, including attending to a stimulus, learning new information, speaking, eating, mating, and evaluating potential threats. The goal of this module is to become familiar with the anatomical structure of neurons and to understand how neurons communicate by electrochemical signals to process sensory information and produce complex behaviors through networks of neurons. Having a basic knowledge of the fundamental structure and function of neurons is a necessary foundation as you move forward in the field of psychology.

LEARNING OBJECTIVES

- Differentiate the functional roles between the two main cell classes in the brain, neurons and glia.
- Describe how the forces of diffusion and electrostatic pressure work collectively to facilitate electrochemical communication.
- Define resting membrane potential, excitatory postsynaptic potentials, inhibitory postsynaptic potentials, and action potentials.
- Explain features of axonal and synaptic communication in neurons.

INTRODUCTION

Imagine trying to string words together into a meaningful sentence without knowing the meaning of each word or its function (i.e., Is it a verb, a noun, or an adjective?). In a similar fashion, to appreciate how groups of cells work together in a meaningful way in the brain as a whole, we must first understand how individual cells in the brain function. Much like words, brain cells, called *neurons*, have an underlying structure that provides the foundation for their functional purpose. Have you ever seen a neuron? Did you know that the basic structure of a neuron is similar whether it is from the brain of a rat or a human? How do the billions of neurons in our brain allow us to do all the fun things we enjoy, such as texting a friend, cheering on our favorite sports team, or laughing?

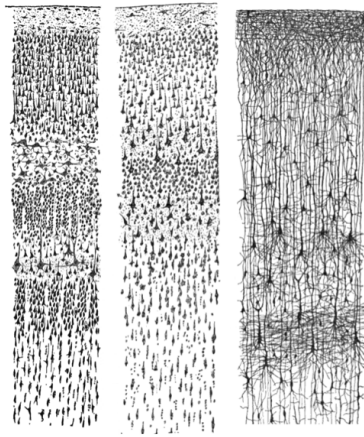


Figure 1. Three drawings by Santiago Ramón y Cajal, taken from "Comparative study of the sensory areas of the human cortex", pages 314, 361, and 363. Left: Nissl-stained visual cortex of a human adult. Middle: Nissl-stained motor cortex of a human adult. Right: Golgi-stained cortex of a 1 1/2 month old infant. [Image: Santiago Ramon y Cajal, <https://goo.gl/zOb2l1>, CCO Public Domain, <https://goo.gl/m25gce>]

Our journey in answering these questions begins more than 100 years ago with a scientist named Santiago Ramón y Cajal. [Ramón y Cajal \(1911\)](#) boldly concluded that discrete individual neurons are the structural and functional units of the nervous system. He based his conclusion on the numerous drawings he made of Golgi-stained tissue, a stain named after the scientist who discovered it, Camillo Golgi. Scientists use several types of stains to visualize cells. Each stain works in a unique way, which causes them to look

differently when viewed under a microscope. For example, a very common Nissl stain labels only the main part of the cell (i.e., the cell body; see left and middle panels of Figure 1). In contrast, a Golgi stain fills the cell body and all the processes that extend outward from it (see right panel of Figure 1). A more notable characteristic of a Golgi stain is that it only stains approximately 1–2% of neurons ([Pasternak & Woolsey, 1975](#); [Smit & Colon, 1969](#)), permitting the observer to distinguish one cell from another. These qualities allowed Cajal to examine the full anatomical structure of individual neurons for the

first time. This significantly enhanced our appreciation of the intricate networks their processes form. Based on his observation of Golgi-stained tissue, Cajal suggested neurons were distinguishable processing units rather than continuous structures. This was in opposition to the dominant theory at the time proposed by Joseph von Gerlach, which stated that the nervous system was composed of a continuous network of nerves (for review see, [Lopez-Munoz, Boya, & Alamo, 2006](#)). Camillo Golgi himself had been an avid supporter of Gerlach's theory. Despite their scientific disagreement, Cajal and Camillo Golgi shared the Nobel Prize for Medicine in 1906 for their combined contribution to the advancement of science and our understanding of the structure of the nervous system. This seminal work paved the pathway to our current understanding of the basic structure of the nervous system described in this module (for review see: [De Carlos & Borrell, 2007](#); [Grant, 2007](#)).

Before moving forward, there will be an introduction to some basic terminology regarding the anatomy of neurons in the section called "The Structure of the Neuron," below. Once we have reviewed this fundamental framework, the remainder of the module will focus on the electrochemical signals through which neurons communicate. While the electrochemical process might sound intimidating, it will be broken down into digestible sections. The first subsection, "Resting Membrane Potential," describes what occurs in a neuron at rest, when it is theoretically not receiving or sending signals. Building upon this knowledge, we will examine the electrical conductance that occurs within a single neuron when it receives signals. Finally, the module will conclude with a description of the electrical conductance, which

results in communication between neurons through a release of chemicals. At the end of the module, you should have a broad concept of how each cell and large groups of cells send and receive information by electrical and chemical signals.

A note of encouragement: This module introduces a vast amount of technical terminology that at times may feel overwhelming. Do not get discouraged or bogged down in the details. Utilize the glossary at the end of the module as a quick reference guide; tab the glossary page so that you can easily refer to it while reading the module. The glossary contains all terms in bold typing. Terms in italics are additional significant terms that may appear in other modules but are not contained within the glossary. On your first read of this module, I suggest focusing on the broader concepts and functional aspects of the terms instead of trying to commit all the terminology to memory. That is right, I said read first! I highly suggest reading this module at least twice, once prior to *and* again following the course lecture on this material. Repetition is the best way to gain clarity and commit to memory the challenging concepts and detailed vocabulary presented here.

THE STRUCTURE OF THE NEURON

BASIC NOMENCLATURE

There are approximately 100 billion neurons in the human brain ([Williams & Herrup, 1988](#)). Each neuron has three main components: dendrites, the soma, and the axon (see Figure 2). **Dendrites** are processes that extend outward from the **soma**, or cell body, of a neuron and typically branch several times. Dendrites receive

information from thousands of other neurons and are the main source of input of the neuron. The **nucleus**, which is located within the soma, contains genetic information, directs protein synthesis, and supplies the energy and the resources the neuron needs to function. The main source of output of the neuron is the **axon**. The axon is a process that extends far away from the soma and carries an important signal called an action potential to another neuron. The place at which the axon of one neuron comes in close contact to the dendrite of another neuron is a **synapse** (see Figures 2–3). Typically, the axon of a neuron is covered with an insulating substance called a **myelin sheath** that allows the signal and communication of one neuron to travel rapidly to another neuron.

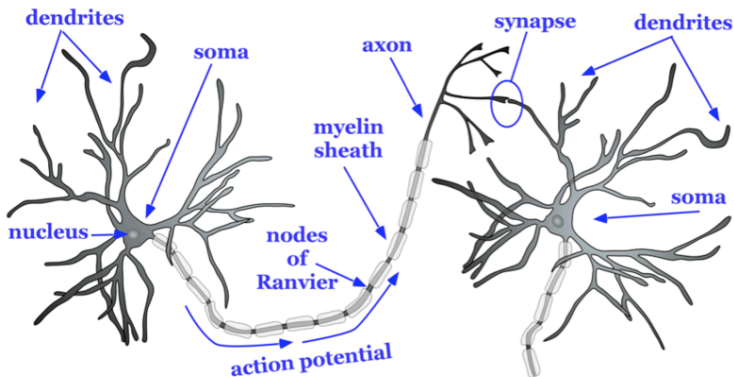


Figure 2. Basic structure of a neuron.

The axon splits many times, so that it can communicate, or synapse, with several other neurons (see Figure 2). At the end of the axon is a **terminal button**, which forms synapses with **spines**, or protrusions, on the dendrites of neurons. Synapses form between the *presynaptic* terminal

button (neuron sending the signal) and the *postsynaptic* membrane (neuron receiving the signal; see Figure 3). Here we will focus specifically on synapses between the terminal button of an axon and a dendritic spine; however, synapses can also form between the terminal button of an axon and the soma or the axon of another neuron.

A very small space called a **synaptic gap** or a synaptic cleft, approximately 5 nm (nanometers), exists between the presynaptic terminal button and the postsynaptic dendritic spine. To give you a better idea of the size, a dime is 1.35 mm (millimeter) thick. There are 1,350,000 nm in the thickness of a dime. In the presynaptic terminal button, there are **synaptic vesicles** that package together groups of chemicals called **neurotransmitters** (see Figure 3). Neurotransmitters are released from the presynaptic terminal button, travel across the synaptic gap, and activate ion channels on the postsynaptic spine by binding to *receptor sites*. We will discuss the role of receptors in more detail later in the module.

TYPES OF CELLS IN THE BRAIN

Not all neurons are created equal! There are neurons that help us receive information about the world around us, *sensory* neurons. There are *motor* neurons that allow us to initiate movement and behavior, ultimately allowing us to interact with the world around us. Finally, there are *interneurons*, which process the sensory input from our environment into meaningful representations, plan the appropriate behavioral response, and connect to the motor neurons to execute these behavioral plans.

There are three main categories of neurons, each

defined by its specific structure. The structures of these three different types of neurons support their unique functions. *Unipolar neurons* are structured in such a way that is ideal for relaying information forward, so they have one neurite (axon) and no dendrites. They are involved in transmission of physiological information from the body's periphery such as communicating body temperature through the spinal cord up to the brain. *Bipolar neurons* are involved in sensory perception such as perception of light in the retina of the eye. They have one axon and one dendrite which help acquire and pass sensory information to various centers in the brain. Finally, *multipolar neurons* are the most common and they communicate sensory and motor information in the brain. For example, their firing causes muscles in the body to contract. Multipolar neurons have one axon and many dendrites which allows them to communicate with other neurons. One of the most prominent neurons is a pyramidal neuron, which falls under the multipolar category. It gets its name from the triangular or pyramidal shape of its soma (for examples see, [Furtak, Moyer, & Brown, 2007](#)).

In addition to neurons, there is a second type of cell in the brain called *glia* cells. Glia cells have several functions, just a few of which we will discuss here. One type of glia cell, called *oligodendroglia*,

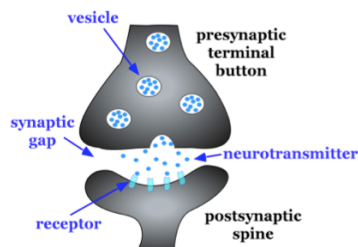


Figure 3. Characteristics of a synapse.

forms the myelin sheaths mentioned above ([Simons & Trotter, 2007](#); see Fig. 2). Oligodendroglia wrap their

dendritic processes around the axons of neurons many times to form the myelin sheath. One cell will form the myelin sheath on several axons. Other types of glia cells, such as *microglia* and *astrocytes*, digest debris of dead neurons, carry nutritional support from blood vessels to the neurons, and help to regulate the ionic composition of the extracellular fluid. While glial cells play a vital role in neuronal support, they do not participate in the communication between cells in the same fashion as neurons do.

COMMUNICATION WITHIN AND BETWEEN NEURONS

Thus far, we have described the main characteristics of neurons, including how their processes come in close contact with one another to form *synapses*. In this section, we consider the conduction of communication within a neuron and how this signal is transmitted to the next neuron. There are two stages of this electrochemical action in neurons. The first stage is the electrical conduction of dendritic input to the initiation of an action potential within a neuron. The second stage is a chemical transmission across the synaptic gap between the presynaptic neuron and the postsynaptic neuron of the synapse. To understand these processes, we first need to consider what occurs within a neuron when it is at a steady state, called *resting membrane potential*.

RESTING MEMBRANE POTENTIAL

The intracellular (inside the cell) fluid and extracellular (outside the cell) fluid of neurons is composed of a combination of ions (electrically charged molecules; see

Figure 4). Cations are positively charged ions, and anions are negatively charged ions. The composition of intracellular and extracellular fluid is similar to salt water, containing sodium (Na^+), potassium (K^+), chloride (Cl^-), and anions (A^-).

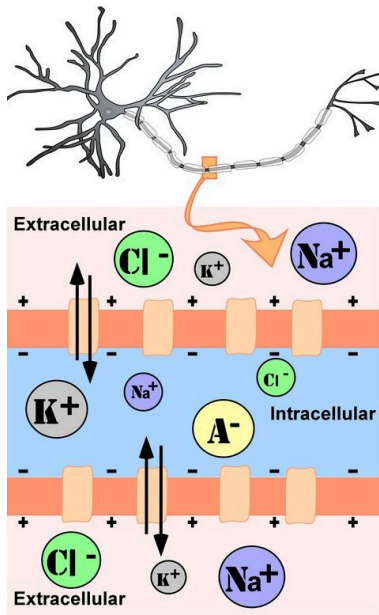


Figure 4. Representation of ion concentrations inside (intracellular) and outside (extracellular) a neuron in the unmyelinated segment of the axon.

The **cell membrane**, which is composed of a lipid bilayer of fat molecules, separates the cell from the surrounding extracellular fluid. There are proteins that span the membrane, forming **ion channels** that allow particular ions to pass between the intracellular and extracellular fluid (see Figure 4). These ions are in different concentrations inside the cell relative to outside the cell, and the ions have different electrical charges. Due to this difference in concentration and charge,

two forces act to maintain a steady state when the cell is at rest: diffusion and electrostatic pressure. **Diffusion** is the force on molecules to move from areas of high concentration to areas of low concentration. **Electrostatic pressure** is the force on two ions with similar charge to repel each other and the force

of two ions with opposite charge to attract to one another. Remember the saying, opposites attract?

Regardless of the ion, there exists a membrane potential at which the force of diffusion is equal and opposite of the force of electrostatic pressure. This voltage, called the *equilibrium potential*, is the voltage at which no ions flow. Since there are several ions that can permeate the cell's membrane, the baseline electrical charge inside the cell compared with outside the cell, referred to as **resting membrane potential**, is based on the collective drive of force on several ions. Relative to the extracellular fluid, the membrane potential of a neuron at rest is negatively charged at approximately -70 mV (see Figure 5). These are very small voltages compared with the voltages of batteries and electrical outlets, which we encounter daily, that range from 1.5 to 240 V.

Let us see how these two forces, diffusion and electrostatic pressure, act on the four groups of ions mentioned above.

1. *Anions (A-)*: Anions are highly concentrated inside the cell and contribute to the negative charge of the resting membrane potential. Diffusion and electrostatic pressure are not forces that determine A⁻ concentration because A⁻ is impermeable to the cell membrane. There are no ion channels that allow for A⁻ to move between the intracellular and extracellular fluid.
2. *Potassium (K+)*: The cell membrane is very permeable to potassium at rest, but potassium remains in high concentrations inside the cell.

Diffusion pushes K^+ outside the cell because it is in high concentration inside the cell. However, electrostatic pressure pushes K^+ inside the cell because the positive charge of K^+ is attracted to the negative charge inside the cell. In combination, these forces oppose one another with respect to K^+ .

3. *Chloride (Cl^-):* The cell membrane is also very permeable to chloride at rest, but chloride remains in high concentration outside the cell. Diffusion pushes Cl^- inside the cell because it is in high concentration outside the cell. However, electrostatic pressure pushes Cl^- outside the cell because the negative charge of Cl^- is attracted to the positive charge outside the cell. Similar to K^+ , these forces oppose one another with respect to Cl^- .
4. *Sodium (Na^+):* The cell membrane is not very permeable to sodium at rest. Diffusion pushes Na^+ inside the cell because it is in high concentration outside the cell. Electrostatic pressure also pushes Na^+ inside the cell because the positive charge of Na^+ is attracted to the negative charge inside the cell. Both of these forces push Na^+ inside the cell; however, Na^+ cannot permeate the cell membrane and remains in high concentration outside the cell. The small amounts of Na^+ inside the cell are removed by a **sodium-potassium pump**, which uses the neuron's energy (adenosine triphosphate, ATP) to pump 3 Na^+ ions out of the cell in exchange for bringing 2 K^+ ions inside the cell.

ACTION POTENTIAL

Now that we have considered what occurs in a neuron at rest, let us consider what changes occur to the resting membrane potential when a neuron receives input, or information, from the presynaptic terminal button of another neuron. Our understanding of the electrical signals or potentials that occurs within a neuron results from the seminal work of Hodgkin and Huxley that began in the 1930s at a well-known marine biology lab in Woodshole, MA. Their work, for which they won the Nobel Prize in Medicine in 1963, has resulted in the general model of electrochemical transduction that is described here ([Hodgkin & Huxley, 1952](#)). Hodgkin and Huxley studied a very large axon in the squid, a common species for that region of the United States. The giant axon of the squid is roughly 100 times larger than that of axons in the mammalian brain, making it much easier to see. Activation of the giant axon is responsible for a withdrawal response the squid uses when trying to escape from a predator, such as large fish, birds, sharks, and even humans. When was the last time you had calamari? The large axon size is no mistake in nature's design; it allows for very rapid transmission of an electrical signal, enabling a swift escape motion in the squid from its predators.

While studying this species, Hodgkin and Huxley noticed that if they applied an electrical stimulus to the axon, a large, transient electrical current conducted down the axon. This transient electrical current is known as an **action potential** (see Figure 5). An action potential is an all-or-nothing response that occurs when there is a change in the charge or potential of the cell from its

resting membrane potential (-70 mV) in a more positive direction, which is a *depolarization* (see Figure 5). What is meant by an all-or-nothing response? I find that this concept is best compared to the binary code used in computers, where there are only two possibilities, 0 or 1. There is no halfway or in-between these possible values; for example, 0.5 does not exist in binary code. There are only two possibilities, either the value of 0 or the value of 1. The action potential is the same in this respect. There is no halfway; it occurs, or it does not occur. There is a specific membrane potential that the neuron must reach to initiate an action potential. This membrane potential, called the **threshold of excitation**, is typically around -50 mV. If the threshold of excitation is reached, then an action potential is triggered.

How is an action potential initiated? At any one time, each neuron is receiving hundreds of inputs from the cells that synapse with it. These inputs can cause several types of fluctuations in the neuron's membrane potentials (see Figure 5):

1. **excitatory postsynaptic potentials (EPSPs):**

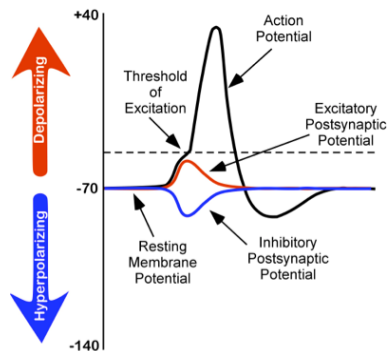


Figure 5. Changes in membrane potentials of neurons.

a *depolarizing* current that causes the membrane

potential to become more positive and closer to the threshold of excitation; or

2. **inhibitory postsynaptic potentials (IPSPs):**
a *hyperpolarizing* current that causes the membrane potential to become more negative and further away from the threshold of excitation.

These postsynaptic potentials, EPSPs and IPSPs, *summate* or add together in time and space. The IPSPs make the membrane potential more negative, but how much so depends on the strength of the IPSPs. The EPSPs make the membrane potential more positive; again, how much more positive depends on the strength of the EPSPs. If you have two small EPSPs at the same time and the same synapse then the result will be a large EPSP. If you have a small EPSP and a small IPSP at the same time and the same synapse then they will cancel each other out. Unlike the action potential, which is an all-or-nothing response, IPSPs and EPSPs are smaller and *graded* potentials, varying in strength. The change in voltage during an action potential is approximately 100 mV. In comparison, EPSPs and IPSPs are changes in voltage between 0.1 to 40 mV. They can be different strengths, or gradients, and they are measured by how far the membrane potentials diverge from the resting membrane potential.

I know the concept of summation can be confusing. As a child, I use to play a game in elementary school with a very large parachute where you would try to knock balls out of the center of the parachute. This game illustrates the properties of summation rather well. In this game, a group of children next to one another would work in unison to produce waves in the parachute in order to

cause a wave large enough to knock the ball out of the parachute. The children would initiate the waves at the same time and in the same direction. The additive result was a larger wave in the parachute, and the balls would bounce out of the parachute. However, if the waves they initiated occurred in the opposite direction or with the wrong timing, the waves would cancel each other out, and the balls would remain in the center of the parachute. EPSPs or IPSPs in a neuron work in the same fashion to the properties of the waves in the parachute; they either add or cancel each other out. If you have two EPSPs, then they sum together and become a larger depolarization. Similarly, if two IPSPs come into the cell at the same time, they will sum and become a larger hyperpolarization in membrane potential. However, if two inputs were opposing one another, moving the potential in opposite directions, such as an EPSP and an IPSP, their sum would cancel each other out.

At any moment in time, each cell is receiving mixed messages, both EPSPs and IPSPs. If the summation of EPSPs is strong enough to depolarize the membrane potential to reach the threshold of excitation, then it initiates an action potential. The action potential then travels down the axon, away from the soma, until it reaches the ends of the axon (the terminal button). In the terminal button, the action potential triggers the release of neurotransmitters from the presynaptic terminal button into the synaptic gap. These neurotransmitters, in turn, cause EPSPs and IPSPs in the postsynaptic dendritic spines of the next cell (see Figures 4 & 6). The neurotransmitter released from the presynaptic terminal button binds with **ionotropic receptors** in a lock-and-key fashion on the post-synaptic dendritic spine.

Ionotropic receptors are receptors on ion channels that open, allowing some ions to enter or exit the cell, depending upon the presence of a particular neurotransmitter. The type of neurotransmitter and the permeability of the ion channel it activates will determine if an EPSP or IPSP occurs in the dendrite of the post-synaptic cell. These EPSPs and IPSPs summate in the same fashion described above and the entire process occurs again in another cell.

THE CHANGE IN MEMBRANE POTENTIAL DURING AN ACTION POTENTIAL

We discussed previously which ions are involved in maintaining the resting membrane potential. Not surprisingly, some of these same ions are involved in the action potential. When the cell becomes depolarized (more positively charged) and reaches the threshold of excitation, this causes a voltage-dependent Na^+ channel to open. A voltage-dependent ion channel is a channel that opens, allowing some ions to enter or exit the cell, depending upon when the cell reaches a particular membrane potential. When the cell is at resting membrane potential, these voltage-dependent Na^+ channels are closed. As we learned earlier, both diffusion and electrostatic pressure are pushing Na^+ inside the cells. However, Na^+ cannot permeate the membrane when the cell is at rest. Now that these channels are open, Na^+ rushes inside the cell, causing the cell to become very positively charged relative to the outside of the cell. This is responsible for the rising or depolarizing phase of the action potential (see Figure 5). The inside of the cell becomes very positively charged, +40mV. At this point,

the Na^+ channels close and become *refractory*. This means the Na^+ channels cannot reopen again until after the cell returns to the resting membrane potential. Thus, a new action potential cannot occur during the refractory period. The refractory period also ensures the action potential can only move in one direction down the axon, away from the soma. As the cell becomes more depolarized, a second type of voltage-dependent channel opens; this channel is permeable to K^+ . With the cell very positive relative to the outside of the cell (depolarized) and the high concentration of K^+ within the cell, both the force of diffusion and the force of electrostatic pressure drive K^+ outside of the cell. The movement of K^+ out of the cell causes the cell potential to return back to the resting membrane potential, the falling or hyperpolarizing phase of the action potential (see Figure 5). A short hyperpolarization occurs partially due to the gradual closing of the K^+ channels. With the Na^+ closed, electrostatic pressure continues to push K^+ out of the cell. In addition, the sodium-potassium pump is pushing Na^+ out of the cell. The cell returns to the resting membrane potential, and the excess extracellular K^+ diffuses away. This exchange of Na^+ and K^+ ions happens very rapidly, in less than 1 msec. The action potential occurs in a wave-like motion down the axon until it reaches the terminal button. Only the ion channels in very close proximity to the action potential are affected.

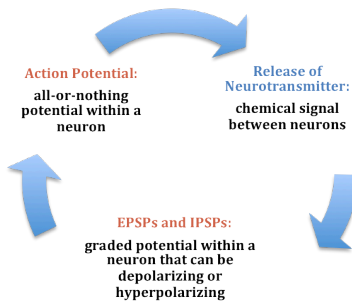


Figure 6. Summary of the electrochemical communication within and between neurons.

Earlier you learned that axons are covered in myelin. Let us consider how myelin speeds up the process of the action potential. There are gaps in the myelin sheaths called *nodes of Ranvier*. The myelin insulates the axon and does not allow any fluid to exist between the myelin and cell

membrane. Under the myelin, when the Na^+ and K^+ channels open, no ions flow between the intracellular and extracellular fluid. This saves the cell from having to expend the energy necessary to rectify or regain the resting membrane potential. (Remember, the pumps need ATP to run.) Under the myelin, the action potential degrades some, but is still large enough in potential to trigger a new action potential at the next node of Ranvier. Thus, the action potential actively jumps from node to node; this process is known as *saltatory conduction*.

In the presynaptic terminal button, the action potential triggers the release of neurotransmitters (see Figure 3). Neurotransmitters cross the synaptic gap and open subtypes of receptors in a lock-and-key fashion (see Figure 3). Depending on the type of neurotransmitter, an EPSP or IPSP occurs in the dendrite of the post-synaptic cell. Neurotransmitters that open Na^+ or calcium (Ca^+) channels cause an EPSP; an example is the NMDA receptors, which are activated by glutamate (the main excitatory neurotransmitter in the brain). In contrast, neurotransmitters that open Cl^- or K^+ channels cause an

IPSP; an example is gamma-aminobutyric acid (GABA) receptors, which are activated by GABA, the main inhibitory neurotransmitter in the brain. Once the EPSPs and IPSPs occur in the postsynaptic site, the process of communication within and between neurons cycles on (see Figure 6). A neurotransmitter that does not bind to receptors is broken down and inactivated by enzymes or glial cells, or it is taken back into the presynaptic terminal button in a process called *reuptake*, which will be discussed further in the module on psychopharmacology.

OUTSIDE RESOURCES

Video Series: Neurobiology/Biopsychology – Tutorial animations of action potentials, resting membrane potentials, and synaptic transmission.

<http://www.sumanasinc.com/webcontent/animations/neurobiology.html>

Video: An animation and an explanation of an action potential



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=55#oembed-1>

Video: An animation of neurotransmitter actions at the synapse



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=55#oembed-2>

Video: An interactive animation that allows students to observe the results of manipulations to excitatory and inhibitory post-synaptic potentials. Also includes animations and explanations of transmission and neural circuits.

<https://apps.childrenshospital.org/clinical/animation/neuron/>

Video: Another animation of an action potential



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=55#oembed-3>

Video: Another animation of neurotransmitter actions at the synapse



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=55#oembed-4>

Video: Domino Action Potential: This hands-on activity helps students grasp the complex process of the action potential, as well as become familiar with the characteristics of transmission (e.g., all-or-none response, refractory period).



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=55#oembed-5>

Video: For perspective on techniques in neuroscience to look inside the brain





One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=55#oembed-6>

Video: The Behaving Brain is the third program in the DISCOVERING PSYCHOLOGY series. This program looks at the structure and composition of the human brain: how neurons function, how information is collected and transmitted, and how chemical reactions relate to thought and behavior.

Video: You can grow new brain cells. Here's how. -Can we, as adults, grow new neurons? Neuroscientist Sandrine Thuret says that we can, and she offers research and practical advice on how we can help our brains better perform neurogenesis—improving mood, increasing memory formation and preventing the decline associated with aging along the way.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=55#oembed-7>

Web: For more information on the Nobel Prize shared

by Ramón y Cajal and Golgi

http://www.nobelprize.org/nobel_prizes/medicine/laureates/1906/

DISCUSSION QUESTIONS

1. What structures of a neuron are the main input and output of that neuron?
2. What does the statement mean that communication within and between cells is an electrochemical process?
3. How does myelin increase speed and efficiency of the action potential?
4. How does diffusion and electrostatic pressure contribute to the resting membrane potential and the action potential?
5. Describe the cycle of communication within and between neurons.

VOCABULARY

Action potential

A transient all-or-nothing electrical current that is conducted down the axon when the membrane potential reaches the threshold of excitation.

Axon

Part of the neuron that extends off the soma, splitting several times to connect with other neurons; main output of the neuron.

Cell membrane

A bi-lipid layer of molecules that separates the cell

from the surrounding extracellular fluid.

Dendrite

Part of a neuron that extends away from the cell body and is the main input to the neuron.

Diffusion

The force on molecules to move from areas of high concentration to areas of low concentration.

Electrostatic pressure

The force on two ions with similar charge to repel each other; the force of two ions with opposite charge to attract to one another.

Excitatory postsynaptic potentials

A depolarizing postsynaptic current that causes the membrane potential to become more positive and move towards the threshold of excitation.

Inhibitory postsynaptic potentials

A hyperpolarizing postsynaptic current that causes the membrane potential to become more negative and move away from the threshold of excitation.

Ion channels

Proteins that span the cell membrane, forming channels that specific ions can flow through between the intracellular and extracellular space.

Ionotropic receptor

Ion channel that opens to allow ions to permeate the cell membrane under specific conditions, such as the presence of a neurotransmitter or a specific membrane potential.

Myelin sheath

Substance around the axon of a neuron that serves as insulation to allow the action potential to conduct rapidly toward the terminal buttons.

Neurotransmitters

Chemical substance released by the presynaptic terminal button that acts on the postsynaptic cell.

Nucleus

Collection of nerve cells found in the brain which typically serve a specific function.

Resting membrane potential

The voltage inside the cell relative to the voltage outside the cell while the cell is at rest (approximately -70 mV).

Sodium-potassium pump

An ion channel that uses the neuron's energy (adenosine triphosphate, ATP) to pump three Na⁺ ions outside the cell in exchange for bringing two K⁺ ions inside the cell.

Soma

Cell body of a neuron that contains the nucleus and genetic information, and directs protein synthesis.

Spines

Protrusions on the dendrite of a neuron that form synapses with terminal buttons of the presynaptic axon.

Synapse

Junction between the presynaptic terminal button of

one neuron and the dendrite, axon, or soma of another postsynaptic neuron.

Synaptic gap

Also known as the synaptic cleft; the small space between the presynaptic terminal button and the postsynaptic dendritic spine, axon, or soma.

Synaptic vesicles

Groups of neurotransmitters packaged together and located within the terminal button.

Terminal button

The part of the end of the axon that form synapses with postsynaptic dendrite, axon, or soma.

Threshold of excitation

Specific membrane potential that the neuron must reach to initiate an action potential.

REFERENCES

- De Carlos, J. A., & Borrell, J. (2007). A historical reflection of the contributions of Cajal and Golgi to the foundations of neuroscience. *Brain Res Rev*, 55(1), 8-16. doi: 10.1016/j.brainresrev.2007.03.010
- Furtak, S. C., Moyer, J. R., Jr., & Brown, T. H. (2007). Morphology and ontogeny of rat perirhinal cortical neurons. *J Comp Neurol*, 505(5), 493-510. doi: 10.1002/cne.21516
- Grant, G. (2007). How the 1906 Nobel Prize in Physiology or Medicine was shared between Golgi and Cajal. *Brain Res Rev*, 55(2), 490-498. doi:

10.1016/j.brainresrev.2006.11.004

- Hodgkin, A. L., & Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol*, 117(4), 500-544.
- Lopez-Munoz, F., Boya, J., & Alamo, C. (2006). Neuron theory, the cornerstone of neuroscience, on the centenary of the Nobel Prize award to Santiago Ramon y Cajal. *Brain Res Bull*, 70(4-6), 391-405. doi: 10.1016/j.brainresbull.2006.07.010
- Pasternak, J. F., & Woolsey, T. A. (1975). On the “selectivity” of the Golgi-Cox method. *J Comp Neurol*, 160(3), 307-312. doi: 10.1002/cne.901600304
- Ramón y Cajal, S. (1911). Histology of the nervous system of man and vertebrates. New York, NY: Oxford University Press.
- Simons, M., & Trotter, J. (2007). Wrapping it up: the cell biology of myelination. *Curr Opin Neurobiol*, 17(5), 533-540. doi: 10.1016/j.conb.2007.08.003
- Smit, G. J., & Colon, E. J. (1969). Quantitative analysis of the cerebral cortex. I. Aselectivity of the Golgi-Cox staining technique. *Brain Res*, 13(3), 485-510.
- Williams, R. W., & Herrup, K. (1988). The control of neuron number. *Annu Rev Neurosci*, 11, 423-453. doi: 10.1146/annurev.ne.11.030188.002231

AUTHORS



Sharon Furtak

Sharon C. Furtak, Assistant Professor in the Department of Psychology at California State University, Sacramento. Dr. Furtak received her Ph.D. from Yale University and her B.A. from the University of Massachusetts, Amherst. Her research focuses on how we learn to fear items and how the environment in which we learn can influence our behavior.

CREATIVE COMMONS LICENSE



NEURONS by [Sharon Furtak](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

PART III.

**UNIT 3: THE NERVOUS
SYSTEM**

CHAPTER 4.

THE NERVOUS SYSTEM

ANEEQ AHMAD

The mammalian nervous system is a complex biological organ, which enables many animals including humans to function in a coordinated fashion. The original design of this system is preserved across many animals through evolution; thus, adaptive physiological and behavioral functions are similar across many animal species. Comparative study of physiological functioning in the nervous systems of different animals lend insights to their behavior and their mental processing and make it easier for us to understand the human brain and behavior. In addition, studying the development of the nervous system in a growing human provides a wealth of information about the change in its form and behaviors that result from this change. The nervous system is divided into central and peripheral nervous systems, and the two heavily interact with one another. The peripheral nervous system controls volitional (somatic nervous system) and nonvolitional (autonomic nervous system) behaviors using cranial and spinal nerves. The central nervous system is divided into forebrain, midbrain, and hindbrain, and each division performs a variety of tasks;

for example, the cerebral cortex in the forebrain houses sensory, motor, and associative areas that gather sensory information, process information for perception and memory, and produce responses based on incoming and inherent information. To study the nervous system, a number of methods have evolved over time; these methods include examining brain lesions, microscopy, electrophysiology, electroencephalography, and many scanning technologies.

LEARNING OBJECTIVES

- Describe and understand the development of the nervous system.
- Learn and understand the two important parts of the nervous system.
- Explain the two systems in the peripheral nervous system and what you know about the different regions and areas of the central nervous system.
- Learn and describe different techniques of studying the nervous system. Understand which of these techniques are important for cognitive neuroscientists.
- Describe the reasons for studying different nervous systems in animals other than human beings. Explain what lessons we learn from the evolutionary history of this organ.

EVOLUTION OF THE NERVOUS SYSTEM

Many scientists and thinkers ([Cajal, 1937](#); [Crick & Koch, 1990](#); [Edelman, 2004](#)) believe that the human nervous system is the most complex machine known to man. Its

complexity points to one undeniable fact—that it has evolved slowly over time from simpler forms. Evolution of the nervous system is intriguing not because we can marvel at this complicated biological structure, but it is fascinating because it inherits a lineage of a long history of many less complex nervous systems (Figure 1), and it documents a record of adaptive behaviors observed in life forms other than humans. Thus, evolutionary study of the nervous system is important, and it is the first step in understanding its design, its workings, and its functional interface with the environment.

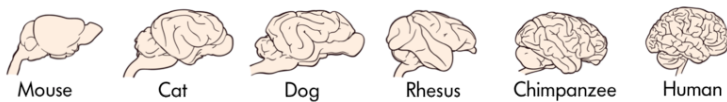


Figure 1. The brains of various animals.

The brains of some animals, like apes, monkeys, and rodents, are structurally similar to humans (Figure 1), while others are not (e.g., invertebrates, single-celled organisms). Does anatomical similarity of these brains suggest that behaviors that emerge in these species are also similar? Indeed, many animals display behaviors that are similar to humans, e.g., apes use nonverbal communication signals with their hands and arms that resemble nonverbal forms of communication in humans ([Gardner & Gardner, 1969](#); [Goodall, 1986](#); [Knapp & Hall, 2009](#)). If we study very simple behaviors, like physiological responses made by individual neurons, then brain-based behaviors of invertebrates ([Kandel & Schwartz, 1982](#)) look very similar to humans, suggesting that from time immemorial such basic behaviors have been conserved in the brains of many simple animal

forms and in fact are the foundation of more complex behaviors in animals that evolved later ([Bullock, 1984](#)).

Even at the micro-anatomical level, we note that individual neurons differ in complexity across animal species. Human neurons exhibit more intricate complexity than other animals; for example, neuronal processes (dendrites) in humans have many more branch points, branches, and spines.

Complexity in the structure of the nervous system, both at the macro- and micro-levels, give rise to complex behaviors. We can observe similar movements of the limbs, as in nonverbal communication, in apes and humans, but the variety and intricacy of nonverbal behaviors using hands in humans surpasses apes. Deaf individuals who use American Sign Language (ASL) express themselves in English nonverbally; they use this language with such fine gradation that many accents of ASL exist ([Walker, 1987](#)). Complexity of behavior with increasing complexity of the nervous system, especially the cerebral cortex, can be observed in the genus *Homo* (Figure 2). If we compare sophistication of material culture in **Homo habilis** (2 million years ago; brain volume ~650 cm³) and **Homo sapiens** (300,000 years to now; brain volume ~1400 cm³), the evidence shows that *Homo habilis* used crude stone tools compared with modern tools used by *Homo sapiens* to erect cities, develop written languages, embark on space travel, and study her own self. All of this is due to increasing complexity of the nervous system.

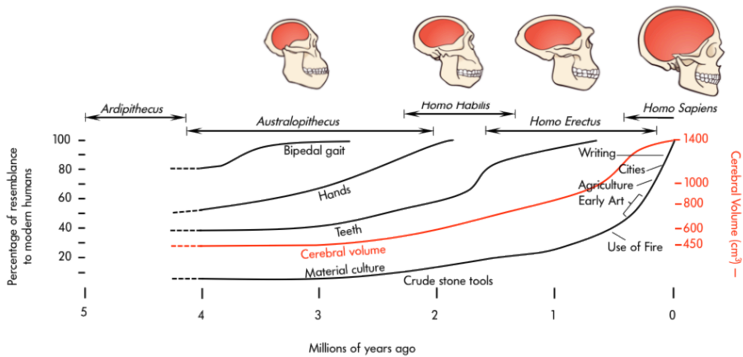


Figure 2. Changes in cerebral volume across evolution.

What has led to the complexity of the brain and nervous system through evolution, to its behavioral and cognitive refinement? Darwin ([1859](#), [1871](#)) proposed two forces of natural and sexual selection as work engines behind this change. He prophesied, “psychology will be based on a new foundation, that of the necessary acquirement of each mental power and capacity by gradation” that is, psychology will be based on evolution ([Rosenzweig, Breedlove, & Leiman, 2002](#)).

DEVELOPMENT OF THE NERVOUS SYSTEM

Where the study of change in the nervous system over eons is immensely captivating, studying the change in a single brain during individual development is no less engaging. In many ways the ontogeny (development) of the nervous system in an individual mimics the evolutionary advancement of this structure observed across many animal species. During development, the nervous tissue emerges from the **ectoderm** (one of the three layers of the mammalian embryo) through the process of **neural induction**. This process causes the formation of the neural tube, which extends in

a **rostrocaudal** (head-to-tail) plane. The tube, which is hollow, seams itself in the rostrocaudal direction. In some disease conditions, the neural tube does not close caudally and results in an abnormality called **spina bifida**. In this pathological condition, the lumbar and sacral segments of the spinal cord are disrupted.

As gestation progresses, the neural tube balloons up (cephalization) at the rostral end, and **forebrain**, midbrain, hindbrain, and the spinal cord can be visually delineated (day 40). About 50 days into gestation, six cephalic areas can be anatomically discerned (also see below for a more detailed description of these areas).

The progenitor cells (**neuroblasts**) that form the lining (**neuroepithelium**) of the neural tube generate all the neurons and glial cells of the central nervous system. During early stages of this development, neuroblasts rapidly divide and specialize into many varieties of neurons and glial cells, but this proliferation of cells is not uniform along the neural tube—that is why we see the forebrain and hindbrain expand into larger cephalic tissues than the midbrain. The neuroepithelium also generates a group of specialized cells that migrate outside the neural tube to form the **neural crest**. This structure gives rise to sensory and autonomic neurons in the peripheral nervous system.

THE STRUCTURE OF THE NERVOUS SYSTEM

The mammalian nervous system is divided into central and peripheral nervous systems.

THE PERIPHERAL NERVOUS SYSTEM

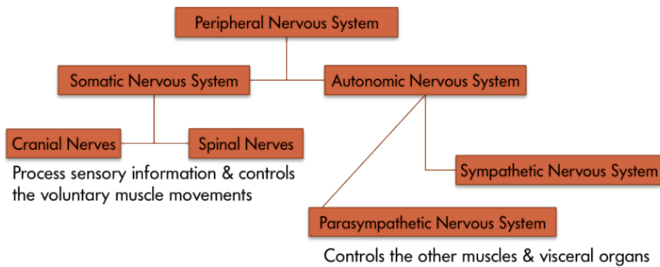


Figure 3. The various components of the peripheral nervous system.

The peripheral nervous system is divided into somatic and **autonomic nervous systems** (Figure 3). Where the **somatic nervous system** consists of cranial nerves (12 pairs) and spinal nerves (31 pairs) and is under the volitional control of the individual in maneuvering bodily muscles, the autonomic nervous system also running through these nerves lets the individual have little control over muscles and glands. Main divisions of the autonomic nervous system that control visceral structures are the sympathetic and **parasympathetic nervous systems**.

At an appropriate cue (say a fear-inducing object like a snake), the sympathetic division generally energizes many muscles (e.g., heart) and glands (e.g., adrenals), causing activity and release of hormones that lead the individual to negotiate the fear-causing snake with fight-or-flight responses. Whether the individual decides to *fight* the snake or *run* away from it, either action requires energy; in short, the **sympathetic nervous system** says “go, go, go.” The parasympathetic nervous system, on the other hand, curtails undue energy mobilization into muscles and glands and modulates the response by saying “stop,

stop, stop.” This push–pull tandem system regulates fight-or-flight responses in all of us.

THE CENTRAL NERVOUS SYSTEM

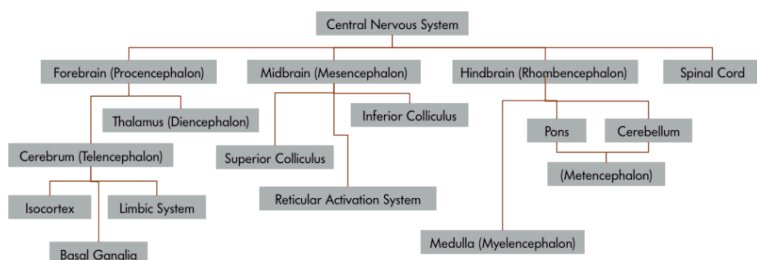


Figure 4. The central nervous system and its components.

The central nervous system is divided into a number of important parts (see Figure 4), including the spinal cord, each specialized to perform a set of specific functions. Telencephalon or **cerebrum** is a *newer* development in the evolution of the mammalian nervous system. In humans, it is about the size of a large napkin and when crumpled into the skull, it forms furrows called sulci (singular form, **sulcus**). The bulges between sulci are called gyri (singular form, **gyrus**). The cortex is divided into two hemispheres, and each hemisphere is further divided into four lobes (Figure 5a), which have specific functions. The division of these lobes is based on two delineating sulci: the **central sulcus** divides the hemisphere into frontal and parietal-occipital lobes and the **lateral sulcus** marks the **temporal lobe**, which lies below.

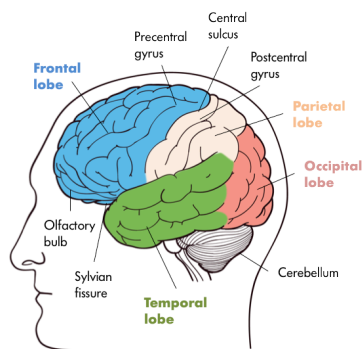


Figure 5a. The lobes of the brain.

Just in front of the central sulcus lies an area called the **primary motor cortex** (precentral gyrus), which connects to the muscles of the body, and on volitional command moves them. From mastication to movements in the genitalia, the body map is represented on this strip (Figure 5b).

Some body parts, like fingers, thumbs, and lips, occupy a greater representation on the strip than, say, the trunk. This disproportionate representation of the body on the primary motor cortex is called the **magnification factor** (Rolls & Cowey, 1970) and is seen in other motor and sensory areas. At the lower end of the central sulcus, close to the lateral sulcus, lies the **Broca's area** (Figure 6b) in the left **frontal lobe**, which is involved with language production. Damage to this part of the brain led Pierre Paul Broca, a French neuroscientist in 1861, to document many different forms of **aphasias**, in which his patients would lose the ability to speak or would retain partial speech impoverished in syntax and grammar (AAAS, 1880). It is no wonder that others have found subvocal rehearsal and central executive processes of **working memory** in this frontal lobe (Smith & Jonides, 1997, 1999).

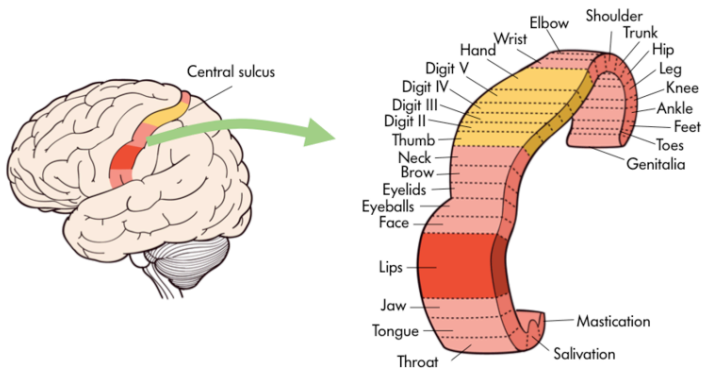


Figure 5b. Specific body parts like the tongue or fingers are mapped onto certain areas of the brain including the primary motor cortex.

Just behind the central gyrus, in the **parietal lobe**, lies the **primary somatosensory cortex** (Figure 6a) on the postcentral gyrus, which represents the whole body receiving inputs from the skin and muscles. The primary somatosensory cortex parallels, abuts, and connects heavily to the primary motor cortex and resembles it in terms of areas devoted to bodily representation. All spinal and some cranial nerves (e.g., the facial nerve) send sensory signals from skin (e.g., touch) and muscles to the primary somatosensory cortex. Close to the lower (ventral) end of this strip, curved inside the parietal lobe, is the taste area (secondary somatosensory cortex), which is involved with taste experiences that originate from the tongue, pharynx, epiglottis, and so forth.

Just below the parietal lobe, and under the caudal end of the lateral fissure, in the temporal lobe, lies the **Wernicke's**

area ([Demonet et al., 1992](#)). This area is involved with language comprehension and is connected to the Broca's area through the **arcuate fasciculus**, nerve fibers that connect these two

regions. Damage to the Wernicke's area (Figure 6b) results in many kinds of **agnosias**; agnosia is defined as an inability to know or understand language and speech-related behaviors. So an individual may show word deafness, which is an inability to recognize spoken language, or word blindness, which is an inability to recognize written or printed language. Close in proximity to the Wernicke's area is the primary auditory cortex, which is involved with audition, and finally the brain region devoted to smell (olfaction) is tucked away inside the primary olfactory cortex (prepyriform cortex).

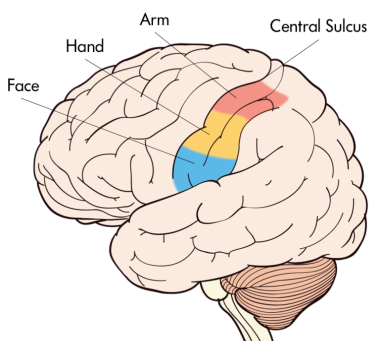


Figure 6a. The Primary Somatosensory Cortex.

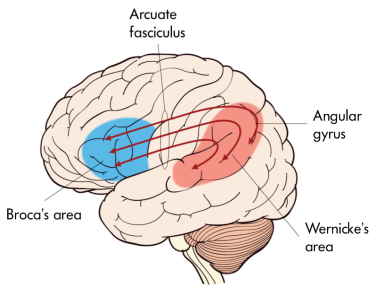


Figure 6b. Wernicke's area.

At the very back of the cerebral cortex lies the **occipital lobe** housing the primary visual cortex. Optic nerves travel all the way to the **thalamus (lateral geniculate nucleus, LGN)** and then to visual cortex, where images that are received on the retina are

projected ([Hubel, 1995](#)).

In the past 50 to 60 years, visual sense and visual pathways have been studied extensively, and our understanding about them has increased manifold. We now understand that all objects that form images on the retina are transformed (**transduction**) in neural language handed down to the visual cortex for further processing. In the visual cortex, all attributes (features) of the image, such as the color, texture, and orientation, are decomposed and processed by different visual cortical modules ([Van Essen, Anderson & Felleman, 1992](#)) and then recombined to give rise to singular perception of the image in question.

If we cut the cerebral hemispheres in the middle, a new set of structures come into view. Many of these perform different functions vital to our being. For example, the **limbic system** contains a number of nuclei that process memory (**hippocampus** and **fornix**) and attention and emotions (**cingulate gyrus**); the **globus pallidus** is involved with motor movements and their coordination; the hypothalamus and thalamus are involved with drives, motivations, and trafficking of

sensory and motor throughputs. The **hypothalamus** plays a key role in regulating endocrine hormones in conjunction with the pituitary gland that extends from the hypothalamus through a stalk (infundibulum).

As we descend down the thalamus, the midbrain comes into view with superior and inferior colliculi, which process visual and auditory information, as does the substantia nigra, which is involved with notorious Parkinson's disease, and

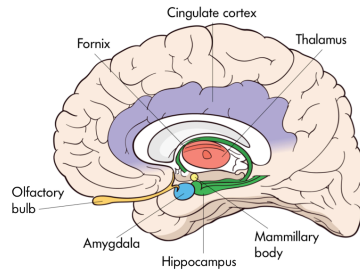


Figure 7. The interior of the brain.

the reticular formation regulating arousal, sleep, and temperature. A little lower, the hindbrain with the **pons** processes sensory and motor information employing the cranial nerves, works as a bridge that connects the cerebral cortex with the medulla, and reciprocally transfers information back and forth between the brain and the spinal cord. The **medulla oblongata** processes breathing, digestion, heart and blood vessel function, swallowing, and sneezing. The **cerebellum** controls motor movement coordination, balance, equilibrium, and muscle tone.

The midbrain and the hindbrain, which make up the brain stem, culminate in the spinal cord. Whereas inside the cerebral cortex, the **gray matter** (neuronal cell bodies) lies outside and **white matter** (myelinated axons) inside; in the spinal cord this arrangement reverses, as the gray matter resides inside and the white matter outside. Paired nerves (ganglia) exit the spinal cord, some closer

in direction towards the back (dorsal) and others towards the front (ventral). The dorsal nerves (afferent) receive sensory information from skin and muscles, and ventral nerves (efferent) send signals to muscles and organs to respond.

STUDYING THE NERVOUS SYSTEM

The study of the nervous system involves anatomical and physiological techniques that have improved over the years in efficiency and caliber. Clearly, gross morphology of the nervous system requires an eye-level view of the brain and the spinal cord. However, to resolve minute components, optical and electron microscopic techniques are needed.

Light microscopes and, later, electron microscopes have changed our understanding of the intricate connections that exist among nerve cells. For example, modern staining procedures (**immunocytochemistry**) make it possible to see selected neurons that are of one type or another or are affected by growth. With better resolution of the electron microscopes, fine structures like the synaptic cleft between the pre- and post-synaptic neurons can be studied in detail.

Along with the neuroanatomical techniques, a number of other methodologies aid neuroscientists in studying the function and physiology of the nervous system. Early on, **lesion studies** in animals (and study of neurological damage in humans) provided information about the function of the nervous system, by ablating (removing) parts of the nervous system or using neurotoxins to destroy them and documenting the effects on behavior or mental processes. Later, more sophisticated

microelectrode techniques were introduced, which led to recording from single neurons in the animal brains and investigating their physiological functions. Such studies led to formulating theories about how sensory and motor information are processed in the brain. To study many neurons (millions of them at a time) electroencephalographic (EEG) techniques were introduced. These methods are used to study how large ensembles of neurons, representing different parts of the nervous system, with (**event-related potentials**) or without stimulation function together. In addition, many scanning techniques that visualize the brain in conjunction with methods mentioned above are used to understand the details of the structure and function of the brain. These include **computerized axial tomography** (CAT), which uses X-rays to capture many pictures of the brain and sandwiches them into 3-D models to study it. The resolution of this method is inferior to **magnetic resonance imaging** (MRI), which is yet another way to capture brain images using large magnets that bobble (precession) hydrogen nuclei in the brain. Although the resolution of MRI scans is much better than CAT scans, they do not provide any functional information about the brain. **Positron Emission Tomography** (PET) involves the acquisition of physiologic (functional) images of the brain based on the detection of positrons. Radio-labeled isotopes of certain chemicals, such as an analog of glucose (fluorodeoxyglucose), enters the active nerve cells and emits positrons, which are captured and mapped into scans. Such scans show how the brain and its many modules become active (or not) when energized with entering glucose analog. Disadvantages of PET scans

include being invasive and rendering poor spatial resolution. The latter is why modern PET machines are coupled with CAT scanners to gain better resolution of the functioning brain. Finally, to avoid the invasiveness of PET, functional MRI (fMRI) techniques were developed. Brain images based on fMRI technique visualize brain function by changes in the flow of fluids (blood) in brain areas that occur over time. These scans provide a wealth of functional information about the brain as the individual may engage in a task, which is why the last two methods of brain scanning are very popular among cognitive neuroscientists.

Understanding the nervous system has been a long journey of inquiry, spanning several hundreds of years of meticulous studies carried out by some of the most creative and versatile investigators in the fields of philosophy, evolution, biology, physiology, anatomy, neurology, neuroscience, cognitive sciences, and psychology. Despite our profound understanding of this organ, its mysteries continue to surprise us, and its intricacies make us marvel at this complex structure unmatched in the universe.

OUTSIDE RESOURCES

Video: Pt. 1 video on the anatomy of the nervous system



One or more interactive elements has been excluded from this version of the

text. You can view them online here:

<https://umssystem.pressbooks.pub/ubb2e/?p=71#oembed-1>

Video: Pt. 2 video on the anatomy of the nervous system



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umssystem.pressbooks.pub/ubb2e/?p=71#oembed-2>

Video: To look at functions of the brain and neurons, watch



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umssystem.pressbooks.pub/ubb2e/?p=71#oembed-3>

Web: To look at different kinds of brains, visit

<http://brainmuseum.org/>

DISCUSSION QUESTIONS

1. Why is it important to study the nervous system in an evolutionary context?
2. How can we compare changes in the nervous system made through evolution to changes made during development?
3. What are the similarities and differences between the somatic and autonomic nervous systems?
4. Describe functions of the midbrain and hindbrain.
5. Describe the anatomy and functions of the forebrain.
6. Compare and contrast electroencephalograms to electrophysiological techniques.
7. Which brain scan methodologies are important for cognitive scientists? Why?

VOCABULARY

Afferent nerves

Nerves that carry messages to the brain or spinal cord.

Agnosias

Due to damage of Wernicke's area. An inability to recognize objects, words, or faces.

Aphasia

Due to damage of the Broca's area. An inability to produce or understand words.

Arcuate fasciculus

A fiber tract that connects Wernicke's and Broca's speech areas.

Autonomic nervous system

A part of the peripheral nervous system that connects to glands and smooth muscles. Consists of sympathetic and parasympathetic divisions.

Broca's area

An area in the frontal lobe of the left hemisphere. Implicated in language production.

Central sulcus

The major fissure that divides the frontal and the parietal lobes.

Cerebellum

A nervous system structure behind and below the cerebrum. Controls motor movement coordination, balance, equilibrium, and muscle tone.

Cerebrum

Consists of left and right hemispheres that sit at the top of the nervous system and engages in a variety of higher-order functions.

Cingulate gyrus

A medial cortical portion of the nervous tissue that is a part of the limbic system.

Computerized axial tomography

A noninvasive brain-scanning procedure that uses X-ray absorption around the head.

Ectoderm

The outermost layer of a developing fetus.

Efferent nerves

Nerves that carry messages from the brain to glands and organs in the periphery.

Electroencephalography

A technique that is used to measure gross electrical activity of the brain by placing electrodes on the scalp.

Event-related potentials

A physiological measure of large electrical change in the brain produced by sensory stimulation or motor responses.

Forebrain

A part of the nervous system that contains the cerebral hemispheres, thalamus, and hypothalamus.

Fornix

(plural form, fornices) A nerve fiber tract that connects the hippocampus to mammillary bodies.

Frontal lobe

The most forward region (close to forehead) of the cerebral hemispheres.

Functional magnetic resonance imaging

(or fMRI) A noninvasive brain-imaging technique that registers changes in blood flow in the brain during a given task (also see magnetic resonance imaging).

Globus pallidus

A nucleus of the basal ganglia.

Gray matter

Composes the bark or the cortex of the cerebrum and consists of the cell bodies of the neurons (see also white matter).

Gyrus

(plural form, gyri) A bulge that is raised between or among fissures of the convoluted brain.

Hippocampus

(plural form, hippocampi) A nucleus inside (medial) the temporal lobe implicated in learning and memory.

Homo habilis

A human ancestor, handy man, that lived two million years ago.

Homo sapiens

Modern man, the only surviving form of the genus Homo.

Hypothalamus

Part of the diencephalon. Regulates biological drives with pituitary gland.

Immunocytochemistry

A method of staining tissue including the brain, using antibodies.

Lateral geniculate nucleus

(or LGN) A nucleus in the thalamus that is

innervated by the optic nerves and sends signals to the visual cortex in the occipital lobe.

Lateral sulcus

The major fissure that delineates the temporal lobe below the frontal and the parietal lobes.

Lesion studies

A surgical method in which a part of the animal brain is removed to study its effects on behavior or function.

Limbic system

A loosely defined network of nuclei in the brain involved with learning and emotion.

Magnetic resonance imaging

Or MRI is a brain imaging noninvasive technique that uses magnetic energy to generate brain images (also see fMRI).

Magnification factor

Cortical space projected by an area of sensory input (e.g., mm of cortex per degree of visual field).

Medulla oblongata

An area just above the spinal cord that processes breathing, digestion, heart and blood vessel function, swallowing, and sneezing.

Neural crest

A set of primordial neurons that migrate outside the neural tube and give rise to sensory and autonomic neurons in the peripheral nervous system.

Neural induction

A process that causes the formation of the neural tube.

Neuroblasts

Brain progenitor cells that asymmetrically divide into other neuroblasts or nerve cells.

Neuroepithelium

The lining of the neural tube.

Occipital lobe

The back part of the cerebrum, which houses the visual areas.

Parasympathetic nervous system

A division of the autonomic nervous system that is slower than its counterpart—that is, the sympathetic nervous system—and works in opposition to it. Generally engaged in “rest and digest” functions.

Parietal lobe

An area of the cerebrum just behind the central sulcus that is engaged with somatosensory and gustatory sensation.

Pons

A bridge that connects the cerebral cortex with the medulla, and reciprocally transfers information back and forth between the brain and the spinal cord.

Positron Emission Tomography

(or PET) An invasive procedure that captures brain images with positron emissions from the brain after the individual has been injected with radio-labeled

isotopes.

Primary Motor Cortex

A strip of cortex just in front of the central sulcus that is involved with motor control.

Primary Somatosensory Cortex

A strip of cerebral tissue just behind the central sulcus engaged in sensory reception of bodily sensations.

Rostrocaudal

A front-back plane used to identify anatomical structures in the body and the brain.

Somatic nervous system

A part of the peripheral nervous system that uses cranial and spinal nerves in volitional actions.

Spina bifida

A developmental disease of the spinal cord, where the neural tube does not close caudally.

Sulcus

(plural form, sulci) The crevices or fissures formed by convolutions in the brain.

Sympathetic nervous system

A division of the autonomic nervous system, that is faster than its counterpart that is the parasympathetic nervous system and works in opposition to it. Generally engaged in “fight or flight” functions.

Temporal lobe

An area of the cerebrum that lies below the lateral sulcus; it contains auditory and olfactory (smell) projection regions.

Thalamus

A part of the diencephalon that works as a gateway for incoming and outgoing information.

Transduction

A process in which physical energy converts into neural energy.

Wernicke's area

A language area in the temporal lobe where linguistic information is comprehended (Also see Broca's area).

White matter

Regions of the nervous system that represent the axons of the nerve cells; whitish in color because of myelination of the nerve cells.

Working memory

Short transitory memory processed in the hippocampus.

REFERENCES

- American Association for the Advancement of Science (AAAS). (1880). Dr. Paul Broca. *Science*, 1(8), 93. <http://www.jstor.org/stable/2900242>
- Bullock, T. H. (1984). Comparative neuroscience holds promise for quiet revolutions. *Science*,

225(4661), 473–478.

- Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in the Neurosciences*, 2, 263–275.
- Darwin, C. (1871). *The descent of man, and the selection in relation to sex*. London: J. Murray.
- Darwin, C. (1859). On the origins of species by means of natural selection, or, *The preservation of favoured races in the struggle for life*. London, UK: J. Murray.
- Demonet, J. F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J. L., Wise, R., . . . Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. *Brain*, 115(6), 1753–1768.
- Edelman, G. (2004). *Wider than the sky: The phenomenal gift of consciousness*. New Haven, CT: Yale University Press.
- Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science*, 165(3894), 664–672.
- Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior*. Cambridge, MA: Harvard University Press.
- Hubel, D. H. (1995). *Eye, brain, and vision*. Freeman & Co., NY: Scientific American Library/Scientific American Books.
- Kandel, E. R., & Schwartz, J. H. (1982). Molecular biology of learning: Modulation of transmitter release. *Science*, 218(4571), 433–443.

- Knapp, M. L., & Hall, J. A. (2009). *Nonverbal communication in human interaction*. Boston, MA: Wadsworth Cengage Learning.
- Ramón y Cahal, S. (1937). Recollections of my life. *Memoirs of the American Philosophical Society*, 8, 1901–1917.
- Rolls, E. T., & Cowey, A. (1970). Topography of the retina and striate cortex and its relationship to visual acuity in rhesus monkeys and squirrel monkeys. *Experimental Brain Research*, 10(3), 298–310.
- Rosenzweig, M. R., Breedlove, S. M., & Leiman, A. L. (2002). *Biological psychology* (3rd ed.). Sunderland, MA: Sinauer Associates.
- Smith, E. E., & Jonides, J. (1999). Storage and executive processes in the frontal lobes. *Science*, 283(5408), 1657–1661.
- Smith, E. E., & Jonides, J. (1997). Working memory: A view from neuroimaging. *Cognitive psychology*, 33, 5–42.
- Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science*, 255(5043), 419–423.
- Walker, L. A. (1987). *A loss for words: The story of deafness in a family*. New York, NY: Harper Perennial.

AUTHORS



Aneeq Ahmad

Aneeq Ahmad, professor at Henderson State University and the president-elect of Southwestern Psychological Association (2014), teaches and conducts a variety of research projects with undergraduate and graduate students all the way from localizing purinergic receptors in the developing retina to prejudice both in the United States and Pakistan.

CREATIVE COMMONS LICENSE



THE NERVOUS SYSTEM by [Aneeq Ahmad](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

CHAPTER 5.

THE BRAIN

DIANE BECK AND EVELINA TAPIA

The human brain is responsible for all behaviors, thoughts, and experiences described in this textbook. This module provides an introductory overview of the brain, including some basic neuroanatomy, and brief descriptions of the neuroscience methods used to study it.

Learning Objectives

- Name and describe the basic function of the brain stem, cerebellum, and cerebral hemispheres.
- Name and describe the basic function of the four cerebral lobes: occipital, temporal, parietal, and frontal cortex.
- Describe a split-brain patient and at least two important aspects of brain function that these patients reveal.
- Distinguish between gray and white matter of

the cerebral hemispheres.

- Name and describe the most common approaches to studying the human brain.
- Distinguish among four neuroimaging methods: PET, fMRI, EEG, and DOI.
- Describe the difference between spatial and temporal resolution with regard to brain function.

INTRODUCTION

Any textbook on psychology would be incomplete without reference to the brain. Every behavior, thought, or experience described in the other modules must be implemented in the brain. A detailed understanding of the human brain can help us make sense of human experience and behavior. For example, one well-established fact about human cognition is that it is limited. We cannot do two complex tasks at once: We cannot read and carry on a conversation at the same time, text and drive, or surf the Internet while listening to a lecture, at least not successfully or safely. We cannot even pat our head and rub our stomach at the same time (with exceptions, see “A Brain Divided”). Why is this? Many people have suggested that such limitations reflect the fact that the behaviors draw on the same resource; if one behavior uses up most of the resource there is not enough resource left for the other. But what might this limited resource be in the brain?

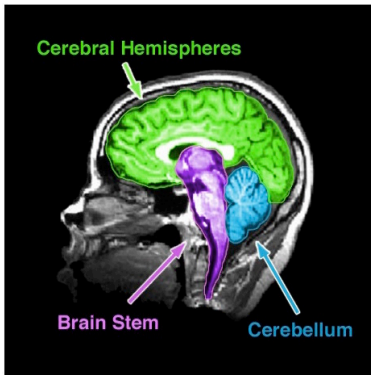


Figure 1. An MRI of the human brain delineating three major structures: the cerebral hemispheres, brain stem, and cerebellum.

The brain uses oxygen and glucose, delivered via the blood **metabolites**. The brain is a large consumer of these metabolites, using 20% of the oxygen and calories we consume despite being only 2% of our total weight. However, as long as we are not oxygen-deprived or malnourished, we have more than enough oxygen and glucose to fuel the

brain. Thus, insufficient “brain fuel” cannot explain our limited capacity. Nor is it likely that our limitations reflect too few neurons. The average human brain contains 86 billion neurons. It is also not the case that we use only 10% of our brain, a myth that was likely started to imply we had untapped potential. Modern neuroimaging (see “Studying the Human Brain”) has shown that we use all parts of brain, just at different times, and certainly more than 10% at any one time.

If we have an abundance of brain fuel and neurons, how can we explain our limited cognitive abilities? Why can’t we do more at once? The most likely explanation is the way these neurons are wired up. We know, for instance, that many neurons in the visual cortex (the part of the brain responsible for processing visual information) are hooked up in such a way as to inhibit each other ([Beck & Kastner, 2009](#)). When one neuron fires, it suppresses the firing of other nearby neurons. If two neurons that are hooked up in an inhibitory way both fire, then neither

neuron can fire as vigorously as it would otherwise. This competitive behavior among neurons limits how much visual information the brain can respond to at the same time. Similar kinds of competitive wiring among neurons may underlie many of our limitations. Thus, although talking about limited resources provides an intuitive description of our limited capacity behavior, a detailed understanding of the brain suggests that our limitations more likely reflect the complex way in which neurons talk to each other rather than the depletion of any specific resource.

THE ANATOMY OF THE BRAIN

There are many ways to subdivide the mammalian brain, resulting in some inconsistent and ambiguous **nomenclature** over the history of neuroanatomy ([Swanson, 2000](#)). For simplicity, we will divide the brain into three basic parts: the brain stem, cerebellum, and cerebral hemispheres (see Figure 1). In Figure 2, however, we depict other prominent groupings ([Swanson, 2000](#)) of the six major subdivisions of the brain ([Kandal, Schwartz, & Jessell, 2000](#)).

BRAIN STEM

The **brain stem** is sometimes referred to as the “trunk” of the brain. It is responsible for many of the neural functions that keep us alive, including regulating our respiration (breathing), heart rate, and digestion. In keeping with its function, if a patient sustains severe damage to the brain stem he or she will require “life support” (i.e., machines are used to keep him or her alive). Because of its vital role in survival, in many countries

a person who has lost brain stem function is said to be “brain dead,” although other countries require significant tissue loss in the cortex (of the cerebral hemispheres), which is responsible for our conscious experience, for the same diagnosis. The brain stem includes the medulla, pons, midbrain, and diencephalon (which consists of thalamus and hypothalamus). Collectively, these regions also are involved in our sleep–wake cycle, some sensory and motor function, as well as growth and other hormonal behaviors.

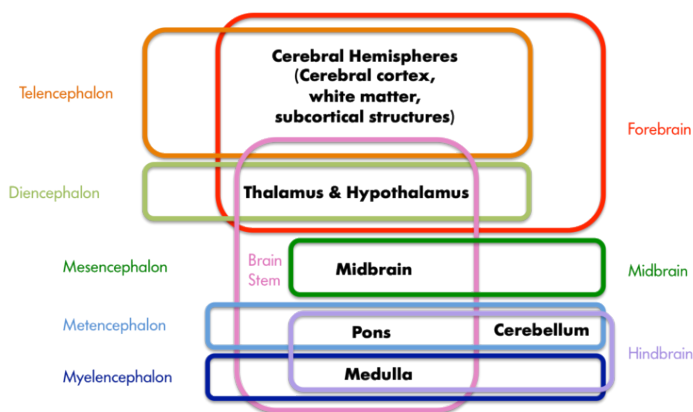


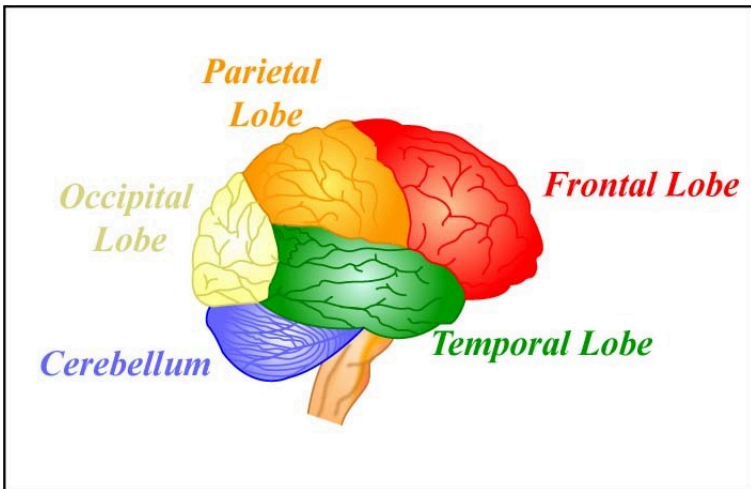
Figure 2. A sample of neuroanatomy nomenclature. The colored boxes indicate the different groupings of the seven structures printed in black, with the labels matching the color of the boxes. The hindbrain, midbrain, and forebrain nomenclature stems from the development of the vertebrate brain; these three areas differentiate early in embryonic development and later give rise to the structures listed in black. These three areas further subdivide into the telencephalon, diencephalon, mesencephalon, metencephalon, and myelencephalon at a later stage of development.

CEREBELLUM

The **cerebellum** is the distinctive structure at the back of the brain. The Greek philosopher and scientist Aristotle

aptly referred to it as the “small brain” (“*parencephalon*” in Greek, “*cerebellum*” in Latin) in order to distinguish it from the “large brain” (“*encephalon*” in Greek, “**cerebrum**” in Latin). The cerebellum is critical for coordinated movement and posture. More recently, neuroimaging studies (see “Studying the Human Brain”) have implicated it in a range of cognitive abilities, including language. It is perhaps not surprising that the cerebellum’s influence extends beyond that of movement and posture, given that it contains the greatest number of neurons of any structure in the brain. However, the exact role it plays in these higher functions is still a matter of further study.

CEREBRAL HEMISPHERES



The four lobes of the brain and the cerebellum. [Image: MIT OpenCourseWare, <https://goo.gl/RwUEVt>, CC BY-NC-SA 2.0, <https://goo.gl/TocOZF>]

The **cerebral hemispheres** are responsible for our cognitive abilities and conscious experience. They consist of the **cerebral cortex** and accompanying white matter

("cerebrum" in Latin) as well as the **subcortical** structures of the **basal ganglia**, amygdala, and hippocampal formation. The cerebral cortex is the largest and most visible part of the brain, retaining the Latin name (cerebrum) for "large brain" that Aristotle coined. It consists of two hemispheres (literally two half spheres) and gives the brain its characteristic gray and convoluted appearance; the folds and grooves of the cortex are called **gyri** and **sulci** (**gyrus** and **sulcus** if referring to just one), respectively.

The two cerebral hemispheres can be further subdivided into four lobes: the occipital, temporal, parietal, and frontal lobes. The **occipital lobe** is responsible for vision, as is much of the temporal lobe. The **temporal lobe** is also involved in auditory processing, memory, and multisensory integration (e.g., the convergence of vision and audition). The **parietal lobe** houses the **somatosensory (body sensations) cortex** and structures involved in visual attention, as well as multisensory convergence zones. The **frontal lobe** houses the **motor cortex** and structures involved in motor planning, language, judgment, and decision-making. Not surprisingly then, the frontal lobe is proportionally larger in humans than in any other animal.

The subcortical structures are so named because they reside beneath the cortex. The **basal ganglia** are critical to voluntary movement and as such make contact with the cortex, the thalamus, and the brain stem. The amygdala and hippocampal formation are part of the **limbic system**, which also includes some cortical structures. The limbic system plays an important role in emotion and, in particular, in aversion and gratification.

A BRAIN DIVIDED

The two cerebral hemispheres are connected by a dense bundle of white matter tracts called the corpus callosum. Some functions are replicated in the two hemispheres. For example, both hemispheres are responsible for sensory and motor function, although the sensory and motor cortices have a **contralateral** (or opposite-side) representation; that is, the left cerebral hemisphere is responsible for movements and sensations on the right side of the body and the right cerebral hemisphere is responsible for movements and sensations on the left side of the body. Other functions are **lateralized**; that is, they reside primarily in one hemisphere or the other. For example, for right-handed and the majority of left-handed individuals, the left hemisphere is most responsible for language.

There are some people whose two hemispheres are not connected, either because the corpus callosum was surgically severed (**callosotomy**) or due to a genetic abnormality. These **split-brain patients** have helped us understand the functioning of the two hemispheres. First, because of the contralateral representation of sensory information, if an object is placed in only the left or only the right **visual hemifield**, then only the right or left hemisphere, respectively, of the split-brain patient will see it. In essence, it is as though the person has two brains in his or her head, each seeing half the world. Interestingly, because language is very often localized in the left hemisphere, if we show the right hemisphere a picture and ask the patient what she saw, she will say she didn't see anything (because only the left hemisphere can speak and it didn't see anything). However, we know

that the right hemisphere sees the picture because if the patient is asked to press a button whenever she sees the image, the left hand (which is controlled by the right hemisphere) will respond despite the left hemisphere's denial that anything was there. There are also some advantages to having disconnected hemispheres. Unlike those with a fully functional corpus callosum, a split-brain patient can simultaneously search for something in his right and left visual fields ([Luck, Hillyard, Mangun, & Gazzaniga, 1989](#)) and can do the equivalent of rubbing his stomach and patting his head at the same time ([Franz, Eliason, Ivry, & Gazzaniga, 1996](#)). In other words, they exhibit less competition between the hemispheres.

GRAY VERSUS WHITE MATTER

The cerebral hemispheres contain both grey and white matter, so called because they appear grayish and whitish in dissections or in an MRI (magnetic resonance imaging; see, "Studying the Human Brain"). The **gray matter** is composed of the neuronal cell bodies (see module, "Neurons"). The cell bodies (or soma) contain the genes of the cell and are responsible for metabolism (keeping the cell alive) and synthesizing proteins. In this way, the cell body is the workhorse of the cell. The **white matter** is composed of the axons of the neurons, and, in particular, axons that are covered with a sheath of **myelin** (fatty support cells that are whitish in color). Axons conduct the electrical signals from the cell and are, therefore, critical to cell communication. People use the expression "use your gray matter" when they want a person to think harder. The "gray matter" in this expression is probably a reference to the cerebral hemispheres more generally; the

gray cortical sheet (the convoluted surface of the cortex) being the most visible. However, both the gray matter and white matter are critical to proper functioning of the mind. Losses of either result in deficits in language, memory, reasoning, and other mental functions. See Figure 3 for MRI slices showing both the inner white matter that connects the cell bodies in the gray cortical sheet.

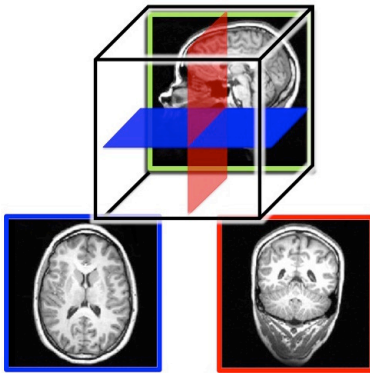


Figure 3. MRI slices of the human brain. Both the outer gray matter and inner white matter are visible in each image. The brain is a three-dimensional (3-D) structure, but an image is two-dimensional (2-D). Here, we show example slices of the three possible 2-D cuts through the brain: a sagittal slice (top image), a horizontal slice (bottom left), which is also known as a transverse or axial slice, and a coronal slice (bottom right). The bottom two images are color coded to match the illustration of the relative orientations of the three slices in the top image.

STUDYING THE HUMAN BRAIN

How do we know what the brain does? We have gathered knowledge about the functions of the brain from many different methods. Each method is useful for answering distinct types of questions, but the strongest evidence for a specific role or function of a particular brain area is **converging evidence**; that is, similar findings reported from multiple studies using different methods.

One of the first organized attempts to study the functions of the brain was **phrenology**, a popular field of study in the first half of the 19th century.

Phrenologists assumed that various features of the brain, such as its uneven surface, are reflected on the skull; therefore, they attempted to correlate bumps and indentations of the skull with specific functions of the brain. For example, they would claim that a very artistic person has ridges on the head that vary in size and location from those of someone who is very good at spatial reasoning. Although the assumption that the skull reflects the underlying brain structure has been proven wrong, phrenology nonetheless significantly impacted current-day neuroscience and its thinking about the functions of the brain. That is, different parts of the brain are devoted to very specific functions that can be identified through scientific inquiry.

NEUROANATOMY

Dissection of the brain, in either animals or cadavers, has been a critical tool of neuroscientists since 340 BC when Aristotle first published his dissections. Since then this method has advanced considerably with the discovery of various staining techniques that can highlight particular cells. Because the brain can be sliced very thinly, examined under the microscope, and particular cells highlighted, this method is especially useful for studying specific groups of neurons or small brain structures; that is, it has a very high **spatial resolution**. Dissections allow scientists to study changes in the brain that occur due to various diseases or experiences (e.g., exposure to drugs or brain injuries).

Virtual dissection studies with living humans are also conducted. Here, the brain is imaged using computerized axial tomography (CAT) or MRI scanners; they reveal

with very high precision the various structures in the brain and can help detect changes in gray or white matter. These changes in the brain can then be correlated with behavior, such as performance on memory tests, and, therefore, implicate specific brain areas in certain cognitive functions.

CHANGING THE BRAIN

Some researchers induce **lesions** or **ablate** (i.e., remove) parts of the brain in animals. If the animal's behavior changes after the lesion, we can infer that the removed structure is important for that behavior. Lesions of human brains are studied in patient populations only; that is, patients who have lost a brain region due to a stroke or other injury, or who have had surgical removal of a structure to treat a particular disease (e.g., a callosotomy to control epilepsy, as in split-brain patients). From such **case studies**, we can infer brain function by measuring changes in the behavior of the patients before and after the lesion.

Because the brain works by generating electrical signals, it is also possible to change brain function with electrical stimulation. **Transcranial magnetic stimulation (TMS)** refers to a technique whereby a brief magnetic pulse is applied to the head that temporarily induces a weak electrical current in the brain. Although effects of TMS are sometimes referred to as temporary virtual lesions, it is more appropriate to describe the induced electricity as interference with neurons' normal communication with each other. TMS allows very precise study of when events in the brain happen so it has a good **temporal resolution**, but its application is limited

only to the surface of the cortex and cannot extend to deep areas of the brain.

Transcranial direct current stimulation (tDCS) is similar to TMS except that it uses electrical current directly, rather than inducing it with magnetic pulses, by placing small electrodes on the skull. A brain area is stimulated by a low current (equivalent to an AA battery) for a more extended period of time than TMS. When used in combination with cognitive training, tDCS has been shown to improve performance of many cognitive functions such as mathematical ability, memory, attention, and coordination (e.g., [Brasil-Neto, 2012](#); [Feng, Bowden, & Kautz, 2013](#); [Kuo & Nitsche, 2012](#)).

NEUROIMAGING

Neuroimaging tools are used to study the brain in action; that is, when it is engaged in a specific task. **Positron emission tomography (PET)** records blood flow in the brain. The PET scanner detects the radioactive substance that is injected into the bloodstream of the participant just before or while he or she is performing some task (e.g., adding numbers). Because active neuron populations require metabolites, more blood and hence more radioactive substance flows into those regions. PET scanners detect the injected radioactive substance in specific brain regions, allowing researchers to infer that those areas were active during the task. **Functional magnetic resonance imaging (fMRI)** also relies on blood flow in the brain. This method, however, measures the changes in oxygen levels in the blood and does not require any substance to be injected into the participant. Both of these tools have good spatial resolution (although

not as precise as dissection studies), but because it takes at least several seconds for the blood to arrive to the active areas of the brain, PET and fMRI have poor temporal resolution; that is, they do not tell us very precisely when the activity occurred.



A researcher looking at the areas of activation in the brain of a study participant who had an fMRI scan – areas of brain activation are determined by the amount of blood flow to a certain area – the more blood flow, the higher the activation of that area of the brain. [Image: National Institute of Mental Health, CCO Public Domain, <https://goo.gl/m25gce>]

Electroencephalography (EEG), on the other hand, measures the electrical activity of the brain, and therefore, it has a much greater temporal resolution (millisecond precision rather than seconds) than PET or fMRI. Like tDCS, electrodes are placed on the participant's head when he or she is performing a task. In this case, however, many more electrodes are used, and they measure rather than produce activity. Because the electrical activity picked up at any particular electrode can be coming from anywhere in the brain, EEG has poor spatial resolution; that is, we have only a rough idea of which part of the brain generates the measured activity.

Diffuse optical imaging (DOI) can give researchers the best of both worlds: high spatial and temporal resolution, depending on how it is used. Here, one shines infrared light into the brain, and measures the light that comes back out. DOI relies on the fact that the properties of the light change when it passes through oxygenated blood, or when it encounters active neurons. Researchers can then infer from the properties of the collected light what regions in the brain were engaged by the task. When DOI is set up to detect changes in blood oxygen levels, the temporal resolution is low and comparable to PET or fMRI. However, when DOI is set up to directly detect active neurons, it has both high spatial and temporal resolution.

Because the spatial and temporal resolution of each tool varies, strongest evidence for what role a certain brain area serves comes from converging evidence. For example, we are more likely to believe that the hippocampal formation is involved in memory if multiple studies using a variety of tasks and different neuroimaging tools provide evidence for this hypothesis. The brain is a complex system, and only advances in brain research will show whether the brain can ever really understand itself.

OUTSIDE RESOURCES

Video: Videos of a split-brain patient



One or more interactive elements has

 been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=22#oembed-1>

Web: Atlas of the Human Brain: interactive demos and brain sections

<http://www.thehumanbrain.info/>

Web: Harvard University Human Brain Atlas: normal and diseased brain scans

<http://www.med.harvard.edu/aanlib/home.html>

DISCUSSION QUESTIONS

1. In what ways does the segmentation of the brain into the brain stem, cerebellum, and cerebral hemispheres provide a natural division?
2. How has the study of split-brain patients been informative?
3. What is behind the expression “use your gray matter,” and why is it not entirely accurate?
4. Why is converging evidence the best kind of evidence in the study of brain function?
5. If you were interested in whether a particular brain area was involved in a specific behavior, what neuroscience methods could you use?
6. If you were interested in the precise time in which a particular brain process occurred, which

neuroscience methods could you use?

VOCABULARY

Ablation

Surgical removal of brain tissue.

Axial plane

See “horizontal plane.”

Basal ganglia

Subcortical structures of the cerebral hemispheres involved in voluntary movement.

Brain stem

The “trunk” of the brain comprised of the medulla, pons, midbrain, and diencephalon.

Callosotomy

Surgical procedure in which the corpus callosum is severed (used to control severe epilepsy).

Case study

A thorough study of a patient (or a few patients) with naturally occurring lesions.

Cerebellum

The distinctive structure at the back of the brain, Latin for “small brain.”

Cerebral cortex

The outermost gray matter of the cerebrum; the distinctive convoluted characteristic of the mammalian brain.

Cerebral hemispheres

The cerebral cortex, underlying white matter, and subcortical structures.

Cerebrum

Usually refers to the cerebral cortex and associated white matter, but in some texts includes the subcortical structures.

Contralateral

Literally “opposite side”; used to refer to the fact that the two hemispheres of the brain process sensory information and motor commands for the opposite side of the body (e.g., the left hemisphere controls the right side of the body).

Converging evidence

Similar findings reported from multiple studies using different methods.

Coronal plane

A slice that runs from head to foot; brain slices in this plane are similar to slices of a loaf of bread, with the eyes being the front of the loaf.

Diffuse optical imaging (DOI)

A neuroimaging technique that infers brain activity by measuring changes in light as it is passed through the skull and surface of the brain.

Electroencephalography (EEG)

A neuroimaging technique that measures electrical brain activity via multiple electrodes on the scalp.

Frontal lobe

The front most (anterior) part of the cerebrum; anterior to the central sulcus and responsible for motor output and planning, language, judgment, and decision-making.

Functional magnetic resonance imaging (fMRI)

Functional magnetic resonance imaging (fMRI): A neuroimaging technique that infers brain activity by measuring changes in oxygen levels in the blood.

Gray matter

The outer grayish regions of the brain comprised of the neurons' cell bodies.

Gyri

(plural) Folds between sulci in the cortex.

Gyrus

A fold between sulci in the cortex.

Horizontal plane

A slice that runs horizontally through a standing person (i.e., parallel to the floor); slices of brain in this plane divide the top and bottom parts of the brain; this plane is similar to slicing a hamburger bun.

Lateralized

To the side; used to refer to the fact that specific functions may reside primarily in one hemisphere or the other (e.g., for the majority individuals, the left hemisphere is most responsible for language).

Lesion

A region in the brain that suffered damage through injury, disease, or medical intervention.

Limbic system

Includes the subcortical structures of the amygdala and hippocampal formation as well as some cortical structures; responsible for aversion and gratification.

Metabolite

A substance necessary for a living organism to maintain life.

Motor cortex

Region of the frontal lobe responsible for voluntary movement; the motor cortex has a contralateral representation of the human body.

Myelin

Fatty tissue, produced by glial cells (see module, “Neurons”) that insulates the axons of the neurons; myelin is necessary for normal conduction of electrical impulses among neurons.

Nomenclature

Naming conventions.

Occipital lobe

The back most (posterior) part of the cerebrum; involved in vision.

Parietal lobe

The part of the cerebrum between the frontal and occipital lobes; involved in bodily sensations, visual

attention, and integrating the senses.

Phrenology

A now-discredited field of brain study, popular in the first half of the 19th century that correlated bumps and indentations of the skull with specific functions of the brain.

Positron emission tomography (PET)

A neuroimaging technique that measures brain activity by detecting the presence of a radioactive substance in the brain that is initially injected into the bloodstream and then pulled in by active brain tissue.

Sagittal plane

A slice that runs vertically from front to back; slices of brain in this plane divide the left and right side of the brain; this plane is similar to slicing a baked potato lengthwise.

Somatosensory (body sensations) cortex

The region of the parietal lobe responsible for bodily sensations; the somatosensory cortex has a contralateral representation of the human body.

Spatial resolution

A term that refers to how small the elements of an image are; high spatial resolution means the device or technique can resolve very small elements; in neuroscience it describes how small of a structure in the brain can be imaged.

Split-brain patient

A patient who has had most or all of his or her

corpus callosum severed.

Subcortical

Structures that lie beneath the cerebral cortex, but above the brain stem.

Sulci

(plural) Grooves separating folds of the cortex.

Sulcus

A groove separating folds of the cortex.

Temporal lobe

The part of the cerebrum in front of (anterior to) the occipital lobe and below the lateral fissure; involved in vision, auditory processing, memory, and integrating vision and audition.

Temporal resolution

A term that refers to how small a unit of time can be measured; high temporal resolution means capable of resolving very small units of time; in neuroscience it describes how precisely in time a process can be measured in the brain.

Transcranial direct current stimulation (tDCS)

A neuroscience technique that passes mild electrical current directly through a brain area by placing small electrodes on the skull.

Transcranial magnetic stimulation (TMS)

A neuroscience technique whereby a brief magnetic pulse is applied to the head that temporarily induces a weak electrical current that interferes with ongoing activity.

Transverse plane

See “horizontal plane.”

Visual hemifield

The half of visual space (what we see) on one side of fixation (where we are looking); the left hemisphere is responsible for the right visual hemifield, and the right hemisphere is responsible for the left visual hemifield.

White matter

The inner whitish regions of the cerebrum comprised of the myelinated axons of neurons in the cerebral cortex.

REFERENCES

- Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research*, 49, 1154–1165.
- Brasil-Neto, J. P. (2012). Learning, memory, and transcranial direct current stimulation. *Frontiers in Psychiatry*, 3(80). doi: 10.3389/fpsyt.2012.00080.
- Feng, W. W., Bowden, M. G., & Kautz, S. (2013). Review of transcranial direct current stimulation in poststroke recovery. *Topics in Stroke Rehabilitation*, 20, 68–77.
- Franz, E. A., Eliassen, J. C., Ivry, R. B., & Gazzaniga, M. S. (1996). Dissociation of spatial and temporal coupling in the bimanual movements of callosotomy patients. *Psychological Science*, 7, 306–310.
- Kandal, E. R., Schwartz, J. H., & Jessell, T. M. (Eds.)

(2000). *Principles of neural science* (Vol. 4). New York, NY: McGraw-Hill.

- Kuo, M. F., & Nitsche, M. A. (2012). Effects of transcranial electrical stimulation on cognition. *Clinical EEG and Neuroscience*, 43, 192–199.
- Luck, S. J., Hillyard, S. A., Mangun, G. R., & Gazzaniga, M. S. (1989). Independent hemispheric attentional systems mediate visual search in split-brain patients. *Nature*, 342, 543–545.
- Swanson, L. (2000). What is the brain? *Trends in Neurosciences*, 23, 519–527.

AUTHORS



Diane Beck

Diane Beck, Associate Professor of Psychology at the University of Illinois, Urbana-Champaign. She uses behavioral methods, fMRI, EEG, optical imaging and transcranial magnetic stimulation to

study visual perception and attention in the human brain.



Evelina Tapia

Evelina Tapia, currently a postdoctoral researcher at the University of Illinois, received her Ph.D. from the University of Houston. She uses both behavioral methods and transcranial magnetic stimulation to study the brain processes involved in visual awareness.

CREATIVE COMMONS LICENSE



THE BRAIN by [Diane Beck and Evelina Tapia](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

PART IV.

**UNIT 4: DEVELOPMENT
AND PLASTICITY OF THE
BRAIN**

CHAPTER 6.

BRAIN PLASTICITY AND BEHAVIOUR IN THE DEVELOPING BRAIN

BRYAN KOLB AND ROBBIN GIBB

Kolb, B., & Gibb, R. (2011). Brain plasticity and behaviour in the developing brain. *Journal of the Canadian Academy of Child and Adolescent Psychiatry = Journal de l'Academie canadienne de psychiatrie de l'enfant et de l'adolescent*, 20(4), 265–276.

- <https://pmc.ncbi.nlm.nih.gov/articles/PMC3222570/>

PART V.

**UNIT 5: VISION &
PERCEPTION**

CHAPTER 7.

VISION

SIMONA BUETTI AND ALEJANDRO LLERAS

Vision is the sensory modality that transforms light into a psychological experience of the world around you, with minimal bodily effort. This module provides an overview of the most significant steps in this transformation and strategies that your brain uses to achieve this visual understanding of the environment.

LEARNING OBJECTIVES

- Describe how the eye transforms light information into neural energy.
- Understand what sorts of information the brain is interested in extracting from the environment and why it is useful.
- Describe how the visual system has adapted to deal with different lighting conditions.
- Understand the value of having two eyes.
- Understand why we have color vision.
- Understand the interdependence between vision and other brain functions.

WHAT IS VISION?

Think about the spectacle of a starry night. You look up at the sky, and thousands of photons from distant stars come crashing into your retina, a light-sensitive structure at the back of your eyeball. These photons are millions of years old and have survived a trip across the universe, only to run into one of your photoreceptors. Tough luck: in one thousandth of a second, this little bit of light energy becomes the fuel to a photochemical reaction known as **photoactivation**. The light energy becomes neural energy and triggers a cascade of neural activity that, a few hundredths of a second later, will result in your becoming aware of that distant star. You and the universe united by photons. That is the amazing power of vision. Light brings the world to you. Without moving, you know what's out there. You can recognize friends coming to meet you before you are able to hear them coming, ripe fruits from green ones on trees without having to taste them and before reaching out to grab them. You can also tell how quickly a ball is moving in your direction (Will it hit you? Can you hit it?).

How does all of that happen? First, light enters the eyeball through a tiny hole known as the pupil and, thanks to the refractive properties of your cornea and lens, this light signal gets projected sharply into the retina (see Outside Resources for links to a more detailed description of the eye structure). There, light is transduced into neural energy by about 200 million photoreceptor cells.

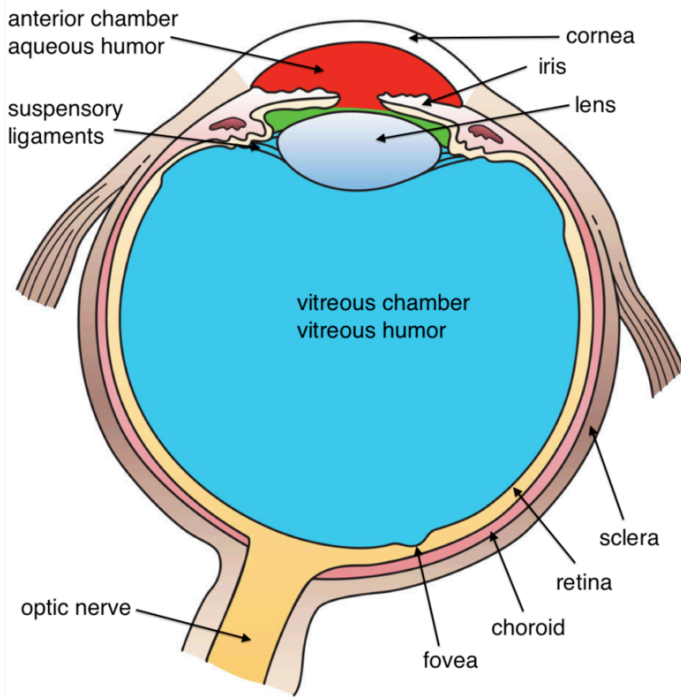


Diagram of the human eye [Image: Holly Fischer, <https://goo.gl/cMOQdh>, CC BY 3.0, <https://goo.gl/EHHVdU>]

This is where the information carried by the light about distant objects and colors starts being encoded by our brain. There are two different types of photoreceptors: **rods** and **cones**. The human eye contains more rods than cones. Rods give us sensitivity under dim lighting conditions and allow us to see at night. Cones allow us to see fine details in bright light and give us the sensation of color. Cones are tightly packed around the fovea (the central region of the retina behind your pupil) and more sparsely elsewhere. Rods populate the periphery (the region surrounding the fovea) and are almost absent from the fovea.

But vision is far more complex than just catching

photons. The information encoded by the photoreceptors undergoes a rapid and continuous set of ever more complex analysis so that, eventually, you can make sense of what's out there. At the fovea, visual information is encoded separately from tiny portions of the world (each about half the width of a human hair viewed at arm's length) so that eventually the brain can reconstruct in great detail fine visual differences from locations at which you are directly looking. This fine level of encoding requires lots of light and it is slow going (neurally speaking). In contrast, in the periphery, there is a different encoding strategy: detail is sacrificed in exchange for sensitivity. Information is summed across larger sections of the world. This aggregation occurs quickly and allows you to detect dim signals under very low levels of light, as well as detect sudden movements in your peripheral vision.

THE IMPORTANCE OF CONTRAST

What happens next? Well, you might think that the eye would do something like record the amount of light at each location in the world and then send this information to the visual-processing areas of the brain (an astounding 30% of the cortex is influenced by visual signals!). But, in fact, that is not what eyes do. As soon as photoreceptors capture light, the nervous system gets busy analyzing *differences* in light, and it is these differences that get transmitted to the brain. The brain, it turns out, cares little about the overall amount of light coming from a specific part of the world, or in the scene overall. Rather, it wants to know: does the light coming from this one point differ from the light coming from the point next

to it? Place your hand on the table in front of you. The contour of your hand is actually determined by the difference in light—the **contrast**—between the light coming from the skin in your hand and the light coming from the table underneath. To find the contour of your hand, we simply need to find the regions in the image where the difference in light between two adjacent points is maximal. Two points on your skin will reflect similar levels of light back to you, as will two points on the table. On the other hand, two points that fall on either side of the boundary contour between your hand and the table will reflect very different light.

The fact that the brain is interested in coding contrast in the world reveals something deeply important about the forces that drove the evolution of our brain: encoding the absolute amount of light in the world tells us little about what is out there. But if your brain can detect the sudden appearance of a *difference* in light somewhere in front of you, then it must be that something new is there. That contrast signal is *information*. That information may represent something that you like (food, a friend) or something dangerous approaching (a tiger, a cliff). The rest of your visual system will work hard to determine what that thing is, but as quickly as 10ms after light enters your eyes, ganglion cells in your retinae have already encoded all the differences in light from the world in front of you.

Contrast is so important that your neurons go out of their way not only to encode differences in light but to exaggerate those differences for you, lest you miss them. Neurons achieve this via a process known as **lateral inhibition**. When a neuron is firing in response to light, it produces two signals: an output signal to pass on to

the next level in vision, and a lateral signal to inhibit all neurons that are next to it. This makes sense on the assumption that nearby neurons are likely responding to the same light coming from nearby locations, so this information is somewhat redundant. The magnitude of the lateral inhibitory signal a neuron produces is proportional to the excitatory input that neuron receives: the more a neuron fires, the stronger the inhibition it produces. Figure 1 illustrates how lateral inhibition amplifies contrast signals at the edges of surfaces.

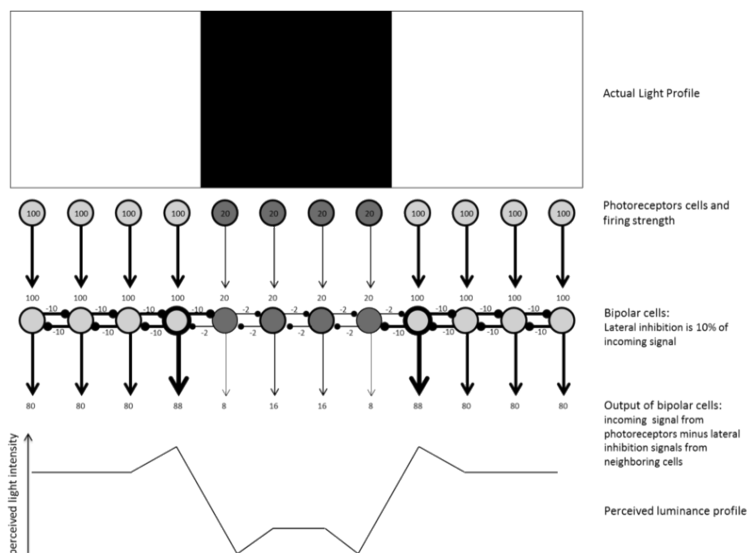


Figure 1. Illustration of Lateral Inhibition at work. The top of the figure shows a black stripe on a white background. The first row of circles illustrates photoreceptors responding in a graded fashion: the more light hits them, the more they fire. The numbers inside the circles represent how much these cells are firing, and the thickness of lines is also meant to illustrate the strength of neural firing. These photoreceptors activate the next layer of neurons in the retina: bipolar cells. These cells produce lateral inhibition signals, depicted by the horizontal lines that end with a small circle. The inhibition signals are proportional (here, 10% for ease) to the excitatory input they receive. Cells receiving 100 units will inhibit their neighbors by 10 units. Cells receiving 20 units will inhibit their neighbors by 2 units. The output of a bipolar cell will be determined by the input it receives minus all the lateral inhibition signals from its neighbors. As a result of the inhibition, notice how on the bright side of the edges, the firing rates are the highest (88) compared to nearby neurons just coding bright light (80). These higher values near the edge occur because these cells receive a comparatively small amount of inhibition from their “dark-side” neighbor (-2). Similarly, on the dark side of the edge, the firing rates are the lowest (8) of all the dark region (16) because these cells receive a comparatively large amount of inhibition from their “bright-side” neighbor (-10). Overall, the image is coded as a black stripe surrounded by brighter light, but now, thanks to lateral inhibition, all the edges in the image have been emphasized (enhanced), as illustrated by the perceived luminance profile at the bottom of the image.

SENSITIVITY TO DIFFERENT LIGHT CONDITIONS



Rods and cones work in tandem to help you adjust when moving between extremes of dark and light. [Image: Pexels, CCO Public Domain, <https://goo.gl/m25gce>]

Let's think for a moment about the range of conditions in which your visual system must operate day in and day out. When you take a walk outdoors on a sunny day, as many as billions of photons enter your eyeballs every second. In contrast, when you wake up in the middle of the night in a dark room, there might be as little as a few hundred photons per

second entering your eyes. To deal with these extremes, the visual system relies on the different properties of the two types of photoreceptors. Rods are mostly responsible for processing light when photons are scarce (just a single photon can make a rod fire!), but it takes time to replenish the visual pigment that rods require for photoactivation. So, under bright conditions, rods are quickly bleached ([Stuart & Brige, 1996](#)) and cannot keep up with the constant barrage of photons hitting them. That's when the cones become useful. Cones require more photons to fire and, more critically, their photopigments replenish much faster than rods' photopigments, allowing them to keep up when photons are abundant.

What happens when you abruptly change lighting conditions? Under bright light, your rods are bleached.

When you move into a dark environment, it will take time (up to 30 minutes) before they chemically recover ([Hurley, 2002](#)). Once they do, you will begin to see things around you that initially you could not. This phenomenon is called **dark adaptation**. When you go from dark to bright light (as you exit a tunnel on a highway, for instance), your rods will be bleached in a blaze and you will be blinded by the sudden light for about 1 second. However, your cones are ready to fire! Their firing will take over and you will quickly begin to see at this higher level of light.

A similar, but more subtle, adjustment occurs when the change in lighting is not so drastic. Think about your experience of reading a book at night in your bed compared to reading outdoors: the room may feel to you fairly well illuminated (enough so you can read) but the light bulbs in your room are not producing the billions of photons that you encounter outside. In both cases, you feel that your experience is that of a well-lit environment. You don't feel one experience as millions of times brighter than the other. This is because vision (as much of perception) is not proportional: seeing twice as many photons does not produce a sensation of seeing twice as bright a light. The visual system tunes into the current experience by favoring a range of contrast values that is most informative in that environment ([Gardner et al., 2005](#)). This is the concept of **contrast gain**: the visual system determines the mean contrast in a scene and represents values around that mean contrast best, while ignoring smaller contrast differences. (See the Outside Resources section for a demonstration.)

THE RECONSTRUCTION PROCESS

What happens once information leaves your eyes and enters the brain? Neurons project first into the thalamus, in a section known as the lateral geniculate nucleus. The information then splits and projects towards two different parts of the brain. Most of the computations regarding reflexive eye movements are computed in subcortical regions, the evolutionarily older part of the brain. Reflexive eye movements allow you to quickly orient your eyes towards areas of interest and to track objects as they move. The more complex computations, those that eventually allow you to have a visual experience of the world, all happen in the cortex, the evolutionarily newer region of the brain. The first stop in the cortex is at the **primary visual cortex** (also known as V1). Here, the “reconstruction” process begins in earnest: based on the contrast information arriving from the eyes, neurons will start computing information about color and simple lines, detecting various orientations and thicknesses. Small-scale motion signals are also computed ([Hubel & Wiesel, 1962](#)).

As information begins to flow towards other “higher” areas of the system, more complex computations are performed. For example, edges are assigned to the object to which they belong, backgrounds are separated from foregrounds, colors are assigned to surfaces, and the global motion of objects is computed. Many of these computations occur in specialized brain areas. For instance, an area called MT processes global-motion information; the parahippocampal place area identifies locations and scenes; the fusiform face area specializes in identifying objects for which fine discriminations are

required, like faces. There is even a brain region specialized in letter and word processing. These visual-recognition areas are located along the *ventral pathway* of the brain (also known as the **What pathway**). Other brain regions along the *dorsal pathway* (or **Where-and-How pathway**) will compute information about self- and object-motion, allowing you to interact with objects, navigate the environment, and avoid obstacles ([Goodale and Milner, 1992](#)).

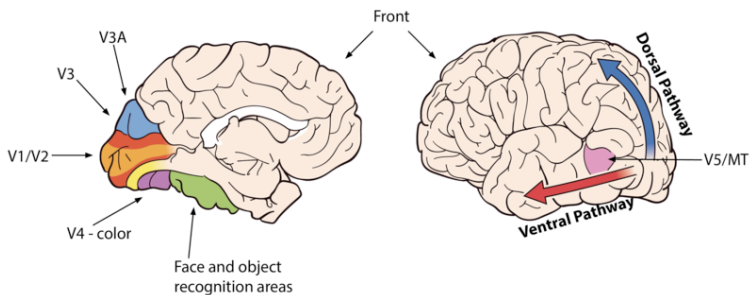


Figure 2. Areas of the brain.

Now that you have a basic understanding of how your visual system works, you can ask yourself the question: why do you have two eyes? Everything that we discussed so far could be computed with information coming from a single eye. So why two? Looking at the animal kingdom gives us a clue. Animals who tend to be prey have eyes located on opposite sides of their skull. This allows them to detect predators whenever one appears anywhere around them. Humans, like most predators, have two eyes pointing in the same direction, encoding almost the exact scene twice. This redundancy gives us a **binocular**

advantage: having two eyes not only provides you with two chances at catching a signal in front of you, but the minute difference in perspective that you get from each eye is used by your brain to reconstruct the sense of three-dimensional space. You can get an estimate of how far distant objects are from you, their size, and their volume. This is no easy feat: the signal in each eye is a two-dimensional projection of the world, like two separate pictures drawn upon your retinae. Yet, your brain effortlessly provides you with a sense of depth by combining those two signals. This 3-D reconstruction process also relies heavily on all the knowledge you acquired through experience about spatial information. For instance, your visual system learns to interpret how the volume, distance, and size of objects change as they move closer or farther from you. (See the Outside Resources section for demonstrations.)

THE EXPERIENCE OF COLOR

Perhaps one of the most beautiful aspects of vision is the richness of the color experience that it provides us. One of the challenges that we have as scientists is to understand why the human color experience is what it is. Perhaps you have heard that dogs only have 2 types of color photoreceptors, whereas humans have 3, chickens have 4, and mantis shrimp have 16. Why is there such variation across species? Scientists believe each species has evolved with different needs and uses color perception to signal information about food, reproduction, and health that are unique to their species. For example, humans have a specific sensitivity that allows you to detect slight changes in skin tone. You can

tell when someone is embarrassed, aroused, or ill. Detecting these subtle signals is adaptive in a social species like ours.

How is color coded in the brain? The two leading theories of color perception were proposed in the mid-19th century, about 100 years before physiological evidence was found to corroborate them both ([Svaetichin, 1956](#)). **Trichromacy theory**, proposed by Young (1802) and Helmholtz (1867), proposed that the eye had three different types of color-sensitive cells based on the observation that any one color can be reproduced by combining lights from three lamps of different hue. If you can adjust separately the intensity of each light, at some point you will find the right combination of the three lights to match any color in the world. This principle is used today on TVs, computer screens, and any colored display. If you look closely enough at a pixel, you will find that it is composed of a blue, a red, and a green light, of varying intensities. Regarding the retina, humans have three types of cones: S-cones, M-cones, and L-cones (also known as blue, green, and red cones, respectively) that are sensitive to three different wavelengths of light.

Around the same time, Hering made a puzzling discovery: some colors are impossible to create. Whereas you can make yellowish greens, bluish reds, greenish blues, and reddish yellows by combining two colors, you can never make a reddish green or a bluish yellow. This observation led Hering (1892) to propose the **Opponent Process theory of color**: color is coded via three opponent channels (red-green, blue-yellow, and black-white). Within each channel, a comparison is constantly computed between the two elements in the pair. In other words, colors are encoded as *differences* between two hues

and not as simple combinations of hues. Again, what matters to the brain is *contrast*. When one element is stronger than the other, the stronger color is perceived and the weaker one is suppressed. You can experience this phenomenon by following the link below.

<https://umsystem.pressbooks.pub/app/uploads/sites/93/2024/09/module-vision-color-demonstration.pdf>

When both colors in a pair are present to equal extents, the color perception is canceled and we perceive a level of grey. This is why you cannot see a reddish green or a bluish yellow: they cancel each other out. By the way, if you are wondering where the yellow signal comes from, it turns out that it is computed by averaging the M- and L-cone signals. Are these colors uniquely human colors? Some think that they are: the red-green contrast, for example, is finely tuned to detect changes in human skin tone so you can tell when someone blushes or becomes pale. So, the next time you go out for a walk with your dog, look at the sunset and ask yourself, what color does my dog see? Probably none of the orange hues you do!

So now, you can ask yourself the question: do all humans experience color in the same way? Color-blind people, as you can imagine, do not see all the colors that the rest of us see, and this is due to the fact that they lack one (or more) cones in their retina. Incidentally, there are a few women who actually have four different sets of cones in their eyes, and recent research suggests that their experience of color can be (but not always is) richer than the one from three-coned people. A slightly different question, though, is whether all three-coned people have the same internal experiences of colors: is the red inside your head the same red inside your mom's head? That is an almost impossible question to answer that has been

debated by philosophers for millennia, yet recent data suggests that there might in fact be cultural differences in the way we perceive color. As it turns out, not all cultures categorize colors in the same way, for example. And some groups “see” different shades of what we in the Western world would call the “same” color, as categorically different colors. The Berinmo tribe in New Guinea, for instance, appear to experience green shades that denote leaves that are alive as belonging to an entirely different color category than the sort of green shades that denote dying leaves. Russians, too, appear to experience light and dark shades of blue as different categories of colors, in a way that most Westerners do not. Further, current brain imaging research suggests that people’s brains change (increase in white-matter volume) when they learn new color categories! These are intriguing and suggestive findings, for certain, that seem to indicate that our cultural environment may in fact have some (small) but definite impact on how people use and experience colors across the globe.

INTEGRATION WITH OTHER MODALITIES

Vision is not an encapsulated system. It interacts with and depends on other sensory modalities. For example, when you move your head in one direction, your eyes reflexively move in the opposite direction to compensate, allowing you to maintain your gaze on the object that you are looking at. This reflex is called the **vestibulo-ocular reflex**. It is achieved by integrating information from both the visual and the vestibular system (which knows about body motion and position). You can experience this compensation quite simply. First, while

you keep your head still and your gaze looking straight ahead, wave your finger in front of you from side to side. Notice how the image of the finger appears blurry. Now, keep your finger steady and look at it while you move your head from side to side. Notice how your eyes reflexively move to compensate the movement of your head and how the image of the finger stays sharp and stable. Vision also interacts with your proprioceptive system, to help you find where all your body parts are, and with your auditory system, to help you understand the sounds people make when they speak. You can learn more about this in the Noba module about multimodal perception (<http://noba.to/cezw4qyn>).

Finally, vision is also often implicated in a blending-of-sensations phenomenon known as **synesthesia**. Synesthesia occurs when one sensory signal gives rise to two or more sensations. The most common type is *grapheme-color* synesthesia. About 1 in 200 individuals experience a sensation of color associated with specific letters, numbers, or words: the number 1 might always be seen as red, the number 2 as orange, etc. But the more fascinating forms of synesthesia blend sensations from entirely different sensory modalities, like taste and color or music and color: the taste of chicken might elicit a sensation of green, for example, and the timbre of violin a deep purple.

CONCLUDING REMARKS

We are at an exciting moment in our scientific understanding of vision. We have just begun to get a functional understanding of the visual system. It is not sufficiently evolved for us to recreate artificial visual

systems (i.e., we still cannot make robots that “see” and understand light signals as we do), but we are getting there. Just recently, major breakthroughs in vision science have allowed researchers to significantly improve retinal prosthetics: photosensitive circuits that can be implanted on the back of the eyeball of blind people that connect to visual areas of the brain and have the capacity to partially restore a “visual experience” to these patients ([Nirenberg & Pandarinath, 2012](#)). And using functional magnetic brain imaging, we can now “decode” from your brain activity the images that you saw in your dreams while you were asleep ([Horikawa, Tamaki, Miyawaki, & Kamitani, 2013](#))! Yet, there is still so much more to understand. Consider this: if vision is a construction process that takes time, whatever we see *now* is no longer what is front of us. Yet, humans can do amazing time-sensitive feats like hitting a 90-mph fastball in a baseball game. It appears then that a fundamental function of vision is not just to know what is happening around you now, but actually to make an accurate inference about what you are about to see *next* ([Enns & Lleras, 2008](#)), so that you can keep up with the world. Understanding how this future-oriented, *predictive* function of vision is achieved in the brain is probably the next big challenge in this fascinating realm of research.

OUTSIDE RESOURCES

Video: Acquired knowledge and its impact on our three-dimensional interpretation of the world – 3D Street Art



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=81#oembed-1>

Video: Acquired knowledge and its impact on our three-dimensional interpretation of the world – Anamorphic Illusions



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=81#oembed-2>

Video: Acquired knowledge and its impact on our three-dimensional interpretation of the world – Optical Illusion



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=81#oembed-3>

Web: Amazing library with visual phenomena and optical illusions, explained

<http://michaelbach.de/ot/index.html>

Web: Anatomy of the eye

<https://www.aao.org/eye-health/anatomy/parts-of-eye>

Web: Demonstration of contrast gain adaptation

<https://michaelbach.de/ot/lum-contrastAdapt/>

Web: Demonstration of illusory contours and lateral inhibition. Mach bands

<http://michaelbach.de/ot/lum-MachBands/index.html>

Web: Demonstration of illusory contrast and lateral inhibition. The Hermann grid

http://michaelbach.de/ot/lum_herGrid/

**Web: Further information regarding what and where/
how pathways**

[http://www.scholarpedia.org/article/
What_and_where_pathways](http://www.scholarpedia.org/article/What_and_where_pathways)

DISCUSSION QUESTIONS

1. When running in the dark, it is recommended that you never look straight at the ground. Why? What would be a better strategy to avoid obstacles?
2. The majority of ganglion cells in the eye specialize in detecting drops in the amount of light coming from a given location. That is, they increase their firing rate when they detect less light coming from a specific location. Why might the absence of light be more important than the presence of light? Why would it be evolutionarily advantageous to code this type of information?
3. There is a hole in each one of your eyeballs called the optic disk. This is where veins enter the eyeball and where neurons (the axons of the ganglion cells) exit the eyeball. Why do you not see two holes in the world all the time? Close one eye now. Why do you not see a hole in the world now? To “experience” a blind spot, follow the instructions in this website:
[http://michaelbach.de/ot/cog_blindSpot/
index.html](http://michaelbach.de/ot/cog_blindSpot/index.html)
4. Imagine you were given the task of testing the color-perception abilities of a newly discovered species of monkeys in the South Pacific. How would you go about it?

5. An important aspect of emotions is that we sense them in ourselves much in the same way as we sense other perceptions like vision. Can you think of an example where the concept of contrast gain can be used to understand people's responses to emotional events?

VOCABULARY

Binocular advantage

Benefits from having two eyes as opposed to a single eye.

Cones

Photoreceptors that operate in lighted environments and can encode fine visual details. There are three different kinds (S or blue, M or green and L or red) that are each sensitive to slightly different types of light. Combined, these three types of cones allow you to have color vision.

Contrast

Relative difference in the amount and type of light coming from two nearby locations.

Contrast gain

Process where the sensitivity of your visual system can be tuned to be most sensitive to the levels of contrast that are most prevalent in the environment.

Dark adaptation

Process that allows you to become sensitive to very small levels of light, so that you can actually see in the near-absence of light.

Lateral inhibition

A signal produced by a neuron aimed at suppressing the response of nearby neurons.

Opponent Process Theory

Theory of color vision that assumes there are four different basic colors, organized into two pairs (red/green and blue/yellow) and proposes that colors in the world are encoded in terms of the opponency (or difference) between the colors in each pair. There is an additional black/white pair responsible for coding light contrast.

Photoactivation

A photochemical reaction that occurs when light hits photoreceptors, producing a neural signal.

Primary visual cortex (V1)

Brain region located in the occipital cortex (toward the back of the head) responsible for processing basic visual information like the detection, thickness, and orientation of simple lines, color, and small-scale motion.

Rods

Photoreceptors that are very sensitive to light and are mostly responsible for night vision.

Synesthesia

The blending of two or more sensory experiences, or the automatic activation of a secondary (indirect) sensory experience due to certain aspects of the primary (direct) sensory stimulation.

Trichromacy theory

Theory that proposes that all of your color perception is fundamentally based on the combination of three (not two, not four) different color signals.

Vestibulo-ocular reflex

Coordination of motion information with visual information that allows you to maintain your gaze on an object while you move.

What pathway

Pathway of neural processing in the brain that is responsible for your ability to recognize what is around you.

Where-and-How pathway

Pathway of neural processing in the brain that is responsible for you knowing where things are in the world and how to interact with them.

REFERENCES

- Enns, J. T., & Lleras, A. (2008). New evidence for prediction in human vision. *Trends in Cognitive Sciences*, 12, 327–333.
- Gardner, J. L., Sun, P., Waggoner, R. A., Ueno, K., Tanaka, K., & Cheng, K. (2005). Contrast adaptation and representation in human early visual cortex. *Neuron*, 47, 607–620.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, 15, 20–25.

- Helmholtz, H. von. (1867). *Handbuch der Physiologischen Optik*. Leipzig: Leopold Voss.
- Hering, E. (1892). *Grundzüge der Lehre vom Lichtsinn*. Berlin, Germany: Springer.
- Horikawa, T., Tamaki, M., Miyawaki, Y., & Kamitani, Y. (2013). Neural decoding of visual imagery during sleep. *Science*, 340(6132), 639–642.
- Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160, 106–154.
- Hurley, J. B. (2002). Shedding light on adaptation. *Journal of General Physiology*, 119, 125–128.
- Nirenberg, S., & Pandarinath, C. (2012). Retinal prosthetic strategy with the capacity to restore normal vision. *Proceedings of the National Academy of Sciences*, 109(37), 15012–15017.
- Stuart, J. A., & Brige, R. R. (1996). Characterization of the primary photochemical events in bacteriorhodopsin and rhodopsin. In A. G. Lee (Ed.), *Rhodopsin and G-protein linked receptors* (Part A, Vol. 2, pp. 33–140). Greenwich, CT: JAI.
- Svaetichin, G. (1956). Spectral response curves from single cones, *Actaphysiologica Scandinavia, Suppl. 134*, 17–46.
- Young, T. (1802). Bakerian lecture: On the theory of light and colours. *Philosophical Transaction of the Royal Society London*, 92, 12–48.

AUTHORS



Simona Buetti

Simona Buetti is a postdoctoral fellow at the University of Illinois, where she teaches sensation and perception. She received a Ph.D. in Psychology (summa cum laude) from the University of Geneva, Switzerland. In 2010, she was awarded a postdoctoral fellowship from the Swiss National Science Foundation to support her postdoctoral work.



Alejandro Lleras

Alejandro Lleras, Associate Professor of

Psychology at the University of Illinois, received a CAREER award for young investigators from the National Science Foundation for his research on perception and visual attention and was named the 2010-2011 Helen Corley Petit Scholar for outstanding achievements as an Assistant Professor.

CREATIVE COMMONS LICENSE



VISION by [Simona Buetti and Alejandro Lleras](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

PART VI.

**UNIT 6: OTHER SENSORY
SYSTEMS & ATTENTION**

CHAPTER 8.

HEARING

ANDREW J. OXENHAM

Hearing allows us to perceive the world of acoustic vibrations all around us, and provides us with our most important channels of communication. This module reviews the basic mechanisms of hearing, beginning with the anatomy and physiology of the ear and a brief review of the auditory pathways up to the auditory cortex. An outline of the basic perceptual attributes of sound, including loudness, pitch, and timbre, is followed by a review of the principles of tonotopic organization, established in the cochlea. An overview of masking and frequency selectivity is followed by a review of the perception and neural mechanisms underlying spatial hearing. Finally, an overview is provided of auditory scene analysis, which tackles the important question of how the auditory system is able to make sense of the complex mixtures of sounds that are encountered in everyday acoustic environments.

LEARNING OBJECTIVES

- Describe the basic auditory attributes of sound.

- Describe the structure and general function of the auditory pathways from the outer ear to the auditory cortex.
- Discuss ways in which we are able to locate sounds in space.
- Describe various acoustic cues that contribute to our ability to perceptually segregate simultaneously arriving sounds.

INTRODUCTION

Hearing forms a crucial part of our everyday life. Most of our communication with others, via speech or music, reaches us through the ears. Indeed, a saying, often attributed to Helen Keller, is that blindness separates us from things, but deafness separates us from people. The ears respond to acoustic information, or sound—tiny and rapid



Hearing provides us with our most important connection to the people around us. [Image: Bindaas Madhavi, <https://goo.gl/Sv6TtR>, CC BY-NC-ND 2.0, <https://goo.gl/62XJAI>]

variations in air pressure. Sound waves travel from the source and produce pressure variations in the listener's ear canals, causing the eardrums (or tympanic membranes) to vibrate. This module provides an overview of the events that follow, which convert these simple mechanical vibrations into our rich experience known as hearing, or auditory perception.

PERCEPTUAL ATTRIBUTES OF SOUND

There are many ways to describe a sound, but the perceptual attributes of a sound can typically be divided into three main categories—namely, loudness, pitch, and timbre. Although all three refer to perception, and not to the physical sounds themselves, they are strongly related to various physical variables.

LOUDNESS

The most direct physical correlate of loudness is sound intensity (or sound pressure) measured close to the eardrum. However, many other factors also influence the loudness of a sound, including its frequency content, its duration, and the context in which it is presented. Some of the earliest psychophysical studies of auditory perception, going back more than a century, were aimed at examining the relationships between perceived loudness, the physical sound intensity, and the just-noticeable differences in loudness ([Fechner, 1860](#); [Stevens, 1957](#)). A great deal of time and effort has been spent refining various measurement methods. These methods involve techniques such as magnitude estimation, where a series of sounds (often sinusoids, or pure tones of single frequency) are presented sequentially at different sound levels, and subjects are asked to assign numbers to each tone, corresponding to the perceived loudness. Other studies have examined how loudness changes as a function of the frequency of a tone, resulting in the international standard iso-loudness-level contours ([ISO, 2003](#)), which are used in many areas of industry to assess noise and annoyance issues. Such studies have led to the development of computational models that are

designed to predict the loudness of arbitrary sounds (e.g., [Moore, Glasberg, & Baer, 1997](#)).

PITCH



Pitch is crucial to our perception and understanding of music and language.
[Image: xroper7, <https://goo.gl/1E4sJY>,
CC BY-NC 2.0, <https://goo.gl/tgFydH>]

Pitch plays a crucial role in acoustic communication. Pitch variations over time provide the basis of melody for most types of music; pitch contours in speech provide us with important prosodic information in non-tone languages, such as English, and help define the meaning of words in tone languages, such as Mandarin Chinese. Pitch

is essentially the perceptual correlate of waveform periodicity, or repetition rate: The faster a waveform repeats over time, the higher is its perceived pitch. The most common pitch-evoking sounds are known as harmonic complex tones. They are complex because they consist of more than one frequency, and they are harmonic because the frequencies are all integer multiples of a common fundamental frequency (F_0). For instance, a harmonic complex tone with a F_0 of 100 Hz would also contain energy at frequencies of 200, 300, 400 Hz, and so on. These higher frequencies are known as harmonics or overtones, and they also play an important role in determining the pitch of a sound. In fact, even if

the energy at the F0 is absent or masked, we generally still perceive the remaining sound to have a pitch corresponding to the F0. This phenomenon is known as the “pitch of the missing fundamental,” and it has played an important role in the formation of theories and models about pitch ([de Cheveigné, 2005](#)). We hear pitch with sufficient accuracy to perceive melodies over a range of F0s from about 30 Hz ([Pressnitzer, Patterson, & Krumbholz, 2001](#)) up to about 4–5 kHz ([Attneave & Olson, 1971](#); [Oxenham, Micheyl, Keebler, Loper, & Santurette, 2011](#)). This range also corresponds quite well to the range covered by musical instruments; for instance, the modern grand piano has notes that extend from 27.5 Hz to 4,186 Hz. We are able to discriminate changes in frequency above 5,000 Hz, but we are no longer very accurate in recognizing melodies or judging musical intervals.

TIMBRE

Timbre refers to the quality of sound, and is often described using words such as bright, dull, harsh, and hollow. Technically, timbre includes anything that allows us to distinguish two sounds that have the same loudness, pitch, and duration. For instance, a violin and a piano playing the same note sound very different, based on their sound quality or timbre.

An important aspect of timbre is the spectral content of a sound. Sounds with more high-frequency energy tend to sound brighter, tinnier, or harsher than sounds with more low-frequency content, which might be described as deep, rich, or dull. Other important aspects of timbre include the temporal envelope (or outline) of the sound,

especially how it begins and ends. For instance, a piano has a rapid onset, or attack, produced by the hammer striking the string, whereas the attack of a clarinet note can be much more gradual. Artificially changing the onset of a piano note by, for instance, playing a recording backwards, can dramatically alter its character so that it is no longer recognizable as a piano note. In general, the overall spectral content and the temporal envelope can provide a good first approximation to any sound, but it turns out that subtle changes in the spectrum over time (or spectro-temporal variations) are crucial in creating plausible imitations of natural musical instruments ([Risset & Wessel, 1999](#)).

AN OVERVIEW OF THE AUDITORY SYSTEM

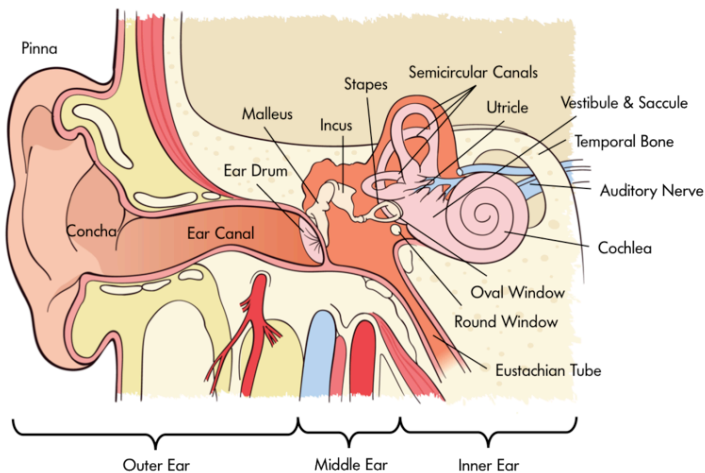


Figure 1. Diagram of the outer, middle, and inner ear.

Our auditory perception depends on how sound is processed through the ear. The ear can be divided into three main parts—the outer, middle, and inner ear (see

Figure 1). The outer ear consists of the **pinna** (the visible part of the ear, with all its unique folds and bumps), the ear canal (or auditory meatus), and the **tympanic membrane**. Of course, most of us have two functioning ears, which turn out to be particularly useful when we are trying to figure out where a sound is coming from. As discussed below in the section on spatial hearing, our brain can compare the subtle differences in the signals at the two ears to localize sounds in space. However, this trick does not always help: for instance, a sound directly in front or directly behind you will not produce a difference between the ears. In these cases, the filtering produced by the pinnae helps us localize sounds and resolve potential front-back and up-down confusions. More generally, the folds and bumps of the pinna produce distinct peaks and dips in the frequency response that depend on the location of the sound source. The brain then learns to associate certain patterns of spectral peaks and dips with certain spatial locations. Interestingly, this learned association remains malleable, or plastic, even in adulthood. For instance, a study that altered the pinnae using molds found that people could learn to use their “new” ears accurately within a matter of a few weeks ([Hofman, Van Riswick, & Van Opstal, 1998](#)). Because of the small size of the pinna, these kinds of acoustic cues are only found at high frequencies, above about 2 kHz. At lower frequencies, the sound is basically unchanged whether it comes from above, in front, or below. The ear canal itself is a tube that helps to amplify sound in the region from about 1 to 4 kHz—a region particularly important for speech communication.

The middle ear consists of an air-filled cavity, which contains the middle-ear bones, known as the incus,

malleus, and stapes, or anvil, hammer, and stirrup, because of their respective shapes. They have the distinction of being the smallest bones in the body. Their primary function is to transmit the vibrations from the tympanic membrane to the oval window of the cochlea and, via a form of lever action, to better match the impedance of the air surrounding the tympanic membrane with that of the fluid within the cochlea.

The inner ear includes the cochlea, encased in the temporal bone of the skull, in which the mechanical vibrations of sound are transduced into neural signals that are processed by the brain. The **cochlea** is a spiral-shaped structure that is filled with fluid. Along the length of the spiral runs the basilar membrane, which vibrates in response to the pressure differences produced by vibrations of the oval window. Sitting on the basilar membrane is the organ of Corti, which runs the entire length of the basilar membrane from the base (by the oval window) to the apex (the “tip” of the spiral). The organ of Corti includes three rows of outer hair cells and one row of inner hair cells. The hair cells sense the vibrations by way of their tiny hairs, or stereocillia. The outer hair cells seem to function to mechanically amplify the sound-induced vibrations, whereas the inner hair cells form synapses with the auditory nerve and transduce those vibrations into action potentials, or neural spikes, which are transmitted along the auditory nerve to higher centers of the auditory pathways.

One of the most important principles of hearing—frequency analysis—is established in the cochlea. In a way, the action of the cochlea can be likened to that of a prism: the many frequencies that make up a complex sound are broken down into their constituent

frequencies, with low frequencies creating maximal basilar-membrane vibrations near the apex of the cochlea and high frequencies creating maximal basilar-membrane vibrations nearer the base of the cochlea. This decomposition of sound into its constituent frequencies, and the frequency-to-place mapping, or “tonotopic” representation, is a major organizational principle of the auditory system, and is maintained in the neural representation of sounds all the way from the cochlea to the primary auditory cortex. The decomposition of sound into its constituent frequency components is part of what allows us to hear more than one sound at a time. In addition to representing frequency by place of excitation within the cochlea, frequencies are also represented by the timing of spikes within the auditory nerve. This property, known as “phase locking,” is crucial in comparing time-of-arrival differences of waveforms between the two ears (see the section on spatial hearing, below).

Unlike vision, where the primary visual cortex (or V1) is considered an early stage of processing, auditory signals go through many stages of processing before they reach the primary auditory cortex, located in the temporal lobe. Although we have a fairly good understanding of the electromechanical properties of the cochlea and its various structures, our understanding of the processing accomplished by higher stages of the auditory pathways remains somewhat sketchy. With the possible exception of spatial localization and neurons tuned to certain locations in space ([Harper & McAlpine, 2004](#); [Knudsen & Konishi, 1978](#)), there is very little consensus on the how, what, and where of auditory feature extraction and representation. There is evidence

for a “pitch center” in the auditory cortex from both human neuroimaging studies (e.g., [Griffiths, Buchel, Frackowiak, & Patterson, 1998](#); [Penagos, Melcher, & Oxenham, 2004](#)) and single-unit physiology studies ([Bendor & Wang, 2005](#)), but even here there remain some questions regarding whether a single area of cortex is responsible for coding single features, such as pitch, or whether the code is more distributed ([Walker, Bizley, King, & Schnupp, 2011](#)).

AUDIBILITY, MASKING, AND FREQUENCY SELECTIVITY

Overall, the human cochlea provides us with hearing over a very wide range of frequencies. Young people with normal hearing are able to perceive sounds with frequencies ranging from about 20 Hz all the way up to 20 kHz. The range of intensities we can perceive is also impressive: the quietest sounds we can hear in the medium-frequency range (between about 1 and 4 kHz) have a sound intensity that is about a factor of 1,000,000,000,000 less intense than the loudest sound we can listen to without incurring rapid and permanent hearing loss. In part because of this enormous dynamic range, we tend to use a logarithmic scale, known as decibels (dB), to describe sound pressure or intensity. On this scale, 0 dB sound pressure level (SPL) is defined as 20 micro-Pascals (μPa), which corresponds roughly to the quietest perceptible sound level, and 120 dB SPL is considered dangerously loud.



When the frequency content of different sounds overlaps, masking occurs. Less intense sounds become difficult or impossible to hear because more intense sounds dominate and interfere. Crowded restaurants or busy city streets full of traffic are typical examples of places where certain sounds can “swamp” others. [Image: Peter van der Sluijs, <https://goo.gl/K8L4c0>, CC BY-SA 3.0, <https://goo.gl/eLCn2O>]

Masking is the process by which the presence of one sound makes another sound more difficult to hear. We all encounter masking in our everyday lives, when we fail to hear the phone ring while we are taking a shower, or when we struggle to follow a conversation in a noisy restaurant. In general, a more intense sound will mask a less intense sound, provided certain conditions are met. The most important condition is that the frequency content of the sounds overlap, such that

the activity in the cochlea produced by a masking sound “swamps” that produced by the target sound. Another type of masking, known as “suppression,” occurs when the response to the masker reduces the neural (and in some cases, the mechanical) response to the target sound. Because of the way that filtering in the cochlea functions, low-frequency sounds are more likely to mask high frequencies than vice versa, particularly at high sound intensities. This asymmetric aspect of masking is known as the “upward spread of masking.” The loss of sharp cochlear tuning that often accompanies cochlear damage leads to broader filtering and more masking—a physiological phenomenon that is likely to contribute to

the difficulties experienced by people with hearing loss in noisy environments ([Moore, 2007](#)).

Although much masking can be explained in terms of interactions within the cochlea, there are other forms that cannot be accounted for so easily, and that can occur even when interactions within the cochlea are unlikely. These more central forms of masking come in different forms, but have often been categorized together under the term “informational masking” ([Durlach et al., 2003](#); [Watson & Kelly, 1978](#)). Relatively little is known about the causes of informational masking, although most forms can be ascribed to a perceptual “fusion” of the masker and target sounds, or at least a failure to segregate the target from the masking sounds. Also relatively little is known about the physiological locus of informational masking, except that at least some forms seem to originate in the auditory cortex and not before ([Gutschalk, Micheyl, & Oxenham, 2008](#)).

SPATIAL HEARING

In contrast to vision, we have a 360° field of hearing. Our auditory acuity is, however, at least an order of magnitude poorer than vision in locating an object in space. Consequently, our auditory localization abilities are most useful in alerting us and allowing us to orient towards sources, with our visual sense generally providing the finer-grained analysis. Of course, there are differences between species, and some, such as barn owls and echolocating bats, have developed highly specialized sound localization systems.

Our ability to locate sound sources in space is an impressive feat of neural computation. The two main sources of information both come from a comparison of the sounds at the two ears. The first is based on interaural time differences (ITD) and relies on the fact that a sound source on the left will generate sound that will reach the left ear slightly before it reaches



Humans are able to locate sound in space to determine whether the source is in front of us or behind us, or whether it is elevated or below us. [Photo: David Goehring, <https://goo.gl/UOLZpB>, CC BY 2.0, <https://goo.gl/BRvSA7>]

the right ear. Although sound is much slower than light, its speed still means that the time of arrival differences between the two ears is a fraction of a millisecond. The largest ITD we encounter in the real world (when sounds are directly to the left or right of us) are only a little over half a millisecond. With some practice, humans can learn to detect an ITD of between 10 and 20 μs (i.e., 20 *millionths* of a second) ([Klump & Eady, 1956](#)).

The second source of information is based in **interaural level differences (ILDs)**. At higher frequencies (higher than about 1 kHz), the head casts an acoustic “shadow,” so that when a sound is presented from the left, the sound level at the left ear is somewhat higher than the sound level at the right ear. At very high frequencies, the ILD can be as much as 20 dB, and we are sensitive to differences as small as 1 dB.

As mentioned briefly in the discussion of the outer ear,

information regarding the elevation of a sound source, or whether it comes from in front or behind, is contained in high-frequency spectral details that result from the filtering effects of the pinnae.

In general, we are most sensitive to ITDs at low frequencies (below about 1.5 kHz). At higher frequencies we can still perceive changes in timing based on the slowly varying temporal envelope of the sound but not the temporal fine structure ([Bernstein & Trahiotis, 2002](#); [Smith, Delgutte, & Oxenham, 2002](#)), perhaps because of a loss of neural phase-locking to the temporal fine structure at high frequencies. In contrast, ILDs are most useful at high frequencies, where the head shadow is greatest. This use of different acoustic cues in different frequency regions led to the classic and very early “duplex theory” of sound localization ([Rayleigh, 1907](#)). For everyday sounds with a broad frequency spectrum, it seems that our perception of spatial location is dominated by interaural time differences in the low-frequency temporal fine structure ([Macpherson & Middlebrooks, 2002](#)).

As with vision, our perception of distance depends to a large degree on context. If we hear someone shouting at a very low sound level, we infer that the shouter must be far away, based on our knowledge of the sound properties of shouting. In rooms and other enclosed locations, the reverberation can also provide information about distance: As a speaker moves further away, the direct sound level decreases but the sound level of the reverberation remains about the same; therefore, the ratio of direct-to-reverberant energy decreases ([Zahorik & Wightman, 2001](#)).

AUDITORY SCENE ANALYSIS

There is usually more than one sound source in the environment at any one time—imagine talking with a friend at a café, with some background music playing, the rattling of coffee mugs behind the counter, traffic outside, and a conversation going on at the table next to yours. All these sources produce sound waves that combine to form a single complex waveform at the eardrum, the shape of which may bear very little relationship to any of the waves produced by the individual sound sources. Somehow the auditory system is able to break down, or decompose, these complex waveforms and allow us to make sense of our acoustic environment by forming separate auditory “objects” or “streams,” which we can follow as the sounds unfold over time ([Bregman, 1990](#)).

A number of heuristic principles have been formulated to describe how sound elements are grouped to form a single object or segregated to form multiple objects. Many of these originate from the early ideas proposed in vision by the so-called Gestalt psychologists, such as Max Wertheimer. According to these rules of thumb, sounds that are in close proximity, in time or frequency, tend to be grouped together. Also, sounds that begin and end at the same time tend to form a single auditory object. Interestingly, spatial location is not always a strong or reliable grouping cue, perhaps because the location information from individual frequency components is often ambiguous due to the effects of reverberation. Several studies have looked into the relative importance of different cues by “trading off” one cue against another. In some cases, this has led to the discovery of interesting auditory illusions, where melodies that are not present

in the sounds presented to either ear emerge in the perception ([Deutsch, 1979](#)), or where a sound element is perceptually “lost” in competing perceptual organizations ([Shinn-Cunningham, Lee, & Oxenham, 2007](#)).

More recent attempts have used computational and neurally based approaches to uncover the mechanisms of auditory scene analysis (e.g., [Elhilali, Ma, Michey, Oxenham, & Shamma, 2009](#)), and the field of computational auditory scene analysis (CASA) has emerged in part as an effort to move towards more principled, and less heuristic, approaches to understanding the parsing and perception of complex auditory scenes (e.g., [Wang & Brown, 2006](#)). Solving this problem will not only provide us with a better understanding of human auditory perception, but may provide new approaches to “smart” hearing aids and cochlear implants, as well as automatic speech recognition systems that are more robust to background noise.

CONCLUSION



An infant with a cochlear implant.

[Image: Bjorn Knetsch, <https://goo.gl/J2wCvJ>, CC BY 2.0, <https://goo.gl/BRvSA7>]

Hearing provides us with our most important connection to the people around us. The intricate physiology of the auditory system transforms the tiny variations in air pressure that reach our ear into the vast array of auditory experiences that we perceive as speech, music, and sounds from the environment around

us. We are only beginning to understand the basic principles of neural coding in higher stages of the auditory system, and how they relate to perception. However, even our rudimentary understanding has improved the lives of hundreds of thousands through devices such as cochlear implants, which re-create some of the ear's functions for people with profound hearing loss.

OUTSIDE RESOURCES

Audio: Auditory Demonstrations from Richard Warren's lab at the University of Wisconsin, Milwaukee

<http://www4.uwm.edu/APL/demonstrations.html>

Audio: Auditory Demonstrations. CD published by the Acoustical Society of America (ASA). You can listen to the demonstrations here

<http://www.feilding.net/sfuad/musi3012-01/demos/audio/>

Web: Demonstrations and illustrations of cochlear mechanics can be found here

<http://lab.rockefeller.edu/hudspeth/graphicalSimulations>

Web: More demonstrations and illustrations of cochlear mechanics

<http://www.neurophys.wisc.edu/animations/>

DISCUSSION QUESTIONS

1. Based on the available acoustic cues, how good do you think we are at judging whether a low-frequency sound is coming from in front of us or behind us? How might we solve this problem in the real world?
2. Outer hair cells contribute not only to amplification but also to the frequency tuning in the cochlea. What are some of the difficulties that might arise for people with cochlear hearing loss, due to these two factors? Why do hearing aids not solve all these problems?
3. Why do you think the auditory system has so many stages of processing before the signals reach the auditory cortex, compared to the visual system? Is there a difference in the speed of processing required?

VOCABULARY

Cochlea

Snail-shell-shaped organ that transduces mechanical vibrations into neural signals.

Interaural differences

Differences (usually in time or intensity) between the two ears.

Pinna

Visible part of the outer ear.

Tympanic membrane

Ear drum, which separates the outer ear from the middle ear.

REFERENCES

- Attneave, F., & Olson, R. K. (1971). Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 84, 147–166.
- Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, 436, 1161–1165.
- Bernstein, L. R., & Trahiotis, C. (2002). Enhancing sensitivity to interaural delays at high frequencies by using “transposed stimuli.” *Journal of the Acoustical Society of America*, 112, 1026–1036.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

- Deutsch, D. (1979). Binaural integration of melodic patterns. *Perception & Psychophysics*, 25, 399–405.
- Durlach, N. I., Mason, C. R., Kidd, G., Jr., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. G. (2003). Note on informational masking. *Journal of the Acoustical Society of America*, 113, 2984–2987.
- Elhilali, M., Ma, L., Micheyl, C., Oxenham, A. J., & Shamma, S. (2009). Temporal coherence in the perceptual organization and cortical representation of auditory scenes. *Neuron*, 61, 317–329.
- Fechner, G. T. (1860). *Elemente der Psychophysik* (Vol. 1). Leipzig, Germany: Breitkopf und Haertl.
- Griffiths, T. D., Buchel, C., Frackowiak, R. S., & Patterson, R. D. (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience*, 1, 422–427.
- Gutschalk, A., Micheyl, C., & Oxenham, A. J. (2008). Neural correlates of auditory perceptual awareness under informational masking. *PLoS Biology*, 6, 1156–1165 (e1138).
- Harper, N. S., & McAlpine, D. (2004). Optimal neural population coding of an auditory spatial cue. *Nature*, 430, 682–686.
- Hofman, P. M., Van Riswick, J. G. A., & Van Opstal, A. J. (1998). Relearning sound localization with new ears. *Nature Neuroscience*, 1, 417–421.

- ISO. (2003). *ISO:226 Acoustics – Normal equal-loudness-level contours*. Geneva, Switzerland: International Organization for Standardization.
- Klump, R. G., & Eady, H. R. (1956). Some measurements of interaural time difference thresholds. *Journal of the Acoustical Society of America*, 28, 859–860.
- Knudsen, E. I., & Konishi, M. (1978). A neural map of auditory space in the owl. *Science*, 200, 795–797.
- Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *Journal of the Acoustical Society of America*, 111, 2219–2236.
- Moore, B. C. J. (2007). *Cochlear hearing loss: Physiological, psychological, and technical issues*. Chichester: Wiley.
- Moore, B. C. J., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society*, 45, 224–240.
- Oxenham, A. J., Micheyl, C., Keebler, M. V., Loper, A., & Santurette, S. (2011). Pitch perception beyond the traditional existence region of pitch. *Proceedings of the National Academy of Sciences USA*, 108, 7629–7634.
- Penagos, H., Melcher, J. R., & Oxenham, A. J. (2004). A neural representation of pitch salience in non-primary human auditory cortex revealed

with fMRI. *Journal of Neuroscience*, 24, 6810–6815.

- Pressnitzer, D., Patterson, R. D., & Krumbholz, K. (2001). The lower limit of melodic pitch. *Journal of the Acoustical Society of America*, 109, 2074–2084.
- Rayleigh, L. (1907). On our perception of sound direction. *Philosophical Magazine*, 13, 214–232.
- Risset, J. C., & Wessel, D. L. (1999). Exploration of timbre by analysis and synthesis. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 113–168): Academic Press.
- Shinn-Cunningham, B. G., Lee, A. K., & Oxenham, A. J. (2007). A sound element gets lost in perceptual competition. *Proceedings of the National Academy of Sciences USA*, 104, 12223–12227.
- Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416, 87–90.
- Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64, 153–181.
- Walker, K. M., Bizley, J. K., King, A. J., & Schnupp, J. W. (2011). Multiplexed and robust representations of sound features in auditory cortex. *Journal of Neuroscience*, 31, 14565–14576.
- Wang, D., & Brown, G. J. (Eds.). (2006). *Computational auditory scene analysis: Principles, algorithms, and applications*. Hoboken, NJ: Wiley.
- Watson, C. S., & Kelly, W. J. (1978). Informational masking in auditory patterns. *Journal of the Acoustical Society of America*, 64, S39.

- Zahorik, P., & Wightman, F. L. (2001). Loudness constancy with varying sound source distance. *Nature Neuroscience*, 4, 78–83.
- de Cheveigné, A. (2005). Pitch perception models. In C. J. Plack, A. J. Oxenham, A. N. Popper, & R. Fay (Eds.), *Pitch: Neural coding and perception* (pp. 169–233). New York, NY: Springer Verlag.

AUTHORS



Andrew J. Oxenham

Andrew Oxenham is a Distinguished McKnight University Professor in the Departments of Psychology and Otolaryngology at the University of Minnesota. He has published over 100 peer-reviewed papers in the field of auditory perception, and has received major awards from the Acoustical Society of America and the National Academy of Sciences.

CREATIVE COMMONS LICENSE



HEARING by [Andrew J. Oxenham](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

CHAPTER 9.

TASTE AND SMELL

LINDA BARTOSHUK AND DEREK SNYDER

Humans are omnivores (able to survive on many different foods). The omnivore's dilemma is to identify foods that are healthy and avoid poisons. Taste and smell cooperate to solve this dilemma. Stimuli for both taste and smell are chemicals. Smell results from a biological system that essentially permits the brain to store rough sketches of the chemical structures of odor stimuli in the environment. Thus, people in very different parts of the world can learn to like odors (paired with calories) or dislike odors (paired with nausea) that they encounter in their worlds. Taste information is preselected (by the nature of the receptors) to be relevant to nutrition. No learning is required; we are born loving sweet and hating bitter. Taste inhibits a variety of other systems in the brain. Taste damage releases that inhibition, thus intensifying sensations like those evoked by fats in foods. Ear infections and tonsillectomies both can damage taste. Adults who have experienced these conditions experience intensified sensations from fats and enhanced palatability of high-fat foods. This may explain why individuals who

have had ear infections or tonsillectomies tend to gain weight.

LEARNING OBJECTIVES

- Explain the salient properties of taste and smell that help solve the omnivore's dilemma.
- Distinguish between the way pleasure/displeasure is produced by smells and tastes.
- Explain how taste damage can have extensive unexpected consequences.

THE OMNIVORE'S DILEMMA

Humans are **omnivores**. We can survive on a wide range of foods, unlike species, such as koalas, that have a highly specialized diet (for koalas, eucalyptus leaves). With our amazing dietary range comes a problem: the omnivore's dilemma ([Pollan, 2006](#); [Rozin & Rozin, 1981](#)). To survive, we must identify healthy food and avoid poisons. The senses of taste and smell cooperate to give us this ability. Smell also has other important functions in lower animals (e.g., avoid predators, identify sexual partners), but these functions are less important in humans. This module will focus on the way taste and smell interact in humans to solve the omnivore's dilemma.



By just examining our teeth, it is apparent that we are animals who evolved to eat a wide range of foods. [Image: Justin Mclean, <https://goo.gl/ffk7ZV>, CC BY-NC 2.0, <https://goo.gl/VnKIK8>]

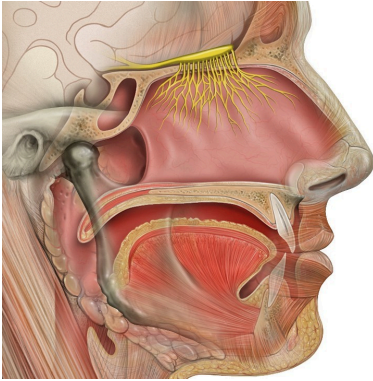
TASTE AND SMELL ANATOMY

Taste (**gustation**) and smell (**olfaction**) are both chemical senses; that is, the stimuli for these senses are chemicals. The more complex sense is olfaction. Olfactory receptors are complex proteins called G protein-coupled receptors (GPCRs). These structures are proteins that weave back and forth across the membranes of olfactory cells seven times, forming structures outside the cell that sense odorant molecules and structures inside the cell that activate the neural message ultimately conveyed to the brain by olfactory neurons. The structures that sense odorants can be thought of as tiny binding pockets with sites that respond to active parts of molecules (e.g., carbon chains). There are about 350 functional olfactory genes in humans; each gene expresses a particular kind of olfactory receptor. All olfactory receptors of a given kind project to structures called glomeruli (paired clusters of cells found on both sides of the brain). For a single

molecule, the pattern of activation across the glomeruli paints a picture of the chemical structure of the molecule. Thus, the olfactory system can identify a vast array of chemicals present in the environment. Most of the odors we encounter are actually mixtures of chemicals (e.g., bacon odor). The olfactory system creates an image for the mixture and stores it in memory just as it does for the odor of a single molecule ([Shepherd, 2005](#)).

Taste is simpler than olfaction. Bitter and sweet utilize GPCRs, just as olfaction does, but the number of different receptors is much smaller. For bitter, 25 receptors are tuned to different chemical structures ([Meyerhof et al., 2010](#)). Such a system allows us to sense many different poisons.

Sweet is even simpler. The primary sweet receptor is composed of two different G protein-coupled receptors; each of these two proteins ends in large structures reminiscent of Venus flytraps. This complex receptor has multiple sites that can bind different structures. The Venus flytrap endings open so that even some very large molecules can fit inside and stimulate the receptor.



Although smell plays a less integral role in our lives than it does for other animals, it is highly important in determining taste. In fact, if you plug your nose while eating chocolate, you will have a difficult time distinguishing it from any other kind of sweet. [Image: Patrick J. Lynch, <https://goo.gl/d5Clff>, CC BY 2.5, <https://goo.gl/OQtWcf>]

Bitter is inclusive (i.e., multiple receptors tuned to very different chemical structures feed into common neurons). Sweet is exclusive. There are many sugars with similar structures, but only three of these are particularly important to humans (sucrose, glucose, and fructose). Thus, our sweet receptor tunes out most sugars, leaving only the most important to stimulate the sweet receptor. However, the ability of the sweet receptor to respond to

some non-sugars presents us with one of the great mysteries of taste. Several non-sugar molecules can stimulate the primary sweet receptor (e.g., saccharine, aspartame, cyclamate). These have given rise to the artificial sweetener industry, but their biological significance is unknown. What biological purpose is served by allowing these non-sugar molecules to stimulate the primary sweet receptor?

Some would have us believe that artificial sweeteners are a boon to those who want to lose weight. It seems like a no-brainer. Sugars have calories; saccharin does not. Theoretically, if we replace sugar with saccharin in our diets, we will lose weight. In fact, recent work showed that rats actually gained weight when saccharin was

substituted for glucose ([Swithers & Davidson, 2008](#)). It turns out that substituting saccharin for sugar can increase appetite so more is eaten later. In addition, eating artificial sweeteners appears to alter metabolism, thus making losing weight even harder. So why did nature give us artificial sweeteners? We don't know.

One more mystery about sweet deserves comment. The discovery of the sweet receptor was met with great excitement because many investigators had searched for it for years. The fact that this complex receptor had multiple sites to which different molecules could bind explained why many different molecules taste sweet. However, this is actually a serious problem. No matter what molecule stimulates this receptor, the neural output from that receptor is the same. This would mean that the sweetness of all sweet substances would have to be the same. Yet artificial sweeteners do not taste exactly like sugar. The answer may lie in the fact that one of the two proteins that makes up the receptor can act alone, but only strong concentrations of sugar stimulate this isolated protein receptor. This permits the brain to distinguish between the sweetness of sugar and the sweetness of non-sugar molecules.

Salty and sour are the simplest tastes; these stimuli ionize (break into positively and negatively charged particles). The first event in the transduction series is the movement of the positively charged particle through channels in the taste cell membrane ([Chaudhari & Roper, 2010](#)).

SOLVING THE OMNIVORE'S DILEMMA: TASTE AFFECT IS HARD-WIRED

The pleasure associated with sweet and salty and the displeasure associated with sour and bitter are hard-wired in the brain. Newborns love sweet (taste of mother's milk) and hate bitter (poisons) immediately.



The receptors mediating salty taste are not mature at birth in humans, but when they are mature a few weeks after birth, the baby likes dilute salt (although more concentrated salt

Just how biologically ingrained is our love for sweets? Newborn infants immediately love sweet tastes, while the taste for salty foods takes longer to develop. [Image: shingleback, <https://goo.gl/fbUH3r>, CC BY 2.0, <https://goo.gl/BRvSA7>]

will evoke stinging sensations that will be avoided). Sour is generally disliked (protecting against tissue damage from acid?), but to the amazement of many parents, some young children appear to actually like the sour candies available today; this may be related to the breadth of their experience with fruits ([Liem & Mennella, 2003](#)). This hard-wired affect is the most salient characteristic of taste and this is why we classify only those taste qualities with hard-wired affect as “basic tastes.”

ANOTHER CONTRIBUTION TO THE OMNIVORE'S DILEMMA: OLFACTORY AFFECT IS LEARNED

The biological functions of olfaction depend on how

odors enter our noses. Sniffing brings odorants through our nostrils. The odorants hit the turbinate bones and a puff of the odorized air rises to the top of the nasal cavity, where it goes through a narrow opening (the olfactory cleft) and arrives at the olfactory mucosa (the tissue that houses the olfactory receptors). Technically, this is called **“orthonasal olfaction.”** Orthonasal olfaction tells us about the world external to our bodies.

When we chew and swallow food, the odorants emitted by the food are forced up behind the palate (roof of the mouth) and enter our noses from the back; this is called **“retronasal olfaction.”** Ortho and retronasal olfaction involve the same odor molecules and the same olfactory receptors; however, the brain can tell the difference between the two and does not send the input to the same areas. Retronasal olfaction and taste project to some common areas where they are presumably integrated into flavor. Flavors tell us about the food we are eating.

If retronasal olfaction is paired with nausea, the food evoking the retronasal olfactory sensation becomes disliked. If retronasal olfaction is paired with situations the brain deems valuable (calories, sweet taste, pleasure from other sources, etc.), the food evoking that sensation becomes liked. These are called **conditioned aversions and preferences** ([Rozin & Vollmecke, 1986](#)).



There are few associations more powerful than the taste of food and an experience of sickness. In fact, many people will go their whole lives without ever trying a food again that once made them sick. [Image: James Palinsad, <https://goo.gl/r2Bph4>, CC BY-ND 2.0, <https://goo.gl/VnHMcZ>]

Those who have experienced a conditioned aversion may have found that the dislike (even disgust) evoked when a flavor is paired with nausea can generalize to the smell of the food alone (orthonasal olfaction). Some years ago, Jeremy Wolfe and Linda Bartoshuk surveyed conditioned aversions among college students and staff that had resulted from consuming foods/beverages associated with nausea ([Bartoshuk &](#)

[Wolfe, 1990](#)). In 29% of the aversions, subjects reported that even the smell of the food/beverage had become aversive. Other properties of food objects can become aversive as well. In one unusual case, an aversion to cheese crackers generalized to vanilla wafers apparently because the containers were similar. Conditioned aversions function to protect us from ingesting a food that our brains associate with illness. Conditioned preferences are harder to form, but they help us learn what is safe to eat.

Is the affect associated with olfaction ever hard-wired? Pheromones are said to be olfactory molecules that evoke specific behaviors. Googling “human pheromone” will take you to websites selling various sprays that are supposed to make one more sexually appealing. However,

careful research does not support such claims in humans or any other mammals ([Doty, 2010](#)). For example, amniotic fluid was at one time believed to contain a pheromone that attracted rat pups to their mother's nipples so they could suckle. Early interest in identifying the molecule that acted as that pheromone gave way to understanding that the behavior was learned when a novel odorant, citral (which smells like lemons), was easily substituted for amniotic fluid ([Pedersen, Williams, & Blass, 1982](#)).

CENTRAL INTERACTIONS: KEY TO UNDERSTANDING TASTE DAMAGE

The integration of retronasal olfaction and taste into flavor is not the only central interaction among the sensations evoked by foods. These integrations in most cases serve important biological functions, but occasionally they go awry and lead to clinical pathologies.

Taste is mediated by three cranial nerves; these are bilateral nerves, each of which innervates one side of the mouth. Since they do not connect in the peripheral nervous system, interactions across the midline must occur in the brain. Incidentally, studying interactions across the midline is a classic way to draw inferences about central interactions. Insights from studies of this type were very important to understanding central processes long before we had direct imaging of brain function.

Taste on the anterior two thirds of the tongue (the part you can stick out) is mediated by the chorda tympani nerve; taste on the posterior one third (the part that stays attached) is mediated by the glossopharyngeal nerve.

Taste buds are tiny clusters of cells (like the segments of an orange) that are buried in the tissue of some papillae, the structures that give the tongue its bumpy appearance. Filiform papillae are the smallest and are distributed all over the tongue; they have no taste buds. In species like the cat, the filiform papillae are shaped like small spoons and help the cat hold liquids on the tongue while lapping (try lapping from a dish and you will see how hard it is without those special filiform papillae). Fungiform papillae (given this name because they resemble small button mushrooms) are larger circular structures on the anterior tongue (innervated by the chorda tympani). They contain about six taste buds each. Fungiform papillae can be seen with the naked eye, but swabbing blue food coloring on the tongue helps. The fungiform papillae do not stain as well as the rest of the tongue so they look like pink circles against a blue background. On some tongues, the spacing of fungiform papillae is like polka dots. Other tongues can have 10 times as many fungiform papillae, spaced so closely that there is little space between them. There is a connection between the density of fungiform papillae and the perception of taste. Those who experience the most intense taste sensations (we call them supertasters) tend to have the most fungiform papillae. Incidentally, this is a rare example in sensory processes of visible anatomical variation that correlates with function. We can look at the tongues of a variety of individuals and predict which of them will experience the most intense taste sensations.

The structures that house taste buds innervated by the glossopharyngeal nerve are called circumvallate papillae. They are relatively large structures arrayed in an inverted

V shape across the back of the tongue. Each of them looks like a small island surrounded by a moat.

Taste nerves project to the brain, where they send inhibitory signals to one another. One of the biological consequences of this inhibition is taste constancy. Damage to one nerve reduces taste input but also reduces inhibition on the other nerves ([Bartoshuk et al 2005](#)). That release of inhibition intensifies the central neural signals from the undamaged nerves, thereby maintaining whole mouth function. Interestingly, this release of inhibition



Approximately 25% of people are “non-tasters,” 50% are “medium tasters,” and another 25% are “super tasters.” The better the “taster” one is, the less he or she can typically tolerate spicy foods, salt, and the taste of alcohol. [Image: Rosie Rogers, <https://goo.gl/j7SepS>, CC BY-NC-SA 2.0, <https://goo.gl/Toc0ZF>]

can be so powerful that it actually increases whole mouth taste. The small effect of limited taste damage is one of the earliest clinical observations. In 1825, Brillat-Savarin described in his book *The Physiology of Taste* an interview with an ex-prisoner who had suffered a horrible punishment: amputation of his tongue. “This man, whom I met in Amsterdam, where he made his living by running errands, had had some education, and it was easy to communicate with him by writing. After I had observed that the forepart of his tongue has been cut off clear to the ligament, I asked him if he still found any flavor in what he ate, and if his sense of taste had survived the cruelty to

which he had been subjected. He replied that ... he still possessed the ability to taste fairly well" ([Brillat-Savarin, 1971, pg. 35](#)). This injury damaged the chorda tympani but spared the glossopharyngeal nerve.

We now know that taste nerves not only inhibit one another but also inhibit other oral sensations. Thus, taste damage can intensify oral touch (fats) and oral burn (chilis). In fact, taste damage appears to be linked to pain in general. Consider an animal injured in the wild. If pain reduced eating, its chance of survival would be diminished. However, nature appears to have wired the brain such that taste input inhibits pain. Eating is reinforced and the animal's chances of survival increase.

TASTE DAMAGE AND WEIGHT GAIN

The effects of taste damage depend on the extent of damage. If only one taste nerve is damaged, then release of inhibition occurs. If the damage is extensive enough, function is lost with one possible exception. Preliminary data suggest that the more extensive the damage to taste, the greater the intensification of pain; this is obviously of clinical interest.

Damage to a single taste nerve can intensify oral touch (e.g., the creamy, viscous sensations evoked by fats). Perhaps most surprising, damage to a single taste nerve can intensify retronasal olfaction; this may occur as a secondary result from the intensification of whole mouth taste.

These sensory changes can alter the palatability of foods; in particular, high-fat foods can be rendered more palatable. Thus one of the first areas we examined was the possibility that mild taste damage could lead to increases

in body mass index. Middle ear infections (otitis media) can damage the chorda tympani nerve; a tonsillectomy can damage the glossopharyngeal nerve. Head trauma damages both nerves, although it tends to take its greatest toll on the chorda tympani nerve. All of these clinical conditions increase body mass index in some individuals. More work is needed, but we suspect a link between the intensification of fat sensations, enhancement of palatability of high-fat foods, and weight gain.

OUTSIDE RESOURCES

Video: Inside the Psychologists Studio with Linda Bartoshuk



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=110#oembed-1>

Video: Linda Bartoshuk at Nobel Conference 46



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=110#oembed-2>

Video: Test your tongue: the science of taste



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=110#oembed-3>

DISCUSSION QUESTIONS

1. In this module, we have defined “basic tastes” in terms of whether or not a sensation produces hard-wired affect. Can you think of any other definitions of basic tastes?
2. Do you think omnivores, herbivores, or carnivores have a better chance at survival?
3. Olfaction is mediated by one cranial nerve. Taste is mediated by three cranial nerves. Why do you think evolution gave more nerves to taste than to smell? What are the consequences of this?

VOCABULARY

Conditioned aversions and preferences

Likes and dislikes developed through associations with pleasurable or unpleasurable sensations.

Gustation

The action of tasting; the ability to taste.

Olfaction

The sense of smell; the action of smelling; the ability to smell.

Omnivore

A person or animal that is able to survive by eating a wide range of foods from plant or animal origin.

Orthonasal olfaction

Perceiving scents/smells introduced via the nostrils.

Retronasal olfaction

Perceiving scents/smells introduced via the mouth/palate.

REFERENCES

- Bartoshuk, L. M., & Wolfe, J. M. (1990). Conditioned taste aversions in humans: Are they olfactory aversions? (abstract). *Chemical Senses*, 15, 551.
- Bartoshuk, L. M., Snyder, D. J., Grushka, M., Berger, A. M., Duffy, V. B., & Kveton, J. F. (2005). Taste damage: previously unsuspected consequences. *Chemical Senses*, 30(Suppl. 1), i218–i219
- Brillat-Savarin, J. A. (1825). *The physiology of taste* (M.F.K. Fisher, Trans., 1971). New York, NY: Alfred A. Knopf.
- Chaudhari, N., & Roper, S. D. (2010). The cell biology of taste. *Journal of Cell Biology*, 190, 285–296.
- Doty, R. L. (2010). *The great pheromone myth*.

Baltimore, MD: Johns Hopkins University Press.

- Liem, D. G., & Mennella, J. A. (2003). Heightened sour preferences during childhood. *Chemical Senses*, 28(2), 173–180.
- Meyerhof, W., Batram, C., Kuhn, C., Brockhoff, A., Chudoba, E., Bufe, B., . . . Behrens, M. (2010). The molecular receptive ranges of human TAS2R bitter taste receptors. *Chemical Senses*, 35, 157–170.
- Pedersen, P.E., Williams, C.L., & Blass, E.M. (1982). Activation and odor conditioning of suckling behavior in 3-day-old albino rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 8(4), 329–341.
- Pollan, M. (2006). *The omnivore's dilemma*. New York, NY: Penguin Books.
- Rozin, E., & Rozin, P. (1981). Culinary themes and variations. *Natural History*, 90, 6–14.
- Rozin, P., & Vollmecke, T.A. (1986). Food likes and dislikes. *Annual Review of Nutrition*, 6, 433-456.
- Shepherd, G. M. (2005). Outline of a theory of olfactory processing and its relevance to humans. *Chemical Senses*, 30(Suppl 1), i3-i5.
- Swithers, S. E., & Davidson, T. L. (2008). A role for sweet taste: Calorie predictive relations in energy regulation by rats. *Behavioral Neuroscience*, 122(1), 161-173.

AUTHORS



Linda Bartoshuk

Linda Bartoshuk, Bushnell Professor, University of Florida, has received the APA Hilgard Award, APS William James Award, AChemS Mozell Award, AAAS McGovern Award and has been elected to the Society for Experimental Psychologists, American Academy of Arts and Sciences and National Academy of Sciences for her research on taste.



Derek Snyder

Derek Snyder holds a faculty appointment in the Division of Occupational Science and Occupational Therapy at the University of Southern California, where he designs neuroscience-related coursework and professional resources. His research focuses on individual differences in human oral sensation in health and disease.

CREATIVE COMMONS LICENSE



TASTE AND SMELL by [Linda Bartoshuk and Derek Snyder](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#).

Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

CHAPTER 10.

TOUCH AND PAIN

GURO E. LØSETH; DAN-MIKAEL ELLINGSON; AND SIRI
LEKNES

The sensory systems of touch and pain provide us with information about our environment and our bodies that is often crucial for survival and well-being. Moreover, touch is a source of pleasure. In this module, we review how information about our environment and our bodies is coded in the periphery and interpreted by the brain as touch and pain sensations. We discuss how these experiences are often dramatically shaped by top-down factors like motivation, expectation, mood, fear, stress, and context. When well-functioning, these circuits promote survival and prepare us to make adaptive decisions. Pathological loss of touch can result in perceived disconnection from the body, and insensitivity to pain can be very dangerous, leading to maladaptive hazardous behavior. On the other hand, chronic pain conditions, in which these systems start signaling pain in response to innocuous touch or even in the absence of any observable sensory stimuli, have tremendous negative impact on the lives of the affected. Understanding how our sensory-processing mechanisms

can be modulated psychologically and physiologically promises to help researchers and clinicians find new ways to alleviate the suffering of chronic-pain patients.

LEARNING OBJECTIVES

- Describe the transduction of somatosensory signals: The properties of the receptor types as well as the difference in the properties of C-afferents and A-afferents and what functions these are thought to have.
- Describe the social touch hypothesis and the role of affective touch in development and bonding.
- Explain the motivation–decision model and descending modulation of pain, and give examples on how this circuitry can promote survival.
- Explain how expectations and context affect pain and touch experiences.
- Describe the concept of chronic pain and why treatment is so difficult.

INTRODUCTION

Imagine a life free of pain. How would it be—calm, fearless, serene? Would you feel invulnerable, invincible? Getting rid of **pain** is a popular quest—a quick search for “pain-free life” on Google returns well over 4 million hits—including links to various bestselling self-help guides promising a pain-free life in only 7 steps, 6 weeks, or 3 minutes. Pain management is a billion-dollar market, and involves much more than just pharmaceuticals. Surely a life with no pain would be a better one?



Figure 1A. Patient with HSAN-V genetic mutation affecting pain nerve growth. Severely affected 12-year-old boy with damages to his left knee and ankles. (Minde et al., 2004) [Used with permission]

Well, consider one of the “lucky few”: 12-year-old “Thomas” has never felt deep pain. Not even when a fracture made him walk around with one leg shorter than the other, so that the bones of his healthy leg were slowly crushed to destruction underneath the knee joint (see Figure 1A). For Thomas and other members of a large Swedish family, life without pain is a harsh reality because of a mutated gene that affects the growth of the nerves conducting deep pain. Most of those affected suffer from joint damage and frequent fractures to

bones in their feet and hands; some end up in wheelchairs even before they reach puberty ([Minde et al., 2004](#)). It turns out pain—generally—serves us well.

Living without a sense of touch sounds less attractive than being free of pain—touch is a source of pleasure and essential to how we feel. Losing the sense of touch has severe implications—something patient G. L. experienced when an antibiotics treatment damaged the type of nerves that signal touch from her skin and the position of her joints and muscles. She reported feeling

like she'd lost her physical self from her nose down, making her “disembodied”—like she no longer had any connection to the body attached to her head. If she didn't look at her arms and legs they could just “wander off” without her knowing—initially she was unable to walk, and even after she relearned this skill she was so dependent on her visual attention that closing her eyes would cause her to land in a hopeless heap on the floor. Only light caresses like those from her children's hands can make her feel she has a body, but even these sensations remain vague and elusive ([Olausson et al., 2002](#); [Sacks, 1985](#)).

SENSATION

CUTANEOUS SENSES OF THE SKIN CONNECT THE BRAIN TO THE BODY AND THE OUTSIDE WORLD

Touch and pain are aspects of the somatosensory system, which provides our brain with information about our own body (**interoception**) and properties of the immediate external world (**exteroception**) ([Craig, 2002](#)). We have somatosensory receptors located all over the body, from the surface of our skin to the depth of our joints. The information they send to the central nervous system is generally divided into four modalities: **cutaneous senses** (senses of the skin), proprioception (body position), kinesthesia (body movement), and **nociception** (pain, discomfort). We are going to focus on the cutaneous senses, which respond to tactile, thermal, and pruritic (itchy) stimuli, and events that cause tissue damage (and hence pain). In addition,

there is growing evidence for a fifth modality specifically channeling *pleasant* touch ([McGlone & Reilly, 2010](#)).

DIFFERENT RECEPTOR TYPES ARE SENSITIVE TO SPECIFIC STIMULI

The skin can convey many sensations, such as the biting cold of a wind, the comfortable pressure of a hand holding yours, or the irritating itch from a woolen scarf. The different types of information activate specific receptors that convert the stimulation of the skin to electrical nerve impulses, a process called **transduction**. There are three main groups of receptors in our skin: *mechanoreceptors*, responding to mechanical stimuli, such as stroking, stretching, or vibration of the skin; *thermoreceptors*, responding to cold or hot temperatures; and *chemoreceptors*, responding to certain types of chemicals either applied externally or released within the skin (such as histamine from an inflammation). For an overview of the different receptor types and their properties, see Box 1. The experience of *pain* usually starts with activation of **nociceptors**—receptors that fire specifically to potentially tissue-damaging stimuli. Most of the nociceptors are subtypes of either chemoreceptors or mechanoreceptors. When tissue is damaged or inflamed, certain chemical substances are released from the cells, and these substances activate the chemosensitive nociceptors. Mechanoreceptive nociceptors have a high threshold for activation—they respond to mechanical stimulation that is so intense it might damage the tissue.

Box 1. Categories of low-threshold mechanoreceptors*

Identity of receptor	Size of receptor*	Type of skin where found	Speed of adaptation*	Adequate stimulus*
Merkel's disks	Small, sharp borders	Glabrous*	Slow	Pressure
Meissner's corpuscles	Small, sharp borders	Glabrous	Rapid	Indentation
Ruffini corpuscles	Large, diffuse borders	Hairy + glabrous	Slow	Stretching
Pacinian corpuscles	Large, diffuse borders	Hairy + glabrous	Rapid	Vibration

* **Adequate stimulus:** The type of stimulus that the receptor is specialized to receive and respond to.

* **Glabrous Skin:** The hairless skin found on our palms and the soles of our feet. This skin has a higher density of receptors of a more complex range, which reflects the fact that we use these areas of our body to actively explore our surroundings and to discriminate tactile properties of objects we're interacting with.

* **Low-threshold mechanoreceptors:** Mechanoreceptors that respond to stimulus that is so light it doesn't threaten to damage the tissue around it. High-threshold mechanoreceptors respond to stimulation of higher intensity, and are a type of nociceptor.

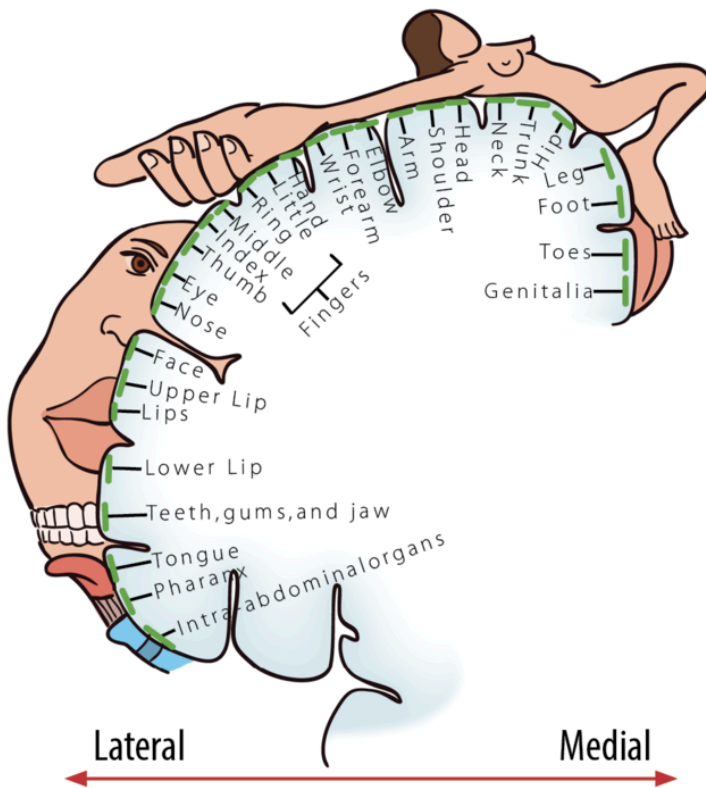
* **Receptive field:** The space of skin or tissue in which stimulation will elicit a response in the receptor. Smaller receptive fields make the receptor more sensitive to details.

* **Speed adaptation:** Slowly adapting mechanoreceptors continue to fire action potentials during sustained stimulation. Rapidly adapting mechanoreceptors fire action potentials in response to stimulus onset and offset (i.e. to stimuli changes), and help detect stimulus movement on the skin.

ACTION POTENTIALS IN THE RECEPTOR CELLS TRAVEL AS NERVE IMPULSES WITH DIFFERENT SPEEDS

When you step on a pin, this activates a host of mechanoreceptors, many of which are nociceptors. You may have noticed that the sensation changes over time. First you feel a sharp stab that propels you to remove your foot, and only then you feel a wave of more aching pain. The sharp stab is signaled via fast-conducting **A-fibers**, which project to the **somatosensory cortex**. This part of the cortex is **somatotopically organized**—that is, the sensory signals are represented according to where in the body they stem from (see Illustration, Figure 2). The unpleasant ache you feel after the sharp pin stab is

a separate, simultaneous signal sent from the nociceptors in your foot via thin **C-pain or A δ -fibers** to the insular cortex and other brain regions involved in processing of emotion and interoception (see Figure 3a for a schematic representation of this pathway). The experience of stepping on a pin is, in other words, composed by two separate signals: one discriminatory signal allowing us to localize the touch stimulus and distinguish whether it's a blunt or a sharp stab; and one affective signal that lets us know that stepping on the pin is bad. It is common to divide pain into sensory–discriminatory and affective–motivational aspects ([Auvray, Myin, & Spence, 2010](#)). This distinction corresponds, at least partly, to how this information travels from the peripheral to the central nervous system and how it is processed in the brain ([Price, 2000](#)).



Somatosensory Map

Figure 2. Somatosensory map: Body parts are represented in specific locations on the somatosensory cortex. Representations map out somatotopically, with the feet located medially and shoulders and arms laterally to the interhemispheric fissure. Facial structures are represented in a different location to the scalp and head; the face oriented upside down with the forehead pointing towards the shoulders.

AFFECTIVE ASPECTS OF TOUCH ARE IMPORTANT FOR DEVELOPMENT AND RELATIONSHIPS

Touch senses are not just there for discrimination or

detection of potentially painful events, as Harlow and Suomi (1970) demonstrated in a series of heartbreaking experiments where baby monkeys were taken from their mothers. The infant monkeys could choose between two artificial surrogate mothers—one “warm” mother without food but with a furry, soft cover; and one cold, steel mother with food. The monkey babies spent most of their time clinging to the soft mother, and only briefly moved over to the hard, steel mother to feed, indicating that touch is of “overpowering importance” to the infant (Harlow & Suomi, 1970, p. 161). Gentle touch is central for creating and maintaining social relationships in primates; they groom each other by stroking the fur and removing parasites—an activity important not only for their individual well-being but also for group cohesion (Dunbar, 2010; Keverne, Martensz, & Tuite, 1989). Although people don’t groom each other in the same way, gentle touch is important for us, too.

The sense of touch is the first to develop while one is in the womb, and human infants crave touch from the moment they’re born. From studies of human orphans, we know that touch is also crucial for human development. In Romanian orphanages where the babies were fed but not given regular attention or physical contact, the children suffered cognitive and neurodevelopmental delay (Simons & Land, 1987). Physical contact helps a crying baby calm down, and the soothing touch a mother gives to her child is thought to reduce the levels of stress hormones such as cortisol. High levels of cortisol have negative effects on neural development, and they can even lead to cell loss (Feldman, Singer, & Zagoory, 2010; Fleming, O’Day, & Kraemer, 1999; Pechtel & Pizzagalli, 2011). Thus, stress reduction

through hugs and caresses might be important not only for children's well-being, but also for the development of the infant brain.

The skin senses are similar across species, likely reflecting the evolutionary advantage of being able to tell what is touching you, where it's happening, and whether or not it's likely to cause tissue damage. An intriguing line of touch research suggests that humans, cats, and other animals have a special, evolutionarily preserved system that promotes gentle touch because it carries social and emotional significance. On a peripheral level, this system consists of a subtype of **C-fibers** that responds not to painful stimuli, but rather to gentle stroking touch—called **C-tactile** fibers. The firing rate of the C-tactile fibers correlates closely with how pleasant the stroking feels—suggesting they are coding specifically for the gentle caresses typical of social affiliative touch ([Löken, Wessberg, Morrison, McGlone, & Olausson, 2009](#)). This finding has led to the **social touch hypothesis**, which proposes that C-tactile fibers form a system for touch perception that supports social bonding ([Morrison, Löken, & Olausson, 2010](#); [Olausson, Wessberg, Morrison, McGlone, & Vallbo, 2010](#)). The discovery of the C-tactile system suggests that touch is organized in a similar way to pain; fast-conducting A-fibers contribute to sensory–discriminatory aspects, while thin C-fibers contribute to affective–motivational aspects ([Löken, Wessberg, Morrison, McGlone, & Olausson, 2009](#)). However, while these “hard-wired” afferent systems often provide us with accurate information about our environment and our bodies, how we experience touch or pain depends very much on top-

down sources like motivation, expectation, mood, fear, and stress.

MODULATION

PAIN IS NECESSARY FOR SURVIVAL, BUT OUR BRAIN CAN STOP IT IF IT NEEDS TO

In April 2003, the climber Aron Ralston found himself at the floor of Blue John Canyon in Utah, forced to make an appalling choice: face a slow but certain death—or amputate his right arm. Five days earlier he fell down the canyon—since then he had been stuck with his right arm trapped between an 800-lb boulder and the steep sandstone wall. Weak from lack of food and water and close to giving up, it occurred to him like an epiphany that if he broke the two bones in his forearm he could manage to cut off the rest with his pocket knife. The thought of freeing himself and surviving made him so excited he spent the next 40 minutes completely engrossed in the task: first snapping his bones using his body as a lever, then sticking his fingers into the arm, pinching bundles of muscle fibers and severing them one by one, before cutting the blue arteries and the pale “noodle-like” nerves. The pain was unimportant. Only cutting through the thick white main nerve made him stop for a minute—the flood of pain, he describes, was like thrusting his entire arm “into a cauldron of magma.” Finally free, he rappelled down a cliff and walked another 7 miles until he was rescued by some hikers ([Ralston, 2010](#)). How is it possible to do something so excruciatingly painful to yourself, and still manage to walk, talk, and think rationally afterwards? The answer lies within the brain, where signals from the body are

interpreted. When we perceive somatosensory and nociceptive signals from the body, the experience is highly subjective and malleable by motivation, attention, emotion, and context.

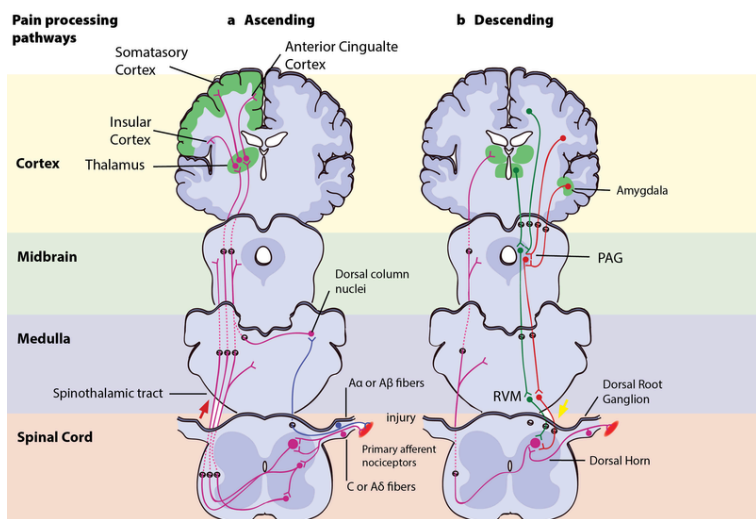


Figure 3. Pain processing pathways Left – Ascending pain pathways: An injury is signaled simultaneously via fast-conducting Aα or Aβ-fibers and slow-conducting C-pain or Aδ-fibers. The fast A-fibers signal pressure, stretching and other tissue movements to the somatosensory cortex via the dorsal column nuclei. The C-pain and Aδ-fibers send pain information from nociceptors in the tissue or skin, and transmit these signals to second order neurons in the dorsal horn of the spinal cord. The second order neurons then cross over to the opposite side, where they form the ascending spinothalamic tract. This tract projects signals to nuclei in the medulla and midbrain on the way up to the thalamus (T). The thalamus relays the information to the somatosensory and insular cortex, as well as cortical regions mediating different aspects of the pain experience such as affective responses in the cingulate cortex. Right – Descending pain modulation pathways: Information from the environment and certain motivational states can activate this top-down pathway. Several areas in the limbic forebrain including the anterior cingulate and insular cortex, nuclei in the amygdala and the hypothalamus (H), project to the midbrain periaqueductal grey (PAG), which then modulates ascending pain transmission from the afferent pain system indirectly through the rostral ventromedial medulla (RVM) in the brainstem. This modulating system produces analgesia by the release of endogenous opioids, and uses ON- and OFF-cells to exert either inhibitory (green) or facilitatory (red) control of nociceptive signals at the spinal dorsal horn.

THE MOTIVATION-DECISION MODEL AND DESCENDING MODULATION OF PAIN

According to *the motivation–decision model*, the brain automatically and continuously evaluates the pros and cons of any situation—weighing impending threats and available rewards ([Fields, 2004, 2006](#)). Anything more important for survival than avoiding the pain activates the brain’s **descending pain modulatory system**—a top-down system involving several parts of the brain and brainstem, which inhibits nociceptive signaling so that the more important actions can be attended to (Figure 3b). In Aron’s extreme case, his actions were likely based on such an unconscious decision process—taking into account his homeostatic state (his hunger, thirst, the inflammation and decay of his crushed hand slowly affecting the rest of his body), the sensory input available (the sweet smell of his dissolving skin, the silence around him indicating his solitude), and his knowledge about the threats facing him (death, or excruciating pain that won’t kill him) versus the potential rewards (survival, seeing his family again). Aron’s story illustrates the evolutionary advantage to being able to shut off pain: The descending pain modulatory system allows us to go through with potentially life-saving actions. However, when one has reached safety or obtained the reward, healing is more important. The very same descending system can then “crank up” nociception from the body to promote healing and motivate us to avoid potentially painful actions. To facilitate or inhibit nociceptive signals from the body, the descending pain modulatory system uses a set of ON- or OFF-cells in the brainstem, which regulates how much of the nociceptive signal reaches the brain. The descending

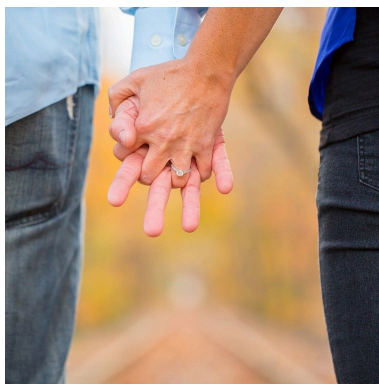
system is dependent on opioid signaling, and **analgesics** like morphine relieve pain via this circuit ([Petrovic, Kalso, Petersson, & Ingvar, 2002](#)).

THE ANALGESIC POWER OF REWARD

Thinking about the good things, like his loved ones and the life ahead of him, was probably pivotal to Aron's survival. The promise of a reward can be enough to relieve pain. Expecting pain relief (getting less pain is often the best possible outcome if you're in pain, i.e., it is a reward) from a medical treatment contributes to the **placebo**

effect—where pain relief

is due at least partly to your brain's descending modulation circuit, and such relief depends on the brain's own opioid system ([Eippert et al., 2009](#); [Eippert, Finsterbusch, Bingel, & Buchel, 2009](#); [Levine, Gordon, & Fields, 1978](#)). Eating tasty food, listening to good music, or feeling pleasant touch on your skin also decreases pain in both animals and humans, presumably through the same mechanism in the brain ([Leknes & Tracey, 2008](#)). In a now classic experiment, Dum and Herz ([1984](#)) either fed rats normal rat food or let them feast on highly rewarding chocolate-covered candy (rats love sweets) while standing on a metal plate until they learned exactly what to expect



Social rewards, like holding the hands or just seeing the picture of a loved one, can reduce sensations of pain. [Image: rogerl01, CCO Public Domain, <https://goo.gl/m25gce>]

when placed there. When the plate was heated up to a noxious/painful level, the rats that expected candy endured the temperature for twice as long as the rats expecting normal chow. Moreover, this effect was completely abolished when the rats' opioid (endorphin) system was blocked with a drug, indicating that the analgesic effect of reward anticipation was caused by **endorphin** release.

For Aron the climber, both the stress from knowing that death was impending and the anticipation of the reward it would be to survive probably flooded his brain with endorphins, contributing to the wave of excitement and euphoria he experienced while he carried out the amputation “like a five-year-old unleashed on his Christmas presents” ([Ralston, 2010](#)). This altered his experience of the pain from the extreme tissue damage he was causing and enabled him to focus on freeing himself. Our brain, it turns out, can modulate the perception of how unpleasant pain is, while still retaining the ability to experience the intensity of the sensation ([Rainville, Duncan, Price, Carrier, & Bushnell, 1997](#); [Rainville, Feine, Bushnell, & Duncan, 1992](#)). Social rewards, like holding the hand of your boyfriend or girlfriend, have pain-reducing effects. Even looking at a picture of him/her can have similar effects—in fact, seeing a picture of a person we feel close to not only reduces subjective pain ratings, but also the activity in pain-related brain areas ([Eisenberger et al., 2011](#)). The most common things to do when wanting to help someone through a painful experience—being present and holding the person's hand—thus seems to have a measurably positive effect.

WHEN TOUCH BECOMES PAINFUL OR PAIN

BECOMES CHRONIC

Chances are you've been sunburned a few times in your life and have experienced how even the lightest pat on the back or the softest clothes can feel painful on your over-sensitive skin. This condition, where innocuous touch gives a burning, tender sensation, is similar to a chronic condition called **allodynia**—where neuronal disease or injury makes touch that is normally pleasant feel unpleasantly painful. In allodynia, neuronal injury in the spinal dorsal horn causes A β -afferents, which are activated by non-nociceptive touch, to access nociceptive pathways ([Liljencrantz et al., 2013](#)). The result is that even gentle touch is interpreted by the brain as painful. While an acute pain response to **noxious stimuli** has a vital protective function, allodynia and other **chronic pain** conditions constitute a tremendous source of unnecessary suffering that affects millions of people. Approximately 100 million Americans suffer from chronic pain, and annual economic cost associated is estimated to be \$560–\$635 billion ([Committee on Advancing Pain Research, Care, & Institute of Medicine, 2011](#)). Chronic pain conditions are highly diverse, and they can involve changes on peripheral, spinal, central, and psychological levels. The mechanisms are far from fully understood, and developing appropriate treatment remains a huge challenge for pain researchers.

Chronic pain conditions often begin with an injury to a peripheral nerve or the tissue surrounding it, releasing hormones and inflammatory molecules that *sensitize* nociceptors. This makes the nerve and neighboring afferents more excitable, so that also uninjured nerves become hyperexcitable and contribute

to the persistence of pain. An injury might also make neurons fire nonstop regardless of external stimuli, providing near-constant input to the pain system. **Sensitization** can also happen in the brain and in the descending modulatory system of the brainstem ([Zambreanu, Wise, Brooks, Iannetti, & Tracey, 2005](#)). Exactly on which levels the pain perception is altered in chronic pain patients can be extremely difficult to pinpoint, making treatment an often exhausting process of trial and error. Suffering from chronic pain has dramatic impacts on the lives of the afflicted. Being in pain over a longer time can lead to depression, anxiety (fear or anticipation of future pain), and immobilization, all of which may in turn exacerbate pain ([Wiech & Tracey, 2009](#)). Negative emotion and attention to pain can increase sensitization to pain, possibly by keeping the descending pain modulatory system in facilitation mode. Distraction is therefore a commonly used technique in hospitals where patients have to undergo painful treatments like changing bandages on large burns. For chronic pain patients, however, diverting attention is not a long-term solution. Positive factors like social support can reduce the risk of chronic pain after an injury, and so they can help to adjust to bodily change as a result of injury. We have already talked about how having a hand to hold might alleviate suffering. Chronic pain treatment should target these emotional and social factors as well as the physiological.

THE POWER OF THE MIND



How powerful is the mind really? Well, the next time you have dental work done, know that some people opt for no Novocain, using only the power of their minds to overcome the pain. [Image: HypnoArt, CCO Public Domain, <https://goo.gl/m25gce>]

The context of pain and touch has a great impact on how we interpret it. Just imagine how different it would feel to Aron if someone amputated his hand against his will and for no discernible reason. Prolonged pain from injuries can be easier to bear if the incident causing them provides a positive context—like a war wound that testifies to a soldier's courage and

commitment—or **phantom pain** from a hand that was cut off to enable life to carry on. The relative meaning of pain is illustrated by a recent experiment, where the same moderately painful heat was administered to participants in two different contexts—one control context where the alternative was a nonpainful heat; and another where the alternative was an intensely painful heat. In the control context, where the moderate heat was the least preferable outcome, it was (unsurprisingly) rated as painful. In the other context it was the best possible outcome, and here the exact same moderately painful heat was actually rated as *pleasant*—because it meant the intensely painful heat had been avoided. This somewhat surprising change in perception—where pain becomes pleasant because it represents relief from something worse—highlights the

importance of the meaning individuals ascribe to their pain, which can have decisive effects in pain treatment ([Leknes et al., 2013](#)). In the case of touch, knowing who or what is stroking your skin can make all the difference—try thinking about slugs the next time someone strokes your skin if you want an illustration of this point. In a recent study, a group of heterosexual males were told that they were about to receive sensual caresses on the leg by either a male experimenter or by an attractive female experimenter ([Gazzola et al., 2012](#)). The study participants could not see who was touching them. Although it was always the female experimenter who performed the caress, the heterosexual males rated the otherwise pleasant sensual caresses as clearly unpleasant when they believed the male experimenter did it. Moreover, brain responses to the “male touch” in somatosensory cortex were reduced, exemplifying how top-down regulation of touch resembles top-down pain inhibition.

Pain and pleasure not only share modulatory systems—another common attribute is that we don’t need to be on the receiving end of it ourselves in order to experience it. How did you feel when you read about Aron cutting through his own tissue, or “Thomas” destroying his own bones unknowingly? Did you cringe? It’s quite likely that some of your brain areas processing affective aspects of pain were active even though the nociceptors in your skin and deep tissue were not firing. Pain can be experienced vicariously, as can itch, pleasurable touch, and other sensations. Tania Singer and her colleagues found in an fMRI study that some of the same brain areas that were active when participants felt pain on their own skin (anterior cingulate and insula)

were also active when they were given a signal that a loved one was feeling the pain. Those who were most “empathetic” also showed the largest brain responses ([Singer et al., 2004](#)). A similar effect has been found for pleasurable touch: The posterior insula of participants watching videos of someone else’s arm being gently stroked shows the same activation as if they were receiving the touch themselves ([Morrison, Bjornsdotter, & Olausson, 2011](#)).

SUMMARY

Sensory experiences connect us to the people around us, to the rest of the world, and to our own bodies. Pleasant or unpleasant, they’re part of being human. In this module, we have seen how being able to inhibit pain responses is central to our survival—and in cases like that of climber Aron Ralston, that ability can allow us to do extreme things. We have also seen how important the ability to feel pain is to our health—illustrated by young “Thomas,” who keeps injuring himself because he simply doesn’t notice pain. While “Thomas” has to learn to avoid harmful activities without the sensory input that normally guides us, G. L. has had to learn how to keep approaching and move about in a world she can hardly feel at all, with a body that is practically disconnected from her awareness. Too little sensation or too much of it leads to no good, no matter how pleasant or unpleasant the sensation usually feels. As long as we have nervous systems that function normally, we are able to adjust the volume of the sensory signals and our behavioral reactions according to the context we’re in. When it comes to sensory signals like touch and pain, we are

interpreters, not measuring instruments. The quest for understanding how our sensory–processing mechanisms can be modulated, psychologically and physiologically, promises to help researchers and clinicians find new ways to alleviate distress from chronic pain.

OUTSIDE RESOURCES

Book: Butler, D. S., Moseley, G. L., & Sunyata. (2003). *Explain pain* (p. 19). Australia: Noigroup.

Book: Kringelbach, M. L., & Berridge, K. C. (Eds.). (2010). *Pleasures of the brain* (p. 343). Oxford, UK: Oxford University Press.

Book: Ralston, A. (2004). *Between a rock and a hard place: The basis of the motion picture 127 Hours*. New York, NY: Atria.

Book: Sacks, O. (1998). *The man who mistook his wife for a hat: And other clinical tales*. New York, NY: Simon & Schuster.

Video: BBC Documentary “Pleasure and Pain with Michael Mosley”

<http://www.bbc.co.uk/programmes/b00y377q>

Video: TEDxAdelaide – Lorimer Moseley – “Why Things Hurt”



One or more interactive elements has been excluded from this version of the

text. You can view them online here:

<https://umssystem.pressbooks.pub/ubb2e/?p=113#oembed-1>

Video: Trailer for the film *127 Hours*, directed by Danny Boyle and released in 2010



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umssystem.pressbooks.pub/ubb2e/?p=113#oembed-2>

Web: Homepage for the International Association for the Study of Pain

<http://www.iasp-pain.org>

Web: Proceedings of the National Academy of Sciences Colloquium “The Neurobiology of Pain”

<http://www.pnas.org/content/96/14.toc#COLLOQUIUM>

Web: Stanford School of Medicine Pain Management Center

<http://paincenter.stanford.edu/>

Website resource aiming to communicate “advances and issues in the clinical sciences as they relate to the

**role of the brain and mind in chronic pain disorders,”
led by Dr. Lorimer Moseley**

<http://www.bodyinmind.org/>

DISCUSSION QUESTIONS

1. Your friend has had an accident and there is a chance the injury might cause pain over a prolonged period. How would you support your friend? What would you say and do to ease the pain, and why do you think it would work?
2. We have learned that touch and pain sensation in many aspects do not reflect “objectively” the outside world or the body state. Rather, these experiences are shaped by various top-down influences, and they can even occur without any peripheral activation. This is similar to the way other sensory systems work, e.g., the visual or auditory systems, and seems to reflect a general way the brain process sensory events. Why do you think the brain interprets the incoming sensory information instead of giving a one-to-one readout the way a thermometer and other measuring instruments would? Imagine you instead had “direct unbiased access” between stimuli and sensation. What would be the advantages and disadvantages of this?
3. Feelings of pain or touch are subjective—they have a particular quality that you perceive subjectively. How can we know whether the pain you feel is similar to the pain I feel? Is it possible that modern scientists can objectively measure such subjective feelings?

VOCABULARY

A-fibers

Fast-conducting sensory nerves with myelinated axons. Larger diameter and thicker myelin sheaths increases conduction speed. A β -fibers conduct touch signals from low-threshold mechanoreceptors with a velocity of 80 m/s and a diameter of 10 μm ; A δ -fibers have a diameter of 2.5 μm and conduct cold, noxious, and thermal signals at 12 m/s. The third and fastest conducting A-fiber is the A α , which conducts proprioceptive information with a velocity of 120 m/s and a diameter of 20 μm .

Allodynia

Pain due to a stimulus that does not normally provoke pain, e.g., when a light, stroking touch feels painful.

Analgesia

Pain relief.

C-fibers

C-fibers: Slow-conducting unmyelinated thin sensory afferents with a diameter of 1 μm and a conduction velocity of approximately 1 m/s. C-pain fibers convey noxious, thermal, and heat signals; C-tactile fibers convey gentle touch, light stroking.

Chronic pain

Persistent or recurrent pain, beyond usual course of acute illness or injury; sometimes present without observable tissue damage or clear cause.

C-pain or A δ -fibers

C-pain fibers convey noxious, thermal, and heat signals

C-tactile fibers

C-tactile fibers convey gentle touch, light stroking

Cutaneous senses

The senses of the skin: tactile, thermal, pruritic (itchy), painful, and pleasant.

Descending pain modulatory system

A top-down pain-modulating system able to inhibit or facilitate pain. The pathway produces analgesia by the release of endogenous opioids. Several brain structures and nuclei are part of this circuit, such as the frontal lobe areas of the anterior cingulate cortex, orbitofrontal cortex, and insular cortex; and nuclei in the amygdala and the hypothalamus, which all project to a structure in the midbrain called the periaqueductal grey (PAG). The PAG then controls ascending pain transmission from the afferent pain system indirectly through the rostral ventromedial medulla (RVM) in the brainstem, which uses ON- and OFF-cells to inhibit or facilitate nociceptive signals at the spinal dorsal horn.

Endorphin

An endogenous morphine-like peptide that binds to the opioid receptors in the brain and body; synthesized in the body's nervous system.

Exteroception

The sense of the external world, of all stimulation

originating from outside our own bodies.

Interoception

The sense of the physiological state of the body. Hunger, thirst, temperature, pain, and other sensations relevant to homeostasis. Visceral input such as heart rate, blood pressure, and digestive activity give rise to an experience of the body's internal states and physiological reactions to external stimulation. This experience has been described as a representation of “the material me,” and it is hypothesized to be the foundation of subjective feelings, emotion, and self-awareness.

Nociception

The neural process of encoding noxious stimuli, the sensory input from nociceptors. Not necessarily painful, and crucially not necessary for the experience of pain.

Nociceptors

High-threshold sensory receptors of the peripheral somatosensory nervous system that are capable of transducing and encoding noxious stimuli. Nociceptors send information about actual or impending tissue damage to the brain. These signals can often lead to pain, but nociception and pain are not the same.

Noxious stimulus

A stimulus that is damaging or threatens damage to normal tissues.

Ocial touch hypothesis

Proposes that social touch is a distinct domain of touch. C-tactile afferents form a special pathway that distinguishes social touch from other types of touch by selectively firing in response to touch of social-affective relevance; thus sending affective information parallel to the discriminatory information from the A β -fibers. In this way, the socially relevant touch stands out from the rest as having special positive emotional value and is processed further in affect-related brain areas such as the insula.

Pain

Defined as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage,” according to the International Association for the Study of Pain.

Phantom pain

Pain that appears to originate in an amputated limb.

Placebo effect

Effects from a treatment that are not caused by the physical properties of a treatment but by the meaning ascribed to it. These effects reflect the brain's own activation of modulatory systems, which is triggered by positive expectation or desire for a successful treatment. Placebo analgesia is the most well-studied placebo effect and has been shown to depend, to a large degree, on opioid mechanisms. Placebo analgesia can be reversed by the pharmacological blocking of μ opioid receptors. The

word “placebo” is probably derived from the Latin word “placebit” (“it will please”).

Sensitization

Increased responsiveness of nociceptive neurons to their normal input and/or recruitment of a response to normally subthreshold inputs. Clinically, sensitization may only be inferred indirectly from phenomena such as hyperalgesia or allodynia. Sensitization can occur in the central nervous system (central sensitization) or in the periphery (peripheral sensitization).

Social touch hypothesis

Proposes that social touch is a distinct domain of touch. C-tactile afferents form a special pathway that distinguishes social touch from other types of touch by selectively firing in response to touch of social-affective relevance; thus sending affective information parallel to the discriminatory information from the A β -fibers. In this way, the socially relevant touch stands out from the rest as having special positive emotional value and is processed further in affect-related brain areas such as the insula.

Somatosensory cortex

Consists of primary sensory cortex (S1) in the postcentral gyrus in the parietal lobes and secondary somatosensory cortex (S2), which is defined functionally and found in the upper bank of the lateral sulcus, called the parietal operculum. Somatosensory cortex also includes parts of the insular cortex.

Somatotopically organized

When the parts of the body that are represented in a particular brain region are organized topographically according to their physical location in the body (see Figure 2 illustration).

Spinothalamic tract

Runs through the spinal cord's lateral column up to the thalamus. C-fibers enter the dorsal horn of the spinal cord and form a synapse with a neuron that then crosses over to the lateral column and becomes part of the spinothalamic tract.

Transduction

The mechanisms that convert stimuli into electrical signals that can be transmitted and processed by the nervous system. Physical or chemical stimulation creates action potentials in a receptor cell in the peripheral nervous system, which is then conducted along the axon to the central nervous system.

REFERENCES

- Auvray, M., Myin, E., & Spence, C. (2010). The sensory-discriminative and affective-motivational aspects of pain. *Neuroscience and Biobehavioral Reviews*, 34(2), 214–223.
- Committee on Advancing Pain Research, Care, & Institute of Medicine. (2011). *Relieving pain in America: A blueprint for transforming prevention, care, education, and research*. Washington, D.C.: National Academy Press.
- Craig, A. D. (2002). How do you feel?

Interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience*, 3(8), 655–666.

- Dum, J., & Herz, A. (1984). Endorphinergic modulation of neural reward systems indicated by behavioral changes. *Pharmacology Biochemistry and Behavior*, 21(2), 259–266.
- Dunbar, R. I. (2010). The social role of touch in humans and primates: behavioural function and neurobiological mechanisms. *Neuroscience & Biobehavioral Reviews*, 34(2), 260–268.
- Eippert, F., Bingel, U., Schoell, E. D., Yacubian, J., Klinger, R., Lorenz, J., & Buchel, C. (2009). Activation of the opioidergic descending pain control system underlies placebo analgesia. *Neuron*, 63(4), 533–543.
- Eippert, F., Finsterbusch, J., Bingel, U., & Buchel, C. (2009). Direct evidence for spinal cord involvement in placebo analgesia. *Science*, 326(5951), 404.
- Eisenberger, N. I., Master, S. L., Inagaki, T. K., Taylor, S. E., Shirinyan, D., Lieberman, M. D., & Naliboff, B. D. (2011). Attachment figures activate a safety signal-related neural region and reduce pain experience. *Proceedings of the National Academy of Sciences of the United States of America*, 108(28), 11721–11726.
- Feldman, R., Singer, M., & Zagoory, O. (2010). Touch attenuates infants' physiological reactivity to stress. *Developmental Science*, 13(2), 271–278.

- Fields, H. L. (2006). A motivation-decision model of pain: The role of opioids. In H. Flor, E. Kalso, & J. O. Dostrovsky (Eds.), *Proceedings of the 11th World Congress on Pain* (pp. 449–459). Seattle, WA: IASP.
- Fields, H. L. (2004). State-dependent opioid control of pain. *Nature Reviews Neuroscience*, 5(7), 565–575.
- Fleming, A. S., O'Day, D. H., & Kraemer, G. W. (1999). Neurobiology of mother-infant interactions: Experience and central nervous system plasticity across development and generations. *Neuroscience and Biobehavioral Reviews*, 23(5), 673–685.
- Gazzola, V., Spezio, M. L., Etzel, J. A., Castelli, F., Adolphs, R., & Keysers, C. (2012). Primary somatosensory cortex discriminates affective significance in social touch. *Proceedings of the National Academy of Sciences of the United States of America*, 109(25), E1657–1666.
- Harlow, H. F., & Suomi, S. J. (1970). Nature of love—simplified. *American Psychologist*, 25(2), 161–168.
- Keverne, E. B., Martensz, N. D., & Tuite, B. (1989). Beta-endorphin concentrations in cerebrospinal fluid of monkeys are influenced by grooming relationships. *Psychoneuroendocrinology*, 14(1–2), 155–161.
- Leknes, S., & Tracey, I. (2008). A common neurobiology for pain and pleasure. *Nature Reviews Neuroscience*, 9(4), 314–320.

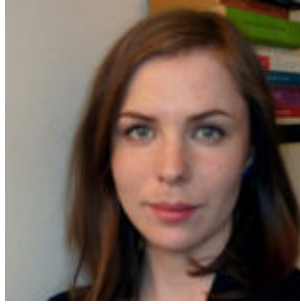
- Leknes, S., Berna, C., Lee, M. C., Snyder, G. D., Biele, G., & Tracey, I. (2013). The importance of context: When relative relief renders pain pleasant. *Pain*, 154(3), 402–410.
- Levine, J. D., Gordon, N. C., & Fields, H. L. (1978). The mechanism of placebo analgesia. *Lancet*, 2(8091), 654–657.
- Liljencrantz, J., Björnsdotter, M., Morrison, I., Bergstrand, S., Ceko, M., Seminowicz, D. A., . . . Olausson, H. (2013). Altered C-tactile processing in human dynamic tactile allodynia. *Pain*, 154(2), 227–234.
- Löken, L. S., Wessberg, J., Morrison, I., McGlone, F., & Olausson, H. (2009). Coding of pleasant touch by unmyelinated afferents in humans. *Nature Neuroscience*, 12(5), 547–548.
- McGlone, F., & Reilly, D. (2010). The cutaneous sensory system. *Neuroscience & Biobehavioral Reviews*, 34(2), 148–159.
- Minde, J., Toolanen, G., Andersson, T., Nennesmo, I., Remahl, I. N., Svensson, O., & Solders, G. (2004). Familial insensitivity to pain (HSAN V) and a mutation in the NGFB gene. A neurophysiological and pathological study. *Muscle & Nerve*, 30(6), 752–760.
- Morrison, I., Björnsdotter, M., & Olausson, H. (2011). Vicarious responses to social touch in posterior insular cortex are tuned to pleasant caressing speeds. *Journal of Neuroscience*, 31(26), 9554–9562.

- Morrison, I., Löken, L. S., & Olausson, H. (2010). The skin as a social organ. *Experimental Brain Research*, 204(3), 305–314.
- Olausson, H., Lamarre, Y., Backlund, H., Morin, C., Wallin, B. G., Starck, G., . . . Bushnell, M. C. (2002). Unmyelinated tactile afferents signal touch and project to insular cortex. *Nature Neuroscience*, 5(9), 900–904.
- Olausson, H., Wessberg, J., Morrison, I., McGlone, F., & Vallbo, Å. (2010). The neurophysiology of unmyelinated tactile afferents. *Neuroscience & Biobehavioral Reviews*, 34(2), 185–191.
- Pechtel, P., & Pizzagalli, D. A. (2011). Effects of early life stress on cognitive and affective function: An integrated review of human literature. *Psychopharmacology*, 214(1), 55–70.
- Petrovic, P., Kalso, E., Petersson, K. M., & Ingvar, M. (2002). Placebo and opioid analgesia—Imaging a shared neuronal network. *Science*, 295(5560), 1737–1740.
- Price, D. D. (2000). Psychological and neural mechanisms of the affective dimension of pain. *Science*, 288(5472), 1769–1772.
- Rainville, P., Duncan, G. H., Price, D. D., Carrier, B., & Bushnell, M. C. (1997). Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science*, 277(5328), 968–971.
- Rainville, P., Feine, J. S., Bushnell, M. C., & Duncan, G. H. (1992). A psychophysical

comparison of sensory and affective responses to four modalities of experimental pain. *Somatosensory and Motor Research*, 9(4), 265–277.

- Ralston, A. (2010, December 31). Aron Ralston's 127 hours: This is going to make one hell of a story. . . Retrieved from The Telegraph: <http://www.telegraph.co.uk/culture/books/8233734/Aron-Ralstons-127-Hours-This-is-going-to-make-one-hell-of-a-story.-.html>
- Sacks, O. (1985). The disembodied lady. In *The man who mistook his wife for a hat and other clinical tales* (pp. 43–54) New York, NY: Simon & Schuster.
- Simons, D. J., & Land, P. W. (1987). Early experience of tactile stimulation influences organization of somatic sensory cortex. *Nature*, 326(6114), 694–697.
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. J., & Frith, C. D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, 303(5661), 1157–1162.
- Wiech, K., & Tracey, I. (2009). The influence of negative emotions on pain: Behavioral effects and neural mechanisms. *NeuroImage*, 47(3), 987–994.
- Zambreanu, L., Wise, R. G., Brooks, J. C. W., Iannetti, G. D., & Tracey, I. (2005). A role for the brainstem in central sensitisation in humans. Evidence from functional magnetic resonance imaging. *Pain*, 114(3), 397–407.

AUTHORS



Guro E. Løseth

Guro Engvig Løseth is a former science journalist who turned to psychology and cognitive neuroscience after realizing that the really interesting questions are better asked on the other side of the microphone. In her current work at the University of Oslo, Norway, she investigates the neurochemistry of affective touch through psychopharmacological studies and psychophysical investigations using self-report and autonomic measures.



Dan-Mikael Ellingsen

Dan-Mikael Ellingsen is a cognitive neuroscientist interested in how hedonic feelings are shaped by top-down influences like expectations, context, and social cues. In his doctoral work at the University of Gothenburg, Sweden, he combines neuroimaging and psychophysics with pharmacological and cognitive manipulations to investigate how pain and pleasant touch experiences are modulated by cognitive and cross-modal influences.



Siri Leknes

Siri Graff Leknes is a cognitive neuroscientist, working as a Research Fellow at the University of Oslo, Norway. Her academic training includes a D.Phil. on 'Pain, Pleasure and Relief' at Oxford University, UK. Her research revolves around the subjective hedonic feelings associated with rewards and punishments, and individual differences in the subjective experience of pain and pleasure.

CREATIVE COMMONS LICENSE



TOUCH AND PAIN by [Guro E. Løseth, Dan-Mikael Ellingson, and Siri Leknes](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

PART VII.

**UNIT 7: THE VESTIBULAR
SYSTEM**

CHAPTER 11.

THE VESTIBULAR SYSTEM

DORA ANGELAKI AND J. DAVID DICKMAN

The vestibular system functions to detect head motion and position relative to gravity and is primarily involved in the fine control of visual gaze, posture, orthostasis, spatial orientation, and navigation. Vestibular signals are highly processed in many regions of the brain and are involved in many essential functions. In this module, we provide an overview of how the vestibular system works and how vestibular signals are used to guide behavior.

LEARNING OBJECTIVES

- Define the basic structures of the vestibular receptor system.
- Describe the neuroanatomy of the vestibuloocular, vestibulospinal, and vestibulo-thalamo-cortical pathways.
- Describe the vestibular commissural system.
- Describe the different multisensory cortical areas for motion perception.

INTRODUCTION

Remember the dizzy feeling you got as a child after you jumped off the merry-go-round or spun around like a top? These feelings result from activation of the **vestibular system**, which detects our movements through space but is not a conscious sense like vision or hearing. In fact, most vestibular functions are imperceptible, but vestibular-related sensations such as motion sickness can pop up rapidly when riding on a roller coaster, having a bumpy plane ride, or a sailing a boat in rough seas. However, these sensations are really side effects and the vestibular system is actually extremely important for everyday activities, with vestibular signals being involved in much of the brain's information processing that controls such fundamental functions as balance, posture, gaze stabilization, spatial orientation, and navigation, to name a few. In many regions of the brain, vestibular information is combined with signals from the other senses as well as with motor information to give rise to motion perception, body awareness, and behavioral control. Here, we will explore the workings of the vestibular system and consider some of the integrated computations the brain performs using vestibular signals to guide our common behavior.



Although you may be most aware of the vestibular system in situations where you experience motion sickness, at other times its working in the background to provide you with essential functions like balance and stable vision. [Image: Leo Reynolds, <https://goo.gl/N8D94w>, CC BY-NC-SA 2.0, <https://goo.gl/Toc0ZF>]

STRUCTURE OF THE VESTIBULAR RECEPTORS

The vestibular receptors lie in the inner ear next to the auditory cochlea. They detect rotational motion (head turns), linear motion (translations), and tilts of the head relative to gravity and transduce these motions into neural signals that can be sent to the brain. There are five vestibular receptors in each ear (Hearing module, Figure 1 - <http://noba.to/jry3cu78>), including three **semicircular canals** (horizontal, anterior, and posterior) that transduce rotational angular accelerations and two **otolith receptors** (utricle and saccule) that

transduce linear accelerations ([Lindeman, 1969](#)). Together, the semicircular canals and otolith organs can respond to head motion and maintained static head position relative to gravity in all directions in 3D space.

These receptors are contained in a series of interconnected fluid filled tubes that are protected by a dense, overlying bone ([Iurato, 1967](#)). Each of the three semicircular canals lies in a plane that is orthogonal to the other two. The horizontal semicircular canal lies in a roughly horizontal head plane, whereas the anterior and posterior semicircular canals lie vertically in the head ([Blanks, Curthoys, Bennett, & Markham, 1985](#)). The semicircular canal receptor cells, termed **hair cells**, are located only in the middle of the circular tubes in a special epithelium, covered by a gelatinous membrane that stretches across the tube to form a fluid-tight seal like the skin of a drum (Figures 1A and 1B). Hair cells are so named due to an array of nearly 100 staggered-height **stereocilia** (like a church pipe organ) that protrude from the top of the cell into the overlying gelatinous membrane ([Wersäll, 1956](#)). The shortest stereocilia are at one end of the cell and the tallest at the other ([Lindeman, 1969](#)). When the head is rotated, the fluid in the semicircular canals lags behind the head motion and pushes on the gelatinous membrane, which bends the stereocilia.

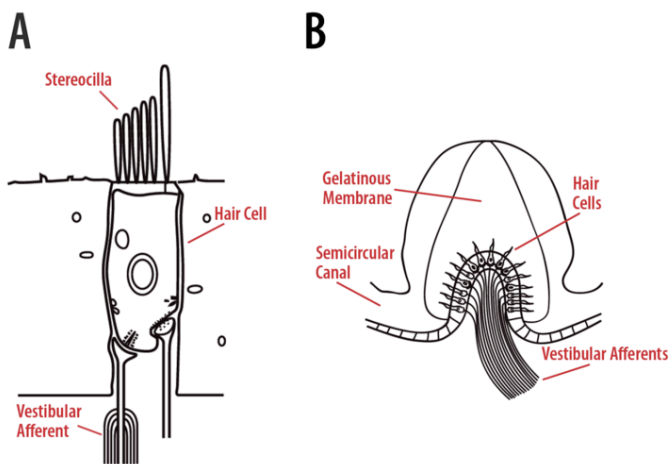


Figure 1. Receptor hair cells and semicircular canal structure. A) Vestibular hair cell in the receptor epithelium with stereocilia in the apical surface of the cell. Innervating afferent and efferent neurons make synaptic contacts with the basal surface of the cell. B) Semicircular canal structure showing the fluid duct, the hair cell stereocilia embedded in a gelatinous membrane on top of the hair cells, and the innervating afferent fibers.

As shown in Figure 2, when the head moves toward the receptor hair cells (e.g., left head turns for the left horizontal semicircular canal), the stereocilia are bent toward the tallest end and special **mechanically gated ion channels** in the tips of the cilia open, which excites (**depolarizes**) the cell ([Shotwell, Jacobs, & Hudspeth, 1981](#)). Head motion in the opposite direction causes bending toward the smallest stereocilia, which closes the channels and inhibits (**hyperpolarizes**) the cell. The left and right ear semicircular canals have opposite polarity, so for example, when you turn your head to the left, the receptors in the left horizontal semicircular canal will be excited while right ear horizontal canal receptors will be inhibited (Figure 3). The same relationship is true for the vertical semicircular canals. Vestibular **afferent nerve**

fibers innervate the base of the hair cell and increase or decrease their neural firing rate as the receptor cell is excited or inhibited ([Dickman and Correia, 1989](#)), respectively, and then carry these signals regarding head rotational motion to the brain as part of the **vestibulocochlear nerve** (Cranial nerve VIII). They enter the brainstem and terminate in the ipsilateral vestibular nuclei, cerebellum, and reticular formation ([Carleton & Carpenter, 1984](#); [Dickman & Fang, 1996](#)). The primary vestibular hair cell and afferent **neurotransmitters** are **glutamate** and **aspartate**. Due to the mechanical properties of the vestibular receptor system, rotational accelerations of the head are integrated into velocity signals ([Van Egmond, Groen, & Jongkess, 1949](#)) that are then encoded by semicircular canal afferents ([Fernandez & Goldberg, 1971](#)). **Detection thresholds** for rotational motion have shown that afferents can discriminate differences in head velocity on the order of 2 deg/sec, but also are sensitive to a broad range of natural head movements up to high head speeds in the hundreds of deg/sec (as you might experience when you make a fast head turn toward a loud sound, or are performing gymnastics; [Sadeghi, Chacron, Taylor, & Cullen, 2007](#); [Yu, Dickman, & Angelaki, 2012](#)).

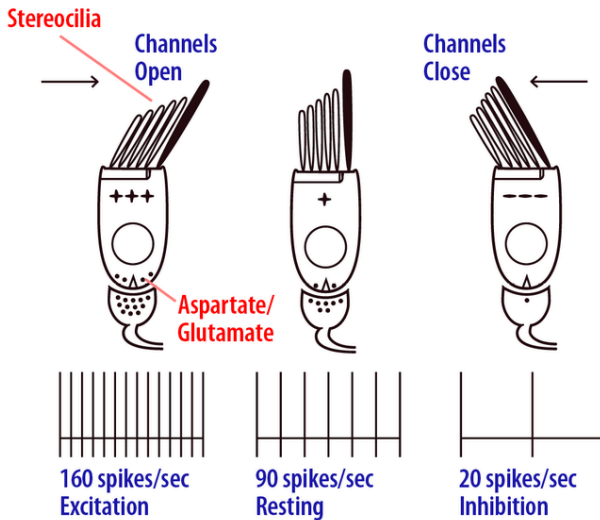


Figure 2. Vestibular receptor cell directional selectivity. Middle) At rest, hair cells release some neurotransmitter, producing a high, spontaneous firing rate in the innervating afferent fibers. Left) When the stereocilia are displaced toward the kinocilium, the cell is depolarized and the afferent firing rate is increased. Right) When the stereocilia are displaced away from the kinocilium, the cell is hyperpolarized and the afferent firing rate is decreased.

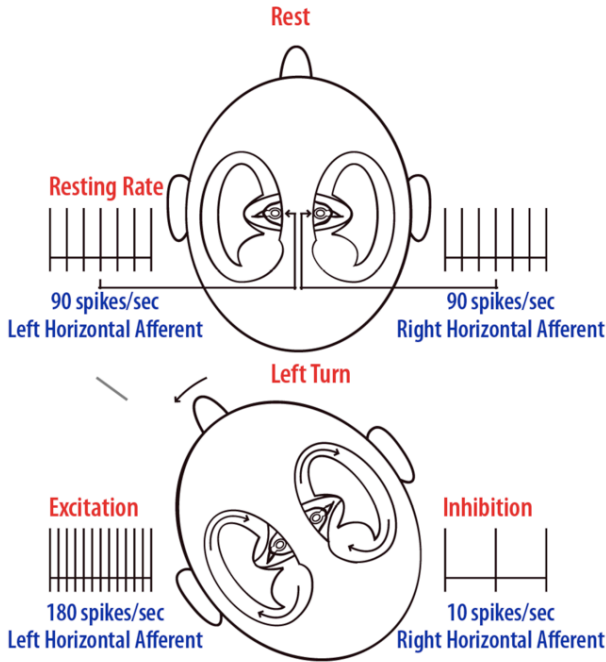


Figure 3. Receptor hair cells in the otolith organs. Receptor cells have stereocilia embedded in the gelatinous membrane, which is covered by thousands of calcium carbonate otoconia. Receptor cells are polarized in opposite directions relative to a central location and are innervated by VIIIth nerve afferent fibers.

Otolith receptors are sensitive to linear accelerations and tilts of the head relative to gravity ([Fernandez & Goldberg, 1976a](#)). The utricle otolith receptor lies parallel to the horizontal semicircular canal and the saccule receptor lies vertical in the head (Hearing module, Figure 1- <http://noba.to/jry3cu78>). As shown in Figure 4, a special otolith epithelium contains receptor hair cells whose stereocilia extend into a gelatin membrane that is covered by a layer of calcium carbonate crystals, termed **otoconia**, like rocks piled up to form a jetty ([Lindeman, 1969](#)). Otoconia are not affected by fluid

movements but instead are displaced by linear accelerations, including translations (e.g., forward/backward or upward/downward motions) or changes in head position relative to gravity. These linear accelerations produce displacements of the otoconia (due to their high mass), much like rocks rolling down a hill or your coffee cup falling off the car dashboard when you push the gas pedal. Movements of the otoconia bend the hair cell stereocilia and open/close channels in a similar way to that described for the semicircular canals. However, otolith hair cells are polarized such that the tallest stereocilia are pointing toward the center of the utricle and away from the center in the saccule, which effectively splits the receptors into two opposing groups ([Flock, 1964](#); [Lindeman, 1969](#)). In this way, some hair cells are excited and some inhibited for each linear motion force or head tilt experienced, with the population of receptors and their innervating afferents being **directionally tuned** to all motions or head tilts in 3D space ([Fernandez & Goldberg, 1976b](#)).

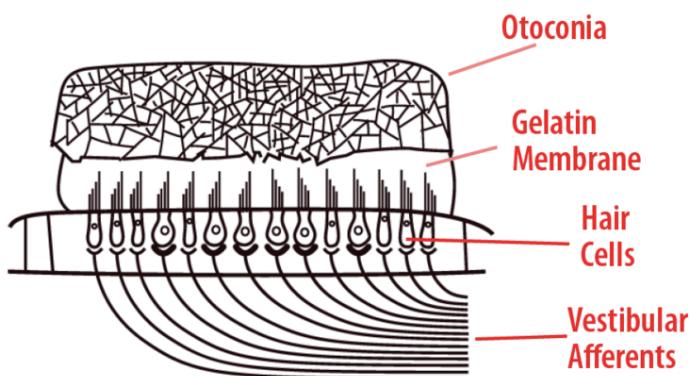


Figure 4. Vestibular response to horizontal plane head rotation. A) When the head is stationary, afferent fibers on both the sides of the head have equivalent firing so there is no sense of motion. B) When the head turns to the left, all of the left horizontal semicircular canal hair cells are excited and afferent fibers increase their firing rate. Conversely, right horizontal canal afferents decrease their firing rate.

All vestibular hair cells and afferents receive connections from **vestibular efferents**, which are fibers projecting from the brain out to the vestibular receptor organs, whose function is not well understood. It is thought that efferents control the sensitivity of the receptor ([Boyle, Carey, & Highstein, 1991](#)). The primary efferents neurotransmitter is **acetylcholine** ([Anniko & Arnold, 1991](#)).

THE VESTIBULAR NUCLEI

The vestibular nuclei comprise a large set of neural elements in the brainstem that receive motion and other multisensory signals, then regulate movement responses and sensory experience. Many vestibular nuclei neurons have reciprocal connections with the cerebellum that

form important regulatory mechanisms for the control of eye movements, head movements, and posture. There are four major vestibular nuclei that lie in the rostral medulla and caudal pons of the brainstem; all receive direct input from vestibular afferents ([Brodal, 1984](#); [Precht & Shimazu, 1965](#)). Many of these nuclei neurons receive convergent motion information from the opposite ear through an inhibitory commissural pathway that uses **gamma-aminobutyric acid** (GABA) as a neurotransmitter ([Kasahara & Uchino, 1974](#); [Shimazu & Precht, 1966](#)). The commissural pathway is highly organized such that cells receiving horizontal excitatory canal signals from the ipsilateral ear will also receive contralateral inhibitory horizontal canal signals from the opposite ear. This fact gives rise to a “push-pull” vestibular function, whereby directional sensitivity to head movement is coded by opposing receptor signals. Because vestibular nuclei neurons receive information from bilateral inner ear receptors and because they maintain a high spontaneous firing rate (nearly 100 impulses/sec), they are thought to act to “compare” the relative discharge rates of left vs. right canal afferent firing activity. For example, during a leftward head turn, left brainstem nuclei neurons receive high firing-rate information from the left horizontal canal and low firing-rate information from the right horizontal canal. The comparison of activity is interpreted as a left head turn. Similar nuclei neuron responses exist when the head is pitched or rolled, with the vertical semicircular canals being stimulated by the rotational motion in their sensitivity planes. However, the opposing push-pull response from the vertical canals occurs with the anterior semicircular canal in one ear and the co-planar posterior semicircular canal

of the opposite ear. Damage or disease that interrupts inner ear signal information from one side of the head can change the normal resting activity in the VIIIth nerve afferent fibers and will be interpreted by the brain as a head rotation, even though the head is stationary. These effects often lead to illusions of spinning or rotating that can be quite upsetting and may produce nausea or vomiting. However, over time the commissural fibers provide for **vestibular compensation**, a process by which the loss of unilateral vestibular receptor function is partially restored centrally and behavioral responses, such as the **vestibuloocular reflex** (VOR) and postural responses, mostly recover ([Beraneck et al., 2003](#); [Fetter & Zee, 1988](#); [Newlands, Hesse, Haque, & Angelaki, 2001](#); [Newlands & Perachio, 1990](#)).

In addition to the commissural pathway, many vestibular nuclei neurons receive **proprioceptive** signals from the spinal cord regarding muscle movement and position, visual signals regarding spatial motion, other multisensory (e.g., trigeminal) signals, and higher order signals from the cortex. It is thought that the cortical inputs regulate fine gaze and postural control, as well as suppress the normal **compensatory reflexes** during motion in order to elicit volitional movements. Of special significance are convergent signals from the semicircular canal and otolith afferents that allow central vestibular neurons to compute specific properties of head motion ([Dickman & Angelaki, 2002](#)). For example, Einstein ([1907](#)) showed that linear accelerations are equivalent whether they arise from translational motion or from tilts of the head relative to gravity. The otolith receptors cannot discriminate between the two, so how is it that we can tell the difference between when we are translating

forward and tilting backward, where the linear acceleration signaled by the otolith afferents is the same? Vestibular nuclei and cerebellar neurons use convergent signals from both the semicircular canals and the otolith receptors to discriminate between tilt and translation, and as a result, some cells encode head tilt ([Zhou, 2006](#)) while other cells encode translational motion ([Angelaki, Shaikh, Green, & Dickman, 2004](#)).

VESTIBULOCULAR SYSTEM

The vestibular system is responsible for controlling **gaze stability** during motion ([Crane & Demer, 1997](#)). For example, if we want to read the sign in a store window while walking by, we must maintain foveal fixation on the words while compensating for the combined rotational and translational head movements incurred during our stride. The vestibular system regulates compensatory eye, neck, spinal, and limb movements in order to maintain gaze ([Keshner & Peterson, 1995](#)). One of the major components contributing to gaze stability is the VOR, which produces reflexive eye movements that are equal in magnitude and opposite in direction to the perceived head motion in 3D space ([Wilson et al., 1995](#)). The VOR is so accurate and fast that it allows people to maintain visual fixation on objects of interest while experiencing demanding motion conditions, such as running, skiing, playing tennis, and driving. In fact, gaze stabilization in humans has been shown to be completely compensatory (essentially perfect) for most natural behaviors. To produce the VOR, vestibular neurons must control each of the six pairs of eye muscles in unison through a specific set of connections to the **oculomotor nuclei** ([Ezure &](#)

[Graf, 1984](#)). The anterior and posterior semicircular canals along with the saccule control vertical and torsional (turning of the eye around the line of sight) eye movements, while the horizontal canals and the utricle control horizontal eye movements.

To understand how the VOR works, let's take the example of the compensatory response for a leftward head turn while reading the words on a computer screen. The basic pathway consists of horizontal semicircular canal afferents that project to specific neurons in the vestibular nuclei. These nuclei cells, in turn, send an excitatory signal to the contralateral **abducens nucleus**, which projects through the sixth cranial nerve to innervate the **lateral rectus muscle** (Figure 5). Some abducens neurons send an excitatory projection back across the midline to a subdivision of cells in the ipsilateral **oculomotor nucleus**, which, in turn, projects through the third cranial nerve to innervate the right (ipsilateral) medial rectus muscle. When a leftward head turn is made, the left horizontal canal vestibular afferents will increase their firing rate and consequently increase the activity of vestibular nuclei neurons projecting to the opposite (contralateral) right abducens nucleus. The abducens neurons produce contraction of the right lateral rectus and, through a separate cell projection to the left oculomotor nucleus, excite the left medial rectus muscles. In addition, matching bilateral inhibitory connections relax the left lateral rectus and right medial rectus eye muscles. The resulting rightward eye movement for both eyes stabilizes the object of interest upon the retina for greatest visual acuity.

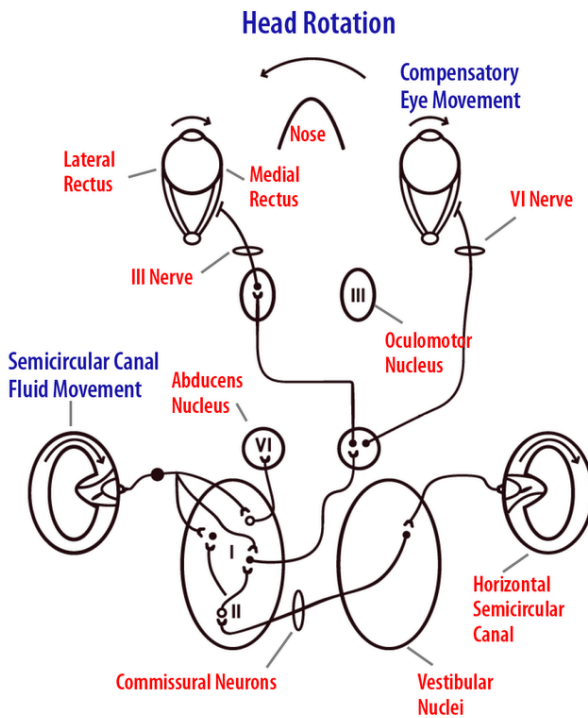


Figure 5. Vestibuloocular reflex. During a leftward head turn, the left horizontal semicircular canal receptors are excited, while the right ear receptors are inhibited. The left excitatory signals excite vestibular nuclei neurons. These cells project across the brain to excite motor neurons in the right abducens nucleus (VI) that excite the lateral rectus muscle of the right eye and to cells in the oculomotor nucleus (III) that excite the medial rectus muscle of the left eye. This moves both eyes to the right to exactly match the leftward head movement and stabilize visual gaze upon a target of interest. The right ear inhibitory signals cross to neurons in the left vestibular nucleus that decrease their firing rate. These cells are inhibitory and decrease their firing rate to further increase the response of rightward motor eye muscle cells.

During linear translations, a different type of VOR also occurs ([Paige & Tomko, 1991](#)). For example, sideways motion to the left results in a horizontal rightward eye movement to maintain visual stability on an object of

interest. In a similar manner, vertical up–down head movements (such as occur while walking or running) elicit oppositely directed vertical eye movements ([Angelaki, McHenry, & Hess, 2000](#)). For these reflexes, the amplitude of the translational VOR depends on viewing distance. This is due to the fact that the **vergence angle** (i.e., the angle between the lines of sight for each eye) varies as a function of the inverse of the distance to the viewed visual object ([Schwarz, Busettoni, & Miles, 1989](#)). Visual objects that are far away (2 meters or more) require no vergence angle, but as the visual objects get closer (e.g., when holding your finger close to your nose), a large vergence angle is needed. During translational motion, the eyes will change their vergence angle as the visual object moves from close to farther away (or vice versa). These responses are a result of activation of the otolith receptors, with connections to the oculomotor nuclei similar to those described above for the rotational vestibuloocular reflex. With tilts of the head, the resulting eye movement is termed **torsion**, and consists of a rotational eye movement around the line of sight that is in the direction opposite to the head tilt. As mentioned above, there are major reciprocal connections between the vestibular nuclei and the cerebellum. It has been well established that these connections are crucial for adaptive motor learning in the vestibuloocular reflex ([Lisberger, Pavelko, & Broussard, 1994](#)).

VESTIBULO-SPINAL NETWORK

There are two vestibular descending pathways that regulate body muscle responses to motion and gravity, consisting of the **lateral vestibulo-spinal tract (LVST)** and the **medial vestibulo-spinal tract (MVST)**. Reflexive control of head and neck muscles arises through the neurons in the medial vestibulospinal tract (MVST). These neurons comprise the rapid vestibulocollic reflex (VCR) that serves to stabilize the head in space and participates in gaze



Similar to how the gyroscope in your cell phone can detect which angle it's at to adjust the screen, so do the MSVT neurons inform your body which direction it's headed in order to make the appropriate bodily adjustments.

[Image: Kleman Gellek, <https://goo.gl/DR9rpR>, CC BY 4.0, <https://goo.gl/QuGXFp>]

control ([Peterson, Goldber, Bilotto, & Fuller, 1985](#)). The MVST neurons receive input from vestibular receptors and the cerebellum, and somatosensory information from the spinal cord. MVST neurons carry both excitatory and inhibitory signals to innervate neck flexor and extensor motor neurons in the spinal cord. For example, if one trips over a crack in the pavement while walking, MVST neurons will receive downward and forward linear acceleration signals from the otolith receptors and forward rotation acceleration signals from the vertical semicircular canals. The VCR will compensate by providing excitatory signals to the dorsal

neck flexor muscles and inhibitory signals to the ventral neck extensor muscles, which moves the head upward and opposite to the falling motion to protect it from impact.

The LVST comprises a topographic organization of vestibular nuclei cells that receive substantial input from the cerebellum, proprioceptive inputs from the spinal cord, and convergent afferent signals from vestibular receptors. LVST fibers project ipsilaterally to many levels of motor neurons in the cord to provide coordination of different muscle groups for postural control ([Shinoda, Sugiuchi, Futami, Ando, & Kawasaki, 1994](#)). LVST neurons contain either acetylcholine or glutamate as a neurotransmitter and exert an excitatory influence upon extensor muscle motor neurons. For example, LVST fibers produce extension of the contralateral axial and limb musculature when the body is tilted sideways. These actions serve to stabilize the body's center of gravity in order to preserve upright posture.

VESTIBULO-AUTONOMIC CONTROL

Some vestibular nucleus neurons send projections to the reticular formation, dorsal pontine nuclei, and nucleus of the solitary tract. These connections regulate breathing and circulation through compensatory vestibular autonomic responses that stabilize respiration and blood pressure during body motion and changes relative to gravity. They may also be important for induction of motion sickness and emesis.

VESTIBULAR SIGNALS IN THE THALAMUS AND CORTEX

The cognitive perception of motion, spatial orientation, and navigation through space arises through multisensory information from vestibular, visual, and somatosensory signals in the thalamus and cortex (Figure 6A). Vestibular nuclei neurons project bilaterally to the several thalamic regions. Neurons in the ventral posterior group respond to either vestibular signals alone, or to vestibular plus somatosensory signals, and projects to primary somatosensory cortex (area 3a, 2v), somatosensory association cortex, posterior parietal cortex (areas 5 and 7), and the insula of the temporal cortex ([Marlinski & McCrea, 2008](#); [Meng, May, Dickman, & Angelaki, 2007](#)). The posterior nuclear group (PO), near the medial geniculate body, receives both vestibular and auditory signals as well as inputs from the superior colliculus and spinal cord, indicating an integration of multiple sensory signals. Some anterior pulvinar neurons also respond to motion stimuli and project to cortical area 3a, the posterior insula, and the temporo-parietal cortex (PIVC). In humans, electrical stimulation of the thalamic areas produces sensations of movement and sometimes dizziness.

Area 2v cells respond to motion, and electrical stimulation of this area in humans produces sensations of moving, spinning, or dizziness. Area 3a lies at the base of the central sulcus adjacent to the motor cortex and is thought to be involved in integrative motor control of the head and body ([Guldin, Akbarian, & Grusser, 1992](#)). Neurons in the PIVC are multisensory, responding to body motion, somatosensory, proprioceptive, and visual

motion stimuli ([Chen, DeAngelis, & Angelaki, 2011](#); [Grusser, Pause, & Schreiter, 1982](#)). PIVC and areas 3a and 2v are heavily interconnected. Vestibular neurons also have been observed in the posterior parietal cortex; in area 7, in the ventral intraparietal area (VIP), the medial intraparietal area (MIP), and the medial superior temporal area (MST). VIP contains multimodal neurons involved in spatial coding. MIP and MST neurons respond to body motion through space by multisensory integration of visual motion and vestibular signals ([Gu, DeAngelis, & Angelaki, 2007](#)) and many MST cells are directly involved in heading perception ([Gu, Watkins, Angelaki, & DeAngelis, 2006](#)). Lesions of the parietal cortical areas can result in confusions in spatial awareness. Finally, areas involved with the control of saccades and pursuit eye movements, including area 6, area 8, and the superior frontal gyrus, receive vestibular signals ([Fukushima, Sato, Fukushima, Shinmei, & Kaneko, 2000](#)). How these different cortical regions contribute to our perception of motion and spatial orientation is still not well understood.

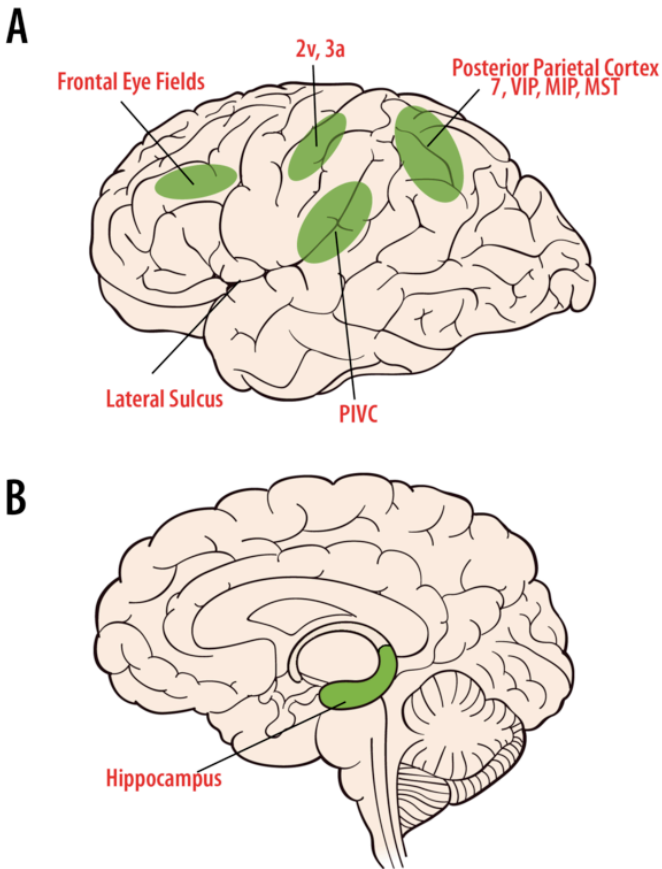


Figure 6. Cortical regions of the brain known to be involved with vestibular processing. A) The frontal eye fields control eye movements and receive vestibular motion information. Areas 2v and 3a are somatosensory areas that map body location and movement signals. Area PIVC responds to body and head motion information. The posterior parietal cortex is involved with motion perception and responds to both visual and vestibular motion cues. B) The hippocampus and parahippocampal regions are involved with spatial orientation and navigation functions. All receive vestibular signals regarding body and head motion.

SPATIAL ORIENTATION AND NAVIGATION

Our ability to know where we are and to navigate different spatial locations is essential for survival. It is believed that a cognitive map of our environment is created through exploration and then used for spatial orientation and navigation, such as driving to the store, or walking through a dark house ([McNaughton, Battaglia, Jensen, Moser, & Moser, 2006](#)). Cells in the limbic system and the hippocampus that contribute to these functions have been identified, including place cells, grid cells, and head direction cells (Figure 6B). Place cells in the hippocampus encode specific locations in the environment ([O'Keefe, 1976](#)). Grid cells in the entorhinal cortex encode spatial maps in a tessellated pattern ([Hafting, Fyhn, Molden, Moser, & Moser, 2005](#)). Head direction cells in the anterior-dorsal thalamus encode heading direction, independent of spatial location ([Taube, 1995](#)). It is thought that these cell types work together to provide for spatial orientation, spatial memory, and our ability to navigate. Both place cells and head direction cells depend upon a functioning vestibular system to maintain their directional and orientation information ([Stackman, Clark, & Taube, 2002](#)). The pathway by which vestibular signals reach the navigation network is not well understood; however, damage to the vestibular system, hippocampus, and dorsal thalamus regions often disrupts our ability to orient in familiar environments, navigate from place to place, or even to find our way home.

MOTION SICKNESS



A common treatment of motion sickness is Dramamine, which helps to reduce the sensitivity of input from your vestibular system to the rest of your body. [Image: Mike Baird, <https://goo.gl/zfeqqr>, CC BY 2.0, <https://goo.gl/BRvSA7>]

Although a number of conditions can produce motion sickness, it is generally thought that it is evoked from a mismatch in sensory cues between vestibular, visual, and proprioceptive signals ([Yates, Miller, & Lucot, 1998](#)). For example, reading a book in a car on a winding road can produce motion sickness, whereby the accelerations experienced by the vestibular system do not match the visual input.

However, if one looks out the window at the scenery going by during the same travel, no sickness occurs because the visual and vestibular cues are in alignment. Sea sickness, a form of motion sickness, appears to be a special case and arises from unusual vertical oscillatory and roll motion. Human studies have found that low frequency oscillations of 0.2 Hz and large amplitudes (such as found in large seas during a storm) are most likely to cause motion sickness, with higher frequencies offering little problems.

SUMMARY

Here, we have seen that the vestibular system transduces and encodes signals about head motion and position with

respect to gravity, information that is then used by the brain for many essential functions and behaviors. We actually understand a great deal regarding vestibular contributions to fundamental reflexes, such as compensatory eye movements and balance during motion. More recent progress has been made toward understanding how vestibular signals combine with other sensory cues, such as vision, in the thalamus and cortex to give rise to motion perception. However, there are many complex cognitive abilities that we know require vestibular information to function, such as spatial orientation and navigation behaviors, but these systems are only just beginning to be investigated. Future research regarding vestibular system function will likely be geared to seeking answers to questions regarding how the brain copes with vestibular signal loss. In fact, according to the National Institutes of Health, nearly 35% of Americans over the age of 40 (69 million people) have reported chronic vestibular-related problems. It is therefore of significant importance to human health to better understand how vestibular cues contribute to common brain functions and how better treatment options for vestibular dysfunction can be realized.

DISCUSSION QUESTIONS

1. If a person sustains loss of the vestibular receptors in one ear due to disease or trauma, what symptoms would the person suffer? Would the symptoms be permanent?
2. Often motion sickness is relieved when a person looks at far distance objects, such as things located on the far horizon. Why does far distance viewing

help in motion sickness while close distance view (like reading a map or book) make it worse?

3. Vestibular signals combine with visual signals in certain areas of cortex and assist in motion perception. What types of cues does the visual system provide for self motion through space? What types of vestibular signals would be consistent with rotational versus translational motion?

VOCABULARY

Abducens nucleus

A group of excitatory motor neurons in the medial brainstem that send projections through the VIth cranial nerve to control the ipsilateral lateral rectus muscle. In addition, abducens interneurons send an excitatory projection across the midline to a subdivision of cells in the ipsilateral oculomotor nucleus, which project through the IIIrd cranial nerve to innervate the ipsilateral medial rectus muscle.

Acetylcholine

An organic compound neurotransmitter consisting of acetic acid and choline. Depending upon the receptor type, acetylcholine can have excitatory, inhibitory, or modulatory effects.

Afferent nerve fibers

Single neurons that innervate the receptor hair cells and carry vestibular signals to the brain as part of the vestibulocochlear nerve (cranial nerve VIII).

Aspartate

An excitatory amino acid neurotransmitter that is widely used by vestibular receptors, afferents, and many neurons in the brain.

Compensatory reflexes

A stabilizing motor reflex that occurs in response to a perceived movement, such as the vestibuloocular reflex, or the postural responses that occur during running or skiing.

Depolarized

When receptor hair cells have mechanically gated channels open, the cell increases its membrane voltage, which produces a release of neurotransmitter to excite the innervating nerve fiber.

Detection thresholds

The smallest amount of head motion that can be reliably reported by an observer.

Directional tuning

The preferred direction of motion that hair cells and afferents exhibit where a peak excitatory response occurs and the least preferred direction where no response occurs. Cells are said to be “tuned” for a best and worst direction of motion, with in-between motion directions eliciting a lesser but observable response.

Gamma-aminobutyric acid

A major inhibitory neurotransmitter in the vestibular commissural system.

Gaze stability

A combination of eye, neck, and head responses that are all coordinated to maintain visual fixation (fovea) upon a point of interest.

Glutamate

An excitatory amino acid neurotransmitter that is widely used by vestibular receptors, afferents, and many neurons in the brain.

Hair cells

The receptor cells of the vestibular system. They are termed hair cells due to the many hairlike cilia that extend from the apical surface of the cell into the gelatin membrane. Mechanical gated ion channels in the tips of the cilia open and close as the cilia bend to cause membrane voltage changes in the hair cell that are proportional to the intensity and direction of motion.

Hyperpolarizes

When receptor hair cells have mechanically gated channels close, the cell decreases its membrane voltage, which produces less release of neurotransmitters to inhibit the innervating nerve fiber.

Lateral rectus muscle

An eye muscle that turns outward in the horizontal plane.

Lateral vestibulo-spinal tract

Vestibular neurons that project to all levels of the spinal cord on the ipsilateral side to control posture

and balance movements.

Mechanically gated ion channels

Ion channels located in the tips of the stereocilia on the receptor cells that open/close as the cilia bend toward the tallest/smallest cilia, respectively. These channels are permeable to potassium ions, which are abundant in the fluid bathing the top of the hair cells.

Medial vestibulo-spinal tract

Vestibular nucleus neurons project bilaterally to cervical spinal motor neurons for head and neck movement control. The tract principally functions in gaze direction and stability during motion.

Neurotransmitters

A chemical compound used to send signals from a receptor cell to a neuron, or from one neuron to another. Neurotransmitters can be excitatory, inhibitory, or modulatory and are packaged in small vesicles that are released from the end terminals of cells.

Oculomotor nuclei

Includes three neuronal groups in the brainstem, the abducens nucleus, the oculomotor nucleus, and the trochlear nucleus, whose cells send motor commands to the six pairs of eye muscles.

Oculomotor nucleus

A group of cells in the middle brainstem that contain subgroups of neurons that project to the medial rectus, inferior oblique, inferior rectus, and

superior rectus muscles of the eyes through the 3rd cranial nerve.

Otoconia

Small calcium carbonate particles that are packed in a layer on top of the gelatin membrane that covers the otolith receptor hair cell stereocilia.

Otolith receptors

Two inner ear vestibular receptors (utricle and saccule) that transduce linear accelerations and head tilt relative to gravity into neural signals that are then transferred to the brain.

Proprioceptive

Sensory information regarding muscle position and movement arising from receptors in the muscles, tendons, and joints.

Semicircular canals

A set of three inner ear vestibular receptors (horizontal, anterior, posterior) that transduce head rotational accelerations into head rotational velocity signals that are then transferred to the brain. There are three semicircular canals in each ear, with the major planes of each canal being orthogonal to each other.

Stereocilia

Hairlike projections from the top of the receptor hair cells. The stereocilia are arranged in ascending height and when displaced toward the tallest cilia, the mechanical gated channels open and the cell is excited (depolarized). When the stereocilia are

displaced toward the smallest cilia, the channels close and the cell is inhibited (hyperpolarized).

Torsion

A rotational eye movement around the line of sight that consists of a clockwise or counterclockwise direction.

Vergence angle

The angle between the line of sight for the two eyes. Low vergence angles indicate far-viewing objects, whereas large angles indicate viewing of near objects.

Vestibular compensation

Following injury to one side of vestibular receptors or the vestibulocochlear nerve, the central vestibular nuclei neurons gradually recover much of their function through plasticity mechanisms. The recovery is never complete, however, and extreme motion environments can lead to dizziness, nausea, problems with balance, and spatial memory.

Vestibular efferents

Nerve fibers originating from a nucleus in the brainstem that project from the brain to innervate the vestibular receptor hair cells and afferent nerve terminals. Efferents have a modulatory role on their targets, which is not well understood.

Vestibular system

Consists of a set of motion and gravity detection receptors in the inner ear, a set of primary nuclei in the brainstem, and a network of pathways carrying

motion and gravity signals to many regions of the brain.

Vestibulocochlear nerve

The VIIIth cranial nerve that carries fibers innervating the vestibular receptors and the cochlea.

Vestibuloocular reflex

Eye movements produced by the vestibular brainstem that are equal in magnitude and opposite in direction to head motion. The VOR functions to maintain visual stability on a point of interest and is nearly perfect for all natural head movements.

REFERENCES

- Angelaki, D. E., McHenry, M. Q., & Hess, B. J. (2000). Primate translational vestibuloocular reflexes. I. High-frequency dynamics and three-dimensional properties during lateral motion. *J. Neurophysiol* 83, 1637–1647.
- Angelaki, D. E., Shaikh, A. G., Green, A. M., & Dickman, J. D. (2004). Neurons compute internal models of the physical laws of motion. *Nature*, 430, 560–564.
- Anniko, M., & Arnold, W. (1991). Acetylcholine receptor localization in human adult cochlear and vestibular hair cells. *AC and Otolar*, 111, 491–499.
- Beraneck, M., Hachemaoui, M., Idoux, E., Ris, L., Uno, A., Godaux, E., . . . Vibert, N. (2003). Long-term plasticity of ipsilesional medial vestibular nucleus neurons after unilateral labyrinthectomy. *J Neurophysiol*, 90, 184–203.

- Blanks, R.H.I., Curthoys, I. S., Bennett, M. L., & Markham, C. H. (1985). Planar relationships of the semicircular canals in rhesus and squirrel monkeys. *Brain Res*, 340, 315–324.
- Boyle, R., Carey, J. P., & Highstein, S.M. (1991). Morphological correlates of response dynamics and efferent stimulation in horizontal semicircular canal afferents of the toadfish, *Opsanus tau*. *J Neurophysiol*, 66, 1504–1511.
- Brodal, A. The vestibular nuclei in the macaque monkey. (1984). *J. Comp. Neurol.*, 227, 252–266.
- Carleton, S. C., & Carpenter, M. B. (1984). Distribution of primary vestibular fibers in the brainstem and cerebellum of the monkey. *Brain Res*, 294, 281–298.
- Chen, A., DeAngelis, G. C., & Angelaki, D. E. (2011). A comparison of vestibular spatiotemporal tuning in macaque parietoinsular vestibular cortex, ventral intraparietal area, and medial superior temporal area. *J. Neurosci.*, 31, 3082–3094.
- Crane, B. T., & Demer, J. L. (1997). Human gaze stabilization during natural activities: translation, rotation, magnification, and target distance effects. *J Neurophysiol*, 78, 2129–2144.
- Dickman, J. D., & Angelaki, D. E. (2002). Vestibular convergence patterns in vestibular nuclei neurons of alert primates. *J. Neurophysiol.*, 88, 3518–3533.

- Dickman, J. D., & Correia, M. J. (1989). Responses of pigeon horizontal semicircular canal afferent fibers. I. Step, trapezoid, and low frequency sinusoid mechanical and rotational stimulation. *J Neurophysiol*, 62, 1090–1101.
- Dickman, J. D., & Fang, Q. (1996). Differential central projections of vestibular afferents in pigeons. *J Comp Neurol*, 367, 110–121.
- Einstein, A. (1907). On the relativity principle and the conclusions drawn from it. *Jahrb. Radioakt*, 4, 411–462.
- Ezure, K., & Graf, K.W. (1984). A quantitative analysis of the spatial organization of the vestibulo-ocular reflexes in lateral- and frontal-eyed animals. I. Orientation of semicircular canals and extraocular muscles. *Neuroscience*, 12, 85–94.
- Fernandez, C., & Goldberg, J. M. (1976b). Physiology of peripheral neurons innervating otolith organs of the squirrel monkey. II. Directional selectivity and force-response relations. *J. Neurophysiol.*, 39, 985–995.
- Fernandez, C., & Goldberg, J. M. (1971). Physiology of peripheral neurons innervating semicircular canals of the squirrel monkey. II. Response to sinusoidal stimulation and dynamics of peripheral vestibular system. *J Neurophysiol*, 34, 661–675.
- Fernandez, C., & Goldberg, J.M. (1976a). Physiology of peripheral neurons innervating otolith organ in the squirrel monkey. I. Response to static tilts and to long-duration centrifugal

force. *J. Neurophysiol.*, 39, 970–984.

- Fetter, M., & Zee, D. S. (1988). Recovery from unilateral labyrinthectomy in rhesus monkey. *J Neurophysiol*, 59, 370–393.
- Flock, A. (1964). Structure of the macula utriculi with special reference to directional interplay of sensory responses as revealed by morphological polarization. *J Cell Biol*, 22, 413–431.
- Fukushima, K., Sato, T., Fukushima, J. Shinmei, Y., and Kaneko C.R. (2000). Activity of smooth pursuit related neurons in the monkey perarcuate cortex during pursuit and passive whole-body rotation. *J. Neurophys.*, 83, 563–587.
- Grusser, O., Pause, M., & Schreier, U. (1982). Neuronal responses in the parieto-insular vestibular cortex of alert java monkeys (*Macaca fascicularis*). In A. Roucoux and M. Crommelinck (Eds.), *Physiological and pathological aspects of eye movements*. The Hague: W. Junk Publishers.
- Gu, Y., DeAngelis, G. C., & Angelaki, D. E. (2007). A functional link between area mstd and heading perception based on vestibular signals. *Nat Neurosci*, 10, 1038–1047.
- Gu, Y., Watkins, P. V., Angelaki, D. E., & DeAngelis, G. C. (2006). Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J Neurosci*, 26, 73–85.
- Guldin, W. O., Akbarian, S., & Grusser, O. J. (1992). Cortico-cortical connections and

cytoarchitectonics of the primate vestibular cortex: A study in squirrel monkeys (*Saimiri sciureus*). *J. Comp. Neurol.*, 326, 375–401.

- Hafting, T., Fyhn, M., Molden, S., Moser, M.-B., Moser, E. I. (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436, 801–806.
- Iurato, S. (1967). *Submicroscopic structure of the inner ear*. London, England: Pergamon Press.
- Kasahara, M., & Uchino, Y. (1974). Bilateral semicircular canal inputs to neurons in cat vestibular nuclei. *Exp Brain Res*, 20, 285–296.
- Keshner, E. A., & Peterson, B. W. (1995). Mechanisms controlling human head stabilization. I. Head-neck dynamics during random rotations in the horizontal plane. *J Neurophysiol*, 73, 2293–2301.
- Lindeman, H. H. (1969). Studies on the morphology of the sensory regions of the vestibular apparatus with 45 figures. *Ergeb Anat Entwicklungsgesch*, 42, 1–113.
- Lisberger, S. G., Pavelko, T. A., & Broussard, D. M. (1994). Neural basis for motor learning in the vestibuloocular reflex of primates. I. Changes in the responses of brain stem neurons. *J Neurophysiol*, 72, 928–973.
- Marlinski, V., & McCrea, R. A. (2008). Activity of ventroposterior thalamus neurons during rotation and translation in the horizontal plane in the alert squirrel monkey. *J Neurophysiol*, 99, 2533–2545.

- McNaughton, B. L., Battaglia, F. P., Jensen, O., Moser, E. I., & Moser, M.-B. (2006). Path integration and the neural basis of the “cognitive map.” *Nature Rev.*, 7, 663–678.
- Meng, H., May, P. J., Dickman, J. D., & Angelaki, D. E. (2007). Vestibular signals in primate thalamus: Properties and origins. *J Neurosci*, 27, 13590–13602.
- Newlands, S. D., & Perachio, A. A. (1990). Compensation of horizontal canal related activity in the medial vestibular nucleus following unilateral labyrinth ablation in the decerebrate gerbil. i. type i neurons. *Exp Brain Res*, 82, 359–372.
- Newlands, S. D., Hesse, S. V., Haque, A., & Angelaki, D. E. (2001). Head unrestrained horizontal gaze shifts after unilateral labyrinthectomy in the rhesus monkey. *Exp Brain Res*, 140, 25–33.
- O’Keefe, J. (1976). Place units in the hippocampus of the freely moving rat. *Exp Neurol*, 51, 78–109.
- Paige, G. D., & Tomko, D. L. (1991). Eye movement responses to linear head motion in the squirrel monkey. I. Basic characteristics. *J Neurophysiol*, 65, 1170–1182.
- Peterson, B. W., Goldberg, J., Bilotto, G., & Fuller, J. H. (1985). Cervicocollic reflex: Its dynamic properties and interaction with vestibular reflexes. *J Neurophysiol*, 54, 90–109.
- Precht, W., & Shimazu, H. (1965). Functional

connections of tonic and kinetic vestibular neurons with primary vestibular afferents. *J Neurophysiol*, 28, 1014–1028.

- Sadeghi, S. G., Chacron, M. J., Taylor, M. C., & Cullen, K. E. (2007). Neural variability, detection thresholds, and information transmission in the vestibular system. *J. Neurosci.*, 27, 771–781.
- Schwarz, U., Busetini, C., & Miles, F. A. (1989). Ocular responses to linear motion are inversely proportional to viewing distance. *Science*, 245, 1394–1396.
- Shimazu, H., & Precht, W. (1966). Inhibition of central vestibular neurons from the contralateral labyrinth and its mediating pathway. *J Neurophysiol*, 29, 467–492.
- Shinoda, Y., Sugiuchi, Y., Futami, T., Ando, N., & Kawasaki, T. (1994). Input patterns and pathways from the six semicircular canals to motoneurons of neck muscles. I. The multifidus muscle group. *J. Neurophysiol.*, 72, 2691–2802.
- Shotwell, S. L., Jacobs, R., & Hudspeth, A. J. (1981). Directional sensitivity of individual vertebrate hair cells to controlled deflection of their hair bundles. *ANYAS*, 374, 1–10.
- Stackman, R. W., Clark, A. S., & Taube, J. (2002). Hippocampal spatial representations require vestibular input. *Hippocampus*, 12, 291–303.
- Taube, J. S. (1995). Head direction cells recorded in the anterior thalamic nuclei of freely moving rats. *J. Neurosci.*, 15, 70–86.

- Van Egmond, A. A. J., Groen, J. J., & Jongkess, L. B. W. (1949). The mechanics of the semicircular canals. *J Physiol Lond*, 110, 1–17.
- Wersäll, J. (1956). Studies on the structure and innervation of the sensory epithelium of the cristae ampullares in the guinea pig; a light and electron microscopic investigation. *Acta Otolaryngol Supp*, 126, 1–85.
- Wilson, V. J., Boyle, R., Fukushima, K., Rose, P. K., Shinoda, Y., Sugiuchi, Y., & Uchino, Y. (1995). The vestibulocollic reflex. *J Vestib Res*, 5, 147–170.
- Yates, B. J., Miller A. D., & Lucot, J. B. (1998). Physiological basis and pharmacology of motion sickness: An update. *Brain Res Bull*, 47, 395–406.
- Yu, X. J., Dickman, J. D., Angelaki, D. E. (2012). Detection thresholds of macaque otolith afferents. *J Neurosci.*, 32, 8306–8316.
- Zhou, W., Tang, B.F., Newlands, S.D., King, W.M. (2006). Responses of monkey vestibular-only neurons to translation and angular rotation. *J Neurophysiol* 96, 2915-2930.

AUTHORS



Dora Angelaki

Dora Angelaki, Professor & Chair of the Department of Neuroscience, Baylor College of Medicine in Houston TX, holds a PhD degree in Biomedical Engineering from the University of Minnesota and is interested in computational, cognitive and systems neuroscience.



J. David Dickman

J. David Dickman, Professor in the Department

of Neuroscience, Baylor College of Medicine, and Neuroscience Program Director at Rice University in Houston TX, received his PhD in psychology from the University of Wyoming and studies the neural processing of motion, navigation, and magnetoreception.

CREATIVE COMMONS LICENSE



THE VESTIBULAR SYSTEM by [Dora Angelaki and J. David Dickman](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

PART VIII.

**UNIT 8: STATES OF
CONSCIOUSNESS**

CHAPTER 12.

ATTENTION

FRANCES FRIEDRICH

We use the term “attention” all the time, but what processes or abilities does that concept really refer to? This module will focus on how attention allows us to select certain parts of our environment and ignore other parts, and what happens to the ignored information. A key concept is the idea that we are limited in how much we can do at any one time. So we will also consider what happens when someone tries to do several things at once, such as driving while using electronic devices.

LEARNING OBJECTIVES

- Understand why selective attention is important and how it can be studied.
- Learn about different models of when and how selection can occur.
- Understand how divided attention or multitasking is studied, and implications of multitasking in situations such as distracted driving.

WHAT IS ATTENTION?

Before we begin exploring attention in its various forms, take a moment to consider how you think about the concept. How would you define attention, or how do you use the term? We certainly use the word very frequently in our everyday language: “ATTENTION! USE ONLY AS DIRECTED!” warns the label on the medicine bottle, meaning be alert to possible danger. “Pay attention!” pleads the weary seventh-grade teacher, not warning about danger (with possible exceptions, depending on the teacher) but urging the students to focus on the task at hand. We may refer to a child who is easily distracted as having an attention disorder, although we also are told that Americans have an attention span of about 8 seconds, down from 12 seconds in 2000, suggesting that we *all* have trouble sustaining concentration for any amount of time (from www.Statisticbrain.com). How that number was determined is not clear from the Web site, nor is it clear how attention span in the goldfish—9 seconds!—was measured, but the fact that our average span reportedly is less than that of a goldfish is intriguing, to say the least.

William James wrote extensively about attention in the late 1800s. An often quoted passage ([James, 1890/1983](#)) beautifully captures how intuitively obvious the concept of attention is, while it remains very difficult to define in measurable, concrete terms:

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. (pp. 381–382)



Are you reading these words right here right now? If so, it's only because you directed your attention toward them.

[Image: CC BY 2.0, <https://goo.gl/BRvSA7>]

Notice that this description touches on the conscious nature of attention, as well as the notion that what is in consciousness is often controlled voluntarily but can also be determined by events that capture our attention. Implied in this description is the idea that we seem to have a **limited capacity** for information processing, and that we can only attend to or be

consciously aware of a small amount of information at any given time.

Many aspects of attention have been studied in the field of psychology. In some respects, we define different types of attention by the nature of the task used to study it. For example, a crucial issue in World War II was how long an individual could remain highly alert and accurate while watching a radar screen for enemy planes, and this problem led psychologists to study how attention works under such conditions. When watching for a rare event, it is easy to allow concentration to lag. (This continues to be a challenge today for TSA agents, charged with looking at images of the contents of your carry-on items in search of knives, guns, or shampoo bottles larger than 3 oz.) Attention in the context of this type of search task

refers to the level of *sustained attention* or *vigilance* one can maintain. In contrast, **divided attention** tasks allow us to determine how well individuals can attend to many sources of information at once. *Spatial attention* refers specifically to how we focus on one part of our environment and how we move attention to other locations in the environment. These are all examples of different aspects of attention, but an implied element of most of these ideas is the concept of **selective attention**; some information is attended to while other information is intentionally blocked out. This module will focus on important issues in selective and divided attention, addressing these questions:

- Can we pay attention to several sources of information at once, or do we have a limited capacity for information?
- How do we select what to pay attention to?
- What happens to information that we try to ignore?
- Can we learn to divide attention between multiple tasks?

SELECTIVE ATTENTION

THE COCKTAIL PARTY

Selective attention is *the ability to select certain stimuli in the environment to process, while ignoring distracting information*. One way to get an intuitive sense of how attention works is to consider situations in which attention is used. A party provides an excellent example for our purposes. Many people may be milling around, there is a dazzling variety of colors and sounds and



Beyond just hearing your name from the clamor at a party, other words or concepts, particularly unusual or significant ones to you, can also snag your attention. [Image: Ross, <https://goo.gl/TVDfTn>, CC BY-NC-SA 2.0, <https://goo.gl/TocOZF>]

smells, the buzz of many conversations is striking. There are so many conversations going on; how is it possible to select just one and follow it? You don't have to be looking at the person talking; you may be listening with great interest to some gossip while pretending not to hear. However, once you are engaged in conversation with someone, you quickly become aware that you cannot also listen to other conversations at the same time. You also are probably *not* aware of how tight your shoes feel or of the smell of a nearby flower arrangement. On the other hand, if someone behind you mentions your name, you typically notice it immediately and may start attending to that (much more interesting) conversation. This situation

highlights an interesting set of observations. We have an amazing ability to select and track one voice, visual object, etc., even when a million things are competing for our attention, but at the same time, we seem to be limited in how much we can attend to at one time, which in turn suggests that attention is crucial in selecting what is important. How does it all work?

DICHOTIC LISTENING STUDIES

This cocktail party scenario is the quintessential example of selective attention, and it is essentially what some early researchers tried to replicate under controlled laboratory conditions as a starting point for understanding the role of attention in perception (e.g., [Cherry, 1953](#); [Moray, 1959](#)). In particular, they used **dichotic listening** and **shadowing** tasks to evaluate the selection process. Dichotic listening simply refers to the situation when two messages are presented simultaneously to an individual, with one message in each ear. In order to control which message the person attends to, the individual is asked to repeat back or “shadow” one of the messages as he hears it. For example, let’s say that a story about a camping trip is presented to John’s left ear, and a story about Abe Lincoln is presented to his right ear. The typical dichotic listening task would have John repeat the story presented to one ear as he hears it. Can he do that without being distracted by the information in the other ear?

People can become pretty good at the shadowing task, and they can easily report the content of the message that they attend to. But what happens to the ignored message? Typically, people can tell you if the ignored message was

a man's or a woman's voice, or other physical characteristics of the speech, but they cannot tell you what the message was about. In fact, many studies have shown that people in a shadowing task were not aware of a change in the language of the message (e.g., from English to German; [Cherry, 1953](#)), and they didn't even notice when the same word was repeated in the unattended ear more than 35 times ([Moray, 1959](#))! Only the basic physical characteristics, such as the pitch of the unattended message, could be reported.

On the basis of these types of experiments, it seems that we can answer the first question about how much information we can attend to very easily: not very much. We clearly have a limited capacity for processing information for meaning, making the selection process all the more important. The question becomes: How does this selection process work?

MODELS OF SELECTIVE ATTENTION

Broadbent's Filter Model. Many researchers have investigated how selection occurs and what happens to ignored information. Donald Broadbent was one of the first to try to characterize the selection process. His Filter Model was based on the dichotic listening tasks described above as well as other types of experiments ([Broadbent, 1958](#)). He found that people select information on the basis of *physical features*: the sensory channel (or ear) that a message was coming in, the pitch of the voice, the color or font of a visual message. People seemed vaguely aware of the physical features of the unattended information, but had no knowledge of the meaning. As a result, Broadbent argued that selection occurs *very early*, with no additional

processing for the unselected information. A flowchart of the model might look like this:

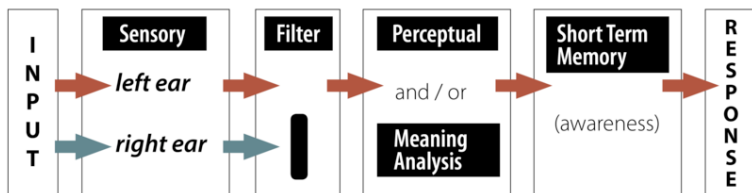


Figure 1. This figure shows information going in both the left and right ears. Some basic sensory information, such as pitch, is processed, but the filter only allows the information from one ear to be processed further. Only the information from the left ear is transferred to short-term memory (STM) and conscious awareness, and then further processed for meaning. That means that the ignored information never makes it beyond a basic physical analysis.

TREISMAN'S ATTENUATION MODEL

Broadbent's model makes sense, but if you think about it you already know that it cannot account for all aspects of the Cocktail Party Effect. What doesn't fit? The fact is that you tend to hear your own name when it is spoken by someone, even if you are deeply engaged in a conversation. We mentioned earlier that people in a shadowing experiment were unaware of a word in the unattended ear that was repeated many times—and yet many people noticed their own name in the unattended ear even it occurred only once.

Anne Treisman (1960) carried out a number of dichotic listening experiments in which she presented two different stories to the two ears. As usual, she asked people to shadow the message in one ear. As the stories progressed, however, she switched the stories to the opposite ears. Treisman found that individuals

spontaneously followed the story, or the content of the message, when it shifted from the left ear to the right ear. Then they realized they were shadowing the wrong ear and switched back.

Results like this, and the fact that you tend to hear meaningful information even when you aren't paying attention to it, suggest that we *do* monitor the unattended information to some degree on the basis of its meaning. Therefore, the filter theory can't be right to suggest that unattended information is completely blocked at the sensory analysis level. Instead, Treisman suggested that selection *starts* at the physical or perceptual level, but that the unattended information is not blocked completely, it is just weakened or *attenuated*. As a result, highly meaningful or pertinent information in the unattended ear will get through the filter for further processing at the level of meaning. The figure below shows information going in both ears, and in this case there is no filter that completely blocks nonselected information. Instead, selection of the left ear information strengthens that material, while the nonselected information in the right ear is weakened. However, if the preliminary analysis shows that the nonselected information is especially pertinent or meaningful (such as your own name), then the Attenuation Control will instead strengthen the more meaningful information.

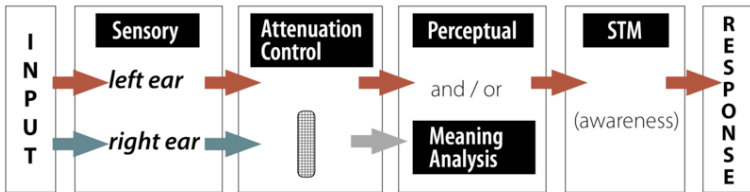


Figure 2.

LATE SELECTION MODELS

Other selective attention models have been proposed as well. *Late selection* or *response selection* model proposed by Deutsch and Deutsch (1963) suggests that all information in the unattended ear is processed on the basis of meaning, not just the selected or highly pertinent information. However, only the information that is relevant for the task response gets into conscious awareness. This model is consistent with ideas of **subliminal perception**; in other words, that you don't have to be aware of or attending a message for it to be fully processed for meaning.

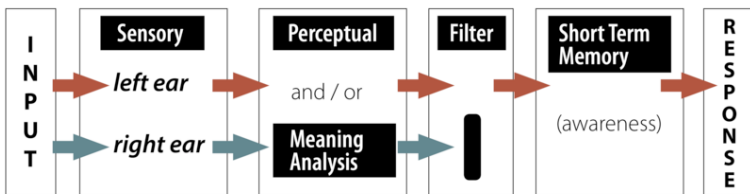


Figure 3.

You might notice that this figure looks a lot like that of the Early Selection model—only the location of the selective filter has changed, with the assumption that analysis of

meaning occurs *before* selection occurs, but only the selected information becomes conscious.

MULTIMODE MODEL

Why did researchers keep coming up with different models? Because no model really seemed to account for all the data, some of which indicates that nonselected information is blocked completely, whereas other studies suggest that it can be processed for meaning. The multimode model addresses this apparent inconsistency, suggesting that the stage at which selection occurs can change depending on the task. Johnston and Heinz (1978) demonstrated that under some conditions, we can select what to attend to at a very early stage and we do not process the content of the unattended message very much at all. Analyzing physical information, such as attending to information based on whether it is a male or female voice, is relatively easy; it occurs automatically, rapidly, and doesn't take much effort. Under the right conditions, we can select what to attend to on the basis of the meaning of the messages. However, the late selection option—processing the content of all messages before selection—is more difficult and requires more effort. The benefit, though, is that we have the flexibility to change how we deploy our attention depending upon what we are trying to accomplish, which is one of the greatest strengths of our cognitive system.

This discussion of selective attention has focused on experiments using auditory material, but the same principles hold for other perceptual systems as well. Neisser (1979) investigated some of the same questions with visual materials by superimposing two semi-

transparent video clips and asking viewers to attend to just one series of actions. As with the auditory materials, viewers often were unaware of what went on in the other clearly visible video. Twenty years later, Simons and Chabris (1999) explored and expanded these findings using similar techniques, and triggered a flood of new work in an area referred to as **inattention blindness**. We touch on those ideas below, and you can also refer to another Noba Module, [Failures of Awareness: The Case of Inattention Blindness](#) for a more complete discussion.

Focus Topic 1: Subliminal Perception

The idea of subliminal perception—that stimuli presented below the threshold for awareness can influence thoughts, feelings, or actions—is a fascinating and kind of creepy one. Can messages you are unaware of, embedded in movies or ads or the music playing in the grocery store, really influence what you buy? Many such claims of the power of subliminal perception have been made. One of the most famous came from a market researcher who claimed that the message “Eat Popcorn” briefly flashed throughout a movie increased popcorn sales by more than 50%, although he later admitted that the study was made up ([Merikle, 2000](#)). Psychologists have worked hard to investigate whether this is a valid phenomenon. Studying subliminal perception is more difficult than it might seem, because of the

difficulty of establishing what the threshold for consciousness is or of even determining what type of threshold is important; for example, Cheesman and Merikle ([1984](#), [1986](#)) make an important distinction between objective and subjective thresholds. The bottom line is that there is some evidence that individuals can be influenced by stimuli they are not aware of, but how complex the stimuli can be or the extent to which unconscious material can affect behavior is not settled (e.g., [Bargh & Morsella, 2008](#); [Greenwald, 1992](#); [Merikle, 2000](#)).

DIVIDED ATTENTION AND MULTITASKING

In spite of the evidence of our limited capacity, we all like to think that we can do several things at once. Some people claim to be able to multitask without any problem: reading a textbook while watching television and talking with friends; talking on the phone while playing computer games; texting while driving. The fact is that we sometimes can *seem* to juggle several things at once, but the question remains whether dividing attention in this way impairs performance.

Is it possible to overcome the limited capacity that we experience when engaging in cognitive tasks? We know that with extensive practice, we can acquire skills that do not appear to require conscious attention. As we walk down the street, we don't need to think consciously about

what muscle to contract in order to take the next step. Indeed, paying attention to automated skills can lead to a breakdown in performance, or “choking” (e.g., [Beilock & Carr, 2001](#)). But what about higher level, more mentally demanding tasks: Is it possible to learn to perform two complex tasks at the same time?

DIVIDED ATTENTION TASKS



Unless a task is fully automated, some researchers suggest that “multi-tasking” doesn’t really exist; you are just rapidly switching your attention back and forth between tasks. [Image: CCO Public Domain, <https://goo.gl/m25gce>]

In a classic study that examined this type of divided attention task, two participants were trained to take dictation for spoken words while reading unrelated material for comprehension ([Spelke, Hirst, & Neisser, 1976](#)). In divided attention tasks such as these, each task is evaluated separately, in order to determine baseline performance when the individual can

allocate as many cognitive resources as necessary to one task at a time. Then performance is evaluated when the two tasks are performed simultaneously. A decrease in performance for either task would suggest that even if attention can be divided or switched between the tasks, the cognitive demands are too great to avoid disruption of performance. (We should note here that divided attention tasks are designed, in principle, to see if two

tasks can be carried out simultaneously. A related research area looks at *task switching* and how well we can switch back and forth among different tasks [e.g., [Monsell, 2003](#)]. It turns out that switching itself is cognitively demanding and can impair performance.)

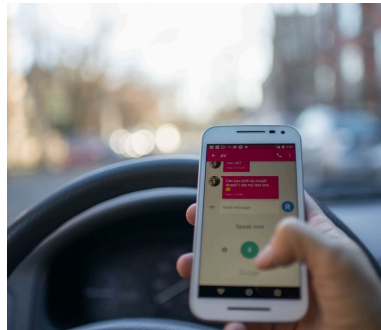
The focus of the Spelke et al. ([1976](#)) study was whether individuals could learn to perform two relatively complex tasks concurrently, without impairing performance. The participants received plenty of practice—the study lasted 17 weeks and they had a 1-hour session each day, 5 days a week. These participants were able to learn to take dictation for lists of words and read for comprehension without affecting performance in either task, and the authors suggested that perhaps there are not fixed limits on our attentional capacity. However, changing the tasks somewhat, such as reading aloud rather than silently, impaired performance initially, so this multitasking ability may be specific to these well-learned tasks. Indeed, not everyone could learn to perform two complex tasks without performance costs ([Hirst, Neisser, & Spelke, 1978](#)), although the fact that some can is impressive.

DISTRACTED DRIVING

More relevant to our current lifestyles are questions about multitasking while texting or having cell phone conversations. Research designed to investigate, under controlled conditions, multitasking while driving has revealed some surprising results. Certainly there are many possible types of distractions that could impair driving performance, such as applying makeup using the rearview mirror, attempting (usually in vain) to stop the kids in the backseat from fighting, fiddling with the CD

player, trying to negotiate a handheld cell phone, a cigarette, and a soda all at once, eating a bowl of cereal while driving (!). But we tend to have a strong sense that we CAN multitask while driving, and cars are being built with more and more technological capabilities that encourage multitasking. How good are we at dividing attention in these cases?

Most people acknowledge the distraction caused by texting while driving and the reason seems obvious: Your eyes are off the road and your hands and at least one hand (often both) are engaged while texting. However, the problem is not simply one of occupied hands or eyes, but rather that the *cognitive* demands on our limited capacity systems can seriously



If you look at your phone for just 5 seconds while driving at 55mph, that means you have driven the length of a football field without looking at the road. [Image: CCO Public Domain, <https://goo.gl/m25gce>]

impair driving performance ([Strayer, Watson, & Drews, 2011](#)). The effect of a cell phone conversation on performance (such as not noticing someone's brake lights or responding more slowly to them) is just as significant when the individual is having a conversation with a hands-free device as with a handheld phone; the same impairments do not occur when listening to the radio or a book on tape ([Strayer & Johnston, 2001](#)). Moreover, studies using eye-tracking devices have shown that drivers are less likely to later recognize objects that they *did* look at when using a cell phone while driving

([Strayer & Drews, 2007](#)). These findings demonstrate that cognitive distractions such as cell phone conversations can produce inattention blindness, or a lack of awareness of what is right before your eyes (see also, [Simons & Chabris, 1999](#)). Sadly, although we all like to think that we can multitask while driving, in fact the percentage of people who can truly perform cognitive tasks without impairing their driving performance is estimated to be about 2% ([Watson & Strayer, 2010](#)).

SUMMARY

It may be useful to think of attention as a mental resource, one that is needed to focus on and fully process important information, especially when there is a lot of distracting “noise” threatening to obscure the message. Our selective attention system allows us to find or track an object or conversation in the midst of distractions. Whether the selection process occurs early or late in the analysis of those events has been the focus of considerable research, and in fact how selection occurs may very well depend on the specific conditions. With respect to divided attention, in general we can only perform one cognitively demanding task at a time, and we may not even be aware of unattended events even though they might seem too obvious to miss (check out some examples in the Outside Resources below). This type of inattention blindness can occur even in well-learned tasks, such as driving while talking on a cell phone. Understanding how attention works is clearly important, even for our everyday lives.

OUTSIDE RESOURCES

Video: Here’s a wild example of how much we fail to

notice when our attention is captured by one element of a scene.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=119#oembed-1>

Video: Try this test to see how well you can focus on a task in the face of a lot of distraction.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=119#oembed-2>

DISCUSSION QUESTIONS

1. Discuss the implications of the different models of selective attention for everyday life. For instance, what advantages and disadvantages would be associated with being able to filter out all unwanted information at a very early stage in processing? What are the implications of processing all ignored information fully, even if you aren't consciously aware of that information?

2. Think of examples of when you feel you can successfully multitask and when you can't. Discuss what aspects of the tasks or the situation seem to influence divided attention performance. How accurate do you think you are in judging your own multitasking ability?
3. What are the public policy implications of current evidence of inattention blindness as a result of distracted driving? Should this evidence influence traffic safety laws? What additional studies of distracted driving would you propose?

VOCABULARY

Dichotic listening

An experimental task in which two messages are presented to different ears.

Divided attention

The ability to flexibly allocate attentional resources between two or more concurrent tasks.

Inattention blindness

The failure to notice a fully visible object when attention is devoted to something else.

Limited capacity

The notion that humans have limited mental resources that can be used at a given time.

Selective attention

The ability to select certain stimuli in the environment to process, while ignoring distracting information.

Shadowing

A task in which the individual is asked to repeat an auditory message as it is presented.

Subliminal perception

The ability to process information for meaning when the individual is not consciously aware of that information.

REFERENCES

- Bargh, J., & Morsella, E. (2008). The unconscious mind. *Perspectives on Psychological Science*, 3(1), 73–79.
- Beilock, S. L., & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General*, 130, 701–725.
- Broadbent, D. A. (1958). *Perception and communication*. London, England: Pergamon Press.
- Cheesman, J., & Merikle, P. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology*, 40, 343–367.
- Cheesman, J., & Merikle, P. (1984). Priming with and without awareness. *Perception and Psychophysics*, 36, 387–395.
- Cherry, E. C. (1953). Experiments on the recognition of speech with one and two ears. *Journal of the Acoustical Society of America*, 25, 975–979.
- Deutsch, J. A., & Deutsch, D. (1963). Attention:

some theoretical considerations. *Psychological Review*, 70, 80–90.

- Greenwald, A. G. (1992). New Look 3: Unconscious cognition reclaimed. *American Psychologist*, 47, 766–779.
- Hirst, W. C., Neisser, U., & Spelke, E. S. (1978). Divided attention. *Human Nature*, 1, 54–61.
- James, W. (1983). *The principles of psychology*. Cambridge, MA: Harvard University Press. (Original work published 1890)
- Johnston, W. A., & Heinz, S. P. (1978). Flexibility and capacity demands of attention. *Journal of Experimental Psychology: General*, 107, 420–435.
- Merikle, P. (2000). Subliminal perception. In A. E. Kazdin (Ed.), *Encyclopedia of psychology* (Vol. 7, pp. 497–499). New York, NY: Oxford University Press.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Science*, 7(3), 134–140.
- Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, 11, 56–60.
- Neisser, U. (1979). The control of information pickup in selective looking. In A. D. Pick (Ed.), *Perception and its development: A tribute to Eleanor J. Gibson* (pp. 201–219). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattention blindness for

dynamic events. *Perception*, 28, 1059–1074.

- Spelke, E. S., Hirst, W. C., & Neisser, U. (1976). Skills of divided attention. *Cognition*, 4, 215–250.
- Strayer, D. L., & Drews, F. A. (2007). Cell-phone induced inattention blindness. *Current Directions in Psychological Science*, 16, 128–131.
- Strayer, D. L., & Johnston, W. A. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone. *Psychological Science*, 12, 462–466.
- Strayer, D. L., Watson, J. M., & Drews, F. A. (2011). Cognitive distraction while multitasking in the automobile. In Brian Ross (Ed.), *The Psychology of Learning and Motivation* (Vol. 54, pp. 29–58). Burlington, VT: Academic Press.
- Treisman, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12, 242–248.
- Watson, J. M., & Strayer, D. L. (2010). Supertaskers: Profiles in extraordinary multitasking ability. *Psychonomic Bulletin & Review*, 17, 479–485.

AUTHORS



Frances Friedrich

Frances J. Friedrich is an Associate Professor of Psychology at the University of Utah, Salt Lake City, Utah. She is a cognitive neuropsychologist interested in behavioral and neural aspects of attention and perception. Her research focuses on spatial and executive attentional processes.

CREATIVE COMMONS LICENSE



ATTENTION by [Frances Friedrich](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

CHAPTER 13.

STATES OF CONSCIOUSNESS

ROBERT BISWAS-DIENER AND JAKE TEENY

No matter what you're doing—solving homework, playing a video game, simply picking out a shirt—all of your actions and decisions relate to your consciousness. But as frequently as we use it, have you ever stopped to ask yourself: What really is consciousness? In this module, we discuss the different levels of consciousness and how they can affect your behavior in a variety of situations. As well, we explore the role of consciousness in other, “altered” states like hypnosis and sleep.

LEARNING OBJECTIVES

- Define consciousness and distinguish between high and low conscious states
- Explain the relationship between consciousness and bias
- Understand the difference between popular portrayals of hypnosis and how it is currently used therapeutically

INTRODUCTION

Have you ever had a fellow motorist stopped beside you at a red light, singing his brains out, or picking his nose, or otherwise behaving in ways he might not normally do in public? There is something about being alone in a car that encourages people to zone out and forget that others can see them. Although these little lapses of attention are amusing for the rest of us, they are also instructive when it comes to the topic of consciousness.



This guy is singing his heart out in his one-man mobile music studio. Have you ever done this? [Image: Joshua Ommen, <https://goo.gl/Za97c3>, CC BY-NC-SA 2.0, <https://goo.gl/Toc0ZF>]

Consciousness is a term meant to indicate awareness. It includes awareness of the self, of bodily sensations, of thoughts and of the environment. In English, we use the opposite word “unconscious” to indicate senselessness or a barrier to awareness, as in the case of “Theresa fell off the ladder and hit her head, knocking herself

unconscious.” And yet, psychological theory and research suggest that consciousness and unconsciousness are more complicated than falling off a ladder. That is, consciousness is more than just being “on” or “off.” For instance, Sigmund Freud (1856 – 1939)—a psychological theorist—understood that even while we are awake, many things lay outside the realm of our conscious awareness (like being in the car and forgetting the rest of the world can see into your windows). In response to this notion,

Freud introduced the concept of the “subconscious” ([Freud, 2001](#)) and proposed that some of our memories and even our basic motivations are not always accessible to our conscious minds.

Upon reflection, it is easy to see how slippery a topic consciousness is. For example, are people conscious when they are daydreaming? What about when they are drunk? In this module, we will describe several levels of consciousness and then discuss altered states of consciousness such as hypnosis and sleep.

LEVELS OF AWARENESS

In 1957, a marketing researcher inserted the words “Eat Popcorn” onto one frame of a film being shown all across the United States. And although that frame was only projected onto the movie screen for 1/24th of a second—a speed too fast to be perceived by conscious awareness—the researcher reported an increase in popcorn sales by nearly 60%. Almost immediately, all forms of “subliminal messaging” were regulated in the US and banned in countries such as Australia and the United Kingdom. Even though it was later shown that the researcher had made up the data (he hadn’t even inserted the words into the film), this fear about influences on our subconscious persists. At its heart, this issue pits various levels of awareness against one another. On the one hand, we have the “low awareness” of subtle, even subliminal influences. On the other hand, there is you—the conscious thinking, feeling you which includes all that you are currently aware of, even reading this sentence. However, when we consider these different levels of

awareness separately, we can better understand how they operate.

LOW AWARENESS

You are constantly receiving and evaluating sensory information. Although each moment has too many sights, smells, and sounds for them all to be consciously considered, our brains are nonetheless processing all that information. For example, have you ever been at a party, overwhelmed by all the people and conversation, when out of nowhere you hear your name called? Even though you have no idea what else the person is saying, you are somehow conscious of your name (for more on this, “the cocktail party effect,” see Noba’s Module on Attention). So, even though you may not be *aware* of various stimuli in your environment, your brain is paying closer attention than you think.

Similar to a reflex (like jumping when startled), some **cues**, or significant sensory information, will automatically elicit a response from us even though we never consciously perceive it. For example, Öhman and Soares ([1994](#)) measured subtle variations in sweating of participants with a fear of snakes. The researchers flashed pictures of different objects (e.g., mushrooms, flowers, and most importantly, snakes) on a screen in front of them, but did so at speeds that left the participant clueless as to what he or she had actually seen. However, when snake pictures were flashed, these participants started sweating more (i.e., a sign of fear), even though they had no idea what they’d just viewed!

Although our brains perceive some stimuli without our conscious awareness, do they really affect our subsequent

thoughts and behaviors? In a landmark study, [Bargh, Chen, and Burrows \(1996\)](#) had participants solve a word search puzzle where the answers pertained to words about the elderly (e.g., “old,” “grandma”) or something random (e.g., “notebook,” “tomato”). Afterward, the researchers secretly measured how fast the participants walked down the hallway exiting the experiment. And although none of the participants were aware of a theme to the answers, those who had solved a puzzle with elderly words (vs. those with other types of words) walked more slowly down the hallway!

Priming Studies and Replication

If the results of priming studies sound too fantastic to be believed, you are not alone in your skepticism. Recently, many studies in psychology—including many priming studies—have come under scrutiny because they do not “replicate.” This means that when later researchers have attempted to recreate certain studies, they have not always gotten the same—or even similar—results. Non-replication does not suggest that the original researchers “faked” the results, but that there may have been flaws in the original sampling or research methods. Fortunately, researchers are very aware of the problem of non-replication and have taken steps to address the issue. For an in-depth exploration of the so-called “replication crisis” in psychology, please see the Noba module on that topic.

This effect is called **priming** (i.e., readily “activating” certain concepts and associations from one’s memory) has been found in a number of other studies. For example, priming people by having them drink from a warm glass

(vs. a cold one) resulted in behaving more “warmly” toward others ([Williams & Bargh, 2008](#)). Although all of these influences occur beneath one’s conscious awareness, they still have a significant effect on one’s subsequent thoughts and behaviors.

In the last two decades, researchers have made advances in studying aspects of psychology that exist beyond conscious awareness. As you can understand, it is difficult to use self-reports and surveys to ask people about motives or beliefs that they, themselves, might not even be aware of! One way of side-stepping this difficulty can be found in the **implicit associations test**, or IAT ([Greenwald, McGhee & Schwartz, 1998](#)). This research method uses computers to assess people’s reaction times to various stimuli and is a very difficult test to fake because it records automatic reactions that occur in milliseconds. For instance, to shed light on deeply held biases, the IAT might present photographs of European American faces and Asian faces while asking research participants to click buttons indicating either “good” or “bad” as quickly as possible. Even if the participant clicks “good” for every face shown, the IAT can still pick up tiny delays in responding. Delays are associated with more mental effort needed to process information. When information is processed quickly—as in the example of white faces being judged as “good”—it can be contrasted with slower processing—as in the example of Asian faces being judged as “good”—and the difference in processing speed is reflective of bias. In this regard, the IAT has been used for investigating stereotypes ([Nosek, Banaji & Greenwald, 2002](#)) as well as self-esteem ([Greenwald & Farnam, 2000](#)). This method can help uncover non-

conscious biases as well as those that we are motivated to suppress.



An actual screenshot from an IAT (Implicit Association Test) that a person might take to test their own mental representations of various cognitive constructs. In this particular case, this is an item testing an individual's unconscious reaction towards members of various ethnic groups. [Image: Courtesy of Anthony Greenwald from Project Implicit]

HIGH AWARENESS

Just because we may be influenced by these “invisible” factors, it doesn’t mean we are helplessly controlled by them. The other side of the awareness continuum is known as “high awareness.” This includes effortful

attention and careful decision making. For example, when you listen to a funny story on a date, or consider which class schedule would be preferable, or complete a complex math problem, you are engaging a state of consciousness that allows you to be highly aware of and focused on particular details in your environment.

Mindfulness is a state of higher consciousness that includes an awareness of the thoughts passing through one's head. For example, have you ever snapped at someone in frustration, only to take a moment and reflect on why you responded so aggressively? This more effortful consideration of your thoughts could be described as an expansion of your conscious awareness as you take the time to consider the possible influences on your thoughts. Research



Meditation has been practiced for centuries in religious contexts. In the past 50 years it has become increasingly popular as a secular practice. Scientific studies have linked meditation to lower stress and higher well-being. [Image: Indrek Torilo, <https://goo.gl/Bc5lwm>, CC BY-NC 2.0, <https://goo.gl/Fllc2e>]

has shown that when you engage in this more deliberate consideration, you are less persuaded by irrelevant yet biasing influences, like the presence of a celebrity in an advertisement ([Petty & Cacioppo, 1986](#)). Higher awareness is also associated with recognizing when you're using a stereotype, rather than fairly evaluating another person ([Gilbert & Hixon, 1991](#)).

Humans alternate between low and high thinking

states. That is, we shift between focused attention and a less attentive default state, and we have neural networks for both ([Raichle, 2015](#)). Interestingly, the less we're paying attention, the more likely we are to be influenced by non-conscious stimuli ([Chaiken, 1980](#)). Although these subtle influences may affect us, we can use our higher conscious awareness to protect against external influences. In what's known as the **Flexible Correction Model** ([Wegener & Petty, 1997](#)), people who are aware that their thoughts or behavior are being influenced by an undue, outside source, can correct their attitude against the bias. For example, you might be aware that you are influenced by mention of specific political parties. If you were motivated to consider a government policy you can take your own biases into account to attempt to consider the policy in a fair way (on its own merits rather than being attached to a certain party).

To help make the relationship between lower and higher consciousness clearer, imagine the brain is like a journey down a river. In low awareness, you simply float on a small rubber raft and let the currents push you. It's not very difficult to just drift along but you also don't have total control. Higher states of consciousness are more like traveling in a canoe. In this scenario, you have a paddle and can steer, but it requires more effort. This analogy applies to many states of consciousness, but not all. What about other states such as like sleeping, daydreaming, or hypnosis? How are these related to our conscious awareness?

	Costs	Benefits
Low Awareness	<i>Influenced by subtle factors</i>	<i>Saves mental effort</i>
High Awareness	<i>Uses mental effort</i>	<i>Can overcome some biases</i>

Table 1: States of Consciousness

OTHER STATES OF CONSCIOUSNESS

HYPNOSIS

If you've ever watched a stage hypnotist perform, it may paint a misleading portrait of this state of consciousness. The hypnotized people on stage, for example, appear to be in a state similar to sleep. However, as the hypnotist continues with the show, you would recognize some profound differences between sleep and hypnosis. Namely, when you're asleep, hearing the word "strawberry" doesn't make you flap your arms like a chicken. In stage performances, the hypnotized participants appear to be highly suggestible, to the point that they are seemingly under the hypnotist's control. Such performances are entertaining but have a way of sensationalizing the true nature of hypnotic states.



People being hypnotized on stage.

[Image: New Media Expo,

<https://goo.gl/FWgBqs>, CC BY-NC-SA

2.0, <https://goo.gl/Fllc2e>]

Hypnosis is an actual, documented

phenomenon—one that has been studied and debated for over 200 years ([Pekala et al., 2010](#)). Franz Mesmer (1734 – 1815) is often credited as among the first people to “discover” hypnosis, which he used to treat members of elite society who were experiencing psychological distress. It is

from Mesmer’s name that we get the English word, “mesmerize” meaning “to entrance or transfix a person’s attention.” Mesmer attributed the effect of hypnosis to “animal magnetism,” a supposed universal force (similar to gravity) that operates through all human bodies. Even at the time, such an account of hypnosis was not scientifically supported, and Mesmer himself was frequently the center of controversy.

Over the years, researchers have proposed that **hypnosis** is a mental state characterized by reduced peripheral awareness and increased focus on a singular stimulus, which results in an enhanced susceptibility to suggestion ([Kihlstrom, 2003](#)). For example, the hypnotist will usually induce hypnosis by getting the person to pay attention only to the hypnotist’s voice. As the individual focuses more and more on that, s/he begins to forget the context of the setting and responds to the hypnotist’s suggestions as if they were his or her own. Some people are naturally more suggestible, and therefore more

“hypnotizable” than are others, and this is especially true for those who score high in empathy ([Wickramasekera II & Szlyk, 2003](#)). One common “trick” of stage hypnotists is to discard volunteers who are less suggestible than others.

Dissociation is the separation of one’s awareness from everything besides what one is centrally focused on. For example, if you’ve ever been daydreaming in class, you were likely so caught up in the fantasy that you didn’t hear a word the teacher said. During hypnosis, this dissociation becomes even more extreme. That is, a person concentrates so much on the words of the hypnotist that s/he loses perspective of the rest of the world around them. As a consequence of dissociation, a person is less effortful, and less self-conscious in consideration of his or her own thoughts and behaviors. Similar to low awareness states, where one often acts on the first thought that comes to mind, so, too, in hypnosis does the individual simply follow the first thought that comes to mind, i.e., the hypnotist’s suggestion. Still, just because one is more susceptible to suggestion under hypnosis, it doesn’t mean s/he will do anything that’s ordered. To be hypnotized, you must first *want* to be hypnotized (i.e., you can’t be hypnotized against your will; [Lynn & Kirsh, 2006](#)), and once you are hypnotized, you won’t do anything you wouldn’t also do while in a more natural state of consciousness ([Lynn, Rhue, & Weekes, 1990](#)).

Today, **hypnotherapy** is still used in a variety of formats, and it has evolved from Mesmer’s early tinkering with the concept. Modern hypnotherapy often uses a combination of relaxation, suggestion, motivation and expectancies to create a desired mental or behavioral state. Although there is mixed evidence on whether

hypnotherapy can help with addiction reduction (e.g., quitting smoking; [Abbot et al., 1998](#)) there is some evidence that it can be successful in treating sufferers of acute and chronic pain ([Ewin, 1978](#); [Syrjala et al., 1992](#)). For example, one study examined the treatment of burn patients with either hypnotherapy, pseudo-hypnosis (i.e., a placebo condition), or no treatment at all. Afterward, even though people in the placebo condition experienced a 16% decrease in pain, those in the actual hypnosis condition experienced a reduction of nearly 50% ([Patterson et al., 1996](#)). Thus, even though hypnosis may be sensationalized for television and movies, its ability to disassociate a person from their environment (or their pain) in conjunction with increased suggestibility to a clinician's recommendations (e.g., "you will feel less anxiety about your chronic pain") is a documented practice with actual medical benefits.

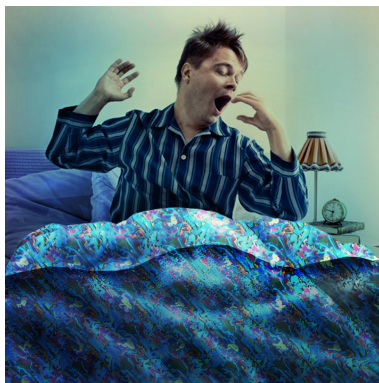
Now, similar to hypnotic states, **trance states** also involve a dissociation of the self; however, people in a trance state are said to have less voluntary control over their behaviors and actions. Trance states often occur in religious ceremonies, where the person believes he or she is "possessed" by an otherworldly being or force. While in trance, people report anecdotal accounts of a "higher consciousness" or communion with a greater power. However, the body of research investigating this phenomenon tends to reject the claim that these experiences constitute an "altered state of consciousness."

Most researchers today describe both hypnosis and trance states as "subjective" alterations of consciousness, not an actually distinct or evolved form ([Kirsch & Lynn, 1995](#)). Just like you feel different when you're in a state of deep relaxation, so, too, are hypnotic and trance states

simply shifts from the standard conscious experience. Researchers contend that even though both hypnotic and trance states appear and feel wildly different than the normal human experience, they can be explained by standard socio-cognitive factors like imagination, expectation, and the interpretation of the situation.

SLEEP

You may have experienced the sensation— as you are falling asleep— of falling and then found yourself physically jerking forward and grabbing out as if you were really falling. Sleep is a unique state of consciousness; it lacks full awareness but the brain is still active. People generally follow a “biological clock” that impacts when they naturally become drowsy,



Sleep is necessary in order for people to function well. [Image: jaci XIII, <https://goo.gl/pog6Fr>, CC BY-NC 2.0, <https://goo.gl/Fllc2e>]

when they fall asleep, and the time they naturally awaken. The hormone **melatonin** increases at night and is associated with becoming sleepy. Your natural daily rhythm, or **Circadian Rhythm**, can be influenced by the amount of daylight to which you are exposed as well as your work and activity schedule. Changing your location, such as flying from Canada to England, can disrupt your natural sleep rhythms, and we call this **jet lag**. You can overcome jet lag by synchronizing yourself to the local

schedule by exposing yourself to daylight and forcing yourself to stay awake even though you are naturally sleepy.

Interestingly, sleep itself is more than shutting off for the night (or for a nap). Instead of turning off like a light with a flick of a switch, your shift in consciousness is reflected in your brain's electrical activity. While you are awake and alert your brain activity is marked by *beta* waves. Beta waves are characterized by being high in frequency but low in intensity. In addition, they are the most inconsistent brain wave and this reflects the wide variation in sensory input that a person processes during the day. As you begin to relax these change to *alpha* waves. These waves reflect brain activity that is less frequent, more consistent and more intense. As you slip into actual sleep you transition through many stages. Scholars differ on how they characterize sleep stages with some experts arguing that there are four distinct stages ([Manoach et al., 2010](#)), while others recognize five ([Šušmáková, & Krakovská, 2008](#)) but they all distinguish between those that include rapid eye movement (REM) and those that are non-rapid eye movement (NREM). In addition, each stage is typically characterized by its own unique pattern of brain activity:

- Stage 1 (called NREM 1, or N1) is the “falling asleep” stage and is marked by theta waves.
- Stage 2 (called NREM 2, or N2) is considered a light sleep. Here, there are occasional “sleep spindles,” or very high intensity brain waves. These are thought to be associated with the processing of memories. NREM 2 makes up about 55% of all sleep.

- Stage 3 (called NREM 3, or N3) makes up between 20-25% of all sleep and is marked by greater muscle relaxation and the appearance of delta waves.
- Finally, REM sleep is marked by rapid eye movement (REM). Interestingly, this stage—in terms of brain activity—is similar to wakefulness. That is, the brain waves occur less intensely than in other stages of sleep. REM sleep accounts for about 20% of all sleep and is associated with dreaming.

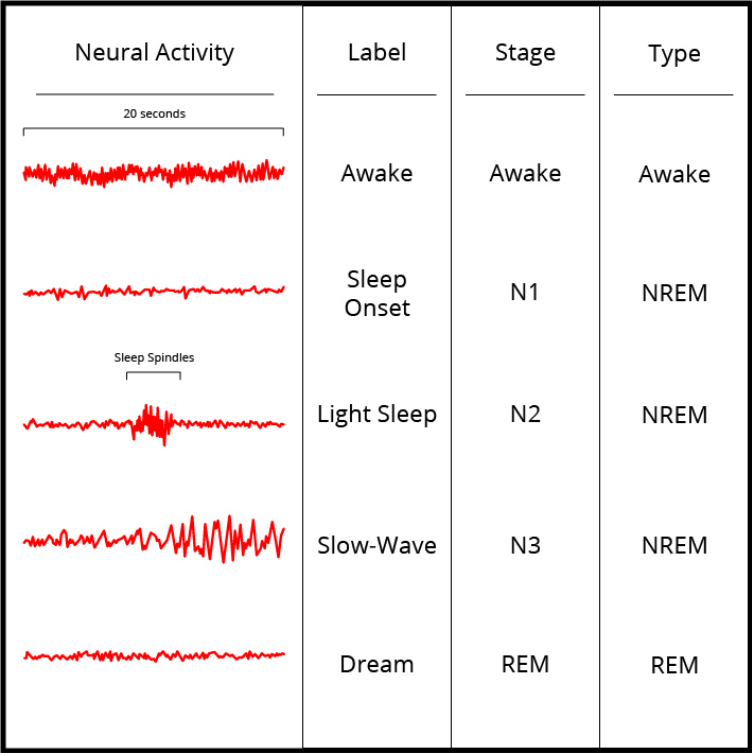


Figure 1. Changes in brain activity or brainwaves across different stages of consciousness – from being awake and throughout various stages of sleep.
 [Image: Noba]

Dreams are, arguably, the most interesting aspect of sleep. Throughout history dreams have been given special importance because of their unique, almost mystical nature. They have been thought to be predictions of the future, hints of hidden aspects of the self, important lessons about how to live life, or opportunities to engage in impossible deeds like flying. There are several competing theories of why humans dream. One is that it is our nonconscious attempt to make sense of our daily experiences and learning. Another, popularized by Freud, is that dreams represent taboo or troublesome wishes

or desires. Regardless of the specific reason we know a few facts about dreams: all humans dream, we dream at every stage of sleep, but dreams during REM sleep are especially vivid. One under-explored area of dream research is the possible social functions of dreams: we often share our dreams with others and use them for entertainment value.

Sleep serves many functions, one of which is to give us a period of mental and physical restoration. Children generally need more sleep than adults since they are developing. It is so vital, in fact, that a lack of sleep is associated with a wide range of problems. People who do not receive adequate sleep are more irritable, have slower reaction time, have more difficulty sustaining attention, and make poorer decisions. Interestingly, this is an issue relevant to the lives of college students. In one highly cited study researchers found that 1 in 5 students took more than 30 minutes to fall asleep at night, 1 in 10 occasionally took sleep medications, and more than half reported being “mostly tired” in the mornings ([Buboltz, et al, 2001](#)).

PSYCHOACTIVE DRUGS

On April 16, 1943, Albert Hoffman—a Swiss chemist working in a pharmaceutical company—accidentally ingested a newly synthesized drug. The drug—lysergic acid diethylimide (LSD)—turned out to be a powerful hallucinogen. Hoffman went home and later reported the effects of the drug, describing them as seeing the world through a “warped mirror” and experiencing visions of “extraordinary shapes with intense, kaleidoscopic play of colors.” Hoffman had discovered what members of many

traditional cultures around the world already knew: there are substances that, when ingested, can have a powerful effect on perception and on consciousness.

Drugs operate on human physiology in a variety of ways and researchers and medical doctors tend to classify drugs according to their effects. Here we will briefly cover 3 categories of drugs: hallucinogens, depressants, and stimulants.

HALLUCINOGENS

It is possible that hallucinogens are the substance that have, historically, been used the most widely. Traditional societies have used plant-based hallucinogens such as peyote, ebene, and psilocybin mushrooms in a wide range of religious ceremonies. **Hallucinogens** are substances that alter a person's perceptions, often by creating visions or hallucinations that are not real. There are a wide range of hallucinogens and many are used as recreational substances in industrialized societies. Common examples include marijuana, LSD, and MDMA (also known as "ecstasy"). Marijuana is the dried flowers of the hemp plant and is often smoked to produce **euphoria**. The active ingredient in marijuana is called THC and can produce distortions in the perception of time, can create a sense of rambling, unrelated thoughts, and is sometimes associated with increased hunger or excessive laughter. The use and possession of marijuana is illegal in most places but this appears to be a trend that is changing. Uruguay, Bangladesh, and several of the United States, have recently legalized marijuana. This may be due, in part, to changing public attitudes or to the fact that

marijuana is increasingly used for medical purposes such as the management of nausea or treating glaucoma.

DEPRESSANTS

Depressants are substances that, as their name suggests, slow down the body's physiology and mental processes. Alcohol is the most widely used depressant. Alcohol's effects include the reduction of inhibition, meaning that intoxicated people are more likely to act in ways they would otherwise be reluctant to. Alcohol's psychological effects are the result of it increasing the neurotransmitter GABA. There are also physical effects, such as loss of balance and coordination, and these stem from the way that alcohol interferes with the coordination of the visual and motor systems of the brain. Despite the fact that alcohol is so widely accepted in many cultures it is also associated with a variety of dangers. First, alcohol is toxic, meaning that it acts like a poison because it is possible to drink more alcohol than the body can effectively remove from the bloodstream. When a person's **blood alcohol content (BAC)** reaches .3 to .4% there is a serious risk of death. Second, the lack of judgment and physical control associated with alcohol is associated with more risk taking behavior or dangerous behavior such as drunk driving. Finally, alcohol is addictive and heavy drinkers often experience significant interference with their ability to work effectively or in their close relationships.

Other common depressants include opiates (also called "narcotics"), which are substances synthesized from the poppy flower. Opiates stimulate endorphin production in the brain and because of this they are often used as pain killers by medical professionals. Unfortunately, because

opiates such as Oxycontin so reliably produce euphoria they are increasingly used—illegally—as recreational substances. Opiates are highly addictive.

STIMULANTS

Stimulants are substances that “speed up” the body’s physiological and mental processes. Two commonly used stimulants are caffeine—the drug found in coffee and tea—and nicotine, the active drug in cigarettes and other tobacco products. These substances are both legal and relatively inexpensive, leading to their widespread use. Many people are attracted to stimulants because they feel more alert when under the influence of these drugs. As with any drug there are health risks associated with consumption. For example, excessive consumption of these types of stimulants can result in anxiety, headaches, and insomnia. Similarly, smoking cigarettes—the most common means of ingesting nicotine—is associated with higher risks of cancer. For instance, among heavy smokers 90% of lung cancer is directly attributable to smoking ([Stewart & Kleihues, 2003](#)).

There are other stimulants such as cocaine and methamphetamine (also known as “crystal meth” or “ice”)



Caffeine is the most widely consumed stimulant in the world. Be honest, how many cups of coffee, tea, or energy drinks have you had today? [Image: Personeelsnet, <https://goo.gl/h0GQ3R>, CC BY-SA 2.0, <https://goo.gl/iZlxAE>]

that are illegal substances that are commonly used. These substances act by blocking “re-uptake” of dopamine in the brain. This means that the brain does not naturally clear out the dopamine and that it builds up in the synapse, creating euphoria and alertness. As the effects wear off it stimulates strong cravings for more of the drug. Because of this these powerful stimulants are highly addictive.

CONCLUSION

When you think about your daily life it is easy to get lulled into the belief that there is one “setting” for your conscious thought. That is, you likely believe that you hold the same opinions, values, and memories across the day and throughout the week. But “you” are like a dimmer switch on a light that can be turned from full darkness increasingly on up to full brightness. This switch is consciousness. At your brightest setting you are fully alert and aware; at dimmer settings you are day dreaming; and sleep or being knocked unconscious represent dimmer settings still. The degree to which you are in high, medium, or low states of conscious awareness affect how susceptible you are to persuasion, how clear your judgment is, and how much detail you can recall. Understanding levels of awareness, then, is at the heart of understanding how we learn, decide, remember and many other vital psychological processes.

OUTSIDE RESOURCES

Book: A wonderful book about how little we know about ourselves: Wilson, T. D. (2004). *Strangers to ourselves*. Cambridge, MA: Harvard University Press.

<http://www.hup.harvard.edu/>

catalog.php?isbn=9780674013827

Book: Another wonderful book about free will—or its absence?: Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

<https://mitpress.mit.edu/books/illusion-conscious-will>

Information on alcoholism, alcohol abuse, and treatment:

<http://www.niaaa.nih.gov/alcohol-health/support-treatment>

The American Psychological Association has information on getting a good night's sleep as well as on sleep disorders

<http://www.apa.org/helpcenter/sleep-disorders.aspx>

The LSD simulator: This simulator uses optical illusions to simulate the hallucinogenic experience of LSD. Simply follow the instructions in this two minute video. After looking away you may see the world around you in a warped or pulsating way similar to the effects of LSD. The effect is temporary and will disappear in about a minute.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=122#oembed-1>

The National Sleep Foundation is a non-profit with videos on insomnia, sleep training in children, and other topics

<https://sleepfoundation.org/video-library>

Video: An artist who periodically took LSD and drew self-portraits:

<http://www.openculture.com/2013/10/artist-draws-nine-portraits-on-lsd-during-1950s-research-experiment.html>

Video: An interesting video on attention:

<http://www.dansimons.com/videos.html>

Video: Clip on out-of-body experiences induced using virtual reality.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=122#oembed-2>

Video: Clip on the rubber hand illusion, from the BBC science series “Horizon.”



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=122#oembed-3>

Video: Demonstration of motion-induced blindness – Look steadily at the blue moving pattern. One or more of the yellow spots may disappear:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=122#oembed-4>

Video: Imaging the Brain, Reading the Mind – A talk by Marsel Mesulam.

http://video.at.northwestern.edu/lores/SO_marsel.m4v

Video: Lucas Handwerker – a stage hypnotist discusses the therapeutic aspects of hypnosis:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=122#oembed-5>

Video: Ted Talk – Simon Lewis: Don't take consciousness for granted

https://www.ted.com/talks/simon_lewis_don_t_take_consciousness_for_granted

Video: TED Talk on Dream Research:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=122#oembed-6>

Video: The mind-body problem – An interview with Ned Block:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=122#oembed-7>

Want a quick demonstration of priming? (Want a quick demonstration of how powerful these effects can be? Check out:



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=122#oembed-8>

Web: A good overview of priming:

[http://en.wikipedia.org/wiki/Priming_\(psychology\)](http://en.wikipedia.org/wiki/Priming_(psychology))

Web: Learn more about motion-induced blindness on Michael Bach's website:

<http://www.michaelbach.de/ot/mot-mib/index.html>

DISCUSSION QUESTIONS

1. If someone were in a coma after an accident, and you wanted to better understand how “conscious” or aware s/he were, how might you go about it?
2. What are some of the factors in daily life that interfere with people’s ability to get adequate sleep? What interferes with your sleep?
3. How frequently do you remember your dreams? Do you have recurring images or themes in your dreams? Why do you think that is?
4. Consider times when you fantasize or let your mind wander? Describe these times: are you more likely to be alone or with others? Are there certain activities you engage in that seem particularly prone to daydreaming?
5. A number of traditional societies use consciousness altering substances in ceremonies. Why do you think they do this?
6. Do you think attitudes toward drug use are changing over time? If so, how? Why do you think these changes occur?
7. Students in high school and college are increasingly using stimulants such as Adderol as study aids and “performance enhancers.” What is your opinion of this trend?

VOCABULARY

Blood Alcohol Content (BAC)

Blood Alcohol Content (BAC): a measure of the percentage of alcohol found in a person’s blood.

This measure is typically the standard used to determine the extent to which a person is intoxicated, as in the case of being too impaired to drive a vehicle.

Circadian Rhythm

Circadian Rhythm: The physiological sleep-wake cycle. It is influenced by exposure to sunlight as well as daily schedule and activity. Biologically, it includes changes in body temperature, blood pressure and blood sugar.

Consciousness

Consciousness: the awareness or deliberate perception of a stimulus

Cues

Cues: a stimulus that has a particular significance to the perceiver (e.g., a sight or a sound that has special relevance to the person who saw or heard it)

Depressants

Depressants: a class of drugs that slow down the body's physiological and mental processes.

Dissociation

Dissociation: the heightened focus on one stimulus or thought such that many other things around you are ignored; a disconnect between one's awareness of their environment and the one object the person is focusing on

Euphoria

Euphoria: an intense feeling of pleasure, excitement or happiness.

Flexible Correction Model

Flexible Correction Model: the ability for people to correct or change their beliefs and evaluations if they believe these judgments have been biased (e.g., if someone realizes they only thought their day was great because it was sunny, they may revise their evaluation of the day to account for this “biasing” influence of the weather)

Hallucinogens

Hallucinogens: substances that, when ingested, alter a person’s perceptions, often by creating hallucinations that are not real or distorting their perceptions of time.

Hypnosis

Hypnosis: the state of consciousness whereby a person is highly responsive to the suggestions of another; this state usually involves a dissociation with one’s environment and an intense focus on a single stimulus, which is usually accompanied by a sense of relaxation

Hypnotherapy

Hypnotherapy: The use of hypnotic techniques such as relaxation and suggestion to help engineer desirable change such as lower pain or quitting smoking.

Implicit Associations Test

Implicit Associations Test (IAT): A computer reaction time test that measures a person’s automatic associations with concepts. For instance, the IAT could be used to measure how quickly a

person makes positive or negative evaluations of members of various ethnic groups.

Jet Lag

Jet Lag: The state of being fatigued and/or having difficulty adjusting to a new time zone after traveling a long distance (across multiple time zones).

Melatonin

Melatonin: A hormone associated with increased drowsiness and sleep.

Mindfulness

Mindfulness: a state of heightened focus on the thoughts passing through one's head, as well as a more controlled evaluation of those thoughts (e.g., do you reject or support the thoughts you're having?)

Priming

Priming: the activation of certain thoughts or feelings that make them easier to think of and act upon

Stimulants

Stimulants: a class of drugs that speed up the body's physiological and mental processes.

Trance States

Trance: a state of consciousness characterized by the experience of "out-of-body possession," or an acute dissociation between one's self and the current, physical environment surrounding them.

REFERENCES

- Abbot, N. C., Stead, L. F., White, A. R., Barnes, J., & Ernst, E. (1998). Hypnotherapy for smoking cessation. *Cochrane Database of Systematic Reviews*, 2.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71(2), 230.
- Buboltz, W., Brown, F. & Soper, B. (2001). Sleep habits and patterns of college students: A preliminary study. *Journal of American College Health*, 50, 131-135.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752.
- Ewin, D. M. (1978). *Clinical use of hypnosis for attenuation of burn depth. Hypnosis at its Bicentennial-Selected Papers from the Seventh International Congress of Hypnosis and Psychosomatic Medicine*. New York: Plenum Press.
- Freud, S. (2001). *The Standard Edition of the Complete Psychological Works of Sigmund Freud: The Interpretation of Dreams (First Part) (Vol. 4)*. Random House.
- Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology*,

60(4), 509.

- Greenwald, A. G., & Farnham, S. D. (2000). Using the Implicit Association Test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, 79, 1022-1038.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. K. L. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74, 1464-1480.
- Kihlstrom, J.F. (2003). Hypnosis and memory. In J.F. Byrne (Ed.), *Learning and memory*, 2nd ed. (pp. 240-242). Farmington Hills, Mi.: Macmillan Reference
- Kirsch, I., & Lynn, S. J. (1995). Altered state of hypnosis: Changes in the theoretical landscape. *American Psychologist*, 50(10), 846.
- Lynn S. J., and Kirsch I. (2006). *Essentials of clinical hypnosis*. Washington, DC: American Psychological Association.
- Lynn, S. J., Rhue, J. W., & Weekes, J. R. (1990). Hypnotic involuntariness: A social-cognitive analysis. *Psychological Review*, 97, 169–184.
- Manoach, D. S., Thakkar, K. N., Stroynowski, E., Ely, A., McKinley, S. K., Wamsley, E., ... & Stickgold, R. (2010). Reduced overnight consolidation of procedural learning in chronic medicated schizophrenia is related to specific sleep stages. *Journal of Psychiatric Research*, 44(2), 112-120.

- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration website. *Group Dynamics*, 6(1), 101-115.
- Patterson, D. R., Everett, J. J., Burns, G. L., & Marvin, J. A. (1992). Hypnosis for the treatment of burn pain. *Journal of Consulting and Clinical Psychology*, 60, 713-17
- Pekala, R. J., Kumar, V. K., Maurer, R., Elliott-Carter, N., Moon, E., & Mullen, K. (2010). Suggestibility, expectancy, trance state effects, and hypnotic depth: I. Implications for understanding hypnotism. *American Journal of Clinical Hypnosis*, 52(4), 275-290.
- Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 19, pp. 123-205). New York: Academic Press.
- Raichle, M. E. (2015). The brain's default mode network. *Annual Review of Neuroscience*, 38, 433-447.
- Stewart, B. & Kleinhues, P. (2003). *World cancer report*. World Health Organization.
- Syrjala, K. L., Cummings, C., & Donaldson, G. W. (1992). Hypnosis or cognitive behavioral training for the reduction of pain and nausea during cancer treatment: A controlled clinical trial. *Pain*, 48, 137-46.
- Wegener, D. T., & Petty, R. E. (1997). The flexible

correction model: The role of naive theories of bias in bias correction. *Advances in Experimental Social Psychology*, 29, 142-208.

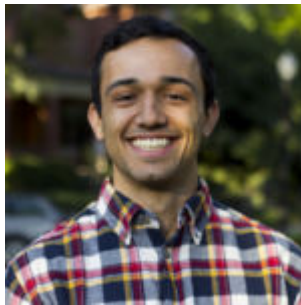
- Wickramasekera II, I. E., & Szlyk, J. (2003). Could empathy be a predictor of hypnotic ability? *International Journal of Clinical and Experimental Hypnosis*, 51(4), 390–399.
- Williams, L. E., & Bargh, J. A. (2008). Experiencing physical warmth promotes interpersonal warmth. *Science*, 322(5901), 606-607.
- Öhman, A., & Soares, J. J. (1994). \"Unconscious anxiety\": phobic responses to masked stimuli. *Journal of Abnormal Psychology*, 103(2), 231.
- Šušmáková, K., & Krakovská, A. (2008). Discrimination ability of individual measures used in sleep stages classification. *Artificial Intelligence in Medicine*, 44(3), 261-277.

AUTHORS



Robert Biswas-Diener

Dr. Robert Biswas-Diener is a part-time instructor at Portland State University and is senior editor of Noba. He has more than 50 publications on happiness and other positive topics in peer-reviewed journals. He is author of *The Upside of Your Dark Side*.



Jake Teeny

Jake Teeny received his M.A. in social

psychology from The Ohio State University, where he now pursuing a PhD. His research examines when and why people advocate their beliefs. In addition to his psychological work, he also writes short fiction, both of which can be found at www.jaketeeny.com

CREATIVE COMMONS LICENSE



STATES OF CONSCIOUSNESS by [Robert Biswas-Diener and Jake Teeny](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

PART IX.

**UNIT 9: REPRODUCTIVE
BEHAVIORS**

CHAPTER 14.

HORMONES & BEHAVIOR

RANDY J. NELSON

The goal of this module is to introduce you to the topic of hormones and behavior. This field of study is also called behavioral endocrinology, which is the scientific study of the interaction between hormones and behavior. This interaction is bidirectional: hormones can influence behavior, and behavior can sometimes influence hormone concentrations. Hormones are chemical messengers released from endocrine glands that travel through the blood system to influence the nervous system to regulate behaviors such as aggression, mating, and parenting of individuals.

LEARNING OBJECTIVES

- Define the basic terminology and basic principles of hormone–behavior interactions.
- Explain the role of hormones in behavioral sex differentiation.
- Explain the role of hormones in aggressive behavior.

- Explain the role of hormones in parental behavior.
- Provide examples of some common hormone–behavior interactions.

INTRODUCTION

This module describes the relationship between hormones and behavior. Many readers are likely already familiar with the general idea that hormones can affect behavior. Students are generally familiar with the idea that sex-hormone concentrations increase in the blood during puberty and decrease as we age, especially after about 50 years of age. Sexual behavior shows a similar pattern. Most people also know about the relationship between aggression and anabolic steroid hormones, and they know that administration of artificial steroid hormones sometimes results in uncontrollable, violent behavior called “roid rage.” Many different hormones can influence several types of behavior, but for the purpose of this module, we will restrict our discussion to just a few examples of hormones and behaviors. For example, are behavioral sex differences the result of hormones, the environment, or some combination of factors? Why are men much more likely than women to commit aggressive acts? Are hormones involved in mediating the so-called maternal “instinct”? Behavioral endocrinologists are interested in how the general physiological effects of hormones alter the development and expression of behavior and how behavior may influence the effects of hormones. This module describes, both phenomenologically and functionally, how hormones affect behavior.

To understand the hormone-behavior relationship, it

is important briefly to describe hormones. **Hormones** are organic chemical messengers produced and released by specialized glands called **endocrine glands**. Hormones are released from these glands into the blood, where they may travel to act on target structures at some distance from their origin. Hormones are similar in function to **neurotransmitters**, the chemicals used by the nervous system in coordinating animals' activities. However, hormones can operate over a greater distance and over a much greater temporal range than neurotransmitters (Focus Topic 1). Examples of hormones that influence behavior include steroid hormones such as **testosterone** (a common type of androgen), estradiol (a common type of estrogen), progesterone (a common type of **progestin**), and cortisol (a common type of glucocorticoid) (Table 1, A-B). Several types of protein or peptide (small protein) hormones also influence behavior, including **oxytocin**, vasopressin, **prolactin**, and leptin.

Focus Topic 1: Neural Transmission versus Hormonal Communication

Although neural and hormonal communication both rely on chemical signals, several prominent differences exist. Communication in the nervous system is analogous to traveling on a train. You can use the train in your travel plans as long as tracks exist between your proposed origin and destination. Likewise, neural messages can travel only to destinations along existing nerve tracts.

Hormonal communication, on the other hand, is like traveling in a car. You can drive to many more destinations than train travel allows because there are many more roads than railroad tracks. Similarly, hormonal messages can travel anywhere in the body via the circulatory system; any cell receiving blood is potentially able to receive a hormonal message.

Neural and hormonal communication differ in other ways as well. To illustrate them, consider the differences between digital and analog technologies. Neural messages are digital, all-or-none events that have rapid onset and offset: neural signals can take place in milliseconds. Accordingly, the nervous system mediates changes in the body that are relatively rapid. For example, the nervous system regulates immediate food intake and directs body movement. In contrast, hormonal messages are analog, graded events that may take seconds, minutes, or even hours to occur. Hormones can mediate long-term processes, such as growth, development, reproduction, and metabolism.

Hormonal and neural messages are both chemical in nature, and they are released and received by cells in a similar manner; however, there are important differences as well. Neurotransmitters, the chemical messengers used by neurons, travel a distance of only 20–30 nanometers (30×10^{-9} m)—to the membrane of the postsynaptic neuron, where they bind with

receptors. Hormones enter the circulatory system and may travel from 1 millimeter to >2 meters before arriving at a target cell, where they bind with specific receptors.

Another distinction between neural and hormonal communication is the degree of voluntary control that can be exerted over their functioning. In general, there is more voluntary control of neural than of hormonal signals. It is virtually impossible to will a change in your thyroid hormone levels, for example, whereas moving your limbs on command is easy.

Although these are significant differences, the division between the nervous system and the endocrine system is becoming more blurred as we learn more about how the nervous system regulates hormonal communication. A better understanding of the interface between the endocrine system and the nervous system, called neuroendocrinology, is likely to yield important advances in the future study of the interaction between hormones and behavior.

Steroid Hormones	
Cortisol	Increases carbohydrate metabolism; mediates stress responses
Estradiol	Uterine and other female tissue development; regulates sexual motivation and performance in females and males
Testosterone	Promotes sperm production and male secondary sexual characteristics; promotes sexual motivation and behavior, typically by being converted to estradiol

Table 1-A: Prominent Hormones That Influence Behavior

Peptides and Protein Hormones	
Oxytocin	Stimulates milk letdown and uterine contractions during birth; Promotes social bonding
Prolactin	Many actions relating to reproduction, water balance, and behavior associated with parental care
Thyroxine	Increases oxidation rates in tissue and affects neural development
Vasopressin	Increases water reabsorption in the kidney and affects learning and memory

Table 1-B: Prominent Hormones That Influence Behavior

Hormones coordinate the physiology and behavior of individuals by regulating, integrating, and controlling bodily functions. Over evolutionary time, hormones have often been co-opted by the nervous system to influence behavior to ensure reproductive success. For example, the same hormones, testosterone and estradiol, that cause gamete (egg or sperm) maturation also promote mating behavior. This dual hormonal function ensures that mating behavior occurs when animals have mature gametes available for fertilization. Another example of endocrine regulation of physiological and behavioral function is provided by pregnancy. Estrogens and progesterone concentrations are elevated during pregnancy, and these hormones are often involved in mediating **maternal behavior** in the mothers.

Not all cells are influenced by each and every hormone.

Rather, any given hormone can directly influence only cells that have specific hormone **receptors** for that particular hormone. Cells that have these specific receptors are called **target cells** for the hormone. The interaction of a hormone with its receptor begins a series of cellular events that eventually lead to activation of enzymatic pathways or, alternatively, turns on or turns off gene activation that regulates protein synthesis. The newly synthesized proteins may activate or deactivate other genes, causing yet another cascade of cellular events. Importantly, sufficient numbers of appropriate hormone receptors must be available for a specific hormone to produce any effects. For example, testosterone is important for male sexual behavior. If men have too little testosterone, then sexual motivation may be low, and it can be restored by testosterone treatment. However, if men have normal or even elevated levels of testosterone yet display low sexual drive, then it might be possible for a lack of receptors to be the cause and treatment with additional hormones will not be effective.

How might hormones affect behavior? In terms of their behavior, one can think of humans and other animals conceptually as comprised of three interacting components: (1) input systems (sensory systems), (2) integrators (the central nervous system), and (3) output systems, or effectors (e.g., muscles). Hormones do not *cause* behavioral changes. Rather, hormones influence these three systems so that specific stimuli are more likely to elicit certain responses in the appropriate behavioral or social context. In other words, hormones change the probability that a particular behavior will be emitted in the appropriate situation ([Nelson, 2011](#)). This is a critical

distinction that can affect how we think of hormone-behavior relationships.

We can apply this three-component behavioral scheme to a simple behavior, singing in zebra finches. Only male zebra finches sing. If the testes of adult male finches are removed, then the birds reduce singing, but castrated finches resume singing if the testes are reimplanted, or if the birds are treated with either testosterone or estradiol. Although we commonly consider androgens to be “male” hormones and estrogens to be “female” hormones, it is common for testosterone to be converted to estradiol in nerve cells (Figure 1). Thus, many male-like behaviors are associated with the actions of estrogens! Indeed, all estrogens must first be converted from androgens because of the typical biochemical synthesis process. If the converting enzyme is low or missing, then it is possible for females to produce excessive androgens and subsequently develop associated male traits. It is also possible for estrogens in the environment to affect the nervous system of animals, including people (e.g., [Kidd et al., 2007](#)). Again, singing behavior is most frequent when blood testosterone or estrogen concentrations are high. Males sing to attract mates or ward off potential competitors from their territories.

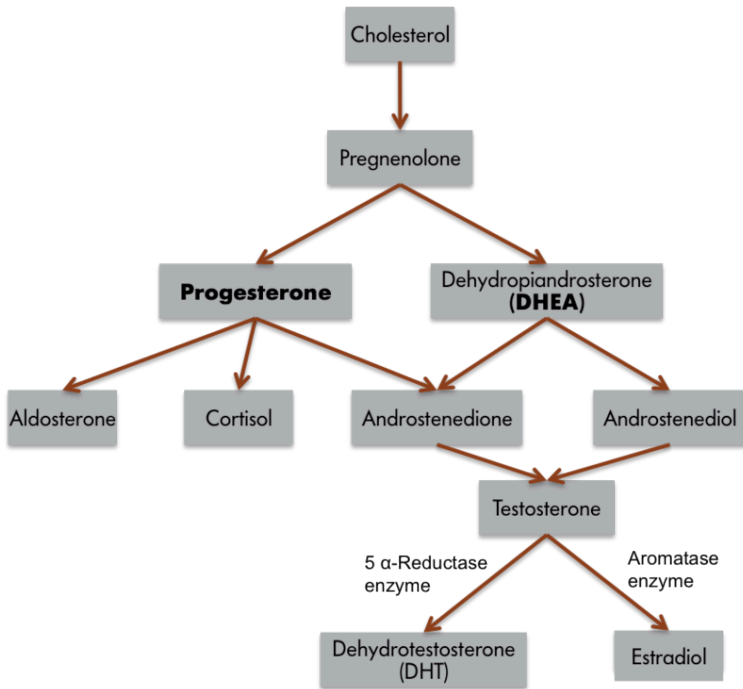


Figure 1. Biochemical Pathway for Steroid Hormone Synthesis: It is important to note that testosterone (an androgen) can be converted to another androgen, DHT, or an estrogen, estradiol. Too much or too little of the converting enzymes can influence brain and behavior.

Although it is apparent from these observations that estrogens are somehow involved in singing, how might the three-component framework just introduced help us to formulate hypotheses to explore estrogen's role in this behavior? By examining input systems, we could determine whether estrogens alter the birds' sensory capabilities, making the environmental cues that normally elicit singing more salient. If this were the case, then females or competitors might be more easily seen or heard. Estrogens also could influence the central nervous system. Neuronal architecture or the speed of neural

processing could change in the presence of estrogens. Higher neural processes (e.g., motivation, attention, or perception) also might be influenced. Finally, the effector organs, muscles in this case, could be affected by the presence of estrogens. Blood estrogen concentrations might somehow affect the muscles of a songbird's syrinx (the vocal organ of birds). Estrogens, therefore, could affect birdsong by influencing the sensory capabilities, central processing system, or effector organs of an individual bird. We do not understand completely how estrogen, derived from testosterone, influences birdsong, but in most cases, hormones can be considered to affect behavior by influencing one, two, or all three of these components, and this three-part framework can aid in the design of hypotheses and experiments to explore these issues.

How might behaviors affect hormones? The birdsong example demonstrates how hormones can affect behavior, but as noted, the reciprocal relation also occurs; that is, behavior can affect hormone concentrations. For example, the sight of a territorial intruder may elevate blood testosterone concentrations in resident male birds and thereby stimulate singing or fighting behavior. Similarly, male mice or rhesus monkeys that lose a fight decrease circulating testosterone concentrations for several days or even weeks afterward. Comparable results have also been reported in humans. Testosterone concentrations are affected not only in humans involved in physical combat, but also in those involved in simulated battles. For example, testosterone concentrations were elevated in winners and reduced in losers of regional chess tournaments.



The expectation of events can influence one's hormonal activity. How do you think yours is affected if you anticipate going on a date with a romantic interest soon? [Image: CCO Public Domain, <https://goo.gl/m25gce>]

People do not have to be directly involved in a contest to have their hormones affected by the outcome of the contest. Male fans of both the Brazilian and Italian teams were recruited to provide saliva samples to be assayed for testosterone before and after the final game of the World Cup soccer match in 1994. Brazil and Italy were tied going into the final game, but Brazil won

on a penalty kick at the last possible moment. The Brazilian fans were elated and the Italian fans were crestfallen. When the samples were assayed, 11 of 12 Brazilian fans who were sampled had increased testosterone concentrations, and 9 of 9 Italian fans had decreased testosterone concentrations, compared with pre-game baseline values ([Dabbs, 2000](#)).

In some cases, hormones can be affected by anticipation of behavior. For example, testosterone concentrations also influence sexual motivation and behavior in women. In one study, the interaction between sexual intercourse and testosterone was compared with other activities (cuddling or exercise) in women ([van Anders, Hamilton, Schmidt, & Watson, 2007](#)). On three separate occasions, women provided a pre-activity, post-activity, and next-morning saliva sample. After analysis, the women's testosterone was determined to be elevated

prior to intercourse as compared to other times. Thus, an anticipatory relationship exists between sexual behavior and testosterone. Testosterone values were higher post-intercourse compared to exercise, suggesting that engaging in sexual behavior may also influence hormone concentrations in women.

SEX DIFFERENCES

Hens and roosters are different. Cows and bulls are different. Men and women are different. Even girls and boys are different. Humans, like many animals, are sexually dimorphic (*di*, “two”; *morph*, “type”) in the size and shape of their bodies, their physiology, and for our purposes, their behavior. The behavior of boys and girls differs in many ways. Girls generally excel in verbal abilities relative to boys; boys are nearly twice as likely as girls to suffer from dyslexia (reading difficulties) and stuttering and nearly 4 times more likely to suffer from autism. Boys are generally better than girls at tasks that require visuospatial abilities. Girls engage in nurturing behaviors more frequently than boys. More than 90% of all anorexia nervosa cases involve young women. Young men are twice as likely as young women to suffer from schizophrenia. Boys are much more aggressive and generally engage in more rough-and-tumble play than girls ([Berenbaum, Martin, Hanish, Briggs, & Fabes, 2008](#)). Many sex differences, such as the difference in aggressiveness, persist throughout adulthood. For example, there are many more men than women serving prison sentences for violent behavior. The hormonal differences between men and women may account for adult sex differences that develop during puberty, but

what accounts for behavioral sex differences among children *prior* to puberty and activation of their gonads? Hormonal secretions from the developing gonads determine whether the individual develops in a male or female manner. The mammalian embryonic testes produce androgens, as well as peptide hormones, that steer the development of the body, central nervous system, and subsequent behavior in a male direction. The embryonic ovaries of mammals are virtually quiescent and do not secrete high concentrations of hormones. In the presence of ovaries, or in the complete absence of any gonads, morphological, neural, and, later, behavioral development follows a female pathway.

Gonadal steroid hormones have organizational (or programming) effects upon brain and behavior ([Phoenix, Goy, Gerall, & Young, 1959](#)).

The organizing effects of steroid hormones are relatively constrained to the early stages of development. An asymmetry exists in the effects of testes and ovaries on the organization of behavior in mammals. Hormone exposure early in life has



Sex differences in appearance are often more pronounced in nonhuman animals than in humans. Male birds particularly, for example roosters, tend to have physical features that differ from the females and also differ significantly in size. [Image: John Cudworth, <https://goo.gl/oopnqM>, CC BY-NC 2.0, <https://goo.gl/VnKIK8>]

organizational effects on subsequent rodent behavior; early steroid hormone treatment causes relatively

irreversible and permanent **masculinization** of rodent behavior (mating and aggressive). These early hormone effects can be contrasted with the reversible behavioral influences of steroid hormones provided in adulthood, which are called activational effects. The activational effects of hormones on adult behavior are temporary and may wane soon after the hormone is metabolized. Thus, typical male behavior requires exposure to androgens during gestation (in humans) or immediately after birth (in rodents) to somewhat masculinize the brain and also requires androgens during or after puberty to activate these neural circuits. Typical female behavior requires a lack of exposure to androgens early in life which leads to **feminization** of the brain and also requires estrogens to activate these neural circuits in adulthood. But this simple dichotomy, which works well with animals with very distinct sexual dimorphism in behavior, has many caveats when applied to people.

If you walk through any major toy store, then you will likely observe a couple of aisles filled with pink boxes and the complete absence of pink packaging of toys in adjacent aisles. Remarkably, you will also see a strong self-segregation of boys and girls in these aisles. It is rare to see boys in the “pink” aisles and vice versa. The toy manufacturers are often accused of making toys that are gender biased, but it seems more likely that boys and girls enjoy playing with specific types and colors of toys. Indeed, toy manufacturers would immediately double their sales if they could sell toys to both sexes. Boys generally prefer toys such as trucks and balls and girls generally prefer toys such as dolls. Although it is doubtful that there are genes that encode preferences for toy cars and trucks on the Y chromosome, it is possible that

hormones might shape the development of a child's brain to prefer certain types of toys or styles of play behavior. It is reasonable to believe that children learn which types of toys and which styles of play are appropriate to their gender. How can we understand and separate the contribution of physiological mechanisms from learning to understand sex differences in human behaviors? To untangle these issues, animal models are often used. Unlike the situation in humans, where sex differences are usually only a matter of degree (often slight), in some animals, members of only one sex may display a particular behavior. As noted, often only male songbirds sing. Studies of such strongly sex-biased behaviors are particularly valuable for understanding the interaction among behavior, hormones, and the nervous system.

A study of vervet monkeys calls into question the primacy of learning in the establishment of toy preferences ([Alexander & Hines, 2002](#)). Female vervet monkeys preferred girl-typical toys, such as dolls or cooking pots, whereas male vervet monkeys preferred boy-typical toys, such as cars or balls. There were no sex differences in preference for gender-neutral toys, such as picture books or stuffed animals. Presumably, monkeys have no prior concept of "boy" or "girl" toys. Young rhesus monkeys also show similar toy preferences.



If you think back to the toys and clothing you played with and wore in your youth, do you think they were more a result of your hormonal activity or the choices that society and your parents made for you? [Image: CCO Public Domain, <https://goo.gl/m25gce>]

What then underlies the sex difference in toy preference? It is possible that certain attributes of toys (or objects) appeal to either boys or girls. Toys that appeal to boys or male vervet or rhesus monkeys, in this case, a ball or toy car, are objects that can be moved actively through space, toys that can be incorporated into active, rough and tumble play. The appeal of toys that girls or female vervet monkeys prefer appears to

be based on color. Pink and red (the colors of the doll and pot) may provoke attention to infants.

Society may reinforce such stereotypical responses to gender-typical toys. The sex differences in toy preferences emerge by 12 or 24 months of age and seem fixed by 36 months of age, but are sex differences in toy preference present during the first year of life? It is difficult to ask pre-verbal infants what they prefer, but in studies where the investigators examined the amount of time that babies looked at different toys, eye-tracking data indicate that infants as young as 3 months showed sex differences in toy preferences; girls preferred dolls, whereas boys preferred trucks. Another result that suggests, but does not prove, that hormones are involved in toy preferences is the observation that girls diagnosed with congenital adrenal hyperplasia (CAH), whose

adrenal glands produce varying amounts of androgens early in life, played with masculine toys more often than girls without CAH. Further, a dose-response relationship between the extent of the disorder (i.e., degree of fetal androgen exposure) and degree of masculinization of play behavior was observed. Are the sex differences in toy preferences or play activity, for example, the inevitable consequences of the differential endocrine environments of boys and girls, or are these differences imposed by cultural practices and beliefs? Are these differences the result of receiving gender-specific toys from an early age, or are these differences some combination of endocrine and cultural factors? Again, these are difficult questions to unravel in people.

Even when behavioral sex differences appear early in development, there seems to be some question regarding the influences of societal expectations. One example is the pattern of human play behavior during which males are more physical; this pattern is seen in a number of other species including nonhuman primates, rats, and dogs. Is the difference in the frequency of rough-and-tumble play between boys and girls due to biological factors associated with being male or female, or is it due to cultural expectations and learning? If there is a combination of biological and cultural influences mediating the frequency of rough-and-tumble play, then what proportion of the variation between the sexes is due to biological factors and what proportion is due to social influences? Importantly, is it appropriate to talk about “normal” sex differences when these traits virtually always arrange themselves along a continuum rather than in discrete categories?

Sex differences are common in humans and in

nonhuman animals. Because males and females differ in the ratio of androgenic and estrogenic steroid hormone concentrations, behavioral endocrinologists have been particularly interested in the extent to which behavioral sex differences are mediated by hormones. The process of becoming female or male is called **sexual differentiation**. The primary step in sexual differentiation occurs at fertilization. In mammals, the ovum (which always contains an X chromosome) can be fertilized by a sperm bearing either a Y or an X chromosome; this process is called **sex determination**. The **chromosomal sex** of homogametic mammals (XX) is female; the chromosomal sex of heterogametic mammals (XY) is male. Chromosomal sex determines **gonadal sex**. Virtually all subsequent sexual differentiation is typically the result of differential exposure to gonadal steroid hormones. Thus, gonadal sex determines hormonal sex, which regulates morphological sex. Morphological differences in the central nervous system, as well as in some effector organs, such as muscles, lead to behavioral sex differences. The process of sexual differentiation is complicated, and the potential for errors is present. Perinatal exposure to androgens is the most common cause of anomalous sexual differentiation among females. The source of androgen may be internal (e.g., secreted by the adrenal glands) or external (e.g., exposure to environmental estrogens). Turner syndrome results when the second X chromosome is missing or damaged; these individuals possess dysgenic ovaries and are not exposed to steroid hormones until puberty. Interestingly, women with Turner syndrome often have impaired spatial memory.

Female mammals are considered the “neutral” sex; additional physiological steps are required for male

differentiation, and more steps bring more possibilities for errors in differentiation. Some examples of male anomalous sexual differentiation include **5 α -reductase** deficiency (in which XY individuals are born with ambiguous genitalia because of a lack of dihydrotestosterone and are reared as females, but masculinization occurs during puberty) and androgen insensitivity syndrome or TFM (in which XY individuals lack receptors for androgens and develop as females). By studying individuals who do not neatly fall into the dichotic boxes of female or male and for whom the process of sexual differentiation is atypical, behavioral endocrinologists glean hints about the process of typical sexual differentiation.

We may ultimately want to know how hormones mediate sex differences in the human brain and behavior (to the extent to which these differences occur). To understand the mechanisms underlying sex differences in the brain and behavior, we return to the birdsong example. Birds provide the best evidence that behavioral sex differences are the result of hormonally induced structural changes in the brain ([Goodson, Saldanha, Hahn, & Soma, 2005](#)). In contrast to mammals, in which structural differences in neural tissues have not been directly linked to behavior, structural differences in avian brains have been directly linked to a sexually behavior: birdsong.

Several brain regions in songbirds display significant sex differences in size. Two major brain circuit pathways, (1) the song production motor pathway and (2) the auditory transmission pathway, have been implicated in the learning and production of birdsong. Some parts of the song production pathway of male zebra finches are 3

to 6 times larger than those of female conspecifics. The larger size of these brain areas reflects that neurons in these nuclei are larger, more numerous, and farther apart. Although castration of adult male birds reduces singing, it does not reduce the size of the brain nuclei controlling song production. Similarly, androgen treatment of adult female zebra finches does not induce changes either in singing or in the size of the song control regions. Thus, activational effects of steroid hormones do not account for the sex differences in singing behavior or brain nucleus size in zebra finches. The sex differences in these structures are organized or programmed in the egg by estradiol (masculinizes) or the lack of steroids (feminizes).

Taken together, estrogens appear to be necessary to activate the neural machinery underlying the song system in birds. The testes of birds primarily produce androgens, which enter the circulation. The androgens enter neurons containing **aromatase**, which converts them to estrogens. Indeed, the brain is the primary source of estrogens, which activate masculine behaviors in many bird species.

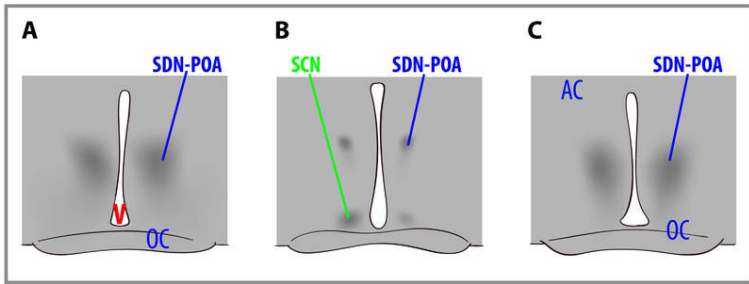


Figure 2. The sexually dimorphic nuclei of the preoptic area (SDN-POA) Gonadal steroid hormones have organizing effects upon brain and behavior. The organizing effects of steroid hormones are relatively constrained to the early stages of development. Exposure to testosterone (which is converted to estradiol) or estradiol causes masculinization of the brain. These are cross-sections through the brains of rats that show a male (left), a female (center), and a female treated with testosterone as a newborn (right). Note that the SDN-POA (the dark cell bodies) of the male are substantially larger than those of the untreated female but are equal in size to those of the testosterone-treated female. The extent that these sex differences in brain structure account for sex differences in behavior remain unspecified in mammals. OC = optic chiasm; SCN = suprachiasmatic nucleus; V = third ventricle.

Sex differences in human brain size have been reported for years. More recently, sex differences in specific brain structures have been discovered (Figure 2). Sex differences in a number of cognitive functions have also been reported. Females are generally more sensitive to auditory information, whereas males are more sensitive to visual information. Females are also typically more sensitive than males to taste and olfactory input. Women display less lateralization of cognitive functions than men. On average, females generally excel in verbal, perceptual, and fine motor skills, whereas males outperform females on quantitative and visuospatial tasks, including map reading and direction finding. Although reliable sex differences can be documented,

these differences in ability are slight. It is important to note that there is more variation *within* each sex than *between* the sexes for most cognitive abilities (Figure 3).

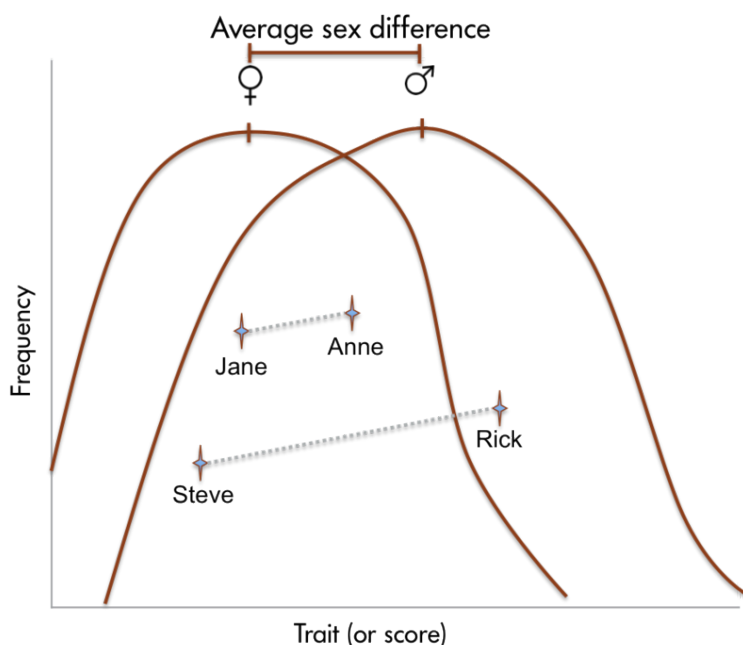


Figure 3. The average sex differences in human performance often reflect significant overlap between the sexes. There are often greater differences in performance between individuals of the same sex (for example, between Steve and Rick in the figure) than between individuals of the opposite sex (for example, between Steve and Jane in the figure).

AGGRESSIVE BEHAVIORS

The possibility for **aggressive behavior** exists whenever the interests of two or more individuals are in conflict (Nelson, 2006). Conflicts are most likely to arise over limited resources such as territories, food, and mates. A

social interaction decides which animal gains access to the contested resource. In many cases, a submissive posture or gesture on the part of one animal avoids the necessity of actual combat over a resource. Animals may also participate in threat displays or ritualized combat in which dominance is determined but no physical damage is inflicted.

There is overwhelming circumstantial evidence that androgenic steroid hormones mediate aggressive behavior across many species. First, seasonal variations in blood plasma concentrations of testosterone and seasonal variations in aggression coincide. For instance, the incidence of aggressive behavior peaks for male deer in autumn, when they are secreting high levels of testosterone. Second, aggressive behaviors increase at the time of puberty, when the testes become active and blood concentrations of androgens rise. Juvenile deer do not participate in the fighting during the mating season. Third, in any given species, males are generally more aggressive than females. This is certainly true of deer; relative to stags, female deer rarely display aggressive behavior, and their rare aggressive acts are qualitatively different from the aggressive behavior of aggressive males. Finally, castration typically reduces aggression in males, and testosterone replacement therapy restores aggression to pre-castration levels. There are some interesting exceptions to these general observations that are outside the scope of this module.

As mentioned, males are generally more aggressive than females. Certainly, human males are much more aggressive than females. Many more men than women are convicted of violent crimes in North America. The sex differences in human aggressiveness appear very early.

At every age throughout the school years, many more boys than girls initiate physical assaults. Almost everyone will acknowledge the existence of this sex difference, but assigning a cause to behavioral sex differences in humans always elicits much debate. It is possible that boys are more aggressive than girls because androgens promote aggressive behavior and boys have higher blood concentrations of androgens than girls. It is possible that boys and girls differ in their aggressiveness because the brains of boys are exposed to androgens prenatally and the “wiring” of their brains is thus organized in a way that facilitates the expression of aggression. It is also possible that boys are encouraged and girls are discouraged by family, peers, or others from acting in an aggressive manner. These three hypotheses are not mutually exclusive, but it is extremely difficult to discriminate among them to account for sex differences in human aggressiveness.

What kinds of studies would be necessary to assess these hypotheses? It is usually difficult to separate out the influences of environment and physiology on the development of behavior in humans. For example, boys and girls differ in their rough-and-tumble play at a very young age, which suggests an early physiological influence on aggression. However, parents interact with their male and female offspring differently; they usually



Researchers have electrically stimulated particular regions in people's brains, and these individuals have burst into aggressive, violent behavior, helping demonstrate that such responses are hardwired into us. [Image: Riccardo Cuppini, <https://goo.gl/b6bHU2>, CC BY-NC-ND 2.0, <https://goo.gl/bhtmlY>]

play more roughly with male infants than with females, which suggests that the sex difference in aggressiveness is partially learned. This difference in parental interaction style is evident by the first week of life. Because of these complexities in the factors influencing human behavior, the study of hormonal effects on sex-differentiated behavior has been pursued in nonhuman animals, for which environmental influences can be held relatively constant. Animal models for which sexual differentiation occurs postnatally are often used so that this process can be easily manipulated experimentally.

Again, with the appropriate animal model, we can address the questions posed above: Is the sex difference in aggression due to higher adult blood concentrations of androgens in males than in females, or are males more

aggressive than females because their brains are organized differently by perinatal hormones? Are males usually more aggressive than females because of an interaction of early and current blood androgen concentrations? If male mice are castrated prior to their sixth day of life, then treated with testosterone propionate in adulthood, they show low levels of aggression. Similarly, females ovariectomized prior to their sixth day but given androgens in adulthood do not express male-like levels of aggression. Treatment of perinatally gonadectomized males or females with testosterone prior to their sixth day life and also in adulthood results in a level of aggression similar to that observed in typical male mice. Thus, in mice, the proclivity for males to act more aggressively than females is organized perinatally by androgens but also requires the presence of androgens after puberty in order to be fully expressed. In other words, aggression in male mice is both organized and activated by androgens. Testosterone exposure in adulthood without prior organization of the brain by steroid hormones does not evoke typical male levels of aggression. The hormonal control of aggressive behavior in house mice is thus similar to the hormonal mediation of heterosexual male mating behavior in other rodent species. Aggressive behavior is both organized and activated by androgens in many species, including rats, hamsters, voles, dogs, and possibly some primate species.

PARENTAL BEHAVIORS

Parental behavior can be considered to be any behavior that contributes directly to the survival of fertilized eggs

or offspring that have left the body of the female. There are many patterns of mammalian parental care. The developmental status of the newborn is an important factor driving the type and quality of parental care in a species. Maternal care is much more common than **paternal** care. The vast majority of research on the hormonal correlates of mammalian parental behavior has been conducted on rats. Rats bear altricial young, and mothers perform a cluster of stereotyped maternal behaviors, including nest building, crouching over the pups to allow nursing and to provide warmth, pup retrieval, and increased aggression directed at intruders. If you expose nonpregnant female rats (or males) to pups, their most common reaction is to huddle far away from them. Rats avoid new things (neophobia). However, if you expose adult rats to pups every day, they soon begin to behave maternally. This process is called concaveation or sensitization and it appears to serve to reduce the adult rats' fear of pups.

Of course a new mother needs to act maternal as soon as her offspring arrive—not in a week. The onset of maternal behavior in rats is mediated by hormones. Several methods of study, such as hormone removal and replacement therapy, have been used to determine the hormonal correlates of rat maternal behavior. A fast decline of blood concentrations of **progesterone** in late pregnancy after sustained high concentrations of this hormone, in combination with high concentrations of estradiol and probably prolactin and oxytocin, induces female rats to behave maternally almost immediately in the presence of pups. This pattern of hormones at parturition overrides the usual fear response of adult rats toward pups, and it permits the onset of maternal

behavior. Thus, the so-called maternal “instinct” requires hormones to increase the approach tendency and lower the avoidance tendency. Laboratory strains of mice and rats are usually docile, but mothers can be quite aggressive toward animals that venture too close to their litter. Progesterone appears to be the primary hormone that induces this maternal aggression in rodents, but species differences exist. The role of maternal aggression in women’s behavior has not been adequately described or tested.



Although cortisol may not directly increase maternal behaviors, the next time your mom gives you a hug you know one hormone to thank. [Image: Maria Grazia Montagnari, <https://goo.gl/LY1Tq0>, CC BY 2.0, <https://goo.gl/BRvSA7>]

A series of elegant experiments by Alison Fleming and her collaborators studied the endocrine correlates of the behavior of human mothers as well as the endocrine correlates of maternal attitudes as expressed in self-report questionnaires. Responses such as patting, cuddling, or kissing the baby were called affectionate behaviors; talking, singing, or cooing to the baby were considered

vocal behaviors. Both affectionate and vocal behaviors were considered approach behaviors. Basic caregiving activities, such as changing diapers and burping the infants, were also recorded. In these studies, no relationship between hormone concentrations and maternal responsiveness, as measured by attitude

questionnaires, was found. For example, most women showed an increasing positive self-image during early pregnancy that dipped during the second half of pregnancy, but recovered after parturition. A related dip in feelings of maternal engagement occurred during late pregnancy, but rebounded substantially after birth in most women. However, when behavior, rather than questionnaire responses, was compared with hormone concentrations, a different story emerged. Blood plasma concentrations of cortisol were positively associated with approach behaviors. In other words, women who had high concentrations of blood cortisol, in samples obtained immediately before or after nursing, engaged in more physically affectionate behaviors and talked more often to their babies than mothers with low cortisol concentrations. Additional analyses from this study revealed that the correlation was even greater for mothers that had reported positive maternal regard (feelings and attitudes) during gestation. Indeed, nearly half of the variation in maternal behavior among women could be accounted for by cortisol concentrations and positive maternal attitudes during pregnancy.

Presumably, cortisol does not induce maternal behaviors directly, but it may act indirectly on the quality of maternal care by evoking an increase in the mother's general level of arousal, thus increasing her responsiveness to infant-generated cues. New mothers with high cortisol concentrations were also more attracted to their infant's odors, were superior in identifying their infants, and generally found cues from infants highly appealing ([Fleming, Steiner, & Corter, 1997](#)).

The medial preoptic area is critical for the expression

of rat maternal behavior. The amygdala appears to tonically inhibit the expression of maternal behavior. Adult rats are fearful of pups, a response that is apparently mediated by chemosensory information. Lesions of the amygdala or afferent sensory pathways from the vomeronasal organ to the amygdala disinhibit the expression of maternal behavior. Hormones or sensitization likely act to disinhibit the amygdala, thus permitting the occurrence of maternal behavior. Although correlations have been established, direct evidence of brain structural changes in human mothers remains unspecified ([Fleming & Gonzalez, 2009](#)).

Considered together, there are many examples of hormones influencing behavior and of behavior feeding back to influence hormone secretion. More and more examples of hormone–behavior interactions are discovered, including hormones in the mediation of food and fluid intake, social interactions, salt balance, learning and memory, stress coping, as well as psychopathology including depression, anxiety disorders, eating disorders, postpartum depression, and seasonal depression. Additional research should reveal how these hormone–behavior interactions are mediated.

OUTSIDE RESOURCES

Book: Adkins-Regan, E. (2005). *Hormones and animal social behavior*. Princeton, NJ: Princeton University Press.

Book: Beach, F. A. (1948). *Hormones and behavior*. New York: Paul Hoeber.

Book: Beach, F. A. (1975). *Behavioral endocrinology*:

An emerging discipline. American Scientist, 63: 178–187.

Book: Nelson, R. J. (2011). An introduction to behavioral endocrinology (4th ed.). Sunderland, MA: Sinauer Associates.

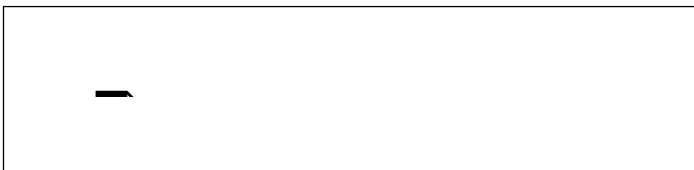
Book: Pfaff, D. W. (2009). Hormones, brain, and behavior (2nd ed.). New York: Academic Press.

Book: Pfaff, D. W., Phillips, I. M., & Rubin, R. T. (2005). Principles of hormone/behavior relations. New York: Academic Press.

Video: Endocrinology Video (Playlist) – This YouTube playlist contains many helpful videos on the biology of hormones, including reproduction and behavior. This would be a helpful resource for students struggling with hormone synthesis, reproduction, regulation of biological functions, and signaling pathways.

<https://www.youtube.com/playlist?list=PLqTetbgey0aemiTfD8QkMsSUq8hQzv-vA>

Video: Paul Zak: Trust, morality – and oxytocin- This Ted talk explores the roles of oxytocin in the body. Paul Zak discusses biological functions of oxytocin, like lactation, as well as potential behavioral functions, like empathy.





One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=232#oembed-1>

Video: The Teenage Brain Explained- This is a great video explaining the roles of hormones during puberty.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=232#oembed-2>

Web: Society for Behavioral Neuroendocrinology – This website contains resources on current news and research in the field of neuroendocrinology.

<http://sbn.org/home.aspx>

DISCUSSION QUESTIONS

1. What are some of the problems associated with attempting to determine causation in a hormone–behavior interaction? What are the best ways to address these problems?

2. Hormones cause changes in the rates of cellular processes or in cellular morphology. What are some ways that these hormonally induced cellular changes might theoretically produce profound changes in behavior?
3. List and describe some behavioral sex differences that you have noticed between boys and girls. What causes girls and boys to choose different toys? Do you think that the sex differences you have noted arise from biological causes or are learned? How would you go about establishing your opinions as fact?
4. Why is it inappropriate to refer to androgens as “male” hormones and estrogens as “female” hormones?
5. Imagine that you discovered that the brains of architects were different from those of non-architects—specifically, that the “drawstraighten nuclei” of the right temporal lobe were enlarged in architects as compared with non-architects. Would you argue that architects were destined to be architects because of their brain organization or that experience as an architect changed their brains? How would you resolve this issue?

VOCABULARY

5 α -reductase

An enzyme required to convert testosterone to 5 α -dihydrotestosterone.

Aggression

A form of social interaction that includes threat,

attack, and fighting.

Aromatase

An enzyme that converts androgens into estrogens.

Chromosomal sex

The sex of an individual as determined by the sex chromosomes (typically XX or XY) received at the time of fertilization.

Defeminization

The removal of the potential for female traits.

Demasculinization

The removal of the potential for male traits.

Dihydrotestosterone (DHT)

A primary androgen that is an androgenic steroid product of testosterone and binds strongly to androgen receptors.

Endocrine gland

A ductless gland from which hormones are released into the blood system in response to specific biological signals.

Estrogen

Any of the C18 class of steroid hormones, so named because of the estrus-generating properties in females. Biologically important estrogens include estradiol and estriol.

Feminization

The induction of female traits.

Gonadal sex

The sex of an individual as determined by the possession of either ovaries or testes. Females have ovaries, whereas males have testes.

Hormone

An organic chemical messenger released from endocrine cells that travels through the blood to interact with target cells at some distance to cause a biological response.

Masculinization

The induction of male traits.

Maternal behavior

Parental behavior performed by the mother or other female.

Neurotransmitter

A chemical messenger that travels between neurons to provide communication. Some neurotransmitters, such as norepinephrine, can leak into the blood system and act as hormones.

Oxytocin

A peptide hormone secreted by the pituitary gland to trigger lactation, as well as social bonding.

Parental behavior

Behaviors performed in relation to one's offspring that contributes directly to the survival of those offspring

Paternal behavior

Parental behavior performed by the father or other

male.

Progesterone

A primary progestin that is involved in pregnancy and mating behaviors.

Progestin

A class of C21 steroid hormones named for their progestational (pregnancy-supporting) effects. Progesterone is a common progestin.

Prohormone

A molecule that can act as a hormone itself or be converted into another hormone with different properties. For example, testosterone can serve as a hormone or as a prohormone for either dihydrotestosterone or estradiol.

Prolactin

A protein hormone that is highly conserved throughout the animal kingdom. It has many biological functions associated with reproduction and synergistic actions with steroid hormones.

Receptor

A chemical structure on the cell surface or inside of a cell that has an affinity for a specific chemical configuration of a hormone, neurotransmitter, or other compound.

Sex determination

The point at which an individual begins to develop as either a male or a female. In animals that have sex chromosomes, this occurs at fertilization. Females are XX and males are XY. All eggs bear X

chromosomes, whereas sperm can either bear X or Y chromosomes. Thus, it is the males that determine the sex of the offspring.

Sex differentiation

The process by which individuals develop the characteristics associated with being male or female. Differential exposure to gonadal steroids during early development causes sexual differentiation of several structures including the brain.

Target cell

A cell that has receptors for a specific chemical messenger (hormone or neurotransmitter).

Testosterone

The primary androgen secreted by the testes of most vertebrate animals, including men.

REFERENCES

- Alexander, G. M. & Hines, M. (2002). Sex differences in response to children's toys in nonhuman primates (*Cercopithecus aethiops sabaeus*). *Evolution and Human Behavior*, 23, 467–479.
- Berenbaum, S. A., Martin, C. L., Hanish, L. D., Briggs, P. T., & Fabes, R. A. (2008). Sex differences in children's play. In J. B. Becker, K. J. Berkley, N. Geary, E. Hampson, J. Herman, & E. Young (Eds.), *Sex differences in the brain: From genes to behavior*. New York: Oxford University Press.
- Dabbs, J. M. (2000). *Heroes, rogues, and lovers: Testosterone and behavior*. Columbus, OH: McGraw

Hill.

- Fleming, A. S., & Gonzalez, A. (2009). Neurobiology of human maternal care. In P. T. Ellison & P. B. Gray (Eds.), *Endocrinology of social relationships* (pp. 294–318). Cambridge, MA: Harvard University Press.
- Fleming, A. S., Steiner, M., & Corter, C. (1997). Cortisol, hedonics, and maternal responsiveness in human mothers. *Hormones and Behavior*, 32, 85–98.
- Goodson, J. L., Saldanha, C. J., Hahn, T. P., Soma, K. K. (2005). Recent advances in behavioral neuroendocrinology: Insights from studies on birds. *Hormones and Behavior*, 48, 461–73.
- Kidd, K. A., Blanchfield, P. J., Mills, K. H., Palace, V. P., Evans, R. E. Lazorchak, J. M. & Flick, R. (2007). Collapse of a fish population following exposure to a synthetic estrogen. *Proceedings of the National Academy of Sciences*, 104, 8897–8901.
- Nelson, R. J. (Ed.) (2006). *Biology of aggression*. New York: Oxford University Press.
- Nelson, R.J. (2011). *An introduction to behavioral endocrinology* (4th ed.). Sunderland, MA: Sinauer Associates.
- Phoenix, C. H., Goy, R. W., Gerall, A. A., & Young, W. C. (1959). Organizing action of prenatally administered testosterone propionate on the tissues mediating mating behavior in the female guinea pig. *Endocrinology*, 65: 369–382.

- van Anders, S., Hamilton, L., Schmidt, N., & Watson, N. (2007). Associations between testosterone secretion and sexual activity in women. *Hormones and Behavior*, 51, 477–82.

AUTHORS



Randy J. Nelson

Randy J. Nelson has two PhDs (Psychology and Endocrinology) and is Professor and Chair of Neuroscience and Professor of Psychology at The Ohio State University. He has been appointed as University Distinguished Professor for his work on hormones, biological rhythms, and behavior. Dr. Nelson has published nearly 400 papers and several books on these topics.

CREATIVE COMMONS LICENSE



HORMONES & BEHAVIOR by [Randy J. Nelson](#) is
licensed under a [Creative Commons Attribution-
NonCommercial-ShareAlike 4.0 International License](#).
Permissions beyond the scope of this license may be
available in our [Licensing Agreement](#).

CHAPTER 15.

BIOCHEMISTRY OF LOVE

SUE CARTER AND STEPHEN PORGES

Love is deeply biological. It pervades every aspect of our lives and has inspired countless works of art. Love also has a profound effect on our mental and physical state. A “broken heart” or a failed relationship can have disastrous effects; bereavement disrupts human physiology and may even precipitate death. Without loving relationships, humans fail to flourish, even if all of their other basic needs are met. As such, love is clearly not “just” an emotion; it is a biological process that is both dynamic and bidirectional in several dimensions. Social interactions between individuals, for example, trigger cognitive and physiological processes that influence emotional and mental states. In turn, these changes influence future social interactions. Similarly, the maintenance of loving relationships requires constant feedback through sensory and cognitive systems; the body seeks love and responds constantly to interactions with loved ones or to the absence of such interactions. The evolutionary principles and ancient hormonal and neural systems that support the beneficial and healing effects of loving relationships are described here.

LEARNING OBJECTIVES

- Understand the role of Oxytocin in social behaviors.
- Articulate the functional differences between Vasopressin and Oxytocin.
- List sex differences in reaction to stress.

INTRODUCTION

Although evidence exists for the healing power of love, only recently has science turned its attention to providing a physiological explanation for love. The study of love in this context offers insight into many important topics, including the biological basis of interpersonal relationships and why and how disruptions in social bonds have such pervasive consequences for behavior and physiology. Some of the answers will be found in our growing knowledge of the neurobiological and endocrinological mechanisms of social behavior and interpersonal engagement.

THE EVOLUTION OF SOCIAL BEHAVIOR

Nothing in biology makes sense except in the light of evolution. Theodosius Dobzhansky's famous dictum also holds true for explaining the evolution of love. Life on earth is fundamentally social: The ability to dynamically interact with other living organisms to support mutual homeostasis, growth, and reproduction evolved very early. Social interactions are present in primitive invertebrates and even among prokaryotes: Bacteria recognize and approach members of their own species. Bacteria also reproduce more successfully in the presence

of their own kind and are able to form communities with physical and chemical characteristics that go far beyond the capabilities of the individual cell ([Ingham & Ben-Jacob, 2008](#)).

As another example, various insect species have evolved particularly complex social systems, known as eusociality. Characterized by a division of labor, eusociality appears to have evolved independently at least 11 times in insects. Research on honeybees indicates that a complex set of genes and their interactions regulate eusociality, and that these resulted from an “accelerated form of evolution” ([Woodard et](#)



Life on earth is essentially social and many species besides humans engage in complex social interactions and have developed complex social systems that shape their behavior. [Image: www.metaphoricalplatypus.com, <https://goo.gl/9nODli>, CC BY 2.0, <https://goo.gl/v4YOZv>]

[al., 2011](#)). In other words, molecular mechanisms favoring high levels of sociality seem to be on an evolutionary fast track.

The evolutionary pathways that led from reptiles to mammals allowed the emergence of the unique anatomical systems and biochemical mechanisms that enable social engagement and selectively reciprocal sociality. Reptiles show minimal parental investment in offspring and form nonselective relationships between individuals. Pet owners may become emotionally attached to their turtle or snake, but this relationship is

not reciprocal. In contrast, most mammals show intense parental investment in offspring and form lasting bonds with their children. Many mammalian species—including humans, wolves, and prairie voles—also develop long-lasting, reciprocal, and selective relationships between adults, with several features of what humans experience as “love.” In turn, these reciprocal interactions trigger dynamic feedback mechanisms that foster growth and health.

WHAT IS LOVE? AN EVOLUTIONARY AND PHYSIOLOGICAL PERSPECTIVE

Human love is more complex than simple feedback mechanisms. Love may create its own reality. The biology of love originates in the primitive parts of the brain—the emotional core of the human nervous system—which evolved long before the cerebral cortex. The brain “in love” is flooded with vague sensations, often transmitted by the **vagus nerve**, and creating much of what we experience as emotion. The modern cortex struggles to interpret love’s primal messages, and weaves a narrative around incoming visceral experiences, potentially reacting to that narrative rather than to reality. It also is helpful to realize that mammalian social behavior is supported by biological components that were repurposed or co-opted over the course of mammalian evolution, eventually permitting lasting relationships between adults.

IS THERE A HORMONE OF LOVE AND OTHER RELATIONSHIPS?

One element that repeatedly appears in the biochemistry

of love is the neuropeptide **oxytocin**. In large mammals, oxytocin adopts a central role in reproduction by helping to expel the big-brained baby from the uterus, ejecting milk and sealing a selective and lasting bond between mother and offspring ([Keverne, 2006](#)). Mammalian offspring crucially depend on their mother's milk for some time after birth. Human mothers also form a strong and lasting bond with their newborns immediately after birth, in a time period that is essential for the nourishment and survival of the baby. However, women who give birth by cesarean section without going through labor, or who opt not to breastfeed, are still able to form a strong emotional bond with their children. Furthermore, fathers, grandparents, and adoptive parents also form lifelong attachments to children. Preliminary evidence suggests that the simple presence of an infant can release oxytocin in adults as well ([Feldman, 2012](#); [Kenkel et al., 2012](#)). The baby virtually forces us to love it.

The case for a major role for oxytocin in love is strong, but until recently was based largely on extrapolation from research on parental behavior ([Feldman, 2012](#)) or social behaviors in animals ([Carter, 1998](#); [Kenkel et al., 2012](#)).

However, recent human experiments have shown that intranasal delivery of oxytocin can facilitate social behaviors, including eye contact and social



How would it feel to learn that love is nothing more than biological processes in the brain? Would it make a difference to you? Or is the fact you feel love all that matters? [Image: CCO Public Domain, <https://goo.gl/m25gce>]

cognition ([Meyer-Lindenberg, Domes, Kirsch, & Heinrichs, 2011](#))—behaviors that are at the heart of love.

Of course, oxytocin is not the molecular equivalent of love. Rather, it is just one important component of a complex neurochemical system that allows the body to adapt to highly emotional situations. The systems necessary for reciprocal social interactions involve extensive neural networks through the brain and autonomic nervous system that are dynamic and constantly changing across the life span of an individual. We also now know that the properties of oxytocin are not predetermined or fixed. Oxytocin's cellular receptors are regulated by other hormones and **epigenetic** factors. These receptors change and adapt based on life experiences. Both oxytocin and the experience of love can change over time. In spite of limitations, new knowledge of the properties of oxytocin has proven useful in explaining several enigmatic features of love.

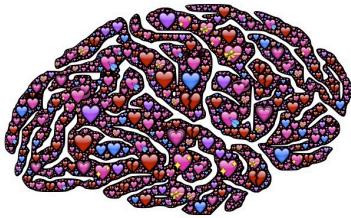
STRESS AND LOVE

Emotional bonds can form during periods of extreme duress, especially when the survival of one individual depends on the presence and support of another. There also is evidence that oxytocin is released in response to acutely stressful experiences, perhaps serving as hormonal “insurance” against overwhelming stress. Oxytocin may help to ensure that parents and others will engage with and care for infants; develop stable, loving relationships; and seek out and receive support from others in times of need.

ANIMAL MODELS AND THE BIOLOGY OF SOCIAL BONDS

To dissect the anatomy and chemistry of love, scientists needed a biological equivalent of the Rosetta Stone. Just as the actual stone helped linguists decipher an archaic language by comparison to a known one, animal models are helping biologists draw parallels between ancient physiology and contemporary behaviors. Studies of socially monogamous mammals that form long-lasting social bonds, such as prairie voles, have been especially helpful to an understanding the biology of human social behavior.

THERE IS MORE TO LOVE THAN OXYTOCIN



Just as complex as our subjective experience of love can be, so, too, is the complexity of brain processes involved with it. [Image: CCO Public Domain, <https://goo.gl/m25gce>]

Research in prairie voles showed that, as in humans, oxytocin plays a major role in social interactions and parental behavior

([Carter, 1998](#); [Carter, Boone, Pournajafi-Nazarloo, & Bales, 2009](#); [Kenkel et al., 2012](#)). Of course, oxytocin

does not act alone. Its release and actions depend on many other

neurochemicals, including endogenous opioids and dopamine ([Aragona & Wang, 2009](#)). Particularly important to social bonding are the interactions of oxytocin with a related neuropeptide known as **vasopressin**. The systems regulated by oxytocin and

vasopressin are sometimes redundant. Both peptides are implicated in behaviors that require social engagement by either males or females, such as huddling over an infant ([Kenkel et al., 2012](#)). For example, it was necessary in voles to block both oxytocin and vasopressin receptors to induce a significant reduction in social engagement, either among adults or between adults and infants. Blocking only one of these two receptors did not eliminate social approach or contact. However, antagonists for either the oxytocin or vasopressin receptor inhibited the selective sociality, which is essential for the expression of a social bond ([Bales, Kim, Lewis-Reese, & Carter, 2004](#); [Cho, DeVries, Williams, & Carter, 1999](#)). If we accept selective social bonds, parenting, and mate protection as proxies for love in humans, research in animals supports the hypothesis that oxytocin and vasopressin interact to allow the dynamic behavioral states and behaviors necessary for love.

Oxytocin and vasopressin have shared functions, but they are not identical in their actions. The specific behavioral roles of oxytocin and vasopressin are especially difficult to untangle because they are components of an integrated neural network with many points of intersection. Moreover, the genes that regulate the production of oxytocin and vasopressin are located on the same chromosome, possibly allowing coordinated synthesis or release of these peptides. Both peptides can bind to and have antagonist or agonist effects on each other's receptors. Furthermore, the pathways necessary for reciprocal social behavior are constantly adapting: These peptides and the systems that they regulate are always in flux. In spite of these difficulties, some of the

different functions of oxytocin and vasopressin have been identified.

FUNCTIONAL DIFFERENCES BETWEEN VASOPRESSIN AND OXYTOCIN

Vasopressin is associated with physical and emotional mobilization, and can help support vigilance and behaviors needed for guarding a partner or territory ([Carter, 1998](#)), as well as other forms of adaptive self-defense ([Ferris, 2008](#)). Vasopressin also may protect against physiologically “shutting down” in the face of danger. In many mammalian species, mothers exhibit agonistic behaviors in defense of their young, possibly through the interactive actions of vasopressin and oxytocin ([Bosch & Neumann, 2012](#)). Prior to mating, prairie voles are generally social, even toward strangers. However, within a day or so of mating, they begin to show high levels of aggression toward intruders ([Carter, DeVries, & Getz, 1995](#)), possibly serving to protect or guard a mate, family, or territory. This mating-induced aggression is especially obvious in males.

Oxytocin, in contrast, is associated with immobility without fear. This includes relaxed physiological states and postures that permit birth, lactation, and consensual sexual behavior. Although not essential for parenting, the increase of oxytocin associated with birth and lactation may make it easier for a woman to be less anxious around her newborn and to experience and express loving feelings for her child ([Carter & Altemus, 1997](#)). In highly social species such as prairie voles ([Kenkel et al., 2013](#)), and presumably in humans, the intricate molecular

dances of oxytocin and vasopressin fine-tune the coexistence of caretaking and protective aggression.

FATHERHOOD ALSO HAS A BIOLOGICAL BASIS

The biology of fatherhood is less well-studied than motherhood is. However, male care of offspring also appears to rely on both oxytocin and vasopressin ([Kenkel et al., 2012](#)), probably acting in part through effects on the autonomic nervous system ([Kenkel et al., 2013](#)). Even sexually naïve male prairie voles show spontaneous parental behavior in the presence of an infant ([Carter et al., 1995](#)). However, the stimuli from infants or the nature of the social interactions that release oxytocin and vasopressin may differ between the sexes ([Feldman, 2012](#)).

AT THE HEART OF THE BENEFITS OF LOVE IS A SENSE OF SAFETY

Parental care and support in a safe environment are particularly important for mental health in social mammals, including humans and prairie voles. Studies of rodents and of lactating women suggest that oxytocin has the important capacity to modulate the behavioral and autonomic distress that typically follows separation from a mother, child, or partner, reducing defensive behaviors and thereby supporting growth and health (Carter, 1998).



Love is a universal feeling of safety and connectedness with another. Research has shown that at the end of people's lives, their primary regret is not spending more time with the people they love. [Image: CCO Public Domain, <https://goo.gl/m25gce>]

THE ABSENCE OF LOVE IN EARLY LIFE CAN BE DETRIMENTAL TO MENTAL AND PHYSICAL HEALTH

During early life in particular, trauma or neglect may produce behaviors and emotional states in humans that are socially pathological. Because the processes involved in creating social behaviors and social emotions are delicately balanced, these may be triggered in inappropriate contexts, leading to aggression toward friends or family. Alternatively, bonds may be formed with prospective partners who fail to provide social support or protection.

SEX DIFFERENCES EXIST IN THE CONSEQUENCES OF EARLY LIFE EXPERIENCES

Males seem to be especially vulnerable to the negative effects of early experiences, possibly helping to explain the increased sensitivity of males to various developmental disorders. The implications of sex differences in the nervous system and in the response to stressful experiences for social behavior are only slowly becoming apparent ([Carter et al., 2009](#)). Both males and females produce vasopressin and oxytocin and are capable of responding to both hormones. However, in brain regions that are involved in defensive aggression, such as the extended amygdala and lateral septum, the production of vasopressin is androgen-dependent. Thus, in the face of a threat, males may be experiencing higher central levels of vasopressin.

Oxytocin and vasopressin pathways, including the peptides and their receptors, are regulated by coordinated genetic, hormonal, and epigenetic factors that influence the adaptive and behavioral functions of these peptides across the animal's life span. As a result, the endocrine and behavioral consequences of a stress or challenge may be different for males and females ([DeVries, DeVries, Taymans, & Carter, 1996](#)). For example, when unpaired prairie voles were exposed to an intense but brief stressor, such as a few minutes of swimming, or injection of the adrenal hormone corticosterone, the males (but not females) quickly formed new pair bonds. These and other experiments suggest that males and females have different coping strategies, and possibly may experience both stressful

experiences, and even love, in ways that are gender-specific.

In the context of nature and evolution, sex differences in the nervous system are important. However, sex differences in brain and behavior also may help to explain gender differences in the vulnerability to mental and physical disorders ([Taylor, et al., 2000](#)). Better understanding these differences will provide clues to the physiology of human mental health in both sexes.

LOVING RELATIONSHIPS IN EARLY LIFE CAN HAVE EPIGENETIC CONSEQUENCES

Love is “epigenetic.” That is, positive experiences in early life can act upon and alter the expression of specific genes. These changes in gene expression may have behavioral consequences through simple biochemical changes, such as adding a methyl group to a particular site within the genome ([Zhang & Meaney, 2010](#)). It is possible that these changes in the genome may even be passed to the next generation.



Although we are all born with a finite set of genes, experiences in childhood will cause some genes to express themselves (e.g., encourage certain personality traits), while other genes will remain dormant. [Image: CCO Public Domain, <https://goo.gl/m25gce>]

Social behaviors, emotional attachment to others, and long-lasting reciprocal relationships also are both plastic and adaptive, and so is the biology upon which they are based. For example, infants of traumatized or highly stressed parents might be chronically exposed to vasopressin, either through their own increased production of the peptide, or through higher levels of

vasopressin in maternal milk. Such increased exposure could sensitize the infant to defensive behaviors or create a lifelong tendency to overreact to threat. Based on research in rats, it seems that in response to adverse early experiences of chronic isolation, the genes for vasopressin receptors can become upregulated ([Zhang et al., 2012](#)), leading to an increased sensitivity to acute stressors or anxiety that may persist throughout life.

Epigenetic programming triggered by early life experiences is adaptive in allowing neuroendocrine systems to project and plan for future behavioral demands. But epigenetic changes that are long-lasting also can create atypical social or emotional behaviors ([Zhang & Meaney, 2010](#)) that may be especially likely to surface in later life, and in the face of social or emotional challenges.

Exposure to exogenous hormones in early life also may

be epigenetic. For example, prairie voles treated postnatally with vasopressin (especially males) were later more aggressive, whereas those exposed to a vasopressin antagonist showed less aggression in adulthood. Conversely, in voles the exposure of infants to slightly increased levels of oxytocin during development increased the tendency to show a pair bond. However, these studies also showed that a single exposure to a higher level of oxytocin in early life could disrupt the later capacity to pair bond ([Carter et al., 2009](#)).

There is little doubt that either early social experiences or the effects of developmental exposure to these neuropeptides holds the potential to have long-lasting effects on behavior. Both parental care and exposure to oxytocin in early life can permanently modify hormonal systems, altering the capacity to form relationships and influence the expression of love across the life span. Our preliminary findings in voles further suggest that early life experiences affect the methylation of the oxytocin receptor gene and its expression ([Connelly, Kenkel, Erickson, & Carter, 2011](#)). Thus, we can plausibly argue that love is epigenetic.

THE ABSENCE OF SOCIAL BEHAVIOR OR ISOLATION ALSO HAS CONSEQUENCES FOR THE OXYTOCIN SYSTEM

Given the power of positive social experiences, it is not surprising that a lack of social relationships also may lead to alterations in behavior as well as changes in oxytocin and vasopressin pathways. We have found that social isolation reduced the expression of the gene for the oxytocin receptor, and at the same time increased the

expression of genes for the vasopressin peptide. In female prairie voles, isolation also was accompanied by an increase in blood levels of oxytocin, possibly as a coping mechanism. However, over time, isolated prairie voles of both sexes showed increases in measures of depression, anxiety, and physiological arousal, and these changes were observed even when endogenous oxytocin was elevated. Thus, even the hormonal insurance provided by endogenous oxytocin in face of the chronic stress of isolation was not sufficient to dampen the consequences of living alone. Predictably, when isolated voles were given additional exogenous oxytocin, this treatment did restore many of these functions to normal ([Grippe, Trahanas, Zimmerman, Porges, & Carter, 2009](#)).

In modern societies, humans can survive, at least after childhood, with little or no human contact. Communication technology, social media, electronic parenting, and many other recent technological advances may reduce social behaviors, placing both children and adults at risk for social isolation and disorders of the autonomic nervous system, including deficits in their capacity for social engagement and love ([Porges, 2011](#)).

Social engagement actually helps us to cope with stress. The same hormones and areas of the brain that increase the capacity of the body to survive stress also enable us to better adapt to an ever-changing social and physical environment. Individuals with strong emotional support and relationships are more resilient in the face of stressors than those who feel isolated or lonely. Lesions in various bodily tissues, including the brain, heal more quickly in animals that are living socially versus in isolation ([Karelina & DeVries, 2011](#)). The protective effects of positive sociality seem to rely on the same

cocktail of hormones that carries a biological message of “love” throughout the body.

CAN LOVE—OR PERHAPS OXYTOCIN—BE A MEDICINE?

Although research has only begun to examine the physiological effects of these peptides beyond social behavior, there is a wealth of new evidence showing that oxytocin can influence physiological responses to stress and injury. As only one example, the molecules associated with love have restorative properties, including the ability to literally heal a “broken heart.” Oxytocin receptors are expressed in the heart, and precursors for oxytocin appear to be critical for the development of the fetal heart ([Danalache, Gutkowska, Slusarz, Berezowska, & Jankowski, 2010](#)). Oxytocin exerts protective and restorative effects in part through its capacity to convert undifferentiated stem cells into cardiomyocytes. Oxytocin can facilitate adult neurogenesis and tissue repair, especially after a stressful experience. We now know that oxytocin has direct anti-inflammatory and antioxidant properties in *in vitro* models of atherosclerosis ([Szeto et al., 2008](#)). The heart seems to rely on oxytocin as part of a normal process of protection and self-healing.

Thus, oxytocin exposure early in life not only regulates our ability to love and form social bonds, it also affects our health and well-being. Oxytocin modulates the hypothalamic–pituitary adrenal (HPA) axis, especially in response to disruptions in homeostasis (Carter, 1998), and coordinates



Researchers are interested in the medical/therapeutic potential of oxytocin. [Image: CCO Public Domain, <https://goo.gl/m25gce>]

demands on the immune system and energy balance. Long-term, secure relationships provide emotional support and down-regulate reactivity of the HPA axis, whereas intense stressors, including birth, trigger activation of the HPA axis and sympathetic nervous system. The ability of oxytocin to regulate these systems probably explains the exceptional capacity of most women to cope with the challenges of childbirth and childrearing.

Dozens of ongoing clinical trials are currently attempting to examine the therapeutic potential of oxytocin in disorders ranging from autism to heart disease. Of course, as in hormonal studies in voles, the effects are likely to depend on the history of the individual and the context, and to be dose-dependent. As this research is emerging, a variety of individual differences and apparent discrepancies in the effects of exogenous oxytocin are being reported. Most of these studies do not include any information on the endogenous hormones, or on the oxytocin or vasopressin

receptors, which are likely to affect the outcome of such treatments.

CONCLUSION

Research in this field is new and there is much left to understand. However, it is already clear that both love and oxytocin are powerful. Of course, with power comes responsibility. Although research into mechanisms through which love—or hormones such as oxytocin—may protect us against stress and disease is in its infancy, this knowledge will ultimately increase our understanding of the way that our emotions impact upon health and disease. The same molecules that allow us to give and receive love also link our need for others with health and well-being.

ACKNOWLEDGMENTS

C. Sue Carter and Stephen W. Porges are both Professors of Psychiatry at the University of North Carolina, Chapel Hill, and also are Research Professors of Psychology at Northeastern University, Boston.

Discussions of “love and forgiveness” with members of the Fetzer Institute’s Advisory Committee on Natural Sciences led to this essay and are gratefully acknowledged here. We are especially appreciative of thoughtful editorial input from Dr. James Harris. Studies from the authors’ laboratories were sponsored by the National Institutes of Health. We also express our gratitude for this support and to our colleagues, whose input and hard work informed the ideas expressed in this article. A version of this paper was previously published in *EMBO Reports* in the series on “Sex and Society”; this paper is

reproduced with the permission of the publishers of that journal.

OUTSIDE RESOURCES

Book: C. S. Carter, L. Ahnert et al. (Eds.), (2006). *Attachment and bonding: A new synthesis*. Cambridge, MA: MIT Press.

Book: Porges, S.W. (2011). *The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication and self-regulation*. New York, NY: Norton.

Web: Database of publicly and privately supported clinical studies of human participants conducted around the world.

<http://www.clinicaltrials.gov>

Web: PubMed comprises over 22 million citations for biomedical literature from MEDLINE, life science journals, and online books. PubMed citations and abstracts include the fields of biomedicine and health, covering portions of the life sciences, behavioral sciences, chemical sciences, and bioengineering. PubMed also provides access to additional relevant web sites and links to the other NCBI molecular biology resources.

<http://www.ncbi.nlm.nih.gov/pubmed>

Web: Website of author Stephen Porges

<http://www.stephenporges.com/>

DISCUSSION QUESTIONS

1. If love is so important in human behavior, why is it so hard to describe and understand?
2. Discuss the role of evolution in understanding what humans call “love” or other forms of prosociality.
3. What are the common biological and neuroendocrine elements that appear in maternal love and adult-adult relationships?
4. Oxytocin and vasopressin are biochemically similar. What are some of the differences between the actions of oxytocin and vasopressin?
5. How may the properties of oxytocin and vasopressin help us understand the biological bases of love?
6. What are common features of the biochemistry of “love” and “safety,” and why are these important to human health?

VOCABULARY

Epigenetics

Heritable changes in gene activity that are not caused by changes in the DNA sequence.

<http://en.wikipedia.org/wiki/Epigenetics>

Oxytocin

A nine amino acid mammalian neuropeptide.

Oxytocin is synthesized primarily in the brain, but also in other tissues such as uterus, heart and thymus, with local effects. Oxytocin is best known

as a hormone of female reproduction due to its capacity to cause uterine contractions and eject milk. Oxytocin has effects on brain tissue, but also acts throughout the body in some cases as an antioxidant or anti-inflammatory.

Vagus nerve

The 10th cranial nerve. The mammalian vagus has an older unmyelinated branch which originates in the dorsal motor complex and a more recently evolved, myelinated branch, with origins in the ventral vagal complex including the nucleus ambiguus. The vagus is the primary source of autonomic-parasympathetic regulation for various internal organs, including the heart, lungs and other parts of the viscera. The vagus nerve is primarily sensory (afferent), transmitting abundant visceral input to the central nervous system.

Vasopressin

A nine amino acid mammalian neuropeptide. Vasopressin is synthesized primarily in the brain, but also may be made in other tissues. Vasopressin is best known for its effects on the cardiovascular system (increasing blood pressure) and also the kidneys (causing water retention). Vasopressin has effects on brain tissue, but also acts throughout the body.

REFERENCES

- Aragona, B. J., & Wang, Z. (2009). Dopamine regulation of social choice in a monogamous rodent species. *Frontiers in Behavioral Neuroscience*,

3, 15.

- Bales, K. L., Kim, A. J., Lewis-Reese, A. D., & Carter, C. S. (2004). Both oxytocin and vasopressin may influence alloparental care in male prairie voles. *Hormones and Behavior*, 44, 454–361.
- Bosch, O. J., & Neumann, I. D. (2012). Both oxytocin and vasopressin are mediators of maternal care and aggression in rodents: from central release to sites of action. *Hormones and Behavior*, 61, 293–303.
- Carter, C. S. (1998). Neuroendocrine perspectives on social attachment and love. *Psychoneuroendocrinology*, 23, 779–818.
- Carter, C. S., & Altemus, M. (1997). Integrative functions of lactational hormones in social behavior and stress management. *Annals of the New York Academy of Sciences, Integrative Neurobiology of Affiliation* 807, 164–174.
- Carter, C. S., Boone, E. M., Pournajafi-Nazarloo, H., & Bales, K. L. (2009). The consequences of early experiences and exposure to oxytocin and vasopressin are sexually-dimorphic. *Developmental Neuroscience*, 31, 332–341.
- Carter, C. S., DeVries, A. C., & Getz, L. L. (1995). Physiological substrates of mammalian monogamy: The prairie vole model. *Neuroscience and Biobehavioral Reviews*, 19, 303–314.
- Cho, M. M., DeVries, A. C., Williams, J. R., Carter, C. S. (1999). The effects of oxytocin and

vasopressin on partner preferences in male and female prairie voles (*Microtus ochrogaster*). *Behavioral Neuroscience*, 113, 1071–1080.

- Connelly, J., Kenkel, W., Erickson, E., & Carter, C. S. (2011). Are birth and oxytocin epigenetic events. *Society for Neuroscience Abstracts*, 388.10.
- Danalache, B. A., Gutkowska, J., Slusarz, M. J., Berezowska, I., & Jankowski, M. (2010). Oxytocin-Gly-Lys-Arg: A novel cardiomyogenic peptide. *PloS One*, 5(10), e13643.
- DeVries, A. C., DeVries, M. B., Taymans, S. E., & Carter, C. S. (1996). Stress has sexually dimorphic effects on pair bonding in prairie voles. *Proceedings of the National Academy of Sciences, USA*, 93, 11980–11984.
- Feldman, R. (2012). Oxytocin and social affiliation in humans. *Hormones and Behavior*, 61, 380–391.
- Ferris, C. F. (2008). Functional magnetic resonance imaging and the neurobiology of vasopressin and oxytocin. *Progress in Brain Research*, 170, 305–320.
- Grippo, A. J., Trahanas, D. M., Zimmerman, R. R., Porges, S. W., & Carter, C. S. (2009). Oxytocin protects against negative behavioral and autonomic consequences of long-term social isolation. *Psychoneuroendocrinology*, 34, 1542–1553.
- Ingham, C. J., & Ben-Jacob, E. (2008). Swarming and complex pattern formation in *Paenibacillus vortex* studied by imaging and tracking cells. *BMC*

Microbiol, 8, 36.

- Karelina, K., & DeVries, A. C. (2011). Modeling social influences on human health. *Psychosomatic Medicine*, 73, 67–74.
- Kenkel, W.M., Paredes, J., Lewis, G.F., Yee, J.R., Pournajafi-Nazarloo, H., Grippo, A.J., Porges, S.W., & Carter, C.S. (2013). Autonomic substrates of the response to pups in male prairie voles. *PlosOne* Aug 5;8(8):e69965. doi: 10.1371/journal.pone.0069965.
- Kenkel, W.M., Paredes, J., Yee, J. R., Pournajafi-Nazarloo, H., Bales, K. L., & Carter, C. S. (2012). Exposure to an infant releases oxytocin and facilitates pair-bonding in male prairie voles. *Journal of Neuroendocrinology*, 24, 874–886.
- Keverne, E. B. (2006). Neurobiological and molecular approaches to attachment and bonding. In C. S. Carter, L. Ahnert et al. (Eds.), *Attachment and bonding: A new synthesis*. Cambridge, MA: MIT Press. Pp. 101-117.
- Meyer-Lindenberg, A., Domes, G., Kirsch, P., & Heinrichs, M. (2011). Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nature: Reviews in Neuroscience* 12, 524–538.
- Porges, S. W. (2011). *The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication and self-regulation*. New York, NY: Norton.
- Szeto, A., Nation, D. A., Mendez, A. J.,

- Dominguez-Bendala, J., Brooks, L. G., Schneiderman, N., & McCabe, P. M. (2008). Oxytocin attenuates NADPH-dependent superoxide activity and IL-6 secretion in macrophages and vascular cells. *American Journal of Physiology, Endocrinology and Metabolism* 295, E1495–1501.
- Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychol Rev*, 107, 411–429.
 - Woodard, S. H., Fischman, B. J., Venkat, A., Hudson, M. E., Varala, K., Cameron S.A., . . . Robinson, G. E. (2011). Genes involved in convergent evolution of eusociality in bees. *Proc Natl Acad Sci USA*, 108, 7472–7477.
 - Zhang, L., Hernandez, V. S., Liu, B., Medina, M. P., Nava-Kopp, A. T., Irles, C., & Morales, M. (2012). Hypothalamic vasopressin system regulation by maternal separation: Its impact on anxiety in rats. *Neuroscience*, 215, 135–148.
 - Zhang, T. Y., & Meaney, M. J. (2010). Epigenetics and the environmental regulation of the genome and its function. *Annual Review of Psychology*, 61, 439–466.

AUTHORS



Sue Carter

Sue Carter currently holds Research Professorships in Psychiatry at the University of North Carolina and in Psychology at Northeastern University, Boston. Dr. Carter pioneered the study of oxytocin and vasopressin in positive social behaviors, and developed the social-monogamous prairie vole as a model for the health benefits of social support.



Stephen Porges

Stephen Porges, Professor of Psychiatry at the University of North Carolina and Professor Emeritus at the University of Illinois at Chicago, discovered the role of vagal regulation of the heart in behavioral neuroscience and is the originator of the Polyvagal Theory, which provides a neurophysiological basis for mind-body sciences.

CREATIVE COMMONS LICENSE



BIOCHEMISTRY OF LOVE by [Sue Carter and Stephen Porges](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

PART X.

**UNIT 10: SOCIAL
NEUROSCIENCE**

CHAPTER 16.

SOCIAL NEUROSCIENCE

TIFFANY A. ITO AND JENNIFER T. KUBOTA

This module provides an overview of the new field of social neuroscience, which combines the use of neuroscience methods and theories to understand how other people influence our thoughts, feelings, and behavior. The module reviews research measuring neural and hormonal responses to understand how we make judgments about other people and react to stress. Through these examples, it illustrates how social neuroscience addresses three different questions: (1) how our understanding of social behavior can be expanded when we consider neural and physiological responses, (2) what the actual biological systems are that implement social behavior (e.g., what specific brain areas are associated with specific social tasks), and (3) how biological systems are impacted by social processes.

LEARNING OBJECTIVES

- Define social neuroscience and describe its three major goals.
- Describe how measures of brain activity such as

EEG and fMRI are used to make inferences about social processes.

- Discuss how social categorization occurs.
- Describe how simulation may be used to make inferences about others.
- Discuss the ways in which other people can cause stress and also protect us against stress.

Psychology has a long tradition of using our brains and body to better understand how we think and act. For example, in 1939 Heinrich Kluver and Paul Bucy removed (i.e. lesioned) the temporal lobes in some rhesus monkeys and observed the effect on behavior. Included in these **lesions** was a subcortical area of the brain called the **amygdala**. After surgery, the monkeys experienced profound behavioral changes, including loss of fear. These results provided initial evidence that the amygdala plays a role in emotional responses, a finding that has since been confirmed by subsequent studies ([Phelps & LeDoux, 2005](#); [Whalen & Phelps, 2009](#)).

WHAT IS SOCIAL NEUROSCIENCE?

Social neuroscience similarly uses the brain and body to understand how we think and act, with a focus on how we think about and act toward other people. More specifically, we can think of social neuroscience as an interdisciplinary field that uses a range of neuroscience measures to understand how other people influence our thoughts, feelings, and behavior. As such, social neuroscience studies the same topics as social psychology, but does so from a multilevel perspective that includes the study of the brain and body. Figure 1 shows the scope

of social neuroscience with respect to the older fields of social psychology and neuroscience. Although the field is relatively new – the term first appeared in 1992 ([Cacioppo & Berntson, 1992](#)) – it has grown rapidly, thanks to technological advances making measures of the brain and body cheaper and more powerful than ever before, and to the recognition that neural and physiological information are critical to understanding how we interact with other people.

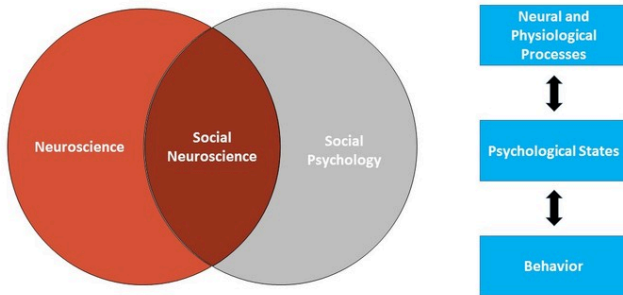


Figure 1. Social neuroscience is the intersection of social psychology and neuroscience. Under this multilevel approach, neural/physiological processes and behavior are two things we can measure or observe. Psychological states cannot be directly observed, but understanding them is the goal. Social neuroscientists use the observable neural/physiological processes and behavioral responses to make inferences about unobservable psychological states. The bidirectional arrows show that all levels of analysis are assumed to influence each other (e.g., psychological states can influence neural responses, and neural responses can influence psychological states).

Social neuroscience can be thought of as both a methodological approach (using measures of the brain and body to study social processes) and a theoretical orientation (seeing the benefits of integrating

neuroscience into the study of social psychology). The overall approach in social neuroscience is to understand the psychological processes that underlie our social behavior. Because those psychological processes are intrapsychic phenomena that cannot be directly observed, social neuroscientists rely on a combination of measureable or observable neural and physiological responses as well as actual overt behavior to make inferences about psychological states (see Figure 1). Using this approach, social neuroscientists have been able to pursue three different types of questions: (1) What more can we learn about social behavior when we consider neural and physiological responses? (2) What are the actual biological systems that implement social behavior (e.g., what specific brain areas are associated with specific social tasks)? and (3) How are biological systems impacted by social processes?

In this module, we review three research questions that have been addressed with social neuroscience that illustrate the different goals of the field. These examples also expose you to some of the frequently used measures.

HOW AUTOMATICALLY DO WE JUDGE OTHER PEOPLE?

Social categorization is the act of mentally classifying someone as belonging in a group. Why do we do this? It is an effective mental shortcut. Rather than effortfully thinking about every detail of every person we encounter, social categorization allows us to rely on information we already know about the person's group. For example, by classifying your restaurant server as a man, you can quickly activate all the information you have stored about

men and use it to guide your behavior. But this shortcut comes with potentially high costs. The stored group beliefs might not be very accurate, and even when they do accurately describe some group members, they are unlikely to be true for every member you encounter. In addition, many beliefs we associate with groups – called **stereotypes** – are negative. This means that relying on social categorization can often lead people to make negative assumptions about others.

The potential costs of social categorization make it important to understand how social categorization occurs. Is it rare or does it occur often? Is it something we can easily stop, or is it hard to override? One difficulty answering these questions is that people are not always consciously aware of what they are doing. In this case, we might not always realize when we are categorizing someone. Another concern is that even when people are aware of their behavior, they can be reluctant to accurately report it to an experimenter. In the case of social categorization, subjects might worry they will look bad if they accurately report classifying someone into a group associated with negative stereotypes. For instance, many racial groups are associated with some negative stereotypes, and subjects may worry that admitting to classifying someone into one of those groups means they believe and use those negative stereotypes.



Figure 2. This man is wearing an elastic electrode cap into which individual electrodes (inside the white circles) are sewn into standardized locations.

*[Image: Hans, CCO Public Domain,
<https://goo.gl/m25gce>]*

Social neuroscience has been useful for studying how social categorization occurs without having to rely on self-report measures, instead measuring brain activity differences that occur when people encounter members of different social groups. Much of this work has been recorded

using the **electroencephalogram**, or **EEG**. EEG is a measure of electrical activity generated by the brain's neurons. Comparing this electrical activity at a given point in time against what a person is thinking and doing at that same time allows us to make inferences about brain activity associated with specific psychological states. One particularly nice feature of EEG is that it provides very precise timing information about when brain activity occurs. EEG is measured non-invasively with small electrodes that rest on the surface of the scalp. This is often done with a stretchy elastic cap, like the one shown in Figure 2, into which the small electrodes are sewn. Researchers simply pull the cap onto the subject's head to get the electrodes into place; wearing it is similar to wearing a swim cap. The subject can then be asked to think about different topics or engage in different tasks as brain activity is measured.

To study social categorization, subjects have been shown pictures of people who belong to different social groups. Brain activity recorded from many individual trials (e.g., looking at lots of different Black individuals) is then averaged together to get an overall idea of how the brain responds when viewing individuals who belong to a particular social group. These studies suggest that social categorization is an **automatic process** – something that happens with little conscious awareness or control – especially for dimensions like gender, race, and age ([Ito & Urland, 2003](#); [Mouchetant-Rostaing & Giard, 2003](#)). The studies specifically show that brain activity differs when subjects view members of different social groups (e.g., men versus women, Black people versus White people), suggesting that the group differences are being encoded and processed by the perceiver. One interesting finding is that these brain changes occur both when subjects are purposely asked to categorize the people into social groups (e.g., to judge whether the person is Black or White), and also when they are asked to do something that draws attention away from group classifications (e.g., making a personality judgment about the person) ([Ito & Urland, 2005](#)). This tells us that we do not have to intend to make group classifications in order for them to happen. It is also very interesting to consider how quickly the changes in brain responses occur. Brain activity is altered by viewing members of different groups within 200 milliseconds of seeing a person's face. That is just two-tenths of a second. Such a fast response lends further support to the idea that social categorization occurs automatically and may not depend on conscious intention.

Overall, this research suggests that we engage in social

categorization very frequently. In fact, it appears to happen automatically (i.e., without us consciously intending for it to happen) in most situations for dimensions like gender, age, and race. Since classifying someone into a group is the first step to activating a group stereotype, this research provides important information about how easily stereotypes can be activated. And because it is hard for people to accurately report on things that happen so quickly, this issue has been difficult to study using more traditional self-report measures. Using EEGs has, therefore, been helpful in providing interesting new insights into social behavior.

DO WE USE OUR OWN BEHAVIOR TO HELP US UNDERSTAND OTHERS?

Classifying someone into a social group then activating the associated stereotype is one way to make inferences about others. However, it is not the only method. Another strategy is to imagine what our own thoughts, feelings, and behaviors would be in a similar situation. Then we can use our simulated reaction as a best guess about how someone else will respond ([Goldman, 2005](#)). After all, we are experts in our own feelings, thoughts, and tendencies. It might be hard to know what other people are feeling and thinking, but we can always ask ourselves how we would feel and act if we were in their shoes.

There has been some debate about whether **simulation** is used to get into the minds of others ([Carruthers & Smith, 1996](#); [Gallese & Goldman, 1998](#)). Social neuroscience research has addressed this question by looking at the brain areas used when people think about themselves and others. If the same brain areas

are active for the two types of judgments, it lends support to the idea that the self may be used to make inferences about others via simulation.

We know that an area in the prefrontal cortex called the **medial prefrontal cortex** (mPFC) – located in the middle of the frontal lobe – is active when people think about themselves ([Kelley, Macrae, Wyland, Caglar, Inati, & Heatherton, 2002](#)). This conclusion comes from studies using **functional magnetic resonance imaging**, or fMRI. While EEG measures the brain's electrical activity, fMRI measures changes in the oxygenation of blood flowing in the brain. When neurons become more active, blood flow to the area increases to bring more oxygen and glucose to the active cells. fMRI allows us to image these changes in oxygenation by placing people in an fMRI machine or scanner (Figure 3), which consists of large magnets that create strong magnetic fields. The magnets affect the alignment of the oxygen molecules within the blood (i.e., how they are tilted). As the oxygen molecules move in and out of alignment with the magnetic fields, their nuclei produce energy that can be detected with special sensors placed close to the head. Recording fMRI involves having the subject lay on a small bed that is then rolled into the scanner. While fMRI does require subjects to lie still within the small scanner and the large magnets involved are noisy, the scanning itself is safe and painless. Like EEG, the subject can then be asked to think about different topics or engage in different tasks as brain activity is measured. If we know what a person is thinking or doing when fMRI detects a blood flow increase to a particular brain area, we can infer that part of the brain is involved with the thought or action. fMRI

is particularly useful for identifying which particular brain areas are active at a given point in time.



Figure 3. Functional magnetic resonance imaging (fMRI) scanner used to image the brain while people perform tasks. The scanner allows researchers to view the changes in blood oxygenation in specific locations in the brain during a task. Images are collected using powerful magnets and radio waves that shift the position of atoms in oxygenated blood that rushes to areas involved in performing the task. MRIs are non-invasive and there are no known risks from exposure to the magnetic fields or radio waves. Subjects lay on the while bed with their head inside the head coil, then the bed is moved into the scanner. [Image: Janne Moren, <https://goo.gl/MKb2jn>, CC BY-NC-SA 2.0, <https://goo.gl/TocOZF>]

The conclusion that the mPFC is associated with the self comes from studies measuring fMRI while subjects think about themselves (e.g., saying whether traits are descriptive of themselves). Using this knowledge, other researchers have looked at whether the same brain area is active when people make inferences about others. Mitchell, Neil Macrae, and Banaji (2005) showed subjects pictures of strangers and had them judge either how pleased the person was to have his or her picture taken

or how symmetrical the face appeared. Judging whether someone is pleased about being photographed requires making an inference about someone's internal feelings – we call this **mentalizing**. By contrast, facial symmetry judgments are based solely on physical appearances and do not involve mentalizing. A comparison of brain activity during the two types of judgments shows more activity in the mPFC when making the mental versus physical judgments, suggesting this brain area is involved when inferring the internal beliefs of others.

There are two other notable aspects of this study. First, mentalizing about others also increased activity in a variety of regions important for many aspects of social processing, including a region important in representing biological motion (**superior temporal sulcus** or STS), an area critical for emotional processing (amygdala), and a region also involved in thinking about the beliefs of others (**temporal parietal junction**, TPJ) ([Gobbini & Haxby, 2007](#); [Schultz, Imamizu, Kawato, & Frith, 2004](#)) (Figure 4). This finding shows that a distributed and interacting set of brain areas is likely to be involved in social processing. Second, activity in the most ventral part of the mPFC (the part closer to the belly rather than toward the top of the head), which has been most consistently associated with thinking about the self, was particularly active when subjects mentalized about people they rated as similar to themselves. Simulation is thought to be most likely for similar others, so this finding lends support to the conclusion that we use simulation to mentalize about others. After all, if you encounter someone who has the same musical taste as you, you will probably assume you have other things in common with him. By contrast, if you learn that someone

loves music that you hate, you might expect him to differ from you in other ways ([Srivastava, Guglielmo, & Beer, 2010](#)). Using a simulation of our own feelings and thoughts will be most accurate if we have reason to think the person's internal experiences are like our own. Thus, we may be most likely to use simulation to make inferences about others if we think they are similar to us.

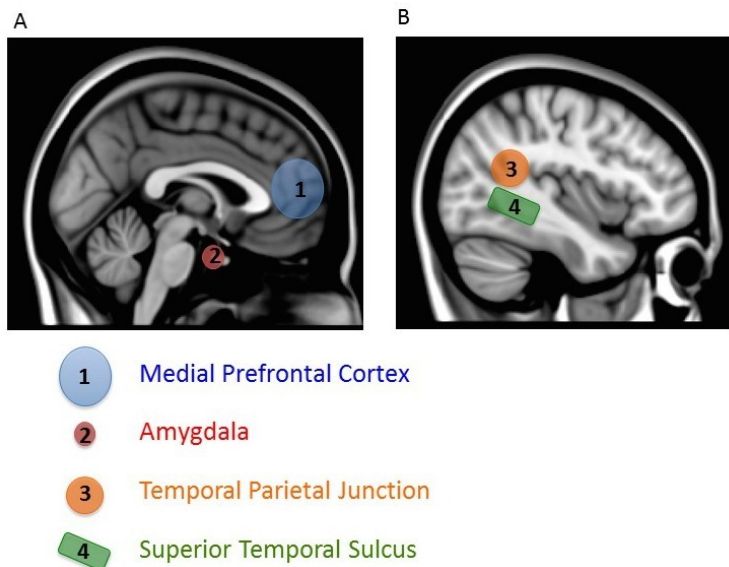


Figure 4. The areas of the brain most commonly associated with processing of self and others. Panel A is a sagittal view, looking at the inside of the brain as if it were sliced in half. Panel B is a lateral view, showing the brain from the outside. The medial prefrontal cortex (mPFC) (Structure 1) is commonly activated when thinking about one's self and when thinking about similar others. Additionally, the amygdala (Structure 2) is important for learning about and detecting important things in our environments and plays an important role in fear learning and expression. The temporal parietal junction (TPJ) (Structure 3), located at the intersection of the parietal and temporal lobes, is activated when people think about the beliefs of others. Activity in the superior temporal sulcus (STS) (Structure 4) is commonly observed when people view biological motion. The common activation of this network of regions when people think about the feelings, thoughts, and intentions of others indicates that the processing of others involves a number of complex psychological processes.

This research is a good example of how social neuroscience is revealing the **functional neuroanatomy** of social behavior. That is, it tells us which brain areas are involved with social behavior. The mPFC (as well as other areas such as the STS, amygdala, and TPJ) is involved in making judgments about the self

and others. This research also provides new information about how inferences are made about others. Whereas some have doubted the widespread use of simulation as a means for making inferences about others, the activation of the mPFC when mentalizing about others, and the sensitivity of this activation to similarity between self and other, provides evidence that simulation occurs.

WHAT IS THE COST OF SOCIAL STRESS?

Stress is an unfortunately frequent experience for many of us. **Stress** – which can be broadly defined as a threat or challenge to our well-being – can result from everyday events like a course exam or more extreme events such as experiencing a natural disaster. When faced with a stressor, **sympathetic nervous system** activity increases in order to prepare our body to respond to the challenge. This produces what Selye (1950) called a **fight or flight response**. The release of **hormones**, which act as messengers from one part of an organism (e.g., a cell or gland) to another part of the organism, is part of the stress response.

A small amount of stress can actually help us stay alert and active. In comparison, sustained stressors, or chronic stress, detrimentally affect our health and impair performance (Al'Absi, Hugdahl, & Lovallo, 2002; Black, 2002; Lazarus, 1974). This happens in part through the chronic secretion of stress-related hormones (e.g., Davidson, Pizzagalli, Nitschke, & Putnam, 2002; Dickerson, Gable, Irwin, Aziz, & Kemeny, 2009). In particular, stress activates the **hypothalamic-pituitary-adrenal (HPA) axis** to release **cortisol** (see Figure 5 for a discussion). Chronic stress, by way of increases in

cortisol, impairs attention, memory, and self-control ([Arnsten, 2009](#)). Cortisol levels can be measured non-invasively in bodily fluids, including blood and saliva. Researchers often collect a cortisol sample before and after a potentially stressful task. In one common collection method, subjects place polymer swabs under their tongue for 1 to 2 minutes to soak up saliva. The saliva samples are then stored and analyzed later to determine the level of cortisol present at each time point.

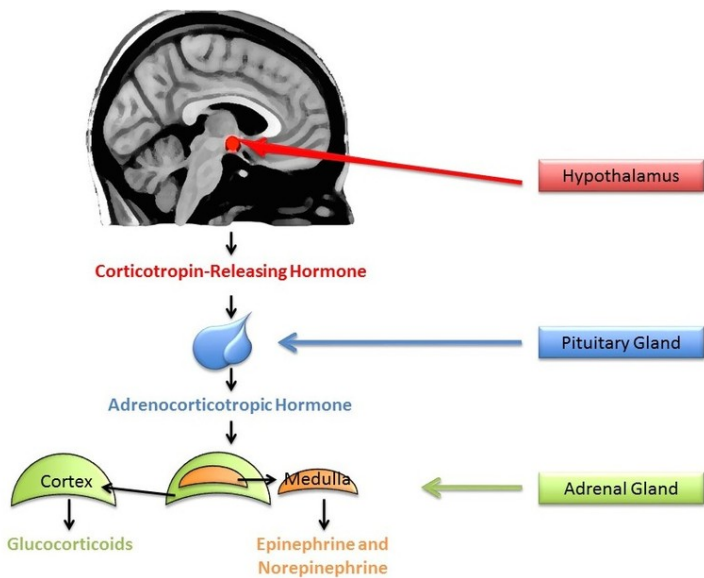


Figure 5: The Hypothalamic-pituitary-adrenal (HPA) axis. Black arrows represent the stress response pathway starting in the brain at the hypothalamus (an area within the brain). Stress triggers neurons in the hypothalamus to release corticotrophin-releasing hormone (CRH). The CRH is transported to the pituitary gland, another area in the brain, that activates the secretion of adrenocorticotrophic hormone (ACTH). In turn, ACTH stimulates the adrenal glands that sit on top of the kidneys. The adrenal glands are composed of the outer adrenal cortex and inner adrenal medulla. The adrenal cortex secretes glucocorticoids (including cortisol) and the medulla secretes epinephrine and norepinephrine. Stress, both psychological and physical, activates the HPA axis and results in the systemic release of cortisol, epinephrine, and norepinephrine.

Whereas early stress researchers studied the effects of physical stressors like loud noises, social neuroscientists have been instrumental in studying how our interactions with other people can cause stress. This question has been addressed through **neuroendocrinology**, or the study of how the brain and hormones act in concert to coordinate the physiology of the body. One contribution of this work has been in understanding the conditions under which

other people can cause stress. In one study, Dickerson, Mycek, and Zaldivar ([2008](#)) asked undergraduates to deliver a speech either alone or to two other people. When the students gave the speech in front of others, there was a marked increase in cortisol compared with when they were asked to give a speech alone. This suggests that like chronic physical stress, everyday social stressors, like having your performance judged by others, induces a stress response. Interestingly, simply giving a speech in the same room with someone who is doing something else did not induce a stress response. This suggests that the mere presence of others is not stressful, but rather it is the potential for them to judge us that induces stress.

Worrying about what other people think of us is not the only source of social stress in our lives. Other research has shown that interacting with people who belong to different social groups than us – what social psychologists call **outgroup** members – can increase physiological stress responses. For example, cardiovascular responses associated with stress like contractility of the heart ventricles and the amount of blood pumped by the heart (what is called cardiac output) are increased when interacting with outgroup as compared with **ingroup** members (i.e., people who belong to the same social group we do) ([Mendes, Blascovich, Likel, & Hunter, 2002](#)). This stress may derive from the expectation that interactions with dissimilar others will be uncomfortable ([Stephan & Stephan, 1985](#)) or concern about being judged as unfriendly and prejudiced if the interaction goes poorly ([Plant & Devine, 2003](#)).

The research just reviewed shows that events in our

social lives can be stressful, but are social interactions always bad for us? No. In fact, while others can be the source of much stress, they are also a major buffer against stress. Research on **social support** shows that relying on a network of individuals in tough times gives us tools for dealing with stress and can ward off loneliness ([Cacioppo & Patrick, 2008](#)). For instance, people who report greater social support show a smaller increase in cortisol when performing a speech in front of two evaluators ([Eisenberger, Taylor, Gable, Hilmert, & Lieberman, 2007](#)).

What determines whether others will increase or decrease stress? What matters is the context of the social interaction. When it has potential to reflect badly on the self, social interaction can be stressful, but when it provides support and comfort, social interaction can protect us from the negative effects of stress. Using neuroendocrinology by measuring hormonal changes in the body has helped researchers better understand how social factors impact our body and ultimately our health.

CONCLUSIONS

Human beings are intensely social creatures – our lives are intertwined with other people and our health and well-being depend on others. Social neuroscience helps us to understand the critical function of how we make sense of and interact with other people. This module provides an introduction to what social neuroscience is and what we have already learned from it, but there is much still to understand. As we move forward, one exciting future direction will be to better understand how different parts of the brain and body interact to produce the numerous and complex patterns of social behavior

that humans display. We hinted at some of this complexity when we reviewed research showing that while the mPFC is involved in mentalizing, other areas such as the STS, amygdala, and TPJ are as well. There are likely additional brain areas involved as well, interacting in ways we do not yet fully understand. These brain areas in turn control other aspects of the body to coordinate our responses during social interactions. Social neuroscience will continue to investigate these questions, revealing new information about how social processes occur, while also increasing our understanding of basic neural and physiological processes.

OUTSIDE RESOURCES

Society for Social Neuroscience

<http://www.s4sn.org>

Video: See a demonstration of fMRI data being collected.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=238#oembed-1>

Video: See an example of EEG data being collected.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=238#oembed-2>

Video: View two tasks frequently used in the lab to create stress – giving a speech in front of strangers, and doing math computations out loud in front of others. Notice how some subjects show obvious signs of stress, but in some situations, cortisol changes suggest that even people who appear calm are experiencing a physiological response associated with stress.

<https://www.youtube.com/watch?v=aYI6lCeeT5g>

Video: Watch a video used by Fritz Heider and Marianne Simmel in a landmark study on social perception published in 1944. Their goal was to investigate how we perceive other people, and they studied it by seeing how readily we apply people-like interpretations to non-social stimuli.



One or more interactive elements has been excluded from this version of the

text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=238#oembed-3>

DISCUSSION QUESTIONS

1. Categorizing someone as a member of a social group can activate group stereotypes. EEG research suggests that social categorization occurs quickly and often automatically. What does this tell us about the likelihood of stereotyping occurring? How can we use this information to develop ways to stop stereotyping from happening?
2. Watch this video, similar to what was used by Fritz Heider and Marianne Simmel in a landmark study on social perception published in 1944, and imagine telling a friend what happened in the video. http://intentionperception.org/wp-content/uploads/2013/02/Heider_Flash.swf. After watching the video, think about the following: Did you describe the motion of the objects solely in geometric terms (e.g., a large triangle moved from the left to the right), or did you describe the movements as actions of animate beings, maybe even of people (e.g., the circle goes into the house and shuts the door)? In the original research, 33 of 34 subjects described the action of the shapes using human terms. What does this tell us about our tendency to mentalize?

3. Consider the types of things you find stressful. How many of them are social in nature (e.g., are related to your interactions with other people)? Why do you think our social relations have such potential for stress? In what ways can social relations be beneficial and serve as a buffer for stress?

VOCABULARY

Amygdala

A region located deep within the brain in the medial area (toward the center) of the temporal lobes (parallel to the ears). If you could draw a line through your eye sloping toward the back of your head and another line between your two ears, the amygdala would be located at the intersection of these lines. The amygdala is involved in detecting relevant stimuli in our environment and has been implicated in emotional responses.

Automatic process

When a thought, feeling, or behavior occurs with little or no mental effort. Typically, automatic processes are described as involuntary or spontaneous, often resulting from a great deal of practice or repetition.

Cortisol

A hormone made by the adrenal glands, within the cortex. Cortisol helps the body maintain blood pressure and immune function. Cortisol increases when the body is under stress.

Electroencephalogram

A measure of electrical activity generated by the brain's neurons.

Fight or flight response

The physiological response that occurs in response to a perceived threat, preparing the body for actions needed to deal with the threat.

Functional magnetic resonance imaging

A measure of changes in the oxygenation of blood flow as areas in the brain become active.

Functional neuroanatomy

Classifying how regions within the nervous system relate to psychology and behavior.

Hormones

Chemicals released by cells in the brain or body that affect cells in other parts of the brain or body.

Hypothalamic-pituitary-adrenal (HPA) axis

A system that involves the hypothalamus (within the brain), the pituitary gland (within the brain), and the adrenal glands (at the top of the kidneys). This system helps maintain homeostasis (keeping the body's systems within normal ranges) by regulating digestion, immune function, mood, temperature, and energy use. Through this, the HPA regulates the body's response to stress and injury.

Ingroup

A social group to which an individual identifies or belongs.

Lesions

Damage or tissue abnormality due, for example, to an injury, surgery, or a vascular problem.

Medial prefrontal cortex

An area of the brain located in the middle of the frontal lobes (at the front of the head), active when people mentalize about the self and others.

Mentalizing

The act of representing the mental states of oneself and others. Mentalizing allows humans to interpret the intentions, beliefs, and emotional states of others.

Neuroendocrinology

The study of how the brain and hormones act in concert to coordinate the physiology of the body.

Outgroup

A social group to which an individual does not identify or belong.

Simulation

Imaginary or real imitation of other people's behavior or feelings.

Social categorization

The act of mentally classifying someone into a social group (e.g., as female, elderly, a librarian).

Social support

A subjective feeling of psychological or physical comfort provided by family, friends, and others.

Stereotypes

The beliefs or attributes we associate with a specific social group. Stereotyping refers to the act of assuming that because someone is a member of a particular group, he or she possesses the group's attributes. For example, stereotyping occurs when we assume someone is unemotional just because he is man, or particularly athletic just because she is African American.

Stress

A threat or challenge to our well-being. Stress can have both a psychological component, which consists of our subjective thoughts and feelings about being threatened or challenged, as well as a physiological component, which consists of our body's response to the threat or challenge (see "fight or flight response").

Superior temporal sulcus

The sulcus (a fissure in the surface of the brain) that separates the superior temporal gyrus from the middle temporal gyrus. Located in the temporal lobes (parallel to the ears), it is involved in perception of biological motion or the movement of animate objects.

Sympathetic nervous system

A branch of the autonomic nervous system that controls many of the body's internal organs. Activity of the SNS generally mobilizes the body's fight or flight response.

Temporal parietal junction

The area where the temporal lobes (parallel to the ears) and parietal lobes (at the top of the head toward the back) meet. This area is important in mentalizing and distinguishing between the self and others.

REFERENCES

- Al'Absi, M., Hugdahl, K., & Lovallo, W. (2002). Adrenocortical stress responses and altered working memory performance. *Psychophysiology*, 39(1), 95–99.
- Arnsten, A. F. T. (2009). Stress signaling pathways that impair prefrontal cortex structure and function. *Nature Neuroscience Reviews*, 10(6), 410–422.
- Black, P. (2002). Stress and the inflammatory response: A review of neurogenic inflammation. **Brain, Behavior, & Immunity*, 16*, 622–653.
- Cacioppo, J. T., & Berntson, G. G. (1992). Social psychological contributions to the decade of the brain: Doctrine of multilevel analysis. *American Psychologist*, 47, 1019–1028.
- Cacioppo, J. T., & Patrick, B. (2008). *Loneliness: Human nature and the need for social connection*. New York, NY: W. W. Norton & Company.
- Carruthers, P. and Smith, P. (1996). *Theories of Theories of Mind*. New York, NY: Cambridge University Press.
- Davidson, R. J., Pizzagalli, D., Nitschke, J. B., &

Putnam, K. (2002). Depression: Perspectives from affective neuroscience. *Annual Review of Psychology*, 53, 545–574.

- Dickerson, S. S., Gable, S. L., Irwin, M. R., Aziz, N., & Kemeny, M. E. (2009). Social-evaluative threat and proinflammatory cytokine regulation in an experimental laboratory investigation. *Psychological Science*, 20, 1237–1244.
- Dickerson, S. S., Mycek, P. J., & Zaldivar, F. (2008). Negative social evaluation, but not mere social presence, elicits cortisol responses to a laboratory stressor task. *Health Psychology*, 27(1), 116–121.
- Eisenberger, N. I., Taylor, S. E., Gable, S. L., Hilmert, C. J., & Lieberman, M. D. (2007). Neural pathways link social support to attenuated neuroendocrine stress responses. *Neuroimage*, 35(4), 1601–1612.
- Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2, 493–501.
- Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, 45(1), 32–41.
- Goldman, A. I. (2005). Imitation, mind reading, and simulation. In S. Hurley & N. Chater (Eds.), *Perspectives on imitation: From neuroscience to social science (Vol. 2: Imitation, human development, and culture*, pp. 79–93). Cambridge, MA: MIT Press.
- Ito, T. A., & Urland, G. R. (2003). Race and gender

on the brain: Electrocortical measures of attention to race and gender of multiply categorizable individuals. *Journal of Personality and Social Psychology*, 85, 616–626.

- Ito, T.A., & Urland, G.R. (2005). The influence of processing objectives on the perception of faces: An ERP study of race and gender perception. *Cognitive, Affective, and Behavioral Neuroscience*, 5, 21–36.
- Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., & Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *Journal of Cognitive Neuroscience*, 14, 785–794.
- Lazarus, R. S., (1974). Psychological stress and coping in adaptation and illness. *International Journal of Psychiatry in Medicine, 5*, 321–333.
- Mendes, W. B., Blascovich, J., Lickel, B., & Hunter, S. (2002). Challenge and threat during social interactions with White and Black men. *Personality and Social Psychology Bulletin*, 28, 939–952.
- Mitchell, J. P., Neil Macrae, C., & Banaji, M. R. (2005). Forming impressions of people versus inanimate objects: social-cognitive processing in the medial prefrontal cortex. *Neuroimage*, 26(1), 251–257.
- Mouchetant-Rostaing, Y., & Giard, M. H. (2003). Electrophysiological correlates of age and gender perception on human faces. *Journal of Cognitive Neuroscience*, 15, 900–910.

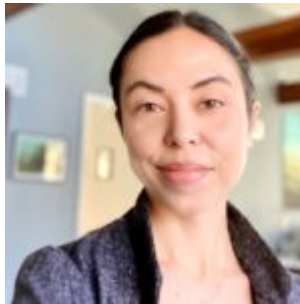
- Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: From animal models to human behavior. *Neuron*, 48, 175.
- Plant, E. A., & Devine, P. G. (2003). The antecedents and implications of interracial anxiety. *Personality and Social Psychology Bulletin, 29*, 790–801.
- Schultz, J., Imamizu, H., Kawato, M., & Frith, C. D. (2004). Activation of the human superior temporal gyrus during observation of goal attribution by intentional objects. *Journal of Cognitive Neuroscience*, 16, 1695–1705.
- Selye, H. (1950). *The physiology and pathology of exposure to stress*. Montreal: Acta Inc.
- Srivastava, S., Guglielmo, S., & Beer, J. S. (2010). Perceiving others' personalities: Examining the dimensionality, assumed similarity to the self, and stability of perceiver effects. *Journal of Personality and Social Psychology*, 98, 520.
- Stephan, W. G., & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues*, 41(3), 157–175.
- Whalen, P. J., & Phelps, E. A. (2009). *The human amygdala*. New York, NY: The Guilford Press.

AUTHORS



Tiffany A. Ito

Tiffany A. Ito is a Professor of Psychology and Neuroscience at the University of Colorado Boulder. Her research integrates neuroscience methods and theories to better understand social processes, with a particular focus on aspects of stereotyping and prejudice.



Jennifer T. Kubota

Jennifer Kubota is an assistant professor at

University of Delaware. She received her Ph.D. in social neuroscience from The University of Colorado Boulder. Her work focuses on the psychological and neural substrates of impression formation and their relation to decision-making.

CREATIVE COMMONS LICENSE



SOCIAL NEUROSCIENCE by [Tiffany A. Ito and Jennifer T. Kubota](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

PART XI.

**UNIT 11: THE BIOLOGY OF
LEARNING AND MEMORY**

CHAPTER 17.

MEMORY (ENCODING, STORAGE, RETRIEVAL)

KATHLEEN B. MCDERMOTT AND HENRY L. ROEDIGER III

“Memory” is a single term that reflects a number of different abilities: holding information briefly while working with it (working memory), remembering episodes of one’s life (episodic memory), and our general knowledge of facts of the world (semantic memory), among other types. Remembering episodes involves three processes: encoding information (learning it, by perceiving it and relating it to past knowledge), storing it (maintaining it over time), and then retrieving it (accessing the information when needed). Failures can occur at any stage, leading to forgetting or to having false memories. The key to improving one’s memory is to improve processes of encoding and to use techniques that guarantee effective retrieval. Good encoding techniques include relating new information to what one already knows, forming mental images, and creating associations among information that needs to be remembered. The key to good retrieval is developing effective cues that will lead the rememberer back to the encoded information. Classic mnemonic systems, known since the time of the

ancient Greeks and still used by some today, can greatly improve one's memory abilities.

LEARNING OBJECTIVES

- Define and note differences between the following forms of memory: working memory, episodic memory, semantic memory, collective memory.
- Describe the three stages in the process of learning and remembering.
- Describe strategies that can be used to enhance the original learning or encoding of information.
- Describe strategies that can improve the process of retrieval.
- Describe why the classic mnemonic device, the method of loci, works so well.

INTRODUCTION

In 2013, Simon Reinhard sat in front of 60 people in a room at Washington University, where he memorized an increasingly long series of digits. On the first round, a computer generated 10 random digits—6 1 9 4 8 5 6 3 7 1—on a screen for 10 seconds. After the series disappeared, Simon typed them into his computer. His recollection was perfect. In the next phase, 20 digits appeared on the screen for 20 seconds. Again, Simon got them all correct. No one in the audience (mostly professors, graduate students, and undergraduate students) could recall the 20 digits perfectly. Then came 30 digits, studied for 30 seconds; once again, Simon didn't misplace even a single digit. For a final trial, 50 digits appeared on the screen for 50 seconds, and again, Simon

got them all right. In fact, Simon would have been happy to keep going. His record in this task—called “forward digit span”—is 240 digits!



In some ways memory is like file drawers where you store mental information. Memory is also a series of processes: how does that information get filed to begin with and how does it get retrieved when needed? [Image: M Cruz, <https://goo.gl/DhOMgp>, CC BY-SA 4.0, <https://goo.gl/SWjq94>]

When most of us witness a performance like that of Simon Reinhard, we think one of two things: First, maybe he’s cheating somehow. (No, he is not.) Second, Simon must have abilities more advanced than the rest of humankind. After all, psychologists established many years ago that the normal memory span for adults is about 7 digits, with some of us able to recall a few more and others a few less ([Miller, 1956](#)). That is why the

first phone numbers were limited to 7 digits—psychologists determined that many errors occurred (costing the phone company money) when the number was increased to even 8 digits. But in normal testing, no one gets 50 digits correct in a row, much less 240. So, does Simon Reinhard simply have a photographic memory? He does not. Instead, Simon has taught himself simple strategies for remembering that have greatly increased his capacity for remembering virtually any type of material—digits, words, faces and names, poetry, historical dates, and so on. Twelve years earlier, before he started training his memory abilities, he

had a digit span of 7, just like most of us. Simon has been training his abilities for about 10 years as of this writing, and has risen to be in the top two of “memory athletes.” In 2012, he came in second place in the World Memory Championships (composed of 11 tasks), held in London. He currently ranks second in the world, behind another German competitor, Johannes Mallow. In this module, we reveal what psychologists and others have learned about memory, and we also explain the general principles by which you can improve your own memory for factual material.

VARIETIES OF MEMORY

For most of us, remembering digits relies on *short-term*

memory, or *working*

memory—the ability to hold information in our minds for a brief time and work with it (e.g., multiplying 24×17 without using paper

would rely on working memory). Another type of memory is **episodic memory**—the ability to remember the episodes of our lives. If you were given the task of recalling everything you did 2 days ago, that would be a test of



To be a good chess player you have to learn to increase working memory so you can plan ahead for several offensive moves while simultaneously anticipating – through use of memory – how the other player could counter each of your planned moves. [Image: karpidis, <https://goo.gl/EhzMKM>, CC BY-SA 2.0, <https://goo.gl/jSSrcO>]

episodic memory; you would be required to mentally travel through the day in your mind and note the main events. **Semantic memory** is our storehouse of more-or-less permanent knowledge, such as the meanings of words in a language (e.g., the meaning of “parasol”) and the huge collection of facts about the world (e.g., there are 196 countries in the world, and 206 bones in your body). *Collective memory* refers to the kind of memory that people in a group share (whether family, community, schoolmates, or citizens of a state or a country). For example, residents of small towns often strongly identify with those towns, remembering the local customs and historical events in a unique way. That is, the community’s collective memory passes stories and recollections between neighbors and to future generations, forming a memory system unto itself.

Psychologists continue to debate the classification of types of memory, as well as which types rely on others ([Tulving, 2007](#)), but for this module we will focus on episodic memory. Episodic memory is usually what people think of when they hear the word “memory.” For example, when people say that an older relative is “losing her memory” due to Alzheimer’s disease, the type of memory-loss they are referring to is the inability to recall events, or episodic memory. (Semantic memory is actually preserved in early-stage Alzheimer’s disease.) Although remembering specific events that have happened over the course of one’s entire life (e.g., your experiences in sixth grade) can be referred to as **autobiographical memory**, we will focus primarily on the episodic memories of more recent events.

THREE STAGES OF THE LEARNING/MEMORY

PROCESS

Psychologists distinguish between three necessary stages in the learning and memory process: **encoding**, **storage**, and **retrieval** ([Melton, 1963](#)). Encoding is defined as the initial learning of information; storage refers to maintaining information over time; retrieval is the ability to access information when you need it. If you meet someone for the first time at a party, you need to encode her name (Lyn Goff) while you associate her name with her face. Then you need to maintain the information over time. If you see her a week later, you need to recognize her face and have it serve as a cue to retrieve her name. Any successful act of remembering requires that all three stages be intact. However, two types of errors can also occur. Forgetting is one type: you see the person you met at the party and you cannot recall her name. The other error is misremembering (false recall or false recognition): you see someone who looks like Lyn Goff and call the person by that name (false recognition of the face). Or, you might see the real Lyn Goff, recognize her face, but then call her by the name of another woman you met at the party (misrecall of her name).

Whenever forgetting or misremembering occurs, we can ask, at which stage in the learning/memory process was there a failure?—though it is often difficult to answer this question with precision. One reason for this inaccuracy is that the three stages are not as discrete as our description implies. Rather, all three stages depend on one another. How we encode information determines how it will be stored and what cues will be effective when we try to retrieve it. And too, the act of retrieval itself also changes the way information is subsequently

remembered, usually aiding later recall of the retrieved information. The central point for now is that the three stages—encoding, storage, and retrieval—affect one another, and are inextricably bound together.

ENCODING

Encoding refers to the initial experience of perceiving and learning information. Psychologists often study recall by having participants study a list of pictures or words. Encoding in these situations is fairly straightforward. However, “real life” encoding is much more challenging. When you walk across campus, for example, you encounter countless sights and sounds—friends passing by, people playing Frisbee, music in the air. The physical and mental environments are much too rich for you to encode all the happenings around you or the internal thoughts you have in response to them. So, an important first principle of encoding is that it is selective: we attend to some events in our environment and we ignore others. A second point about encoding is that it is prolific; we are always encoding the events of our lives—attending to the world, trying to understand it. Normally this presents no problem, as our days are filled with routine occurrences, so we don’t need to pay attention to everything. But if something does happen that seems strange—during your daily walk across campus, you see a giraffe—then we pay close attention and try to understand why we are seeing what we are seeing.

Right after your typical walk across campus (one without the appearance of a giraffe), you would be able to remember the events reasonably well if you were asked. You could say whom you bumped into, what song was playing from a radio, and so on. However, suppose someone asked you to recall the same walk a month later. You wouldn't stand a chance. You would likely be able to recount the basics of a typical walk across campus, but not the precise details of that

particular walk. Yet, if you had seen a giraffe during that walk, the event would have been fixed in your mind for a long time, probably for the rest of your life. You would tell your friends about it, and, on later occasions when you saw a giraffe, you might be reminded of the day you saw one on campus. Psychologists have long pinpointed **distinctiveness**—having an event stand out as quite different from a background of similar events—as a key to remembering events (Hunt, 2003).

In addition, when vivid memories are tinged with strong emotional content, they often seem to leave a permanent mark on us. Public tragedies, such as terrorist attacks, often create vivid memories in those who witnessed them. But even those of us not directly



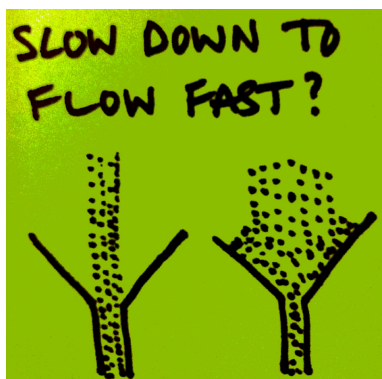
A giraffe in the context of a zoo or its natural habitat may register as nothing more than ordinary, but put it in another setting – in the middle of a campus or a busy city – and its level of distinctiveness increases dramatically. Distinctiveness is a key attribute to remembering events. [Image: Colin J Babb, <https://goo.gl/Cci2yl>, CC BY-SA 2.0, <https://goo.gl/jSSrcO>]

involved in such events may have vivid memories of them, including memories of first hearing about them. For example, many people are able to recall their exact physical location when they first learned about the assassination or accidental death of a national figure. The term **flashbulb memory** was originally coined by Brown and Kulik ([1977](#)) to describe this sort of vivid memory of finding out an important piece of news. The name refers to how some memories seem to be captured in the mind like a flash photograph; because of the distinctiveness and emotionality of the news, they seem to become permanently etched in the mind with exceptional clarity compared to other memories.

Take a moment and think back on your own life. Is there a particular memory that seems sharper than others? A memory where you can recall unusual details, like the colors of mundane things around you, or the exact positions of surrounding objects? Although people have great confidence in flashbulb memories like these, the truth is, our objective accuracy with them is far from perfect ([Talarico & Rubin, 2003](#)). That is, even though people may have great confidence in what they recall, their memories are not as accurate (e.g., what the actual colors were; where objects were truly placed) as they tend to imagine. Nonetheless, all other things being equal, distinctive and emotional events are well-remembered.

Details do not leap perfectly from the world into a person's mind. We might say that we went to a party and remember it, but what we remember is (at best) what we encoded. As noted above, the process of encoding is selective, and in complex situations, relatively few of many possible details are noticed and encoded. The process of encoding always involves **recoding**—that is,

taking the information from the form it is delivered to us and then converting it in a way that we can make sense of it. For example, you might try to remember the colors of a rainbow by using the acronym ROY G BIV (red, orange, yellow, green, blue, indigo, violet). The process of recoding the colors into a name can help us to remember. However, recoding can also introduce errors—when we accidentally add information during encoding, then remember that *new* material as if it had been part of the actual experience (as discussed below).



Although it requires more effort, using images and associations can improve the process of recoding. [Image: psd, <https://goo.gl/9xjcDe>, CC BY 2.0, <https://goo.gl/9uSnqN>]

Psychologists have studied many recoding strategies that can be used during study to improve retention. First, research advises that, as we study, we should think of the meaning of the events ([Craig & Lockhart, 1972](#)), and we should try to relate new events to information we already know. This helps us form associations that we can use to retrieve information later. Second,

imagining events also makes them more memorable; creating vivid images out of information (even verbal information) can greatly improve later recall ([Bower & Reitman, 1972](#)). Creating imagery is part of the technique Simon Reinhard uses to remember huge numbers of digits, but we can all use images to encode information more effectively. The basic concept behind good encoding strategies is to form distinctive memories (ones

that stand out), and to form links or associations among memories to help later retrieval ([Hunt & McDaniel, 1993](#)). Using study strategies such as the ones described here is challenging, but the effort is well worth the benefits of enhanced learning and retention.

We emphasized earlier that encoding is selective: people cannot encode all information they are exposed to. However, recoding can add information that was not even seen or heard during the initial encoding phase. Several of the recoding processes, like forming associations between memories, can happen without our awareness. This is one reason people can sometimes remember events that did not actually happen—because during the process of recoding, details got added. One common way of inducing false memories in the laboratory employs a word-list technique ([Deese, 1959](#); [Roediger & McDermott, 1995](#)). Participants hear lists of 15 words, like *door, glass, pane, shade, ledge, sill, house, open, curtain, frame, view, breeze, sash, screen, and shutter*. Later, participants are given a test in which they are shown a list of words and asked to pick out the ones they'd heard earlier. This second list contains some words from the first list (e.g., *door, pane, frame*) and some words not from the list (e.g., *arm, phone, bottle*). In this example, one of the words on the second list is *window*, which—importantly—does not appear in the first list, but which is related to other words in that list. When subjects were tested with the second list, they were reasonably accurate with the studied words (*door*, etc.), recognizing them 72% of the time. However, when *window* was on the test, they falsely recognized it as having been on the list 84% of the time ([Stadler, Roediger, & McDermott, 1999](#)). The same thing happened with

many other lists the authors used. This phenomenon is referred to as the DRM (for Deese-Roediger-McDermott) effect. One explanation for such results is that, while students listened to items in the list, the words triggered the students to think about *window*, even though *window* was never presented. In this way, people seem to encode events that are not actually part of their experience.

Because humans are creative, we are always going beyond the information we are given: we automatically make associations and infer from them what is happening. But, as with the word association mix-up above, sometimes we make false memories from our inferences—remembering the inferences themselves as if they were actual experiences. To illustrate this, Brewer (1977) gave people sentences to remember that were designed to elicit *pragmatic inferences*. Inferences, in general, refer to instances when something is not explicitly stated, but we are still able to guess the undisclosed intention. For example, if your friend told you that she didn't want to go out to eat, you may infer that she doesn't have the money to go out, or that she's too tired. With *pragmatic inferences*, there is usually *one* particular inference you're likely to make. Consider the statement Brewer (1977) gave her participants: "The karate champion hit the cinder block." After hearing or seeing this sentence, participants who were given a memory test tended to remember the statement as having been, "The karate champion *broke* the cinder block." This remembered statement is not necessarily a *logical* inference (i.e., it is perfectly reasonable that a karate champion could hit a cinder block without breaking it). Nevertheless,

the *pragmatic* conclusion from hearing such a sentence is that the block was likely broken. The participants remembered this inference they made while hearing the sentence in place of the actual words that were in the sentence (see also [McDermott & Chan, 2006](#)).

Encoding—the initial registration of information—is essential in the learning and memory process. Unless an event is encoded in some fashion, it will not be successfully remembered later. However, just because an event is encoded (even if it is encoded well), there's no guarantee that it will be remembered later.

STORAGE

Every experience we have changes our brains. That may seem like a bold, even strange, claim at first, but it's true. We encode each of our experiences within the structures of the nervous system, making new impressions in the process—and each of those impressions involves changes in the brain. Psychologists (and neurobiologists) say that experiences

leave **memory traces**, or **engrams** (the two terms are synonyms). Memories have to be stored somewhere in the brain, so in order to do so, the brain biochemically alters itself and its neural



Memory traces, or engrams, are NOT perfectly preserved recordings of past experiences. The traces are combined with current knowledge to reconstruct what we think happened in the past.

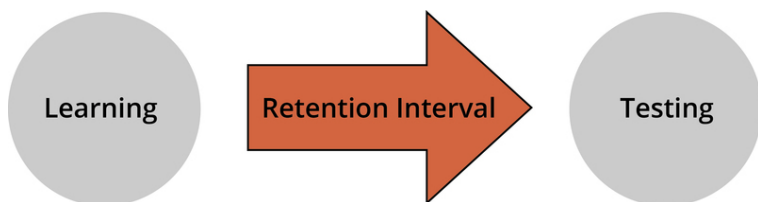
[Simon Bierdwald, <https://goo.gl/JDhdCE>, CC BY-NC-SA 2.0, <https://goo.gl/jSSrcO>]

tissue. Just like you might write yourself a note to remind you of something, the brain “writes” a memory trace, changing its own physical composition to do so. The basic idea is that events (occurrences in our environment) create engrams through a process of **consolidation**: the neural changes that occur after learning to create the memory trace of an experience. Although neurobiologists are concerned with exactly what neural processes change when memories are created, for psychologists, the term *memory trace* simply refers to the physical change in the nervous system (whatever that may be, exactly) that represents our experience.

Although the concept of engram or memory trace is extremely useful, we shouldn’t take the term too literally. It is important to understand that memory traces are not perfect little packets of information that lie dormant in the brain, waiting to be called forward to give an accurate report of past experience. Memory traces are not like video or audio recordings, capturing experience with great accuracy; as discussed earlier, we often have errors in our memory, which would not exist if memory traces were perfect packets of information. Thus, it is wrong to think that remembering involves simply “reading out” a faithful record of past experience. Rather, when we remember past events, we reconstruct them with the aid of our memory traces—but also with our current belief of what happened. For example, if you were trying to recall for the police who started a fight at a bar, you may not have a memory trace of who pushed whom first. However, let’s say you remember that one of the guys held the door open for you. When thinking back to the start of the fight, this knowledge (of how one guy was friendly to you) may unconsciously influence your memory of

what happened in favor of the nice guy. Thus, memory is a construction of what you actually recall and what you believe happened. In a phrase, remembering is reconstructive (we reconstruct our past with the aid of memory traces) not reproductive (a perfect reproduction or recreation of the past).

Psychologists refer to the time between learning and testing as the retention interval. Memories can consolidate during that time, aiding retention. However, experiences can also occur that undermine the memory. For example, think of what you had for lunch yesterday—a pretty easy task. However, if you had to recall what you had for lunch 17 days ago, you may well fail (assuming you don't eat the same thing every day). The 16 lunches you've had since that one have created **retroactive interference**. Retroactive interference refers to new activities (i.e., the subsequent lunches) during the retention interval (i.e., the time between the lunch 17 days ago and now) that interfere with retrieving the specific, older memory (i.e., the lunch details from 17 days ago). But just as newer things can interfere with remembering older things, so can the opposite happen. *Proactive interference* is when past memories interfere with the encoding of new ones. For example, if you have ever studied a second language, often times the grammar and vocabulary of your native language will pop into your head, impairing your fluency in the foreign language.



Retroactive interference is one of the main causes of forgetting ([McGeoch, 1932](#)). In the module *Eyewitness Testimony and Memory Biases* <http://noba.to/uy49tm37> Elizabeth Loftus describes her fascinating work on eyewitness memory, in which she shows how memory for an event can be changed via misinformation supplied during the retention interval. For example, if you witnessed a car crash but subsequently heard people describing it from their own perspective, this new information may interfere with or disrupt your own personal recollection of the crash. In fact, you may even come to remember the event happening exactly as the others described it! This **misinformation effect** in eyewitness memory represents a type of retroactive interference that can occur during the retention interval (see Loftus [[2005](#)] for a review). Of course, if correct information is given during the retention interval, the witness's memory will usually be improved.

Although interference may arise between the occurrence of an event and the attempt to recall it, *the effect itself is always expressed when we retrieve memories*, the topic to which we turn next.

RETRIEVAL

Endel Tulving argued that “the key process in memory is retrieval” ([1991](#), p. 91). Why should retrieval be given

more prominence than encoding or storage? For one thing, if information were encoded and stored but could not be retrieved, it would be useless. As discussed previously in this module, we encode and store thousands of events—conversations, sights and sounds—every day, creating memory traces. However, we later access only a tiny portion of what we’ve taken in. Most of our memories will never be used—in the sense of being brought back to mind, consciously. This fact seems so obvious that we rarely reflect on it. All those events that happened to you in the fourth grade that seemed so important then? Now, many years later, you would struggle to remember even a few. You may wonder if the traces of those memories still exist in some latent form. Unfortunately, with currently available methods, it is impossible to know.

Psychologists distinguish information that is available in memory from that which is accessible ([Tulving & Pearlstone, 1966](#)). *Available* information is the information that is stored in memory—but precisely how much and what types are stored cannot be known. That is, all we can know is what information we can retrieve—*accessible* information. The assumption is that accessible information represents only a tiny slice of the information available in our brains. Most of us have had the experience of trying to remember some fact or event, giving up, and then—all of a sudden!—it comes to us at a later time, even after we’ve stopped trying to remember it. Similarly, we all know the experience of failing to recall a fact, but then, if we are given several choices (as in a multiple-choice test), we are easily able to recognize it.



We can't know the entirety of what is in our memory, but only that portion we can actually retrieve. Something that cannot be retrieved now and which is seemingly gone from memory may, with different cues applied, reemerge.

[Image: Ores2k, <https://goo.gl/1du8Qe>, CC BY-NC-SA 2.0, <https://goo.gl/jSSrcO>]

What factors determine what information can be retrieved from memory? One critical factor is the type of hints, or *cues*, in the environment. You may hear a song on the radio that suddenly evokes memories of an earlier time in your life, even if you were not trying to remember it when the song came on. Nevertheless, the song is closely associated with that time, so it brings the experience to mind.

The general principle that underlies the effectiveness of retrieval cues is the **encoding specificity principle** ([Tulving & Thomson, 1973](#)): when people encode information, they do so in specific ways. For example, take the song on the radio: perhaps you heard it while you were at a terrific party, having a great, philosophical conversation with a friend. Thus, the song became part of that whole complex experience. Years later, even though you haven't thought about that party in ages, when you hear the song on the radio, the whole experience rushes back to you. In general, the encoding specificity principle states that, to the extent a retrieval cue (the song) matches or overlaps the memory trace of an experience (the party, the conversation), it will be effective in evoking the memory. A classic experiment on the encoding specificity principle

had participants memorize a set of words in a unique setting. Later, the participants were tested on the word sets, either in the same location they learned the words or a different one. As a result of encoding specificity, the students who took the test in the same place they learned the words were actually able to recall more words ([Godden & Baddeley, 1975](#)) than the students who took the test in a new setting.

One caution with this principle, though, is that, for the cue to work, it can't match too many other experiences ([Nairne, 2002](#); [Watkins, 1975](#)). Consider a lab experiment. Suppose you study 100 items; 99 are words, and one is a picture—of a penguin, item 50 in the list. Afterwards, the cue “recall the picture” would evoke “penguin” perfectly. No one would miss it. However, if the *word* “penguin” were placed in the same spot among the other 99 words, its memorability would be exceptionally worse. This outcome shows the power of distinctiveness that we discussed in the section on encoding: one picture is perfectly recalled from among 99 words because it stands out. Now consider what would happen if the experiment were repeated, but there were 25 pictures distributed within the 100-item list. Although the picture of the penguin would still be there, the probability that the cue “recall the picture” (at item 50) would be useful for the penguin would drop correspondingly. Watkins ([1975](#)) referred to this outcome as demonstrating the **cue overload principle**. That is, to be effective, a retrieval cue cannot be overloaded with too many memories. For the cue “recall the picture” to be effective, it should only match one item in the target set (as in the one-picture, 99-word case).

To sum up how memory cues function: for a retrieval

cue to be effective, a match must exist between the cue and the desired target memory; furthermore, to produce the best retrieval, the cue-target relationship should be distinctive. Next, we will see how the encoding specificity principle can work in practice.

Psychologists measure memory performance by using production tests (involving recall) or recognition tests (involving the selection of correct from incorrect information, e.g., a multiple-choice test). For example, with our list of 100 words, one group of people might be asked to recall the list in any order (a free recall test), while a different group might be asked to circle the 100 studied words out of a mix with another 100, unstudied words (a recognition test). In this situation, the recognition test would likely produce better performance from participants than the recall test.

We usually think of recognition tests as being quite easy, because the cue for retrieval is a copy of the actual event that was presented for study. After all, what could be a better cue than the exact target (memory) the person is trying to access? In most cases, this line of reasoning is true; nevertheless, recognition tests do not provide perfect indexes of what is stored in memory. That is, you can fail to recognize a target staring you right in the face, yet be able to recall it later with a different set of cues ([Watkins & Tulving, 1975](#)). For example, suppose you had the task of recognizing the surnames of famous authors. At first, you might think that being given the actual last name would always be the best cue. However, research has shown this not necessarily to be true ([Muter, 1984](#)). When given names such as Tolstoy, Shaw, Shakespeare, and Lee, subjects might well say that Tolstoy and Shakespeare are famous authors, whereas Shaw and Lee

are not. But, when given a cued recall test using first names, people often recall items (produce them) that they had failed to recognize before. For example, in this instance, a cue like *George Bernard* _____ often leads to a recall of “Shaw,” even though people initially failed to recognize *Shaw* as a famous author’s name. Yet, when given the cue “William,” people may not come up with Shakespeare, because William is a common name that matches many people (the cue overload principle at work). This strange fact—that recall can sometimes lead to better performance than recognition—can be explained by the encoding specificity principle. As a cue, *George Bernard* _____ matches the way the famous writer is stored in memory better than does his surname, Shaw, does (even though it is the target). Further, the match is quite distinctive with *George Bernard* _____, but _____ the cue *William* _____ is much more overloaded (Prince William, William Yeats, William Faulkner, will.i.am).

The phenomenon we have been describing is called the *recognition failure of recallable words*, which highlights the point that a cue will be most effective depending on how the information has been encoded ([Tulving & Thomson, 1973](#)). The point is, the cues that work best to evoke retrieval are those that recreate the event or name to be remembered, whereas sometimes even the target itself, such as *Shaw* in the above example, is not the best cue. Which cue will be most effective depends on how the information has been encoded.

Whenever we think about our past, we engage in the act of retrieval. We usually think that retrieval is an objective act because we tend to imagine that retrieving a memory

is like pulling a book from a shelf, and after we are done with it, we return the book to the shelf just as it was. However, research shows this assumption to be false; far from being a static repository of data, the memory is constantly changing. In fact, every time we retrieve a memory, it is altered. For example, the act of retrieval itself (of a fact, concept, or event) makes the retrieved memory much more likely to be retrieved again, a phenomenon called the *testing effect* or the *retrieval practice effect* ([Pyc & Rawson, 2009](#); [Roediger & Karpicke, 2006](#)). However, retrieving some information can actually cause us to forget other information related to it, a phenomenon called *retrieval-induced forgetting* ([Anderson, Bjork, & Bjork, 1994](#)). Thus the act of retrieval can be a double-edged sword—strengthening the memory just retrieved (usually by a large amount) but harming related information (though this effect is often relatively small).

As discussed earlier, retrieval of distant memories is reconstructive. We weave the concrete bits and pieces of events in with assumptions and preferences to form a coherent story ([Bartlett, 1932](#)). For example, if during your 10th birthday, your dog got to your cake before you did, you would likely tell that story for years afterward. Say, then, in later years you misremember where the dog actually found the cake, but repeat that error over and over during subsequent retellings of the story. Over time, that inaccuracy would become a basic fact of the event in your mind. Just as retrieval practice (repetition) enhances accurate memories, so will it strengthen errors or false memories ([McDermott, 2006](#)). Sometimes memories can even be manufactured just from hearing a vivid story. Consider the following episode, recounted by Jean Piaget,

the famous developmental psychologist, from his childhood:

One of my first memories would date, if it were true, from my second year. I can still see, most clearly, the following scene, in which I believed until I was about 15. I was sitting in my pram . . . when a man tried to kidnap me. I was held in by the strap fastened round me while my nurse bravely tried to stand between me and the thief. She received various scratches, and I can still vaguely see those on her face. . . . When I was about 15, my parents received a letter from my former nurse saying that she had been converted to the Salvation Army. She wanted to confess her past faults, and in particular to return the watch she had been given as a reward on this occasion. She had made up the whole story, faking the scratches. I therefore must have heard, as a child, this story, which my parents believed, and projected it into the past in the form of a visual memory. . . . Many real memories are doubtless of the same order. ([Norman & Schacter, 1997](#), pp. 187–188)

Piaget's vivid account represents a case of a pure reconstructive memory. He heard the tale told repeatedly, and doubtless told it (and thought about it) himself. The repeated telling cemented the events as though they had really happened, just as we are all open to the possibility of having "many real memories . . . of the same order." The fact that one can remember precise details (the location, the scratches) does not necessarily indicate that the memory is true, a point that has been confirmed in laboratory studies, too (e.g., [Norman & Schacter, 1997](#)).

PUTTING IT ALL TOGETHER: IMPROVING YOUR MEMORY

A central theme of this module has been the importance of the encoding and retrieval processes, and their

interaction. To recap: to improve learning and memory, we need to encode information in conjunction with excellent cues that will bring back the remembered events when we need them. But how do we do this? Keep in mind the two critical principles we have discussed: to maximize retrieval, we should construct *meaningful* cues that remind us of the original experience, and those cues should be *distinctive* and *not associated with other memories*. These two conditions are critical in maximizing cue effectiveness ([Nairne, 2002](#)).

So, how can these principles be adapted for use in many situations? Let's go back to how we started the module, with Simon Reinhard's ability to memorize huge numbers of digits. Although it was not obvious, he applied these same general memory principles, but in a more deliberate way. In fact, all **mnemonic devices**, or memory aids/tricks, rely on these fundamental principles. In a typical case, the person learns a set of cues and then applies these cues to learn and remember information. Consider the set of 20 items below that are easy to learn and remember ([Bower & Reitman, 1972](#)).

1. is a gun. 11 is penny-one, hot dog bun.
2. is a shoe. 12 is penny-two, airplane glue.
3. is a tree. 13 is penny-three, bumble bee.
4. is a door. 14 is penny-four, grocery store.
5. is knives. 15 is penny-five, big beehive.
6. is sticks. 16 is penny-six, magic tricks.
7. is oven. 17 is penny-seven, go to heaven.
8. is plate. 18 is penny-eight, golden gate.

9. is wine. 19 is penny-nine, ball of twine.
10. is hen. 20 is penny-ten, ballpoint pen.

It would probably take you less than 10 minutes to learn this list and practice recalling it several times (remember to use retrieval practice!). If you were to do so, you would have a set of peg words on which you could “hang” memories. In fact, this mnemonic device is called the *peg word technique*. If you then needed to remember some discrete items—say a grocery list, or points you wanted to make in a speech—this method would let you do so in a very precise yet flexible way. Suppose you had to remember bread, peanut butter, bananas, lettuce, and so on. The way to use the method is to form a vivid image of what you want to remember and imagine it interacting with your peg words (as many as you need). For example, for these items, you might imagine a large gun (the first peg word) shooting a loaf of bread, then a jar of peanut butter inside a shoe, then large bunches of bananas hanging from a tree, then a door slamming on a head of lettuce with leaves flying everywhere. The idea is to provide good, distinctive cues (the weirder the better!) for the information you need to remember while you are learning it. If you do this, then retrieving it later is relatively easy. You know your cues perfectly (one is gun, etc.), so you simply go through your cue word list and “look” in your mind’s eye at the image stored there (bread, in this case).



On Old Olympus' Towering Top, A
Finn And German Viewed Some Hip

*Example of a mnemonic system
 created by a student to study cranial
 nerves. [Image: Kelidimari,
<https://goo.gl/kiA1kP>, CC BY-SA 3.0,
<https://goo.gl/SckRfm>]*

This peg word method may sound strange at first, but it works quite well, even with little training (Roediger, 1980). One word of warning, though, is that the items to be remembered need to be presented relatively slowly at first, until you have practice associating each with its cue word. People get faster with time. Another interesting aspect of this technique is that it's just as easy to recall the items in

backwards order as forwards. This is because the peg words provide direct access to the memorized items, regardless of order.

How did Simon Reinhard remember those digits? Essentially he has a much more complex system based on these same principles. In his case, he uses "memory palaces" (elaborate scenes with discrete places) combined with huge sets of images for digits. For example, imagine mentally walking through the home where you grew up and identifying as many distinct areas and objects as possible. Simon has hundreds of such memory palaces that he uses. Next, for remembering digits, he has memorized a set of 10,000 images. Every four-digit number for him immediately brings forth a mental image. So, for example, 6187 might recall Michael Jackson. When Simon hears all the numbers coming at him, he

places an image for every four digits into locations in his memory palace. He can do this at an incredibly rapid rate, faster than 4 digits per 4 seconds when they are flashed visually, as in the demonstration at the beginning of the module. As noted, his record is 240 digits, recalled in exact order. Simon also holds the world record in an event called “speed cards,” which involves memorizing the precise order of a shuffled deck of cards. Simon was able to do this in 21.19 seconds! Again, he uses his memory palaces, and he encodes groups of cards as single images.

Many books exist on how to improve memory using mnemonic devices, but all involve forming distinctive encoding operations and then having an infallible set of memory cues. We should add that to develop and use these memory systems beyond the basic peg system outlined above takes a great amount of time and concentration. The World Memory Championships are held every year and the records keep improving. However, for most common purposes, just keep in mind that to remember well you need to encode information in a distinctive way and to have good cues for retrieval. You can adapt a system that will meet most any purpose.

OUTSIDE RESOURCES

Book: Brown, P.C., Roediger, H. L. & McDaniel, M. A. (2014). *Make it stick: The science of successful learning*. Cambridge, MA: Harvard University Press.
<https://www.amazon.com/Make-Stick-Science-Successful-Learning/dp/0674729013>

Student Video 1: Eureka Foong's – The Misinformation Effect. This is a student-made video

illustrating this phenomenon of altered memory. It was one of the winning entries in the 2014 Noba Student Video Award.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=244#oembed-1>

Student Video 2: Ang Rui Xia & Ong Jun Hao's – The Misinformation Effect. Another student-made video exploring the misinformation effect. Also an award winner from 2014.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=244#oembed-2>

Video: Flashbulb Memories



One or more interactive elements has

— been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=244#oembed-3>

Video: Simon Reinhard breaking the world record in speedcards.



One or more interactive elements has been excluded from this version of the text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=244#oembed-4>

Web: Retrieval Practice, a website with research, resources, and tips for both educators and learners around the memory-strengthening skill of retrieval practice.

<http://www.retrievalpractice.org/>

DISCUSSION QUESTIONS

1. Mnemonists like Simon Reinhard develop mental “journeys,” which enable them to use the method of loci. Develop your own journey, which contains 20 places, in order, that you know well. One example might be: the front walkway to your parents’ apartment; their doorbell; the couch in

their living room; etc. Be sure to use a set of places that you know well and that have a natural order to them (e.g., the walkway comes before the doorbell). Now you are more than halfway toward being able to memorize a set of 20 nouns, in order, rather quickly. As an optional second step, have a friend make a list of 20 such nouns and read them to you, slowly (e.g., one every 5 seconds). Use the method to attempt to remember the 20 items.

2. Recall a recent argument or misunderstanding you have had about memory (e.g., a debate over whether your girlfriend/boyfriend had agreed to something). In light of what you have just learned about memory, how do you think about it? Is it possible that the disagreement can be understood by one of you making a pragmatic inference?
3. Think about what you've learned in this module and about how you study for tests. On the basis of what you have learned, is there something you want to try that might help your study habits?

VOCABULARY

Autobiographical memory

Memory for the events of one's life.

Consolidation

The process occurring after encoding that is believed to stabilize memory traces.

Cue overload principle

The principle stating that the more memories that

are associated to a particular retrieval cue, the less effective the cue will be in prompting retrieval of any one memory.

Distinctiveness

The principle that unusual events (in a context of similar events) will be recalled and recognized better than uniform (nondistinctive) events.

Encoding

The initial experience of perceiving and learning events.

Encoding specificity principle

The hypothesis that a retrieval cue will be effective to the extent that information encoded from the cue overlaps or matches information in the engram or memory trace.

Engrams

A term indicating the change in the nervous system representing an event; also, memory trace.

Episodic memory

Memory for events in a particular time and place.

Flashbulb memory

Vivid personal memories of receiving the news of some momentous (and usually emotional) event.

Memory traces

A term indicating the change in the nervous system representing an event.

Misinformation effect

When erroneous information occurring after an event is remembered as having been part of the original event.

Mnemonic devices

A strategy for remembering large amounts of information, usually involving imaging events occurring on a journey or with some other set of memorized cues.

Recoding

The ubiquitous process during learning of taking information in one form and converting it to another form, usually one more easily remembered.

Retrieval

The process of accessing stored information.

Retroactive interference

The phenomenon whereby events that occur after some particular event of interest will usually cause forgetting of the original event.

Semantic memory

The more or less permanent store of knowledge that people have.

Storage

The stage in the learning/memory process that bridges encoding and retrieval; the persistence of memory over time.

REFERENCES

- Anderson, M. C., Bjork, R., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology-Learning Memory and Cognition*, 20, 1063–1087.
- Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Bower, G. H., & Reitman, J. S. (1972). Mnemonic elaboration in multilist learning. *Journal of Verbal Learning and Verbal Behavior*, 11, 478–485.
- Brewer, W. F. (1977). Memory for the pragmatic implications of sentences. *Memory & Cognition*, 5(6), 673–678.
- Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, 5, 73–99.
- Chan, J.C.K. & McDermott, K.B. (2006). Remembering pragmatic inferences. *Applied Cognitive Psychology*, 20, 633–639.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671–684.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58, 17.
- Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural

environments: On land and underwater. *British Journal of Psychology*, 66 (3), 325-331

- Hunt, R. (2003). Two contributions of distinctive processing to accurate memory. *Journal of Memory and Language*, 48, 811-825.
- Hunt, R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, 32, 421-445.
- Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory*, 12, 361-366.
- McDermott, K. B. (2006). Paradoxical effects of testing: Repeated retrieval attempts enhance the likelihood of later accurate and false recall. *Memory & Cognition*, 34, 261-267.
- McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Review*, 39(4), 352.
- Melton, A. W. (1963). Implications of short-term memory for a general theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 2, 1-21.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Muter, P. (1984). Recognition and recall of words with a single meaning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 198-202.
- Nairne, J. S. (2002). The myth of the encoding-

retrieval match. *Memory*, 10, 389–395.

- Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition*, 25, 838–848.
- Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language*, 60, 437–447.
- Roediger, H. L. (1980). The effectiveness of four mnemonics in ordering recall. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 558.
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249–255.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology-Learning Memory and Cognition*, 21, 803–814.
- Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, 27, 494–500.
- Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science*, 14, 455–461.
- Tulving, E. (2007). Are there 256 different kinds of

memory? In J.S. Nairne (Ed.), *The foundations of remembering: Essays in honor of Henry L. Roediger, III* (pp. 39–52). New York: Psychology Press.

- Tulving, E. (1991). Interview. *Journal of Cognitive Neuroscience*, 3, 89–94
- Tulving, E., & Bower, G. H. (1975). The logic of memory representations. *The psychology of learning and motivation*, 8, 265–301.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381–391.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352–373.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80, 352–373.
- Watkins, M. J. (1975). Inhibition in recall with extralist “cues.” *Journal of Verbal Learning and Verbal Behavior*, 14, 294–303.
- Watkins, M. J., & Tulving, E. (1975). Episodic memory: When recognition fails. *Journal of Experimental Psychology: General*, 104, 5–29.

AUTHORS



Kathleen B. McDermott

Kathleen B. McDermott is Professor of Psychology and Radiology at Washington University in St. Louis. She studies remembering using both behavioral and neuroimaging techniques. She received the Shahin Hashtroudi Memorial Prize for Researcher in Memory from the Association for Psychological Science and the James S. McGuigan Young Investigator Prize from the American Psychological Foundation. She is a Fellow of the Association for Psychological Science.



Henry L. Roediger III

Henry L. Roediger, III is the James S. McDonnell Distinguished University Professor at Washington University in St. Louis who has spent his career studying learning and memory. He has received the Howard Warren Crosby Medal from the Society of Experimental Psychologists and the William James Award for Lifetime Achievements in Psychology from the Association of Psychological Science. He also served as President of APS.

CREATIVE COMMONS LICENSE



MEMORY (ENCODING, STORAGE, RETRIEVAL) by [Kathleen B. McDermott and Henry L. Roediger III](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International](#)

[License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

CHAPTER 18.

CONDITIONING AND LEARNING

MARK E. BOUTON

Basic principles of learning are always operating and always influencing human behavior. This module discusses the two most fundamental forms of learning — classical (Pavlovian) and instrumental (operant) conditioning. Through them, we respectively learn to associate 1) stimuli in the environment, or 2) our own behaviors, with significant events, such as rewards and punishments. The two types of learning have been intensively studied because they have powerful effects on behavior, and because they provide methods that allow scientists to analyze learning processes rigorously. This module describes some of the most important things you need to know about classical and instrumental conditioning, and it illustrates some of the many ways they help us understand normal and disordered behavior in humans. The module concludes by introducing the concept of observational learning, which is a form of learning that is largely distinct from classical and operant conditioning.

LEARNING OBJECTIVES

- Distinguish between classical (Pavlovian) conditioning and instrumental (operant) conditioning.
- Understand some important facts about each that tell us how they work.
- Understand how they work separately and together to influence human behavior in the world outside the laboratory.
- Students will be able to list the four aspects of observational learning according to Social Learning Theory.

TWO TYPES OF CONDITIONING

Although Ivan Pavlov won a Nobel Prize for studying digestion, he is much more famous for something else: working with a dog, a bell, and a bowl of saliva. Many people are familiar with the classic study of “Pavlov’s dog,” but rarely do they understand the significance of its discovery. In fact, Pavlov’s work helps explain why some people get anxious just looking at a crowded bus, why the sound of a morning alarm is so hated, and even why we swear off certain foods we’ve only tried once. Classical (or Pavlovian) conditioning is one of the fundamental ways we learn about the world around us. But it is far more than just a theory of learning; it is also arguably a theory of identity. For, once you understand classical conditioning, you’ll recognize that your favorite music, clothes, even political candidate, might all be a result of the same process that makes a dog drool at the sound of bell.



The Pavlov in All of Us: Does your dog learn to beg for food because you reinforce her by feeding her from the table? [Image: David Mease, <https://goo.gl/R9cQV7>, CC BY-NC 2.0, <https://goo.gl/Filc2e>]

Around the turn of the 20th century, scientists who were interested in understanding the behavior of animals and humans began to appreciate the importance of two very basic forms of learning. One, which was first studied by the Russian physiologist Ivan Pavlov, is known as **classical**, or **Pavlovian conditioning**. In his famous experiment, Pavlov rang a bell and

then gave a dog some food. After repeating this pairing multiple times, the dog eventually treated the bell as a signal for food, and began salivating in anticipation of the treat. This kind of result has been reproduced in the lab using a wide range of signals (e.g., tones, light, tastes, settings) paired with many different events besides food (e.g., drugs, shocks, illness; see below).

We now believe that this same learning process is engaged, for example, when humans associate a drug they've taken with the environment in which they've taken it; when they associate a stimulus (e.g., a symbol for vacation, like a big beach towel) with an emotional event (like a burst of happiness); and when they associate the flavor of a food with getting food poisoning. Although classical conditioning may seem "old" or "too simple" a theory, it is still widely studied today for at least two reasons: First, it is a straightforward test of associative

learning that can be used to study other, more complex behaviors. Second, because classical conditioning is always occurring in our lives, its effects on behavior have important implications for understanding normal and disordered behavior in humans.

In a general way, classical conditioning occurs whenever neutral stimuli are associated with psychologically significant events. With food poisoning, for example, although having fish for dinner may not normally be something to be concerned about (i.e., a “neutral stimuli”), if it causes you to get sick, you will now likely associate that neutral stimuli (the fish) with the psychologically significant event of getting sick. These paired events are often described using terms that can be applied to any situation.

The dog food in Pavlov’s experiment is called the **unconditioned stimulus (US)** because it elicits an **unconditioned response (UR)**. That is, without any kind of “training” or “teaching,” the stimulus produces a natural or instinctual reaction. In Pavlov’s case, the food (US) automatically makes the dog drool (UR). Other examples of unconditioned stimuli include loud noises (US) that startle us (UR), or a hot shower (US) that produces pleasure (UR).

On the other hand, a conditioned stimulus produces a conditioned response. A **conditioned stimulus (CS)** is a signal that has no importance to the organism until it is paired with something that does have importance. For example, in Pavlov’s experiment, the bell is the conditioned stimulus. Before the dog has learned to associate the bell (CS) with the presence of food (US), hearing the bell means nothing to the dog. However, after multiple pairings of the bell with the presentation of food,

the dog starts to drool at the sound of the bell. This drooling in response to the bell is the **conditioned response (CR)**. Although it can be confusing, the conditioned response is almost always the same as the unconditioned response. However, it is called the conditioned response because it is conditional on (or, depends on) being paired with the conditioned stimulus (e.g., the bell). To help make this clearer, consider becoming really hungry when you see the logo for a fast food restaurant. There's a good chance you'll start salivating. Although it is the actual eating of the food (US) that normally produces the salivation (UR), simply seeing the restaurant's logo (CS) can trigger the same reaction (CR).

Another example you are probably very familiar with involves your alarm clock. If you're like most people, waking up early usually makes you unhappy. In this case, waking up early (US) produces a natural sensation of grumpiness (UR). Rather than waking up early on your own, though, you likely have an alarm clock that plays a tone to wake you. Before setting your alarm to that particular tone, let's imagine you had neutral feelings about it (i.e., the tone had no prior meaning for you). However, now that you use it to wake up every morning, you psychologically "pair" that tone (CS) with your feelings of grumpiness in the morning (UR). After enough pairings, this tone (CS) will automatically produce your natural response of grumpiness (CR). Thus, this linkage between the unconditioned stimulus (US; waking up early) and the conditioned stimulus (CS; the tone) is so strong that the unconditioned response (UR; being grumpy) will become a conditioned response (CR; e.g., hearing the tone at any point in the day—whether waking

up or walking down the street—will make you grumpy). Modern studies of classical conditioning use a very wide range of CSs and USs and measure a wide range of conditioned responses.

Although classical conditioning is a powerful explanation for how we learn many different things, there is a second form of conditioning that also helps explain how we learn. First studied by Edward Thorndike, and later extended by B. F. Skinner, this second type of conditioning is known



Receiving a reward can condition you toward certain behaviors. For example, when you were a child, your mother may have offered you this deal: “Don’t make a fuss when we’re in the supermarket and you’ll get a treat on the way out.” [Image: Oliver Hammond, <https://goo.gl/xFKiZL>, CC BY-NC-SA 2.0, <https://goo.gl/Toc0ZF>]

as **instrumental** or **operant conditioning**. Operant conditioning occurs when a *behavior* (as opposed to a stimulus) is associated with the occurrence of a significant event. In the best-known example, a rat in a laboratory learns to press a lever in a cage (called a “Skinner box”) to receive food. Because the rat has no “natural” association between pressing a lever and getting food, the rat has to learn this connection. At first, the rat may simply explore its cage, climbing on top of things, burrowing under things, in search of food. Eventually while poking around its cage, the rat accidentally presses

the lever, and a food pellet drops in. This voluntary behavior is called an **operant** behavior, because it “operates” on the environment (i.e., it is an action that the animal itself makes).

Now, once the rat recognizes that it receives a piece of food every time it presses the lever, the behavior of lever-pressing becomes reinforced. That is, the food pellets serve as **reinforcers** because they strengthen the rat’s desire to engage with the environment in this particular manner. In a parallel example, imagine that you’re playing a street-racing video game. As you drive through one city course multiple times, you try a number of different streets to get to the finish line. On one of these trials, you discover a shortcut that dramatically improves your overall time. You have learned this new path through operant conditioning. That is, by engaging with your environment (operant responses), you performed a sequence of behaviors that that was positively reinforced (i.e., you found the shortest distance to the finish line). And now that you’ve learned how to drive this course, you will perform that same sequence of driving behaviors (just as the rat presses on the lever) to receive your reward of a faster finish.

Operant conditioning research studies how the effects of a behavior influence the probability that it will occur again. For example, the effects of the rat’s lever-pressing behavior (i.e., receiving a food pellet) influences the probability that it will keep pressing the lever. For, according to Thorndike’s **law of effect**, when a behavior has a positive (satisfying) effect or consequence, it is likely to be repeated in the future. However, when a behavior has a negative (painful/annoying) consequence, it is less likely to be repeated in the future. Effects that increase

behaviors are referred to as reinforcers, and effects that decrease them are referred to as **punishers**.

An everyday example that helps to illustrate operant conditioning is striving for a good grade in class—which could be considered a reward for students (i.e., it produces a positive emotional response). In order to get that reward (similar to the rat learning to press the lever), the student needs to modify his/her behavior. For example, the student may learn that speaking up in class gets him/her participation points (a reinforcer), so the student speaks up repeatedly. However, the student also learns that s/he shouldn't speak up about just anything; talking about topics unrelated to school actually costs points. Therefore, through the student's freely chosen behaviors, s/he learns which behaviors are reinforced and which are punished.

An important distinction of operant conditioning is that it provides a method for studying how consequences influence “voluntary” behavior. The rat's decision to press the lever is voluntary, in the sense that the rat is free to make and repeat that response whenever it wants. Classical conditioning, on the other hand, is just the opposite—depending instead on “involuntary” behavior (e.g., the dog doesn't choose to drool; it just does). So, whereas the rat must actively participate and perform some kind of behavior to attain its reward, the dog in Pavlov's experiment is a passive participant. One of the lessons of operant conditioning research, then, is that voluntary behavior is strongly influenced by its consequences.



Classical or Pavlovian
Conditioning



Instrumental or Operant
Conditioning

[Image courtesy of Bernard W. Balleine]

The illustration above summarizes the basic elements of classical and instrumental conditioning. The two types of learning differ in many ways. However, modern thinkers often emphasize the fact that they differ—as illustrated here—in *what* is learned. In classical conditioning, the animal behaves as if it has learned to associate a *stimulus* with a significant event. In operant conditioning, the animal behaves as if it has learned to associate a *behavior* with a significant event. Another difference is that the response in the classical situation (e.g., salivation) is *elicited* by a stimulus that comes before it, whereas the response in the operant case is not elicited by any particular stimulus. Instead, operant responses are said to be *emitted*. The word “emitted” further conveys the idea that operant behaviors are essentially voluntary in nature.

Understanding classical and operant conditioning provides psychologists with many tools for understanding learning and behavior in the world outside the lab. This is in part because the two types of learning occur continuously throughout our lives. It has been said that “much like the laws of gravity, the laws of learning are always in effect” ([Spreat & Spreat, 1982](#)).

USEFUL THINGS TO KNOW ABOUT CLASSICAL CONDITIONING

CLASSICAL CONDITIONING HAS MANY EFFECTS ON BEHAVIOR

A classical CS (e.g., the bell) does not merely elicit a simple, unitary reflex. Pavlov emphasized salivation because that was the only response he measured. But his bell almost certainly elicited a whole *system* of responses that functioned to get the organism ready for the upcoming US (food) (see [Timberlake, 2001](#)). For example, in addition to salivation, CSs (such as the bell) that signal that food is near also elicit the secretion of gastric acid, pancreatic enzymes, and insulin (which gets blood glucose into cells). All of these responses prepare the body for digestion. Additionally, the CS elicits approach behavior and a state of excitement. And presenting a CS for food can also cause animals whose stomachs are full to eat more food if it is available. In fact, food CSs are so prevalent in modern society, humans are likewise inclined to eat or feel hungry in response to cues associated with food, such as the sound of a bag of potato chips opening, the sight of a well-known logo (e.g., Coca-Cola), or the feel of the couch in front of the television.

Classical conditioning is also involved in other aspects of eating. Flavors associated with certain nutrients (such as sugar or fat) can become preferred without arousing any awareness of the pairing. For example, protein is a US that your body automatically craves more of once you start to consume it (UR): since proteins are highly concentrated in meat, the flavor of meat becomes a CS (or cue, that proteins are on the way), which perpetuates the

cycle of craving for yet more meat (this automatic bodily reaction now a CR).

In a similar way, flavors associated with stomach pain or illness become avoided and *disliked*. For example, a person who gets sick after drinking too much tequila may acquire a profound dislike of the taste and odor of tequila—a phenomenon called **taste aversion conditioning**. The fact that flavors are often associated with so many consequences of eating is important for animals (including rats and humans) that are frequently exposed to new foods. And it is clinically relevant. For example, drugs used in chemotherapy often make cancer patients sick. As a consequence, patients often acquire aversions to foods eaten just before treatment, or even aversions to such things as the waiting room of the chemotherapy clinic itself (see [Bernstein, 1991](#); [Scalera & Bavieri, 2009](#)).

Classical conditioning occurs with a variety of significant events. If an experimenter sounds a tone just before applying a mild shock to a rat's feet, the tone will elicit fear or anxiety after one or two pairings. Similar **fear conditioning** plays a role in creating many anxiety disorders in humans, such as phobias and panic disorders, where people associate cues (such as closed spaces, or a shopping mall) with panic or other emotional trauma (see [Mineka & Zinbarg, 2006](#)). Here, rather than a physical response (like drooling), the CS triggers an emotion.

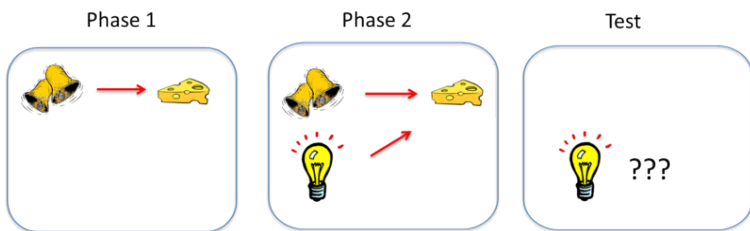
Another interesting effect of classical conditioning can occur when we ingest drugs. That is, when a drug is taken, it can be associated with the cues that are present at the same time (e.g., rooms, odors, drug paraphernalia). In this regard, if someone associates a particular smell with the

sensation induced by the drug, whenever that person smells the same odor afterward, it may cue responses (physical and/or emotional) related to taking the drug itself. But drug cues have an even more interesting property: They elicit responses that often “compensate” for the upcoming effect of the drug (see [Siegel, 1989](#)). For example, morphine itself suppresses pain; however, if someone is used to taking morphine, a cue that signals the “drug is coming soon” can actually make the person more sensitive to pain. Because the person knows a pain suppressant will soon be administered, the body becomes more sensitive, anticipating that “the drug will soon take care of it.” Remarkably, such **conditioned compensatory responses** in turn decrease the impact of the drug on the body—because the body has become more sensitive to pain.

This conditioned compensatory response has many implications. For instance, a drug user will be most “tolerant” to the drug in the presence of cues that have been associated with it (because such cues elicit compensatory responses). As a result, overdose is usually not due to an increase in dosage, but to taking the drug in a new place without the familiar cues—which would have otherwise allowed the user to tolerate the drug (see [Siegel, Hinson, Krank, & McCully, 1982](#)). Conditioned compensatory responses (which include heightened pain sensitivity and decreased body temperature, among others) might also cause discomfort, thus motivating the drug user to continue usage of the drug to reduce them. This is one of several ways classical conditioning might be a factor in drug addiction and dependence.

A final effect of classical cues is that they motivate ongoing operant behavior (see [Balleine, 2005](#)). For

example, if a rat has learned via operant conditioning that pressing a lever will give it a drug, in the presence of cues that signal the “drug is coming soon” (like the sound of the lever squeaking), the rat will work harder to press the lever than if those cues weren’t present (i.e., there is no squeaking lever sound). Similarly, in the presence of food-associated cues (e.g., smells), a rat (or an overeater) will work harder for food. And finally, even in the presence of negative cues (like something that signals fear), a rat, a human, or any other organism will work harder to avoid those situations that might lead to trauma. Classical CSs thus have many effects that can contribute to significant behavioral phenomena.



[Image courtesy of Bernard W. Balleine]

THE LEARNING PROCESS

As mentioned earlier, classical conditioning provides a method for studying basic learning processes. Somewhat counterintuitively, though, studies show that pairing a CS and a US together is not sufficient for an association to be learned between them. Consider an effect called **blocking** (see [Kamin, 1969](#)). In this effect, an animal first learns to associate one CS—call it stimulus A—with a US. In the illustration above, the sound of a bell (stimulus A) is paired with the presentation of food. Once

this association is learned, in a second phase, a second stimulus—stimulus B—is presented alongside stimulus A, such that the two stimuli are paired with the US together. In the illustration, a light is added and turned on at the same time the bell is rung. However, because the animal has already learned the association between stimulus A (the bell) and the food, the animal doesn't learn an association between stimulus B (the light) and the food. That is, the conditioned response only occurs during the presentation of stimulus A, because the earlier conditioning of A “blocks” the conditioning of B when B is added to A. The reason? Stimulus A already predicts the US, so the US is not surprising when it occurs with Stimulus B.

Learning depends on such a surprise, or a discrepancy between what occurs on a conditioning trial and what is already predicted by cues that are present on the trial. To learn something through classical conditioning, there must first be some **prediction error**, or the chance that a conditioned stimulus won't lead to the expected outcome. With the example of the bell and the light, because the bell always leads to the reward of food, there's no “prediction error” that the addition of the light helps to correct. However, if the researcher suddenly requires that the bell and the light both occur in order to receive the food, the bell alone will produce a prediction error that the animal has to learn.

Blocking and other related effects indicate that the learning process tends to take in the most valid predictors of significant events and ignore the less useful ones. This is common in the real world. For example, imagine that your supermarket puts big star-shaped stickers on products that are on sale. Quickly, you learn that items

with the big star-shaped stickers are cheaper. However, imagine you go into a similar supermarket that not only uses these stickers, but also uses bright orange price tags to denote a discount. Because of blocking (i.e., you already know that the star-shaped stickers indicate a discount), you don't have to learn the color system, too. The star-shaped stickers tell you everything you need to know (i.e. there's no prediction error for the discount), and thus the color system is irrelevant.

Classical conditioning is strongest if the CS and US are intense or salient. It is also best if the CS and US are relatively new and the organism hasn't been frequently exposed to them before. And it is especially strong if the organism's biology has prepared it to associate a particular CS and US. For example, rats and humans are naturally inclined to associate an illness with a flavor, rather than with a light or tone. Because foods are most commonly experienced by taste, if there is a particular food that makes us ill, associating the flavor (rather than the appearance—which may be similar to other foods) with the illness will more greatly ensure we avoid that food in the future, and thus avoid getting sick. This sorting tendency, which is set up by evolution, is called **preparedness**.

There are many factors that affect the strength of classical conditioning, and these have been the subject of much research and theory (see [Rescorla & Wagner, 1972](#); [Pearce & Bouton, 2001](#)). Behavioral neuroscientists have also used classical conditioning to investigate many of the basic brain processes that are involved in learning (see [Fanselow & Poulos, 2005](#); [Thompson & Steinmetz, 2009](#)).

ERASING CLASSICAL LEARNING

After conditioning, the response to the CS can be eliminated if the CS is presented repeatedly without the US. This effect is called **extinction**, and the response is said to become “extinguished.” For example, if Pavlov kept ringing the bell but never gave the dog any food afterward, eventually the dog’s CR (drooling) would no longer happen when it heard the CS (the bell), because the bell would no longer be a predictor of food. Extinction is important for many reasons. For one thing, it is the basis for many therapies that clinical psychologists use to eliminate maladaptive and unwanted behaviors. Take the example of a person who has a debilitating fear of spiders: one approach might include systematic exposure to spiders. Whereas, initially the person has a CR (e.g., extreme fear) every time s/he sees the CS (e.g., the spider), after repeatedly being shown pictures of spiders in neutral conditions, pretty soon the CS no longer predicts the CR (i.e., the person doesn’t have the fear reaction when seeing spiders, having learned that spiders no longer serve as a “cue” for that fear). Here, repeated exposure to spiders without an aversive consequence causes extinction.

Psychologists must accept one important fact about extinction, however: it does not necessarily destroy the original learning (see [Bouton, 2004](#)). For example, imagine you strongly associate the smell of chalkboards with the agony of middle school detention. Now imagine that, after years of encountering chalkboards, the smell of them no longer recalls the agony of detention (an example of extinction). However, one day, after entering a new building for the first time, you suddenly catch a whiff of a

chalkboard and WHAM!, the agony of detention returns. This is called **spontaneous recovery**: following a lapse in exposure to the CS after extinction has occurred, sometimes re-exposure to the CS (e.g., the smell of chalkboards) can evoke the CR again (e.g., the agony of detention).

Another related phenomenon is the **renewal effect**: After extinction, if the CS is tested in a new **context**, such as a different room or location, the CR can also return. In the chalkboard example, the action of entering a new building—where you don’t expect to smell chalkboards—suddenly renews the sensations associated with detention. These effects have been interpreted to suggest that extinction *inhibits* rather than erases the learned behavior, and this inhibition is mainly expressed in the context in which it is learned (see “context” in the Key Vocabulary section below).

This does not mean that extinction is a bad treatment for behavior disorders. Instead, clinicians can increase its effectiveness by using basic research on learning to help defeat these relapse effects (see [Craske et al., 2008](#)). For example, conducting extinction therapies in contexts where patients might be most vulnerable to relapsing (e.g., at work), might be a good strategy for enhancing the therapy’s success.

USEFUL THINGS TO KNOW ABOUT INSTRUMENTAL CONDITIONING

Most of the things that affect the strength of classical conditioning also affect the strength of instrumental learning—whereby we learn to associate our actions with their outcomes. As noted earlier, the “bigger” the

reinforcer (or punisher), the stronger the learning. And, if an instrumental behavior is no longer reinforced, it will also be extinguished. Most of the rules of associative learning that apply to classical conditioning also apply to instrumental learning, but other facts about instrumental learning are also worth knowing.

INSTRUMENTAL RESPONSES COME UNDER STIMULUS CONTROL

As you know, the classic operant response in the laboratory is lever-pressing in rats, reinforced by food. However, things can be arranged so that lever-pressing only produces pellets when a particular stimulus is present. For example, lever-pressing can be reinforced only when a light in the Skinner box is turned on; when the light is off, no food is released from lever-pressing. The rat soon learns to discriminate between the light-on and light-off conditions, and presses the lever only in the presence of the light (responses in light-off are extinguished). In everyday life, think about waiting in the turn lane at a traffic light. Although you know that green means go, only when you have the green *arrow* do you turn. In this regard, the operant behavior is now said to be under **stimulus control**. And, as is the case with the traffic light, in the real world, stimulus control is probably the rule.

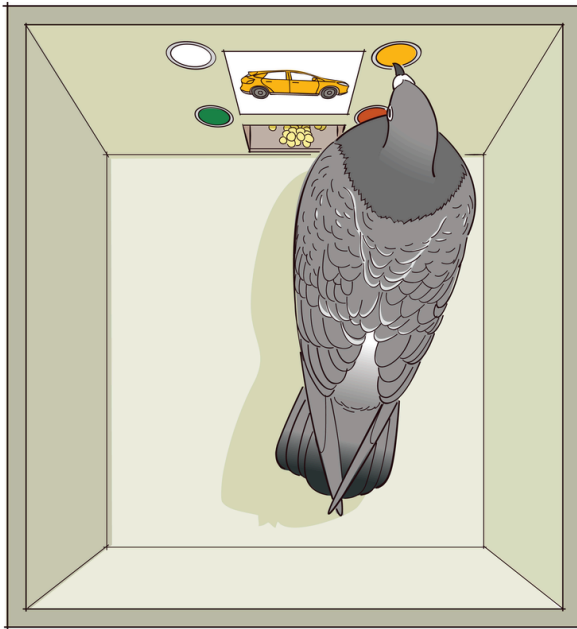
The stimulus controlling the operant response is called a **discriminative stimulus**. It can be associated directly with the response, or the reinforcer (see below). However, it usually does not elicit the response the way a classical CS does. Instead, it is said to “set the occasion for” the operant response. For example, a canvas put in front of an

artist does not elicit painting behavior or compel her to paint. It allows, or sets the occasion for, painting to occur.

Stimulus-control techniques are widely used in the laboratory to study perception and other psychological processes in animals. For example, the rat would not be able to respond appropriately to light-on and light-off conditions if it could not see the light. Following this logic, experiments using stimulus-control methods have tested how well animals see colors, hear ultrasounds, and detect magnetic fields. That is, researchers pair these discriminative stimuli with those they know the animals already understand (such as pressing the lever). In this way, the researchers can test if the animals can learn to press the lever only when an ultrasound is played, for example.

These methods can also be used to study “higher” cognitive processes. For example, pigeons can learn to peck at different buttons in a Skinner box when pictures of flowers, cars, chairs, or people are shown on a miniature TV screen (see [Wasserman, 1995](#)). Pecking button 1 (and no other) is reinforced in the presence of a flower image, button 2 in the presence of a chair image, and so on. Pigeons can learn the discrimination readily, and, under the right conditions, will even peck the correct buttons associated with pictures of *new* flowers, cars, chairs, and people they have never seen before. The birds have learned to **categorize** the sets of stimuli. Stimulus-control methods can be used to study how such categorization is learned.

OPERANT CONDITIONING INVOLVES CHOICE



Another thing to know about operant conditioning is that the response always requires choosing one behavior over others. The student who goes to the bar on Thursday night chooses to drink instead of staying at home and studying. The rat chooses to press the lever instead of sleeping or scratching its ear in the back of the box. The alternative behaviors are each associated with their own reinforcers. And the tendency to perform a particular action depends on both the reinforcers earned for it and the reinforcers earned for its alternatives.

To investigate this idea, choice has been studied in the Skinner box by making two levers available for the rat (or two buttons available for the pigeon), each of which has its own reinforcement or payoff rate. A thorough study of choice in situations like this has led to a rule called

the **quantitative law of effect** (see [Herrnstein, 1970](#)), which can be understood without going into quantitative detail: The law acknowledges the fact that the effects of reinforcing one behavior depend crucially on how much reinforcement is earned for the behavior's alternatives. For example, if a pigeon learns that pecking one light will reward two food pellets, whereas the other light only rewards one, the pigeon will only peck the first light. However, what happens if the first light is more strenuous to reach than the second one? Will the cost of energy outweigh the bonus of food? Or will the extra food be worth the work? In general, a given reinforcer will be less reinforcing if there are many alternative reinforcers in the environment. For this reason, alcohol, sex, or drugs may be less powerful reinforcers if the person's environment is full of other sources of reinforcement, such as achievement at work or love from family members.

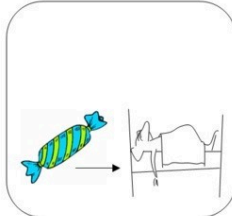
COGNITION IN INSTRUMENTAL LEARNING

Modern research also indicates that reinforcers do more than merely strengthen or “stamp in” the behaviors they are a consequence of, as was Thorndike's original view. Instead, animals learn about the specific consequences of each behavior, and will perform a behavior depending on how much they currently want—or “value”—its consequence.

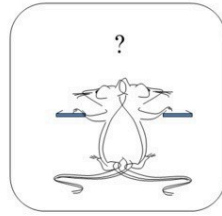
1. Instrumental Learning



2. Taste aversion learning



3. Test



[Image courtesy of Bernard W. Balleine]

This idea is best illustrated by a phenomenon called the **reinforcer devaluation effect** (see [Colwill & Rescorla, 1986](#)). A rat is first trained to perform two instrumental actions (e.g., pressing a lever on the left, and on the right), each paired with a different reinforcer (e.g., a sweet sucrose solution, and a food pellet). At the end of this training, the rat tends to press both levers, alternating between the sucrose solution and the food pellet. In a second phase, one of the reinforcers (e.g., the sucrose) is then separately paired with illness. This conditions a taste aversion to the sucrose. In a final test, the rat is returned to the Skinner box and allowed to press either lever freely. No reinforcers are presented during this test (i.e., no sucrose or food comes from pressing the levers), so behavior during testing can only result from the rat's memory of what it has learned earlier. Importantly here, the rat chooses *not* to perform the response that once produced the reinforcer that it now has an aversion to (e.g., it won't press the sucrose lever). This means that the rat has learned and remembered the reinforcer associated with each response, and can combine that knowledge with the knowledge that the reinforcer is now "bad." Reinforcers do not merely stamp in responses; the animal learns much more than that. The behavior is said to be

“goal-directed” (see [Dickinson & Balleine, 1994](#)), because it is influenced by the current value of its associated goal (i.e., how much the rat wants/doesn't want the reinforcer).

Things can get more complicated, however, if the rat performs the instrumental actions frequently and repeatedly. That is, if the rat has spent many months learning the value of pressing each of the levers, the act of pressing them becomes automatic and routine. And here, this once goal-directed action (i.e., the rat pressing the lever for the goal of getting sucrose/food) can become a **habit**. Thus, if a rat spends many months performing the lever-pressing behavior (turning such behavior into a habit), even when sucrose is again paired with illness, the rat will continue to press that lever (see [Holland, 2004](#)). After all the practice, the instrumental response (pressing the lever) is no longer sensitive to reinforcer devaluation. The rat continues to respond automatically, regardless of the fact that the sucrose from this lever makes it sick.

Habits are very common in human experience, and can be useful. You do not need to relearn each day how to make your coffee in the morning or how to brush your teeth. Instrumental behaviors can eventually become habitual, letting us get the job done while being free to think about other things.

PUTTING CLASSICAL AND INSTRUMENTAL CONDITIONING TOGETHER

Classical and operant conditioning are usually studied separately. But outside of the laboratory they almost always occur at the same time. For example, a person who is reinforced for drinking alcohol or eating excessively

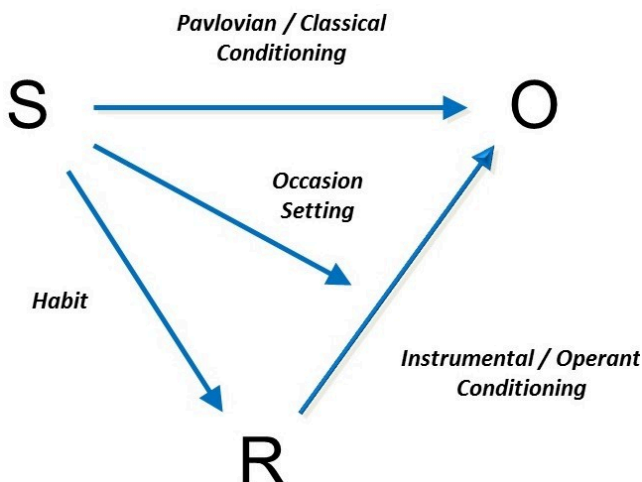
learns these behaviors in the presence of certain stimuli—a pub, a set of friends, a restaurant, or possibly the couch in front of the TV. These stimuli are also available for association with the reinforcer. In this way, classical and operant conditioning are always intertwined.

The figure below summarizes this idea, and helps review what we have discussed in this module. Generally speaking, any reinforced or punished operant response (R) is paired with an outcome (O) in the presence of some stimulus or set of stimuli (S).

The figure illustrates the types of associations that can be learned in this very general scenario. For one thing, the organism will learn to associate the response *and* the outcome (R – O). This is instrumental conditioning. The learning process here is probably similar to classical conditioning, with all its emphasis on surprise and prediction error. And, as we discussed while considering the reinforcer devaluation effect, once R – O is learned, the organism will be ready to perform the response if the outcome is desired or valued. The value of the reinforcer can also be influenced by other reinforcers earned for other behaviors in the situation. These factors are at the heart of instrumental learning.

Second, the organism can also learn to associate the stimulus with the reinforcing outcome (S – O). This is the classical conditioning component, and as we have seen, it can have many consequences on behavior. For one thing, the stimulus will come to evoke a system of responses that help the organism prepare for the reinforcer (not shown in the figure): The drinker may undergo changes in body temperature; the eater may salivate and have an increase in insulin secretion. In addition, the stimulus will

evoke approach (if the outcome is positive) or retreat (if the outcome is negative). Presenting the stimulus will also prompt the instrumental response.



The third association in the diagram is the one between the stimulus and the response (S – R). As discussed earlier, after a lot of practice, the stimulus may begin to elicit the response directly. This is habit learning, whereby the response occurs relatively automatically, without much mental processing of the relation between the action and the outcome and the outcome’s current value.

The final link in the figure is between the stimulus and the response-outcome association [S – (R – O)]. More than just entering into a simple association with the R or the O, the stimulus can signal that the R – O relationship is now in effect. This is what we mean when we say that the stimulus can “set the occasion” for the operant response: It sets the occasion for the response-reinforcer relationship. Through this mechanism, the painter might begin to paint when given the right tools and the opportunity enabled by the canvas. The canvas

theoretically signals that the behavior of painting will now be reinforced by positive consequences.

The figure provides a framework that you can use to understand almost any learned behavior you observe in yourself, your family, or your friends. If you would like to understand it more deeply, consider taking a course on learning in the future, which will give you a fuller appreciation of how classical learning, instrumental learning, habit learning, and occasion setting actually work and interact.

OBSERVATIONAL LEARNING

Not all forms of learning are accounted for entirely by classical and operant conditioning. Imagine a child walking up to a group of children playing a game on the playground. The game looks fun, but it is new and unfamiliar. Rather than joining the game immediately, the child opts to sit back and watch the other children play a round or two. Observing the others, the child takes note of the ways in which they behave while playing the game. By watching the behavior of the other kids, the child can figure out the rules of the game and even some strategies for doing well at the game. This is called **observational learning**.



Children observing a social model (an experienced chess player) to learn the rules and strategies of the game of chess. [Image: David R. Tribble, <https://goo.gl/nWsgxl>, CC BY-SA 3.0, <https://goo.gl/uhHOLA>]

Observational learning is a component of Albert Bandura's **Social**

Learning

Theory ([Bandura, 1977](#)), which posits that individuals can learn novel responses via observation of key others' behaviors. Observational learning does not necessarily require reinforcement, but instead hinges on the presence of others, referred to as **social models**. Social

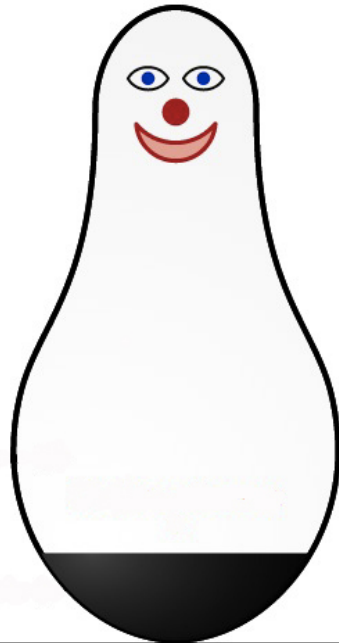
models are typically of higher status or authority compared to the observer, examples of which include parents, teachers, and police officers. In the example above, the children who already know how to play the game could be thought of as being authorities—and are therefore social models—even though they are the same age as the observer. By observing how the social models behave, an individual is able to learn how to act in a certain situation. Other examples of observational learning might include a child learning to place her napkin in her lap by watching her parents at the dinner table, or a customer learning where to find the ketchup and mustard after observing other customers at a hot dog stand.

Bandura theorizes that the observational learning process consists of four parts. The first is *attention*—as, quite simply, one must pay attention to what s/he is

observing in order to learn. The second part is *retention*: to learn one must be able to retain the behavior s/he is observing in memory. The third part of observational learning, *initiation*, acknowledges that the learner must be able to execute (or initiate) the learned behavior. Lastly, the observer must possess the *motivation* to engage in observational learning. In our vignette, the child must want to learn how to play the game in order to properly engage in observational learning.

Researchers have conducted countless experiments designed to explore observational learning, the most famous of which is Albert Bandura's "Bobo doll experiment."

In this experiment ([Bandura, Ross & Ross 1961](#)), Bandura had children individually observe an adult social model interact with a clown doll ("Bobo"). For one group of children, the adult interacted aggressively with Bobo: punching it, kicking it, throwing it, and even hitting it in the face with a toy mallet. Another group of children watched the adult interact with other toys, displaying no aggression toward Bobo. In both instances the adult left and the children were



Bobo [Image: © Sémhur / Wikimedia Commons / CC-BY-SA-3.0 (or Free Art License), <https://goo.gl/uhHOLA>]

allowed to interact with Bobo on their own. Bandura found that children exposed to the aggressive social model were significantly more likely to behave aggressively toward Bobo, hitting and kicking him, compared to those exposed to the non-aggressive model. The researchers concluded that the children in the aggressive group used their observations of the adult social model's behavior to determine that aggressive behavior toward Bobo was acceptable.

While reinforcement was not required to elicit the children's behavior in Bandura's first experiment, it is important to acknowledge that consequences do play a role within observational learning. A future adaptation of this study ([Bandura, Ross, & Ross, 1963](#)) demonstrated that children in the aggression group showed less aggressive behavior if they witnessed the adult model receive punishment for aggressing against Bobo. Bandura referred to this process as **vicarious reinforcement**, as the children did not experience the reinforcement or punishment directly, yet were still influenced by observing it.

CONCLUSION

We have covered three primary explanations for how we learn to behave and interact with the world around us. Considering your own experiences, how well do these theories apply to you? Maybe when reflecting on your personal sense of fashion, you realize that you tend to select clothes others have complimented you on (operant conditioning). Or maybe, thinking back on a new restaurant you tried recently, you realize you chose it because its commercials play happy music (classical

conditioning). Or maybe you are now always on time with your assignments, because you saw how others were punished when they were late (observational learning). Regardless of the activity, behavior, or response, there's a good chance your "decision" to do it can be explained based on one of the theories presented in this module.

OUTSIDE RESOURCES

Article: Rescorla, R. A. (1988). Pavlovian conditioning: It's not what you think it is. *American Psychologist*, 43, 151–160.

Book: Bouton, M. E. (2007). *Learning and behavior: A contemporary synthesis*. Sunderland, MA: Sinauer Associates.

Book: Bouton, M. E. (2009). Learning theory. In B. J. Sadock, V. A. Sadock, & P. Ruiz (Eds.), *Kaplan & Sadock's comprehensive textbook of psychiatry* (9th ed., Vol. 1, pp. 647–658). New York, NY: Lippincott Williams & Wilkins.

Book: Domjan, M. (2010). *The principles of learning and behavior* (6th ed.). Belmont, CA: Wadsworth.

Video: Albert Bandura discusses the Bobo Doll Experiment.



One or more interactive elements has been excluded from this version of the

text. You can view them online here:

<https://umsystem.pressbooks.pub/ubb2e/?p=247#oembed-1>

DISCUSSION QUESTIONS

1. Describe three examples of Pavlovian (classical) conditioning that you have seen in your own behavior, or that of your friends or family, in the past few days.
2. Describe three examples of instrumental (operant) conditioning that you have seen in your own behavior, or that of your friends or family, in the past few days.
3. Drugs can be potent reinforcers. Discuss how Pavlovian conditioning and instrumental conditioning can work together to influence drug taking.
4. In the modern world, processed foods are highly available and have been engineered to be highly palatable and reinforcing. Discuss how Pavlovian and instrumental conditioning can work together to explain why people often eat too much.
5. How does blocking challenge the idea that pairings of a CS and US are sufficient to cause Pavlovian conditioning? What is important in creating Pavlovian learning?
6. How does the reinforcer devaluation effect

challenge the idea that reinforcers merely “stamp in” the operant response? What does the effect tell us that animals actually learn in operant conditioning?

7. With regards to social learning do you think people learn violence from observing violence in movies? Why or why not?
8. What do you think you have learned through social learning? Who are your social models?

VOCABULARY

Blocking

In classical conditioning, the finding that no conditioning occurs to a stimulus if it is combined with a previously conditioned stimulus during conditioning trials. Suggests that information, surprise value, or prediction error is important in conditioning.

Categorize

To sort or arrange different items into classes or categories.

Classical conditioning

The procedure in which an initially neutral stimulus (the conditioned stimulus, or CS) is paired with an unconditioned stimulus (or US). The result is that the conditioned stimulus begins to elicit a conditioned response (CR). Classical conditioning is nowadays considered important as both a behavioral phenomenon and as a method to study simple associative learning. Same as Pavlovian

conditioning.

Conditioned compensatory response

In classical conditioning, a conditioned response that opposes, rather than is the same as, the unconditioned response. It functions to reduce the strength of the unconditioned response. Often seen in conditioning when drugs are used as unconditioned stimuli.

Conditioned response (CR)

The response that is elicited by the conditioned stimulus after classical conditioning has taken place.

Conditioned stimulus (CS)

An initially neutral stimulus (like a bell, light, or tone) that elicits a conditioned response after it has been associated with an unconditioned stimulus.

Context

Stimuli that are in the background whenever learning occurs. For instance, the Skinner box or room in which learning takes place is the classic example of a context. However, “context” can also be provided by internal stimuli, such as the sensory effects of drugs (e.g., being under the influence of alcohol has stimulus properties that provide a context) and mood states (e.g., being happy or sad). It can also be provided by a specific period in time—the passage of time is sometimes said to change the “temporal context.”

Discriminative stimulus

In operant conditioning, a stimulus that signals

whether the response will be reinforced. It is said to “set the occasion” for the operant response.

Extinction

Decrease in the strength of a learned behavior that occurs when the conditioned stimulus is presented without the unconditioned stimulus (in classical conditioning) or when the behavior is no longer reinforced (in instrumental conditioning). The term describes both the procedure (the US or reinforcer is no longer presented) as well as the result of the procedure (the learned response declines). Behaviors that have been reduced in strength through extinction are said to be “extinguished.”

Fear conditioning

A type of classical or Pavlovian conditioning in which the conditioned stimulus (CS) is associated with an aversive unconditioned stimulus (US), such as a foot shock. As a consequence of learning, the CS comes to evoke fear. The phenomenon is thought to be involved in the development of anxiety disorders in humans.

Goal-directed behavior

Instrumental behavior that is influenced by the animal’s knowledge of the association between the behavior and its consequence and the current value of the consequence. Sensitive to the reinforcer devaluation effect.

Habit

Instrumental behavior that occurs automatically in the presence of a stimulus and is no longer

influenced by the animal's knowledge of the value of the reinforcer. Insensitive to the reinforcer devaluation effect.

Instrumental conditioning

Process in which animals learn about the relationship between their behaviors and their consequences. Also known as operant conditioning.

Law of effect

The idea that instrumental or operant responses are influenced by their effects. Responses that are followed by a pleasant state of affairs will be strengthened and those that are followed by discomfort will be weakened. Nowadays, the term refers to the idea that operant or instrumental behaviors are lawfully controlled by their consequences.

Observational learning

Learning by observing the behavior of others.

Operant

A behavior that is controlled by its consequences. The simplest example is the rat's lever-pressing, which is controlled by the presentation of the reinforcer.

Operant conditioning

See instrumental conditioning.

Pavlovian conditioning

See classical conditioning.

Prediction error

When the outcome of a conditioning trial is different from that which is predicted by the conditioned stimuli that are present on the trial (i.e., when the US is surprising). Prediction error is necessary to create Pavlovian conditioning (and associative learning generally). As learning occurs over repeated conditioning trials, the conditioned stimulus increasingly predicts the unconditioned stimulus, and prediction error declines. Conditioning works to correct or reduce prediction error.

Preparedness

The idea that an organism's evolutionary history can make it easy to learn a particular association. Because of preparedness, you are more likely to associate the taste of tequila, and not the circumstances surrounding drinking it, with getting sick. Similarly, humans are more likely to associate images of spiders and snakes than flowers and mushrooms with aversive outcomes like shocks.

Punisher

A stimulus that decreases the strength of an operant behavior when it is made a consequence of the behavior.

Quantitative law of effect

A mathematical rule that states that the effectiveness of a reinforcer at strengthening an operant response depends on the amount of reinforcement earned for all alternative behaviors. A reinforcer is less effective if there is a lot of reinforcement in the environment

for other behaviors.

Reinforcer

Any consequence of a behavior that strengthens the behavior or increases the likelihood that it will be performed it again.

Reinforcer devaluation effect

The finding that an animal will stop performing an instrumental response that once led to a reinforcer if the reinforcer is separately made aversive or undesirable.

Renewal effect

Recovery of an extinguished response that occurs when the context is changed after extinction. Especially strong when the change of context involves return to the context in which conditioning originally occurred. Can occur after extinction in either classical or instrumental conditioning.

Social Learning Theory

The theory that people can learn new responses and behaviors by observing the behavior of others.

Social models

Authorities that are the targets for observation and who model behaviors.

Spontaneous recovery

Recovery of an extinguished response that occurs with the passage of time after extinction. Can occur after extinction in either classical or instrumental conditioning.

Stimulus control

When an operant behavior is controlled by a stimulus that precedes it.

Taste aversion learning

The phenomenon in which a taste is paired with sickness, and this causes the organism to reject—and dislike—that taste in the future.

Unconditioned response (UR)

In classical conditioning, an innate response that is elicited by a stimulus before (or in the absence of) conditioning.

Unconditioned stimulus (US)

In classical conditioning, the stimulus that elicits the response before conditioning occurs.

Vicarious reinforcement

Learning that occurs by observing the reinforcement or punishment of another person.

REFERENCES

- Balleine, B. W. (2005). Neural basis of food-seeking: Affect, arousal, and reward in corticostriatal limbic circuits. *Physiology & Behavior*, 86, 717–730.
- Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall
- Bandura, A., Ross, D., Ross, S (1963). Imitation of film-mediated aggressive models. *Journal of Abnormal and Social Psychology* 66(1), 3 – 11.

- Bandura, A.; Ross, D.; Ross, S. A. (1961). "Transmission of aggression through the imitation of aggressive models". *Journal of Abnormal and Social Psychology* 63(3), 575–582.
- Bernstein, I. L. (1991). Aversion conditioning in response to cancer and cancer treatment. *Clinical Psychology Review*, 11, 185–191.
- Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learning & Memory*, 11, 485–494.
- Colwill, R. M., & Rescorla, R. A. (1986). Associative structures in instrumental learning. In G. H. Bower (Ed.), *The psychology of learning and motivation*, (Vol. 20, pp. 55–104). New York, NY: Academic Press.
- Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, 46, 5–27.
- Dickinson, A., & Balleine, B. W. (1994). Motivational control of goal-directed behavior. *Animal Learning & Behavior*, 22, 1–18.
- Fanselow, M. S., & Poulos, A. M. (2005). The neuroscience of mammalian associative learning. *Annual Review of Psychology*, 56, 207–234.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, 13, 243–266.
- Holland, P. C. (2004). Relations between

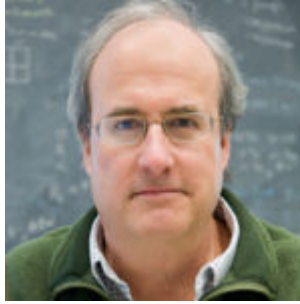
Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes*, 30, 104–117.

- Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279–296). New York, NY: Appleton-Century-Crofts.
- Mineka, S., & Zinbarg, R. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: It's not what you thought it was. *American Psychologist*, 61, 10–26.
- Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, 52, 111–139.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton-Century-Crofts.
- Scalera, G., & Bavieri, M. (2009). Role of conditioned taste aversion on the side effects of chemotherapy in cancer patients. In S. Reilly & T. R. Schachtman (Eds.), *Conditioned taste aversion: Behavioral and neural processes* (pp. 513–541). New York, NY: Oxford University Press.
- Siegel, S. (1989). Pharmacological conditioning and drug effects. In A. J. Goudie & M. Emmett-Oglesby (Eds.), *Psychoactive drugs* (pp. 115–180).

Clifton, NY: Humana Press.

- Siegel, S., Hinson, R. E., Krank, M. D., & McCully, J. (1982). Heroin “overdose” death: Contribution of drug associated environmental cues. *Science*, 216, 436–437.
- Spreat, S., & Spreat, S. R. (1982). Learning principles. In V. Voith & P. L. Borchelt (Eds.), *Veterinary clinics of North America: Small animal practice* (pp. 593–606). Philadelphia, PA: W. B. Saunders.
- Thompson, R. F., & Steinmetz, J. E. (2009). The role of the cerebellum in classical conditioning of discrete behavioral responses. *Neuroscience*, 162, 732–755.
- Timberlake, W. L. (2001). Motivational modes in behavior systems. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 155–210). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Wasserman, E. A. (1995). The conceptual abilities of pigeons. *American Scientist*, 83, 246–255.

AUTHORS



Mark E. Bouton

Mark E. Bouton is the Lawson Professor of Psychology at the University of Vermont. His research on learning and extinction is internationally known. He is a Fellow of several scientific organizations, including APA, APS, and the Society of Experimental Psychologists. He was recently awarded the Gantt Medal from the Pavlovian Society.

CREATIVE COMMONS LICENSE



CONDITIONING AND LEARNING by [Mark E. Bouton](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#).

Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

CHAPTER 19.

FORGETTING AND AMNESIA

NICOLE DUDUKOVIC AND BRICE KUHL

This module explores the causes of everyday forgetting and considers pathological forgetting in the context of amnesia. Forgetting is viewed as an adaptive process that allows us to be efficient in terms of the information we retain.

LEARNING OBJECTIVES

- Identify five reasons we forget and give examples of each.
- Describe how forgetting can be viewed as an adaptive process.
- Explain the difference between anterograde and retrograde amnesia.

INTRODUCTION

Chances are that you have experienced memory lapses and been frustrated by them. You may have had trouble remembering the definition of a key term on an exam or found yourself unable to recall the name of an actor

from one of your favorite TV shows. Maybe you forgot to call your aunt on her birthday or you routinely forget where you put your cell phone. Oftentimes, the bit of information we are searching for comes back to us, but sometimes it does not. Clearly, forgetting seems to be a natural part of life. Why do we forget? And is forgetting always a bad thing?



Forgetting can often be obnoxious or even embarrassing. But as we explore this module, you'll learn that forgetting is important and necessary for everyday functionality. [Image: jazbeck, <https://goo.gl/nkRrJy>, CC BY 2.0, <https://goo.gl/BRvSA7>]

CAUSES OF FORGETTING

One very common and obvious reason why you cannot remember a piece of information is because you did not learn it in the first place. If you fail to encode information into memory, you are not going to remember it later on. Usually, **encoding** failures occur because we are distracted or are not paying attention to specific details.

For example, people have a lot of trouble recognizing an actual penny out of a set of drawings of very similar pennies, or lures, even though most of us have had a lifetime of experience handling pennies ([Nickerson & Adams, 1979](#)). However, few of us have studied the features of a penny in great detail, and since we have not attended to those details, we fail to recognize them later. Similarly, it has been well documented that distraction during learning impairs later memory (e.g., [Craik, Govoni, Naveh-Benjamin, & Anderson, 1996](#)). Most of the time this is not problematic, but in certain situations, such as when you are studying for an exam, failures to encode due to distraction can have serious repercussions.

Another proposed reason why we forget is that memories fade, or **decay**, over time. It has been known since the pioneering work of Hermann Ebbinghaus ([1885/1913](#)) that as time passes, memories get harder to recall. Ebbinghaus created more than 2,000 nonsense syllables, such as *dax*, *bap*, and *rif*, and studied his own memory for them, learning as many as 420 lists of 16 nonsense syllables for one experiment. He found that his memories diminished as time passed, with the most forgetting happening early on after learning. His observations and subsequent research suggested that if we do not rehearse a memory and the neural representation of that memory is not reactivated over a long period of time, the memory representation may disappear entirely or fade to the point where it can no longer be accessed. As you might imagine, it is hard to definitively prove that a memory has decayed as opposed to it being inaccessible for another reason. Critics argued that forgetting must be due to processes other than simply the passage of time, since disuse of a memory

does not always guarantee forgetting ([McGeoch, 1932](#)). More recently, some memory theorists have proposed that recent memory traces may be degraded or disrupted by new experiences ([Wixted, 2004](#)). Memory traces need to be **consolidated**, or transferred from the hippocampus to more durable representations in the cortex, in order for them to last ([McGaugh, 2000](#)). When the consolidation process is interrupted by the encoding of other experiences, the memory trace for the original experience does not get fully developed and thus is forgotten.



At times, we will completely blank on something we're certain we've learned – people we went to school with years ago for example. However, once we get the right retrieval cue (a name perhaps), the memory (faces or experiences) rushes back to us like it was there all along. [Image: sbhsclass84, <https://goo.gl/sHZyQI>, CC BY-SA 2.0, <https://goo.gl/rxiUsF>]

Both encoding failures and decay account for more permanent forms of forgetting, in which the memory trace does not exist, but forgetting may also occur when a memory exists yet we temporarily cannot access it. This type of forgetting may occur when we lack the

appropriate **retrieval** cues for bringing the memory to mind. You have probably had the frustrating experience of forgetting your password for an online site. Usually,

the password has not been permanently forgotten; instead, you just need the right reminder to remember what it is. For example, if your password was “pizza0525,” and you received the password hints “favorite food” and “Mom’s birthday,” you would easily be able to retrieve it. Retrieval hints can bring back to mind seemingly forgotten memories ([Tulving & Pearlstone, 1966](#)). One real-life illustration of the importance of retrieval cues comes from a study showing that whereas people have difficulty recalling the names of high school classmates years after graduation, they are easily able to recognize the names and match them to the appropriate faces ([Bahrick, Bahrick, & Wittinger, 1975](#)). The names are powerful enough retrieval cues that they bring back the memories of the faces that went with them. The fact that the presence of the right retrieval cues is critical for remembering adds to the difficulty in proving that a memory is permanently forgotten as opposed to temporarily unavailable.

Retrieval failures can also occur because other memories are blocking or getting in the way of recalling the desired memory. This blocking is referred to as **interference**. For example, you may fail to remember the name of a town you visited with your family on summer vacation because the names of other towns you visited on that trip or on other trips come to mind instead. Those memories then prevent the desired memory from being retrieved. Interference is also relevant to the example of forgetting a password: passwords that we have used for other websites may come to mind and interfere with our ability to retrieve the desired password. Interference can be either proactive, in which old memories block the learning of

new related memories, or retroactive, in which new memories block the retrieval of old related memories. For both types of interference, competition between memories seems to be key ([Mensink & Raaijmakers, 1988](#)). Your memory for a town you visited on vacation is unlikely to interfere with your ability to remember an Internet password, but it is likely to interfere with your ability to remember a different town's name. Competition between memories can also lead to forgetting in a different way. Recalling a desired memory in the face of competition may result in the inhibition of related, competing memories ([Levy & Anderson, 2002](#)). You may have difficulty recalling the name of Kennebunkport, Maine, because other Maine towns, such as Bar Harbor, Winterport, and Camden, come to mind instead. However, if you are able to recall Kennebunkport despite strong competition from the other towns, this may actually change the competitive landscape, weakening memory for those other towns' names, leading to forgetting of them instead.

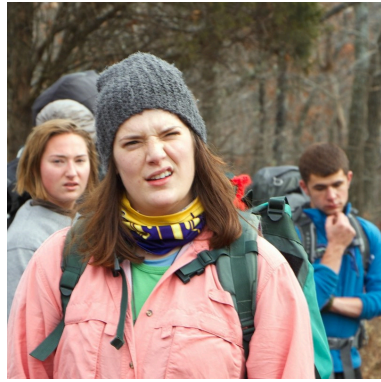
Box 1. Five Impediments to Remembering

1. Encoding failures - we don't learn the information in the first place
2. Decay - memories fade over time
3. Inadequate retrieval cues - we lack sufficient reminders
4. Interference - other memories get in the way
5. Trying not to remember - we deliberately attempt to keep things out of mind

Finally, some memories may be forgotten because *we deliberately attempt to keep them out of mind*. Over time, by actively trying not to remember an event, we can sometimes successfully keep the undesirable memory from being retrieved either by inhibiting the undesirable memory or generating diversionary thoughts ([Anderson & Green, 2001](#)). Imagine that you slipped and fell in your high school cafeteria during lunch time, and everyone at the surrounding tables laughed at you. You would likely wish to avoid thinking about that event and might try to prevent it from coming to mind. One way that you could accomplish this is by thinking of other, more positive, events that are associated with the cafeteria. Eventually, this memory may be suppressed to the point that it would only be retrieved with great difficulty ([Hertel & Calcaterra, 2005](#)).

ADAPTIVE FORGETTING

We have explored five different causes of forgetting. Together they can account for the day-to-day episodes of forgetting that each of us experience. Typically, we think of these episodes in a negative light and view forgetting as a memory failure. Is forgetting ever good? Most people would reason that forgetting that occurs in response to a deliberate attempt to keep an event out of mind is a good thing. No one wants to be constantly reminded

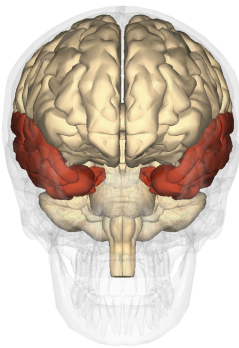


Could you imagine being unable to forget every path you have taken while hiking? Each new trip, you would be walking around the forest for days, incapable of distinguishing today's path from the prior ones. [Image: Dan Trew, <https://goo.gl/8fJWWE>, CC BY-SA 2.0, <https://goo.gl/rxiUsF>]

of falling on their face in front of all of their friends. However, beyond that, it can be argued that forgetting is adaptive, allowing us to be efficient and hold onto only the most relevant memories ([Bjork, 1989](#); [Anderson & Milson, 1989](#)). Shereshevsky, or “S,” the mnemonist studied by Alexander Luria ([1968](#)), was a man who almost never forgot. His memory appeared to be virtually limitless. He could memorize a table of 50 numbers in under 3 minutes and recall the numbers in rows, columns, or diagonals with ease. He could recall lists of words and passages that he had memorized over a decade before. Yet Shereshevsky found it difficult to function in his everyday life because he was constantly distracted by

a flood of details and associations that sprung to mind. His case history suggests that remembering everything is not always a good thing. You may occasionally have trouble remembering where you parked your car, but imagine if every time you had to find your car, every single former parking space came to mind. The task would become impossibly difficult to sort through all of those irrelevant memories. Thus, forgetting is adaptive in that it makes us more efficient. The price of that efficiency is those moments when our memories seem to fail us ([Schacter, 1999](#)).

AMNESIA



Patients with damage to the temporal lobes may experience anterograde amnesia and/or retrograde amnesia.

*[Image: en:Anatomography,
<https://goo.gl/ALPAu6>, CC BY-SA 2.1
JP, <https://goo.gl/BDF2Z4>]*

Clearly, remembering everything would be maladaptive, but what would it be like to remember nothing? We will now consider a profound form of forgetting called amnesia that is distinct from more ordinary forms of forgetting. Most of us have had exposure to the concept of amnesia through popular movies and television. Typically, in these fictionalized

portrayals of amnesia, a character suffers some type of blow to the head and suddenly has no idea who they are and can no longer recognize their family or remember

any events from their past. After some period of time (or another blow to the head), their memories come flooding back to them. Unfortunately, this portrayal of amnesia is not very accurate. What does amnesia typically look like?

The most widely studied amnesic patient was known by his initials H. M. ([Scoville & Milner, 1957](#)). As a teenager, H. M. suffered from severe epilepsy, and in 1953, he underwent surgery to have both of his medial temporal lobes removed to relieve his epileptic seizures. The **medial temporal lobes** encompass the hippocampus and surrounding cortical tissue. Although the surgery was successful in reducing H. M.'s seizures and his general intelligence was preserved, the surgery left H. M. with a profound and permanent memory deficit. From the time of his surgery until his death in 2008, H. M. was unable to learn new information, a memory impairment called **anterograde amnesia**. H. M. could not remember any event that occurred since his surgery, including highly significant ones, such as the death of his father. He could not remember a conversation he had a few minutes prior or recognize the face of someone who had visited him that same day. He could keep information in his short-term, or working, memory, but when his attention turned to something else, that information was lost for good. It is important to note that H. M.'s memory impairment was restricted to **declarative memory**, or conscious memory for facts and events. H. M. could learn new motor skills and showed improvement on motor tasks even in the absence of any memory for having performed the task before ([Corkin, 2002](#)).

In addition to anterograde amnesia, H. M. also suffered from **temporally graded retrograde amnesia**. **Retrograde amnesia** refers to an inability to

retrieve old memories that occurred before the onset of amnesia. Extensive retrograde amnesia in the absence of anterograde amnesia is very rare ([Kopelman, 2000](#)). More commonly, retrograde amnesia co-occurs with anterograde amnesia and shows a temporal gradient, in which memories closest in time to the onset of amnesia are lost, but more remote memories are retained ([Hodges, 1994](#)). In the case of H. M., he could remember events from his childhood, but he could not remember events that occurred a few years before the surgery.

Amnesiac patients with damage to the hippocampus and surrounding medial temporal lobes typically manifest a similar clinical profile as H. M. The degree of anterograde amnesia and retrograde amnesia depend on the extent of the medial temporal lobe damage, with greater damage associated with a more extensive impairment ([Reed & Squire, 1998](#)). Anterograde amnesia provides evidence for the role of the hippocampus in the formation of long-lasting declarative memories, as damage to the hippocampus results in an inability to create this type of new memory. Similarly, temporally graded retrograde amnesia can be seen as providing further evidence for the importance of memory consolidation ([Squire & Alvarez, 1995](#)). A memory depends on the hippocampus until it is consolidated and transferred into a more durable form that is stored in the cortex. According to this theory, an amnesiac patient like H. M. could remember events from his remote past because those memories were fully consolidated and no longer depended on the hippocampus.

The classic amnesiac syndrome we have considered here is sometimes referred to as organic amnesia, and it is distinct from functional, or dissociative, amnesia.

Functional amnesia involves a loss of memory that cannot be attributed to brain injury or any obvious brain disease and is typically classified as a mental disorder rather than a neurological disorder ([Kihlstrom, 2005](#)). The clinical profile of dissociative amnesia is very different from that of patients who suffer from amnesia due to brain damage or deterioration. Individuals who experience **dissociative amnesia** often have a history of trauma. Their amnesia is retrograde, encompassing autobiographical memories from a portion of their past. In an extreme version of this disorder, people enter a dissociative fugue state, in which they lose most or all of their autobiographical memories and their sense of personal identity. They may be found wandering in a new location, unaware of who they are and how they got there. Dissociative amnesia is controversial, as both the causes and existence of it have been called into question. The memory loss associated with dissociative amnesia is much less likely to be permanent than it is in organic amnesia.

CONCLUSION

Just as the case study of the mnemonist Shereshevsky illustrates what a life with a near perfect memory would be like, amnesiac patients show us what a life without memory would be like. Each of the mechanisms we discussed that explain everyday forgetting—encoding failures, decay, insufficient retrieval cues, interference, and intentional attempts to forget—help to keep us highly efficient, retaining the important information and for the most part, forgetting the unimportant. Amnesiac patients allow us a glimpse into what life would be like if we

suffered from profound forgetting and perhaps show us that our everyday lapses in memory are not so bad after all.

OUTSIDE RESOURCES

Web: Brain Case Study: Patient HM

<https://bigpictureeducation.com/brain-case-study-patient-hm>

Web: Self-experiment, Penny demo

<http://www.indiana.edu/~p1013447/dictionary/penny.htm>

Web: The Man Who Couldn't Remember

<http://www.pbs.org/wgbh/nova/body/corkin-hm-memory.html>

DISCUSSION QUESTIONS

1. Is forgetting good or bad? Do you agree with the authors that forgetting is an adaptive process? Why or why not?
2. Can we ever prove that something is forgotten? Why or why not?
3. Which of the five reasons for forgetting do you think explains the majority of incidences of forgetting? Why?
4. How is real-life amnesia different than amnesia that is portrayed on TV and in film?

VOCABULARY

Anterograde amnesia

Inability to form new memories for facts and events after the onset of amnesia.

Consolidation

Process by which a memory trace is stabilized and transformed into a more durable form.

Decay

The fading of memories with the passage of time.

Declarative memory

Conscious memories for facts and events.

Dissociative amnesia

Loss of autobiographical memories from a period in the past in the absence of brain injury or disease.

Encoding

Process by which information gets into memory.

Interference

Other memories get in the way of retrieving a desired memory

Medial temporal lobes

Inner region of the temporal lobes that includes the hippocampus.

Retrieval

Process by which information is accessed from memory and utilized.

Retrograde amnesia

Inability to retrieve memories for facts and events acquired before the onset of amnesia.

Temporally graded retrograde amnesia

Inability to retrieve memories from just prior to the onset of amnesia with intact memory for more remote events.

REFERENCES

- Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. **Psychological Review**, 96, 703–719.
- Anderson, M. C., & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature*, 410, 366–369.
- Bahrick, H. P., Bahrick, P. O., & Wittinger, R. P. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General*, 104, 54–75.
- Bjork, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger, III, & F. I. M. Craik (Eds.), *Varieties of Memory and Consciousness* (pp. 309– 330). Hillsdale, NJ: Erlbaum.
- Corkin, S. (2002). What's new with the amnesic patient H. M.? **Nature Reviews Neuroscience**, 3, 153–160.
- Craik, F. I. M., Govoni, R., Naveh-Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in

human memory. *Journal of Experimental Psychology: General*, 125, 159–180.

- Ebbinghaus, H. (1913). *Memory. A contribution to experimental psychology*. New York: Teachers College/Columbia University (Engl. ed.). (Original work published in 1885.)
- Hertel, P. T., & Calcaterra, G. (2005). Intentional forgetting benefits from thought substitution. *Psychonomic Bulletin & Review*, 12, 484–489.
- Hodges, J. R. (1994). Retrograde amnesia. In A. Baddeley, B. A. Wilson, & F. Watts (Eds.), *Handbook of Memory Disorders* (pp. 81–107). New York: Wiley.
- Kihlstrom, J. F. (2005). Dissociative disorders. *Annual Review of Clinical Psychology*, 1, 227–253.
- Kopelman, M. (2000). Focal retrograde amnesia and the attribution of causality: An exceptionally critical review. *Cognitive Neuropsychology*, 17, 585–621.
- Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Sciences*, 6, 299–305.
- Luria, A. R. (1968). *The mind of a mnemonist: A little book about a vast memory* (L. Solataroff, Trans.). New York: Basic Books.
- McGaugh, J. L. (2000). Memory: A century of consolidation. *Science*, 287, 248–251.
- McGeoch, J. A. (1932). Forgetting and the law of disuse. *Psychological Reviews*, 39, 352–370.

- Mensink, G., & Raaijmakers, J. G. (1988). A model for interference and forgetting. **Psychological Review**, 95, 434–455.
- Nickerson, R. S., & Adams, M. J. (1979). Long-term memory for a common object. **Cognitive Psychology**, 11, 287–307.
- Reed, J. M. & Squire, L. R. (1998). Retrograde amnesia for facts and events: Findings from four new cases. *Journal of Neuroscience*, 18, 3943–3954.
- Schacter, D. L. (1999). The seven sins of memory: Insights from psychology and cognitive neuroscience. *American Psychologist*, 54, 182–203.
- Scoville, W. B. & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of Neurology, Neurosurgery, & Psychiatry*, 20, 11–21.
- Squire, L. R., & Alvarez, P. (1995). Retrograde amnesia and memory consolidation: A neurobiological perspective. *Current Opinions in Neurobiology*, 5, 169–177.
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5, 381–391.
- Wixted, J. T. (2004). The psychology and neuroscience of forgetting. **Annual Reviews of Psychology**, 55, 235–269.

AUTHORS



Nicole Dudukovic

Nicole Dudukovic earned her Ph.D. in Psychology from Stanford University and has taught courses at Stanford, Trinity College, and New York University. Her research explores interactions between attention and memory, and she is interested in applying memory research to other fields, particularly education.



Brice Kuhl

Brice Kuhl earned his Ph.D. in Psychology from Stanford University and completed postdoctoral work at Yale University. He is currently an Assistant Professor of Psychology at New York University. His research explores the neural mechanisms of memory and causes of forgetting.

CREATIVE COMMONS LICENSE



FORGETTING AND AMNESIA by [Nicole Dudukovic and Brice Kuhl](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

PART XII.

**UNIT 12: PSYCHOLOGICAL
DISORDERS**

CHAPTER 20.

MOOD DISORDERS

ANDA GERSHON AND RENEE THOMPSON

Everyone feels down or euphoric from time to time, but this is different from having a mood disorder such as major depressive disorder or bipolar disorder. Mood disorders are extended periods of depressed, euphoric, or irritable moods that in combination with other symptoms cause the person significant distress and interfere with his or her daily life, often resulting in social and occupational difficulties. In this module, we describe major mood disorders, including their symptom presentations, general prevalence rates, and how and why the rates of these disorders tend to vary by age, gender, and race. In addition, biological and environmental risk factors that have been implicated in the development and course of mood disorders, such as heritability and stressful life events, are reviewed. Finally, we provide an overview of treatments for mood disorders, covering treatments with demonstrated effectiveness, as well as new treatment options showing promise.

LEARNING OBJECTIVES

- Describe the diagnostic criteria for mood disorders.
- Understand age, gender, and ethnic differences in prevalence rates of mood disorders.
- Identify common risk factors for mood disorders.
- Know effective treatments of mood disorders.



Perinatal depression following child birth afflicts about 5% of all mothers. An unfortunate social stigma regarding this form of depression compounds the problem for the women who suffer its effects. [Image: CCO Public Domain]

The actress Brooke Shields published a memoir titled [Down Came the Rain: My Journey through Postpartum Depression](#) in which she described her struggles with depression following the birth of her daughter. Despite the fact that about one in 20 women experience depression after the birth of a baby

(American Psychiatric Association [\[APA\], 2013](#)), postpartum depression—recently renamed “perinatal depression”—continues to be veiled by stigma, owing in part to a widely held expectation that motherhood should be a time of great joy. In an opinion piece in the *New York Times*, Shields revealed that entering motherhood was a profoundly overwhelming experience

for her. She vividly describes experiencing a sense of “doom” and “dread” in response to her newborn baby. Because motherhood is conventionally thought of as a joyous event and not associated with sadness and hopelessness, responding to a newborn baby in this way can be shocking to the new mother as well as those close to her. It may also involve a great deal of shame for the mother, making her reluctant to divulge her experience to others, including her doctors and family. Feelings of shame are not unique to perinatal depression. Stigma applies to other types of depressive and bipolar disorders and contributes to people not always receiving the necessary support and treatment for these disorders. In fact, the World Health Organization ranks both major depressive disorder (MDD) and bipolar disorder (BD) among the top 10 leading causes of disability worldwide. Further, MDD and BD carry a high risk of suicide. It is estimated that 25%–50% of people diagnosed with BD will attempt suicide at least once in their lifetimes ([Goodwin & Jamison, 2007](#)).

WHAT ARE MOOD DISORDERS?

MOOD EPISODES

Everyone experiences brief periods of sadness, irritability, or euphoria. This is different than having a mood disorder, such as MDD or BD, which are characterized by a constellation of symptoms that causes people significant distress or impairs their everyday functioning.

Major Depressive Episode

A major depressive episode (MDE) refers to symptoms that co-occur for at least two weeks and cause significant distress or impairment in functioning, such as interfering with work, school, or relationships. Core symptoms include feeling down or depressed or experiencing **anhedonia**—loss of interest or pleasure in things that one typically enjoys. According to the fifth edition of the *Diagnostic and Statistical Manual (DSM-5; APA, 2013)*, the criteria for an MDE require five or more of the following nine symptoms, including one or both of the first two symptoms, for most of the day, nearly every day:

1. depressed mood
2. diminished interest or pleasure in almost all activities
3. significant weight loss or gain or an increase or decrease in appetite
4. insomnia or hypersomnia
5. **psychomotor agitation or retardation**
6. fatigue or loss of energy
7. feeling worthless or excessive or inappropriate guilt
8. diminished ability to concentrate or indecisiveness
9. recurrent thoughts of death, **suicidal ideation**, or a suicide attempt

These symptoms cannot be caused by physiological

effects of a substance or a general medical condition (e.g., hypothyroidism).

Manic or Hypomanic Episode

The core criterion for a manic or hypomanic episode is a distinct period of abnormally and persistently euphoric, expansive, or irritable mood and persistently increased goal-directed activity or energy. The mood disturbance must be present for one week or longer in mania (unless hospitalization is required) or four days or longer in hypomania. Concurrently, at least three of the following symptoms must be present in the context of euphoric mood (or at least four in the context of irritable mood):

1. inflated self-esteem or **grandiosity**
2. increased goal-directed activity or psychomotor agitation
3. reduced need for sleep
4. racing thoughts or flight of ideas
5. distractibility
6. increased talkativeness
7. excessive involvement in risky behaviors

Manic episodes are distinguished from hypomanic episodes by their duration and associated impairment; whereas manic episodes must last one week and are defined by a significant impairment in functioning, hypomanic episodes are shorter and not necessarily accompanied by impairment in functioning.

MOOD DISORDERS

Unipolar Mood Disorders

Two major types of unipolar disorders described by the *DSM-5* ([APA, 2013](#)) are major depressive disorder and persistent depressive disorder (PDD; dysthymia). MDD is defined by one or more MDEs, but no history of manic or hypomanic episodes. Criteria for PDD are feeling depressed most of the day for more days than not, for at least two years. At least two of the following symptoms are also required to meet criteria for PDD:

1. poor appetite or overeating
2. insomnia or hypersomnia
3. low energy or fatigue
4. low self-esteem
5. poor concentration or difficulty making decisions
6. feelings of hopelessness

Like MDD, these symptoms need to cause significant distress or impairment and cannot be due to the effects of a substance or a general medical condition. To meet criteria for PDD, a person cannot be without symptoms for more than two months at a time. PDD has overlapping symptoms with MDD. If someone meets criteria for an MDE during a PDD episode, the person will receive diagnoses of PDD and MDD.

Bipolar Mood Disorders

Three major types of BDs are described by the *DSM-5* (APA, 2013). Bipolar I Disorder (BD I), which was previously known as manic-depression, is characterized by a single (or recurrent) manic episode. A depressive episode is not necessary but commonly present for the diagnosis of BD I.



Bipolar disorders are characterized by cycles of high energy and depression.

[Image: Brett Whaley, <https://goo.gl/k4HTR7>, CC BY-NC 2.0, <https://goo.gl/VnKIK8>]

Bipolar II Disorder is characterized by single (or recurrent) hypomanic episodes and depressive episodes. Another type of BD is cyclothymic disorder, characterized by numerous and alternating periods of hypomania and depression, lasting at least two years. To qualify for cyclothymic disorder, the periods of depression cannot meet full diagnostic criteria for an MDE; the person must experience symptoms at least half the time with no more than two consecutive symptom-free months; and the symptoms must cause significant distress or impairment.

Box 1. Specifiers

Both MDEs and manic episodes can be further described using standardized tags based on the timing of, or other symptoms that are occurring during, the mood episode, to increase diagnostic specificity and inform treatment. Psychotic features is specified when the episodes are accompanied by delusions (rigidly held beliefs that are false) or hallucinations (perceptual disturbances that are not based in reality). Seasonal pattern is specified when a mood episode occurs at the same time of the year for two consecutive years—most commonly occurring in the fall and winter. Peripartum onset is specified when a mood episode has an onset during pregnancy or within four weeks of the birth of a child. Approximately 3%–6% of women who have a child experience an MDE with peripartum onset (APA, 2013). This is less frequent and different from the baby blues or when women feel transient mood symptoms usually within 10 days of giving birth, which are experienced by most women (Nolen-Hoeksema & Hilt, 2009).

It is important to note that the *DSM-5* was published in 2013, and findings based on the updated manual will be forthcoming. Consequently, the research presented below was largely based on a similar, but not identical, conceptualization of mood disorders drawn from the *DSM-IV* ([APA, 2000](#)).

HOW COMMON ARE MOOD DISORDERS? WHO DEVELOPS MOOD DISORDERS?

DEPRESSIVE DISORDERS

In a nationally representative sample, lifetime prevalence rate for MDD is 16.6% ([Kessler, Berglund, Demler, Jin, Merikangas, & Walters, 2005](#)). This means that nearly one in five Americans will meet the criteria for MDD during their lifetime. The 12-month prevalence—the proportion of people who meet criteria for a disorder during a 12-month period—for PDD is approximately 0.5% ([APA, 2013](#)).

Although the onset of MDD can occur at any time throughout the lifespan, the average age of onset is mid-20s, with the age of onset decreasing with people born more recently ([APA, 2000](#)). Prevalence of MDD among older adults is much lower than it is for younger cohorts ([Kessler, Birnbaum, Bromet, Hwang, Sampson, & Shahly, 2010](#)). The duration of MDEs varies widely. Recovery begins within three months for 40% of people with MDD and within 12 months for 80% ([APA, 2013](#)). MDD tends to be a recurrent disorder with about 40%–50% of those who experience one MDE experiencing a second MDE ([Monroe & Harkness, 2011](#)). An earlier age of onset predicts a worse course. About 5%–10% of people who experience an MDE will later experience a manic episode ([APA, 2000](#)), thus no longer meeting criteria for MDD but instead meeting them for BD I. Diagnoses of other disorders across the lifetime are common for people with MDD: 59% experience an anxiety disorder; 32% experience an impulse control

disorder, and 24% experience a substance use disorder ([Kessler, Merikangas, & Wang, 2007](#)).

Women experience two to three times higher rates of MDD than do men ([Nolen-Hoeksema & Hilt, 2009](#)). This gender difference emerges during puberty ([Conley & Rudolph, 2009](#)). Before puberty, boys exhibit similar or higher prevalence rates of MDD than do girls ([Twenge & Nolen-Hoeksema, 2002](#)). MDD is inversely correlated with **socioeconomic status** (SES), a person's economic and social position based on income, education, and occupation. Higher prevalence rates of MDD are associated with lower SES ([Lorant, Deliege, Eaton, Robert, Philippot, & Ansseau, 2003](#)), particularly for adults over 65 years old ([Kessler et al., 2010](#)). Independent of SES, results from a nationally representative sample found that European Americans had a higher prevalence rate of MDD than did African Americans and Hispanic Americans, whose rates were similar ([Breslau, Aguilar-Gaxiola, Kendler, Su, Williams, & Kessler, 2006](#)). The course of MDD for African Americans is often more severe and less often treated than it is for European Americans, however ([Williams et al., 2007](#)). Native Americans have a higher prevalence rate than do European Americans, African Americans, or Hispanic Americans ([Hasin, Goodwin, Stinson & Grant, 2005](#)). Depression is not limited to industrialized or western cultures; it is found in all countries that have been examined, although the symptom presentation as well as prevalence rates vary across cultures ([Chentsova-Dutton & Tsai, 2009](#)).

BIPOLAR DISORDERS



Adolescents experience a higher incidence of bipolar spectrum disorders than do adults. Making matters worse, those who are diagnosed with BD at a younger age seem to suffer symptoms more intensely than those with adult onset. [Image: CCO Public Domain]

The lifetime prevalence rate of bipolar spectrum disorders in the general U.S. population is estimated at approximately 4.4%, with BD I constituting about 1% of this rate ([Merikangas et al., 2007](#)). Prevalence estimates, however, are highly dependent on the diagnostic procedures used (e.g., interviews vs. self-report) and whether or not sub-threshold forms of the disorder are

included in the estimate. BD often co-occurs with other psychiatric disorders. Approximately 65% of people with BD meet diagnostic criteria for at least one additional psychiatric disorder, most commonly anxiety disorders and substance use disorders ([McElroy et al., 2001](#)). The co-occurrence of BD with other psychiatric disorders is associated with poorer illness course, including higher rates of suicidality ([Leverich et al., 2003](#)). A recent cross-national study sample of more than 60,000 adults from 11 countries, estimated the worldwide prevalence of BD at 2.4%, with BD I constituting 0.6% of this rate ([Merikangas et al., 2011](#)). In this study, the prevalence of BD varied somewhat by country. Whereas the United States had the highest lifetime prevalence (4.4%), India had the lowest

(0.1%). Variation in prevalence rates was not necessarily related to SES, as in the case of Japan, a high-income country with a very low prevalence rate of BD (0.7%).

With regard to ethnicity, data from studies not confounded by SES or inaccuracies in diagnosis are limited, but available reports suggest rates of BD among European Americans are similar to those found among African Americans ([Blazer et al., 1985](#)) and Hispanic Americans ([Breslau, Kendler, Su, Gaxiola-Aguilar, & Kessler, 2005](#)). Another large community-based study found that although prevalence rates of mood disorders were similar across ethnic groups, Hispanic Americans and African Americans with a mood disorder were more likely to remain persistently ill than European Americans ([Breslau et al., 2005](#)). Compared with European Americans with BD, African Americans tend to be underdiagnosed for BD (and over-diagnosed for schizophrenia) ([Kilbourne, Haas, Mulsant, Bauer, & Pincus, 2004](#); [Minsky, Vega, Miskimen, Gara, & Escobar, 2003](#)), and Hispanic Americans with BD have been shown to receive fewer psychiatric medication prescriptions and specialty treatment visits ([Gonzalez et al., 2007](#)). Misdiagnosis of BD can result in the underutilization of treatment or the utilization of inappropriate treatment, and thus profoundly impact the course of illness.

As with MDD, adolescence is known to be a significant risk period for BD; mood symptoms start by adolescence in roughly half of BD cases ([Leverich et al., 2007](#); [Perlis et al., 2004](#)). Longitudinal studies show that those diagnosed with BD prior to adulthood experience a more pernicious course of illness relative to those with adult onset, including more episode recurrence, higher rates of suicidality, and profound social, occupational, and

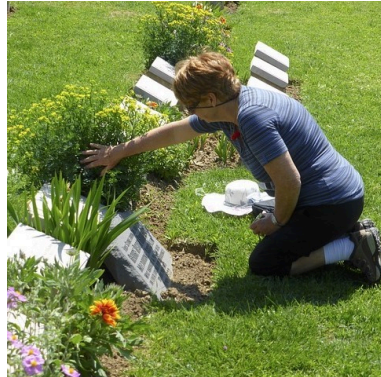
economic repercussions (e.g., [Lewinsohn, Seeley, Buckley, & Klein, 2002](#)). The prevalence of BD is substantially lower in older adults compared with younger adults (1% vs. 4%) ([Merikangas et al., 2007](#)).

WHAT ARE SOME OF THE FACTORS IMPLICATED IN THE DEVELOPMENT AND COURSE OF MOOD DISORDERS?

Mood disorders are complex disorders resulting from multiple factors. Causal explanations can be attempted at various levels, including biological and psychosocial levels. Below are several of the key factors that contribute to onset and course of mood disorders are highlighted.

DEPRESSIVE DISORDERS

Research across family and twin studies has provided support that genetic factors are implicated in the development of MDD. Twin studies suggest that familial influence on MDD is mostly due to genetic effects and that individual-specific environmental effects (e.g., romantic relationships) play an important role, too. By contrast, the contribution of shared environmental effect by siblings is negligible ([Sullivan, Neale & Kendler, 2000](#)). The mode of



Romantic relationships can affect mood as in the case of divorce or the death of a spouse. [Image: CCO Public Domain]

inheritance is not fully understood although no single genetic variation has been found to increase the risk of MDD significantly. Instead, several genetic variants and environmental factors most likely contribute to the risk for MDD ([Lohoff, 2010](#)).

One environmental stressor that has received much support in relation to MDD is stressful life events. In particular, severe stressful life events—those that have long-term consequences and involve loss of a significant relationship (e.g., divorce) or economic stability (e.g., unemployment) are strongly related to depression ([Brown & Harris, 1989](#); [Monroe et al., 2009](#)). Stressful life events are more likely to predict the first MDE than subsequent episodes ([Lewinsohn, Allen, Seeley, & Gotlib, 1999](#)). In contrast, minor events may play a larger role in subsequent episodes than the initial episodes ([Monroe & Harkness, 2005](#)).

Depression research has not been limited to examining reactivity to stressful life events. Much research, particularly brain imagining research using functional magnetic resonance imaging (fMRI), has centered on examining neural circuitry—the interconnections that allow multiple brain regions to perceive, generate, and encode information in concert. A meta-analysis of neuroimaging studies showed that when viewing negative stimuli (e.g., picture of an angry face, picture of a car accident), compared with healthy control participants, participants with MDD have greater activation in brain regions involved in stress response and reduced activation of brain regions involved in positively motivated behaviors ([Hamilton, Etkin, Furman, Lemus, Johnson, & Gotlib, 2012](#)).

Other environmental factors related to increased risk

for MDD include experiencing **early adversity** (e.g., childhood abuse or neglect; [Widom, DuMont, & Czaja, 2007](#)), **chronic stress** (e.g., poverty) and interpersonal factors. For example, marital dissatisfaction predicts increases in depressive symptoms in both men and women. On the other hand, depressive symptoms also predict increases in marital dissatisfaction ([Whisman & Uebelacker, 2009](#)). Research has found that people with MDD generate some of their interpersonal stress ([Hammen, 2005](#)). People with MDD whose relatives or spouses can be described as critical and emotionally overinvolved have higher relapse rates than do those living with people who are less critical and emotionally overinvolved ([Butzlaff & Hooley, 1998](#)).

People's **attributional styles** or their general ways of thinking, interpreting, and recalling information have also been examined in the etiology of MDD ([Gotlib & Joormann, 2010](#)). People with a pessimistic attributional style tend to make internal (versus external), global (versus specific), and stable (versus unstable) attributions to negative events, serving as a vulnerability to developing MDD. For example, someone who when he fails an exam thinks that it was his fault (internal), that he is stupid (global), and that he will always do poorly (stable) has a pessimistic attribution style. Several influential theories of depression incorporate attributional styles ([Abramson, Metalsky, & Alloy, 1989](#); [Abramson Seligman, & Teasdale, 1978](#)).

BIPOLAR DISORDERS

Although there have been important advances in research on the etiology, course, and treatment of BD, there

remains a need to understand the mechanisms that contribute to episode onset and relapse. There is compelling evidence for biological causes of BD, which is known to be highly heritable ([McGuffin, Rijdsdijk, Andrew, Sham, Katz, & Cardno, 2003](#)). It may be argued that a high rate of heritability demonstrates that BD is fundamentally a biological phenomenon. However, there is much variability in the course of BD both within a person across time and across people ([Johnson, 2005](#)). The triggers that determine how and when this genetic vulnerability is expressed are not yet understood; however, there is evidence to suggest that psychosocial triggers may play an important role in BD risk (e.g., [Johnson et al., 2008](#); [Malkoff-Schwartz et al., 1998](#)).

In addition to the genetic contribution, biological explanations of BD have also focused on brain function. Many of the studies using fMRI techniques to characterize BD have focused on the processing of emotional stimuli based on the idea that BD is fundamentally a disorder of emotion ([APA, 2000](#)). Findings show that regions of the brain thought to be involved in emotional processing and regulation are activated differently in people with BD relative to healthy controls (e.g., [Altshuler et al., 2008](#); [Hassel et al., 2008](#); [Lennox, Jacob, Calder, Lupson, & Bullmore, 2004](#)).

However, there is little consensus as to whether a particular brain region becomes more or less active in response to an emotional stimulus among people with BD compared with healthy controls. Mixed findings are in part due to samples consisting of participants who are at various phases of illness at the time of testing (manic, depressed, inter-episode). Sample sizes tend to be relatively small, making comparisons between subgroups

difficult. Additionally, the use of a standardized stimulus (e.g., facial expression of anger) may not elicit a sufficiently strong response. Personally engaging stimuli, such as recalling a memory, may be more effective in inducing strong emotions ([Isacowitz, Gershon, Allard, & Johnson, 2013](#)).

Within the psychosocial level, research has focused on the environmental contributors to BD. A series of studies show that environmental stressors, particularly severe stressors (e.g., loss of a significant relationship), can adversely impact the course of BD. People with BD have substantially increased risk of relapse ([Ellicott, Hammen, Gitlin, Brown, & Jamison, 1990](#)) and suffer more depressive symptoms ([Johnson, Winett, Meyer, Greenhouse, & Miller, 1999](#)) following a severe life stressor. Interestingly, positive life events can also adversely impact the course of BD. People with BD suffer more manic symptoms after life events involving attainment of a desired goal ([Johnson et al., 2008](#)). Such findings suggest that people with BD may have a hypersensitivity to rewards.

Evidence from the life stress literature has also suggested that people with mood disorders may have a circadian vulnerability that renders them sensitive to stressors that disrupt their sleep or rhythms. According to **social zeitgeber** theory ([Ehlers, Frank, & Kupfer, 1988](#); [Frank et al., 1994](#)), stressors that disrupt sleep, or that disrupt the daily routines that entrain the biological clock (e.g., meal times) can trigger episode relapse. Consistent with this theory, studies have shown that life events that involve a disruption in sleep and daily routines, such as overnight travel, can increase bipolar

symptoms in people with BD ([Malkoff-Schwartz et al., 1998](#)).

WHAT ARE SOME OF THE WELL-SUPPORTED TREATMENTS FOR MOOD DISORDERS?

DEPRESSIVE DISORDERS



A number of medications are effective in treating mood disorders. Meditation, exercise, counseling and other therapies also show effectiveness for some disorders. [Image: CCO Public Domain]

There are many treatment options available for people with MDD. First, a number of antidepressant medications are available, all of which target one or more of the neurotransmitters implicated in depression. The earliest antidepressant medications were monoamine oxidase inhibitors (MAOIs). MAOIs inhibit

monoamine oxidase, an enzyme involved in deactivating dopamine, norepinephrine, and serotonin. Although effective in treating depression, MAOIs can have serious side effects. Patients taking MAOIs may develop dangerously high blood pressure if they take certain drugs (e.g., antihistamines) or eat foods containing tyramine, an amino acid commonly found in foods such as aged cheeses, wine, and soy sauce. Tricyclics, the second-oldest class of antidepressant medications, block the reabsorption of norepinephrine, serotonin, or dopamine at synapses, resulting in their increased

availability. Tricyclics are most effective for treating vegetative and somatic symptoms of depression. Like MAOIs, they have serious side effects, the most concerning of which is being cardiotoxic. Selective serotonin reuptake inhibitors (SSRIs; e.g., Fluoxetine) and serotonin and norepinephrine reuptake inhibitors (SNRIs; e.g., Duloxetine) are the most recently introduced antidepressant medications. SSRIs, the most commonly prescribed antidepressant medication, block the reabsorption of serotonin, whereas SNRIs block the reabsorption of serotonin and norepinephrine. SSRIs and SNRIs have fewer serious side effects than do MAOIs and tricyclics. In particular, they are less cardiotoxic, less lethal in overdose, and produce fewer cognitive impairments. They are not, however, without their own side effects, which include but are not limited to difficulty having orgasms, gastrointestinal issues, and insomnia. It should be noted that anti-depressant medication may not work equally for all people. This approach to treatment often involves experimentation with several medications and dosages, and may be more effective when paired with physical exercise and psychotherapy.

Other biological treatments for people with depression include electroconvulsive therapy (ECT), transcranial magnetic stimulation (TMS), and deep brain stimulation. ECT involves inducing a seizure after a patient takes muscle relaxants and is under general anesthesia. ECT is a viable treatment for patients with severe depression or who show resistance to antidepressants although the mechanisms through which it works remain unknown. A common side effect is confusion and memory loss, usually short-term ([Schulze-Rauschenbach, Harms, Schlaepfer, Maier, Falkai, & Wagner, 2005](#)). Repetitive

TMS is a noninvasive technique administered while a patient is awake. Brief pulsating magnetic fields are delivered to the cortex, inducing electrical activity. TMS has fewer side effects than ECT ([Schulze-Rauschenbach et al., 2005](#)), and while outcome studies are mixed, there is evidence that TMS is a promising treatment for patients with MDD who have shown resistance to other treatments ([Rosa et al., 2006](#)). Most recently, deep brain stimulation is being examined as a treatment option for patients who did not respond to more traditional treatments like those already described. Deep brain stimulation involves implanting an electrode in the brain. The electrode is connected to an implanted neurostimulator, which electrically stimulates that particular brain region. Although there is some evidence of its effectiveness ([Mayberg et al., 2005](#)), additional research is needed.

Several psychosocial treatments have received strong empirical support, meaning that independent investigations have achieved similarly positive results—a high threshold for examining treatment outcomes. These treatments include but are not limited to behavior therapy, cognitive therapy, and interpersonal therapy. Behavior therapies focus on increasing the frequency and quality of experiences that are pleasant or help the patient achieve mastery. Cognitive therapies primarily focus on helping patients identify and change distorted automatic thoughts and assumptions (e.g., [Beck, 1967](#)). Cognitive-behavioral therapies are based on the rationale that thoughts, behaviors, and emotions affect and are affected by each other. Interpersonal Therapy for Depression focuses largely on improving interpersonal relationships by targeting problem areas, specifically unresolved grief,

interpersonal role disputes, role transitions, and interpersonal deficits. Finally, there is also some support for the effectiveness of Short-Term Psychodynamic Therapy for Depression ([Leichsenring, 2001](#)). The short-term treatment focuses on a limited number of important issues, and the therapist tends to be more actively involved than in more traditional psychodynamic therapy.

BIPOLAR DISORDERS

Patients with BD are typically treated with pharmacotherapy. Antidepressants such as SSRIs and SNRIs are the primary choice of treatment for depression, whereas for BD, lithium is the first line treatment choice. This is because SSRIs and SNRIs have the potential to induce mania or hypomania in patients with BD. Lithium acts on several neurotransmitter systems in the brain through complex mechanisms, including reduction of excitatory (dopamine and glutamate) neurotransmission, and increasing of inhibitory (GABA) neurotransmission ([Lenox & Hahn, 2000](#)). Lithium has strong efficacy for the treatment of BD ([Geddes, Burgess, Hawton, Jamison, & Goodwin, 2004](#)). However, a number of side effects can make lithium treatment difficult for patients to tolerate. Side effects include impaired cognitive function ([Wingo, Wingo, Harvey, & Baldessarini, 2009](#)), as well as physical symptoms such as nausea, tremor, weight gain, and fatigue ([Dunner, 2000](#)). Some of these side effects can improve with continued use; however, medication noncompliance remains an ongoing concern in the treatment of patients with BD. Anticonvulsant

medications (e.g., carbamazepine, valproate) are also commonly used to treat patients with BD, either alone or in conjunction with lithium.

There are several adjunctive treatment options for people with BD. Interpersonal and social rhythm therapy (IPSRT; [Frank et al., 1994](#)) is a psychosocial intervention focused on addressing the mechanism of action posited in social *zeitgeber* theory to predispose patients who have BD to relapse, namely sleep disruption. A growing body of literature provides support for the central role of sleep dysregulation in BD ([Harvey, 2008](#)). Consistent with this literature, IPSRT aims to increase rhythmicity of patients' lives and encourage vigilance in maintaining a stable rhythm. The therapist and patient work to develop and maintain a healthy balance of activity and stimulation such that the patient does not become overly active (e.g., by taking on too many projects) or inactive (e.g., by avoiding social contact). The efficacy of IPSRT has been demonstrated in that patients who received this treatment show reduced risk of episode recurrence and are more likely to remain well ([Frank et al., 2005](#)).

CONCLUSION

Everyone feels down or euphoric from time to time. For some people, these feelings can last for long periods of time and can also co-occur with other symptoms that, in combination, interfere with their everyday lives. When people experience an MDE or a manic episode, they see the world differently. During an MDE, people often feel hopeless about the future, and may even experience suicidal thoughts. During a manic episode, people often behave in ways that are risky or place them in danger.

They may spend money excessively or have unprotected sex, often expressing deep shame over these decisions after the episode. MDD and BD cause significant problems for people at school, at work, and in their relationships and affect people regardless of gender, age, nationality, race, religion, or sexual orientation. If you or someone you know is suffering from a mood disorder, it is important to seek help. Effective treatments are available and continually improving. If you have an interest in mood disorders, there are many ways to contribute to their understanding, prevention, and treatment, whether by engaging in research or clinical work.

OUTSIDE RESOURCES

Books: Recommended memoirs include *A Memoir of Madness* by William Styron (MDD); *Noonday Demon: An Atlas of Depression* by Andrew Solomon (MDD); and *An Unquiet Mind: A Memoir of Moods and Madness* by Kay Redfield (BD).

Web: Visit the Association for Behavioral and Cognitive Therapies to find a list of the recommended therapists and evidence-based treatments.

<http://www.abct.org>

Web: Visit the Depression and Bipolar Support Alliance for educational information and social support options.

<http://www.dbsalliance.org/>

DISCUSSION QUESTIONS

1. What factors might explain the large gender

difference in the prevalence rates of MDD?

2. Why might American ethnic minority groups experience more persistent BD than European Americans?
3. Why might the age of onset for MDD be decreasing over time?
4. Why might overnight travel constitute a potential risk for a person with BD?
5. What are some reasons positive life events may precede the occurrence of manic episode?

VOCABULARY

Anhedonia

Loss of interest or pleasure in activities one previously found enjoyable or rewarding.

Attributional style

The tendency by which a person infers the cause or meaning of behaviors or events.

Chronic stress

Discrete or related problematic events and conditions which persist over time and result in prolonged activation of the biological and/or psychological stress response (e.g., unemployment, ongoing health difficulties, marital discord).

Early adversity

Single or multiple acute or chronic stressful events, which may be biological or psychological in nature (e.g., poverty, abuse, childhood illness or injury), occurring during childhood and resulting in a

biological and/or psychological stress response.

Grandiosity

Inflated self-esteem or an exaggerated sense of self-importance and self-worth (e.g., believing one has special powers or superior abilities).

Hypersomnia

Excessive daytime sleepiness, including difficulty staying awake or napping, or prolonged sleep episodes.

Psychomotor agitation

Increased motor activity associated with restlessness, including physical actions (e.g., fidgeting, pacing, feet tapping, handwringing).

Psychomotor retardation

A slowing of physical activities in which routine activities (e.g., eating, brushing teeth) are performed in an unusually slow manner.

Social zeitgeber

Zeitgeber is German for “time giver.” Social zeitgebers are environmental cues, such as meal times and interactions with other people, that entrain biological rhythms and thus sleep-wake cycle regularity.

Socioeconomic status (SES)

A person’s economic and social position based on income, education, and occupation.

Suicidal ideation

Recurring thoughts about suicide, including

considering or planning for suicide, or preoccupation with suicide.

REFERENCES

- Abramson, L. Y, Seligman, M. E. P., & Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology*, 87, 49–74. doi: 10.1037/0021-843X.87.1.49
- Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, 96, 358–373. doi: 10.1037/0022-3514.56.3.431
- Altshuler, L., Bookheimer, S., Townsend, J., Proenza, M. A., Sabb, F., Mintz, J., & Cohen, M. S. (2008). Regional brain changes in bipolar I depression: A functional magnetic resonance imaging study. *Bipolar Disorders*, 10, 708–717. doi: 10.1111/j.1399-5618.2008.00617.x
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Beck, A. T. (1967). *Depression: Clinical, experimental, and theoretical aspects*. New York, NY: Hoeber.
- Blazer, D., George, L. K., Landerman, R., Pennybacker, M., Melville, M. L., Woodbury, M.,

- et al. (1985). Psychiatric disorders. A rural/urban comparison. *Archives of General Psychiatry*, 42, 651–656. PMID: 4015306. doi: 10.1001/archpsyc.1985.01790300013002
- Breslau, J., Aguilar-Gaxiola, S., Kendler, K. S., Su, M., Williams, D., & Kessler, R. C. (2006). Specifying race-ethnic differences in risk for psychiatric disorder in a US national sample. *Psychological Medicine*, 36, 57–68. doi: 10.1017/S0033291705006161
 - Breslau, J., Kendler, K. S., Su, M., Gaxiola-Aguilar, S., & Kessler, R. C. (2005). Lifetime risk and persistence of psychiatric disorders across ethnic groups in the United States. *Psychological Medicine*, 35, 317–327. doi: 10.1017/S0033291704003514
 - Brown, G. W., & Harris, T. O. (1989). *Life events and illness*. New York, NY: Guilford Press.
 - Butzlaff, R. L., & Hooley, J. M. (1998). Expressed emotion and psychiatric relapse: A meta-analysis. *Archives of General Psychiatry*, 55, 547–552. doi: 10.1001/archpsyc.55.6.547
 - Chentsova-Dutton, Y. E., & Tsai, J. L. (2009). Understanding depression across cultures. In I. H. Gotlib & C.L. Hammen (Eds.), *Handbook of depression* (2nd ed., pp. 363–385). New York, NY: Guilford Press.
 - Conley, C. S., & Rudolph, K. D. (2009). The emerging sex difference in adolescent depression: Interacting contributions of puberty and peer stress. *Development and Psychopathology*, 21, 593–620. doi: 10.1017/S0954579409000327

- Dunner, D. L. (2000). Optimizing lithium treatment. *Journal of Clinical Psychiatry*, 61(S9), 76–81.
- Ehlers, C. L., Frank, E., & Kupfer, D. J. (1988). Social zeitgebers and biological rhythms: a unified approach to understanding the etiology of depression. *Archives of General Psychiatry*, 45, 948–952. doi: 10.1001/archpsyc.1988.01800340076012
- Ellicott, A., Hammen, C., Gitlin, M., Brown, G., & Jamison, K. (1990). Life events and the course of bipolar disorder. *American Journal of Psychiatry*, 147, 1194–1198.
- Frank, E., Kupfer, D. J., Ehlers, C. L., Monk, T., Cornes, C., Carter, S., et al. (1994). Interpersonal and social rhythm therapy for bipolar disorder: Integrating interpersonal and behavioral approaches. *Behavior Therapy*, 17, 143–149.
- Frank, E., Kupfer, D. J., Thase, M. E., Mallinger, A. G., Swartz, H. A., Fagiolini, A. M., et al. (2005). Two-year outcomes for interpersonal and social rhythm therapy in individuals with bipolar I disorder. *Archives of General Psychiatry*, 62, 996–1004. doi: 10.1001/archpsyc.62.9.996
- Geddes, J. R., Burgess, S., Hawton, K., Jamison, K., & Goodwin, G. M. (2004). Long-term lithium therapy for bipolar disorder: systematic review and meta-analysis of randomized controlled trials. *American Journal of Psychiatry*, 161, 217–222. doi: 10.1176/appi.ajp.161.2.217
- Gonzalez, J. M., Perlick, D. A., Miklowitz, D. J.,

- Kaczynski, R., Hernandez, M., Rosenheck, R. A., et al. (2007). Factors associated with stigma among caregivers of patients with bipolar disorder in the STEP-BD study. *Psychiatric Services*, 58, 41–48. doi: 10.1176/appi.ps.58.1.41
- Goodwin, F. K., & Jamison, K. R. (2007). *Manic-depressive illness: Bipolar disorders and recurrent depression*. New York, NY: Oxford University Press.
 - Gotlib, I. H., & Joormann, J. (2010). Cognition and depression: Current status and future directions. *Annual Review of Clinical Psychology*, 6, 285–312. doi: 10.1146/annurev.clinpsy.121208.131305
 - Hamilton, J. P., Etkin, A., Furman, D. F., Lemus, M. G., Johnson, R. F., & Gotlib, I. H. (2012). Functional neuroimaging of major depressive disorder: A meta-analysis and new integration of baseline activation and neural response data. *American Journal of Psychiatry*, 169, 693–703.
 - Hammen, C. (2005). Stress and depression. *Annual Review of Clinical Psychology*, 1, 293–319. doi: 10.1146/annurev.clinpsy.1.102803.143938
 - Harvey, A. G. (2008). Sleep and Circadian Rhythms in Bipolar Disorder: Seeking synchrony, harmony and regulation. *American Journal of Psychiatry*, 165, 820–829. doi: 10.1176/appi.ajp.2008.08010098
 - Hasin, D. S., Goodwin, R. D., Sintson, F. S., & Grant, B. F. (2005). Epidemiology of major depressive disorder: Results from the National

Epidemiological Survey on Alcoholism and Related Conditions. *Archives of General Psychiatry*, 62, 1097–1106. doi: 10.1001/archpsyc.62.10.1097

- Hassel, S., Almeida, J. R., Kerr, N., Nau, S., Ladouceur, C. D., Fissell, K., et al. (2008). Elevated striatal and decreased dorsolateral prefrontal cortical activity in response to emotional stimuli in euthymic bipolar disorder: No associations with psychotropic medication load. *Bipolar Disorders*, 10, 916–927. doi: 10.1111/j.1399-5618.2008.00641.x
- Isaacowitz, D. M., Gershon, A., Allard, E. S., & Johnson, S. L. (2013). Emotion in aging and bipolar disorder: Similarities, differences and lessons for further research. *Emotion Review*, 5, 312–320. doi: 10.1177/1754073912472244
- Johnson, S. L. (2005). Mania and dysregulation in goal pursuit: A review. *Clinical Psychology Review*, 25, 241–262. doi: 10.1016/j.cpr.2004.11.002
- Johnson, S. L., Cueller, A. K., Ruggero, C., Winett-Perlman, C., Goodnick, P., White, R., et al. (2008). Life events as predictors of mania and depression in bipolar I disorder. *Journal of Abnormal Psychology*, 117, 268–277. doi: 10.1037/0021-843X.117.2.268
- Johnson, S. L., Winett, C. A., Meyer, B., Greenhouse, W. J., & Miller, I. (1999). Social support and the course of bipolar disorder. *Journal of Abnormal Psychology*, 108, 558–566. doi: 10.1037/0021-843X.108.4.558
- Kessler, R. C., Berglund, P., Demler, O., Jim, R.,

- Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, 62, 593–602. doi: 10.1001/archpsyc.62.6.593
- Kessler, R. C., Birnbaum, H., Bromet, E., Hwang, I., Sampson, N., & Shahly, V. (2010). Age differences in major depression: Results from the National Comorbidity Surveys Replication (NCS-R). *Psychological Medicine*, 40, 225–237. doi: 10.1017/S0033291709990213
 - Kessler, R. C., Merikangas, K. R., & Wang, P. S. (2007). Prevalence, comorbidity, and service utilization for mood disorders in the United States at the beginning of the 21st century. *Annual Review of Clinical Psychology*, 3, 137–158. doi: 10.1146/annurev.clinpsy.3.022806.091444
 - Kilbourne, A. M., Haas, G. L., Mulsant, B. H., Bauer, M. S., & Pincus, H. A. (2004) Concurrent psychiatric diagnoses by age and race among persons with bipolar disorder. *Psychiatric Services*, 55, 931–933. doi: 10.1176/appi.ps.55.8.931
 - Leichenring, F. (2001). Comparative effects of short-term psychodynamic psychotherapy and cognitive-behavioral therapy in depression: A meta-analytic approach. *Clinical Psychology Review*, 21, 401–419. doi: 10.1016/S0272-7358(99)00057-4
 - Lennox, B. R., Jacob, R., Calder, A. J., Lupson, V., & Bullmore, E. T. (2004). Behavioural and

neurocognitive responses to sad facial affect are attenuated in patients with mania. *Psychological Medicine*, 34, 795–802. doi: 10.1017/S0033291704002557

- Lenox, R. H., & Hahn C. G. (2000). Overview of the mechanism of action of lithium in the brain: fifty-year update. *Journal of Clinical Psychiatry*, 61 (S9), 5–15.
- Leverich, G. S., Altshuler, L. L., Frye, M. A., Suppes, T., Keck, P. E. Jr, McElroy, S. L., et al. (2003). Factors associated with suicide attempts in 648 patients with bipolar disorder in the Stanley Foundation Bipolar Network. *Journal of Clinical Psychiatry*, 64, 506–515. doi: 10.4088/JCP.v64n0503
- Leverich, G. S., Post, R. M., Keck, P. E. Jr, Altshuler, L. L., Frye, M. A., Kupka, R. W., et al. (2007). The poor prognosis of childhood-onset bipolar disorder. *Journal of Pediatrics*, 150, 485–490. PMID: 17452221. doi: 10.1016/j.jpeds.2006.10.070
- Lewinsohn, P. M., Allen, N. B., Seeley, J. R., & Gotlib, I. H. (1999). First onset versus recurrence of depression: differential processes of psychosocial risk. *Journal of Abnormal Psychology*, 108, 483–489. doi: 10.1037/0021-843X.108.3.483
- Lewinsohn, P. M., Seeley, J. R., Buckley, M. E., & Klein, D. N. (2002). Bipolar disorder in adolescence and young adulthood. *Child & Adolescent Psychiatric Clinics of North America*, 11, 461–475. doi: 10.1016/S1056-4993(02)00005-6

- Lohoff, F. W. (2010). Overview of genetics of major depressive disorder. *Current Psychiatry Reports*, 12, 539–546. doi: 10.1007/s11920-010-0150-6
- Lorant, V., Deliege, D., Eaton, W., Robert, A., Philippot, P., & Ansseau, A. (2003). Socioeconomic inequalities in depression: A meta-analysis. *American Journal of Epidemiology*, 157, 98–112. doi: 10.1093/aje/kwf182
- Malkoff-Schwartz, S., Frank, E., Anderson, B. P., Sherrill, J. T., Siegel, L., Patterson, D., et al. (1998). Stressful life events and social rhythm disruption in the onset of manic and depressive bipolar episodes: a preliminary investigation. *Archives of General Psychiatry*, 55, 702–707. doi: 10.1001/archpsyc.55.8.702
- Mayberg, H. S., Lozano, A. M., Voon, V., McNeely, H. E., Seminowicz, D., Hamani, C., Schwab, J. M., & Kennedy, S. H. (2005). Deep brain stimulation for treatment-resistant depression. *Neuron*, 45, 651–660. doi: 10.1016/j.neuron.2005.02.014
- McElroy, S. L., Altshuler, L. L., Suppes, T., Keck, P. E. Jr, Frye, M. A., Denicoff, K. D., et al. (2001). Axis I psychiatric comorbidity and its relationship to historical illness variables in 288 patients with bipolar disorder. *American Journal of Psychiatry*, 158, 420–426. doi: 10.1176/appi.ajp.158.3.420
- McGuffin, P., Rijsdijk, F., Andrew, M., Sham, P., Katz, R., Cardno, A. (2003). The heritability of bipolar affective disorder and the genetic relationship to unipolar depression. *Archives of*

General Psychiatry, 60, 497–502. doi: 10.1001/archpsyc.60.5.497

- Merikangas, K. R., Akiskal, H. S., Angst, J., Greenberg, P. E., Hirschfeld, R. M., Petukhova, M., et al. (2007). Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication. *Archives of General Psychiatry*, 64, 543–552. doi: 10.1001/archpsyc.64.5.543
- Merikangas, K. R., Jin, R., He, J. P., Kessler, R. C., Lee, S., Sampson, N. A., et al. (2011). Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. *Archives of General Psychiatry*, 68, 241–251. doi: 10.1001/archgenpsychiatry.2011.12
- Minsky, S., Vega, W., Miskimen, T., Gara, M., & Escobar, J. (2003). Diagnostic patterns in Latino, African American, and European American psychiatric patients. *Archives of General Psychiatry*, 60, 637–644. doi: 10.1001/archpsyc.60.6.637
- Monroe, S. M., & Harkness, K. L. (2011). Recurrence in major depression: A conceptual analysis. *Psychological Review*, 118, 655–674. doi: 10.1037/a0025190
- Monroe, S. M., & Harkness, K. L. (2005). Life stress, the “Kindling” hypothesis, and the recurrence of depression: Considerations from a life stress perspective. *Psychological Review*, 112, 417–445. doi: 10.1037/0033-295X.112.2.417
- Monroe, S.M., Slavich, G.M., Georgiades, K. (2009). The social environment and life stress in

depression. In Gotlib, I.H., Hammen, C.L (Eds.) *Handbook of depression* (2nd ed., pp. 340-360). New York, NY: Guilford Press.

- Nolen-Hoeksema, S., & Hilt, L. M. (2009). Gender differences in depression. In I. H. Gotlib & Hammen, C. L. (Eds.), *Handbook of depression* (2nd ed., pp. 386–404). New York, NY: Guilford Press.
- Perlis, R. H., Miyahara, S., Marangell, L. B., Wisniewski, S. R., Ostacher, M., DelBello, M. P., et al. (2004). Long-term implications of early onset in bipolar disorder: data from the first 1000 participants in the systematic treatment enhancement program for bipolar disorder (STEP-BD). *Biological Psychiatry*, 55, 875–881. PMID: 15110730. doi: 10.1016/j.psychresns.2007.10.003
- Rosa, M. A., Gattaz, W. F., Pascual-Leone, A., Fregni, F., Rosa, M. O., Rumi, D. O., ... Marcolin, M. A. (2006). Comparison of repetitive transcranial magnetic stimulation and electroconvulsive therapy in unipolar non-psychotic refractory depression: a randomized, single-blind study. *International Journal of Neuropsychopharmacology*, 9, 667–676. doi: 10.1017/S1461145706007127
- Schulze-Rauschenbach, S. C., Harms, U., Schlaepfer, T. E., Maier, W., Falkai, P., & Wagner, M. (2005). Distinctive neurocognitive effects of repetitive transcranial magnetic stimulation and electroconvulsive therapy in major depression. *British Journal of Psychiatry*, 186,

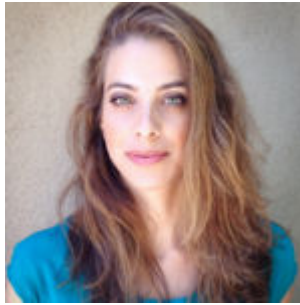
410–416. doi: 10.1192/bjp.186.5.410

- Shields, B. (2005). *Down Came the Rain: My Journey Through Postpartum Depression*. New York: Hyperion.
- Sullivan, P., Neale, M. C., & Kendler, K. S. (2000). Genetic epidemiology of major depression: Review and meta-analysis. *American Journal of Psychiatry*, 157, 1552–1562. doi: 10.1176/appi.ajp.157.10.1552
- Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, SES, and birth cohort differences on the Children's Depression Inventory: A meta-analysis. *Journal of Abnormal Psychology*, 111, 578–588. doi: 10.1037/0021-843X.111.4.578
- Whisman, M. A., & Uebelacker, L. A. (2009). Prospective associations between marital discord and depressive symptoms in middle-aged and older adults. *Psychology and Aging*, 24, 184–189. doi: 10.1037/a0014759
- Widom, C. S., DuMont, K., & Czaja, S. J. (2007). A prospective investigation of major depressive disorder and comorbidity in abused and neglected children grown up. *Archives of General Psychiatry*, 64, 49–56. doi: 10.1001/archpsyc.64.1.49
- Williams, D. R., Gonzalez, H. M., Neighbors, H., Nesse, R., Abelson, J. M., Sweetman, J., & Jackson, J. S. (2007). Prevalence and distribution of major depressive disorder in African Americans, Caribbean blacks, and non-Hispanic whites: Results from the National Survey of American Life. *Archives of General Psychiatry*, 64, 305–315.

doi: 10.1001/archpsyc.64.3.305

- Wingo, A. P., Wingo, T. S., Harvey, P. D., & Baldessarini, R. J. (2009). Effects of lithium on cognitive performance: a meta-analysis. *Journal of Clinical Psychiatry*, 70, 1588–1597. doi: 10.4088/JCP.08r04972

AUTHORS



Anda Gershon

Anda Gershon is an Instructor at the Department of Psychiatry and Behavioral Sciences at Stanford University. Her research focuses on the contributions of life stress and sleep disturbances to the onset and maintenance of mood disorders, using longitudinal and ecologically sensitive methods.



Renee Thompson

Renee J. Thompson is an assistant professor at the Department of Psychology at Washington University in St. Louis. Her research centers on understanding the emotional experience in individuals with depression and anxiety. She uses a multi-method approach, including ecological momentary assessment, to examine temporal dynamics of emotional experience.

CREATIVE COMMONS LICENSE



MOOD DISORDERS by [Anda Gershon and Renee Thompson](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond the scope of this license may be available in our [Licensing Agreement](#).

CHAPTER 21.

SCHIZOPHRENIA SPECTRUM DISORDERS

DEANNA M. BARCH

Schizophrenia and the other psychotic disorders are some of the most impairing forms of psychopathology, frequently associated with a profound negative effect on the individual's educational, occupational, and social function. Sadly, these disorders often manifest right at time of the transition from adolescence to adulthood, just as young people should be evolving into independent young adults. The spectrum of psychotic disorders includes schizophrenia, schizoaffective disorder, delusional disorder, schizotypal personality disorder, schizophreniform disorder, brief psychotic disorder, as well as psychosis associated with substance use or medical conditions. In this module, we summarize the primary clinical features of these disorders, describe the known cognitive and neurobiological changes associated with schizophrenia, describe potential risk factors and/or causes for the development of schizophrenia, and describe currently available treatments for schizophrenia.

LEARNING OBJECTIVES

- Describe the signs and symptoms of schizophrenia and related psychotic disorders.
- Describe the most well-replicated cognitive and neurobiological changes associated with schizophrenia.
- Describe the potential risk factors for the development of schizophrenia.
- Describe the controversies associated with “clinical high risk” approaches to identifying individuals at risk for the development of schizophrenia.
- Describe the treatments that work for some of the symptoms of schizophrenia.

THE PHENOMENOLOGY OF SCHIZOPHRENIA AND RELATED PSYCHOTIC DISORDERS

Most of you have probably had the experience of walking down the street in a city and seeing a person you thought was acting oddly. They may have been dressed in an unusual way, perhaps disheveled or wearing an unusual collection of clothes, makeup, or jewelry that did not seem to fit any particular group or subculture. They may have been talking to themselves or yelling at someone you could not see. If you tried to speak to them, they may have been difficult to follow or understand, or they may have acted paranoid or started telling a bizarre story about the people who were plotting against them. If so, chances are that you have encountered an individual with schizophrenia or another type of psychotic disorder. If

you have watched the movie *A Beautiful Mind* or *The Fisher King*, you have also seen a portrayal of someone thought to have schizophrenia. Sadly, a few of the individuals who have committed some of the recently highly publicized mass murders may have had schizophrenia, though most people who commit such crimes do not have schizophrenia. It is also likely that you have met people with schizophrenia without ever knowing it, as they may suffer in silence or stay isolated to protect themselves from the horrors they see, hear, or believe are operating in the outside world. As these examples begin to illustrate, psychotic disorders involve many different types of symptoms, including delusions, hallucinations, disorganized speech and behavior, abnormal motor behavior (including **catatonia**), and negative symptoms such **anhedonia/amotivation** and blunted affect/reduced speech.

Delusions are false beliefs that are often fixed, hard to change even when the person is presented with conflicting information, and are often culturally influenced in their content (e.g., delusions involving Jesus in Judeo-Christian cultures, delusions involving Allah in Muslim cultures). They can be terrifying for the person, who may remain convinced that they are true even when loved ones and friends present them with clear information that they cannot be true. There are many different types or themes to delusions.



Under Surveillance: Abstract groups like the police or the government are commonly the focus of a schizophrenic's persecutory delusions.
 [Image: Thomas Hawk, <https://goo.gl/qsrqiR>, CC BY-NC 2.0, <https://goo.gl/VnKIK8>]

The most common delusions are persecutory and involve the belief that individuals or groups are trying to hurt, harm, or plot against the person in some way. These can be people that the person knows (people at work, the neighbors, family members), or more abstract groups (the FBI, the CIA, aliens, etc.). Other types of delusions include grandiose delusions, where the person believes that they

have some special power or ability (e.g., I am the new Buddha, I am a rock star); referential delusions, where the person believes that events or objects in the environment have special meaning for them (e.g., that song on the radio is being played *specifically* for me); or other types of delusions where the person may believe that others are controlling their thoughts and actions, their thoughts are being broadcast aloud, or that others can read their mind (or they can read other people's minds).

When you see a person on the street talking to themselves or shouting at other people, they are experiencing **hallucinations**. These are perceptual experiences that occur even when there is no stimulus in the outside world generating the experiences. They can be auditory, visual, olfactory (smell), gustatory (taste), or somatic (touch). The most common hallucinations in

psychosis (at least in adults) are auditory, and can involve one or more voices talking about the person, commenting on the person's behavior, or giving them orders. The content of the hallucinations is frequently negative ("you are a loser," "that drawing is stupid," "you should go kill yourself") and can be the voice of someone the person knows or a complete stranger. Sometimes the voices sound as if they are coming from outside the person's head. Other times the voices seem to be coming from inside the person's head, but are not experienced the same as the person's inner thoughts or inner speech.

Talking to someone with schizophrenia is sometimes difficult, as their speech may be difficult to follow, either because their answers do not clearly flow from your questions, or because one sentence does not logically follow from another. This is referred to as **disorganized speech**, and it can be present even when the



People who suffer from schizophrenia may see the world differently. This can include hallucinations, delusions, and disorganized thinking. [Image: Noba Project CCBYNCSA 4.0 <https://tinyurl.com/y3k6qoz4>]

person is writing. **Disorganized behavior** can include odd dress, odd makeup (e.g., lipstick outlining a mouth for 1 inch), or unusual rituals (e.g., repetitive hand gestures). Abnormal motor behavior can include catatonia, which refers to a variety of behaviors that seem to reflect a reduction in responsiveness to the external environment. This can include holding unusual postures for long periods of time, failing to respond to verbal or

motor prompts from another person, or excessive and seemingly purposeless motor activity.

Some of the most debilitating symptoms of schizophrenia are difficult for others to see. These include what people refer to as “negative symptoms” or the absence of certain things we typically expect most people to have. For example, anhedonia or amotivation reflect a lack of apparent interest in or drive to engage in social or recreational activities. These symptoms can manifest as a great amount of time spent in physical immobility. Importantly, anhedonia and amotivation do not seem to reflect a lack of enjoyment in pleasurable activities or events ([Cohen & Minor, 2010](#); [Kring & Moran, 2008](#); [Llerena, Strauss, & Cohen, 2012](#)) but rather a reduced drive or ability to take the steps necessary to obtain the potentially positive outcomes ([Barch & Dowd, 2010](#)). **Flat affect** and reduced speech (**alogia**) reflect a lack of showing emotions through facial expressions, gestures, and speech intonation, as well as a reduced amount of speech and increased pause frequency and duration.

In many ways, the types of symptoms associated with psychosis are the most difficult for us to understand, as they may seem far outside the range of our normal experiences. Unlike depression or anxiety, many of us may not have had experiences that we think of as on the same continuum as psychosis. However, just like many of the other forms of **psychopathology** described in this book, the types of psychotic symptoms that characterize disorders like schizophrenia are on a continuum with “normal” mental experiences. For example, work by Jim van Os in the Netherlands has shown that a surprisingly large percentage of the general population (10%+)

experience psychotic-like symptoms, though many fewer have multiple experiences and most will not continue to experience these symptoms in the long run ([Verdoux & van Os, 2002](#)). Similarly, work in a general population of adolescents and young adults in Kenya has also shown that a relatively high percentage of individuals experience one or more psychotic-like experiences (~19%) at some point in their lives ([Mamah et al., 2012](#); [Ndetei et al., 2012](#)), though again most will not go on to develop a full-blown psychotic disorder.

Schizophrenia is the primary disorder that comes to mind when we discuss “psychotic” disorders (see Table 1 for diagnostic criteria), though there are a number of other disorders that share one or more features with schizophrenia. In the remainder of this module, we will use the terms “psychosis” and “schizophrenia” somewhat interchangeably, given that most of the research has focused on schizophrenia. In addition to schizophrenia (see Table 1), other psychotic disorders include schizophreniform disorder (a briefer version of schizophrenia), schizoaffective disorder (a mixture of psychosis and depression/mania symptoms), delusional disorder (the experience of only delusions), and brief psychotic disorder (psychotic symptoms that last only a few days or weeks).

Schizophrenia (Lifetime prevalence about 0.3% to 0.7% [APA, 2013])
• Two or more of the following for at least 1 month: hallucinations, delusions, disorganized speech, grossly disorganized or catatonic behavior, negative symptoms. • Impairment in one or more areas of function (social, occupational, educational self-care) for a significant period of time since the onset of the illness. • Continuous signs of the illness for at least 6 months (this can include prodromal or residual symptoms, which are attenuated forms of the symptoms described above).
Schizophreniform Disorder (Lifetime prevalence similar to Schizophrenia [APA, 2013])
• The same symptoms of schizophrenia described above that are present for at least 1 month but less than 6 months.
Schizoaffective Disorder (Lifetime prevalence about 0.3% [APA, 2013])
• A period of illness where the person has both the psychotic symptoms necessary to meet criteria for schizophrenia and either a major depression or manic episode. • The person experiences either delusions or hallucinations for at least 2 weeks when they are not having a depressive or manic episode. • The symptoms that meet criteria for depressive or manic episodes are present for over half of the illness duration.
Delusional Disorder (Lifetime prevalence about 0.2% [APA, 2013])
• The presence of at least one delusion for at least a month. • The person has never met criteria for schizophrenia. • The person's function is not impaired outside the specific impact of the delusion. • The duration of any depressive or manic episodes have been brief relative to the duration of the delusion(s).
Brief Psychotic Disorder (Lifetime prevalence unclear [APA, 2013])
• One or more of the following symptoms present for at least 1 day but less than 1 month: delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior.
Attenuated Psychotic Disorder (In Section III of the [APA, 2013]-V, Lifetime presence unclear [APA, 2013])
• One or more of the following symptoms in an "attenuated" form: delusions, hallucinations, or disorganized speech. • The symptoms must have occurred at least once a week for the past month and must have started or gotten worse in the past year. • The symptoms must be severe enough to distress or disable the individual or to suggest to others that the person needs clinical help. • The person has never met the diagnostic criteria for a psychotic disorder, and the symptoms are not better attributed to another disorder, to substance use, or to a medical condition.

Table 1: Types of Psychotic Disorders (Simplified from the Diagnostic and Statistical Manual – 5th Edition (DSM-5) (APA, 2013)

THE COGNITIVE NEUROSCIENCE OF SCHIZOPHRENIA

As described above, when we think of the core symptoms of psychotic disorders such as schizophrenia, we think of people who hear voices, see visions, and have false beliefs about reality (i.e., delusions). However, problems in cognitive function are also a critical aspect of psychotic disorders and of schizophrenia in particular. This emphasis on cognition in schizophrenia is in part due to the growing body of research suggesting that cognitive

problems in schizophrenia are a major source of disability and loss of **functional capacity** ([Green, 2006](#); [Nuechterlein et al., 2011](#)). The cognitive deficits that are present in schizophrenia are widespread and can include problems with **episodic memory** (the ability to learn and retrieve new information or episodes in one's life), **working memory** (the ability to maintain information over a short period of time, such as 30 seconds), and other tasks that require one to "control" or regulate one's behavior ([Barch & Ceaser, 2012](#); [Bora, Yucel, & Pantelis, 2009a](#); [Fioravanti, Carlone, Vitale, Cinti, & Clare, 2005](#); [Forbes, Carrick, McIntosh, & Lawrie, 2009](#); [Mesholam-Gately, Giuliano, Goff, Faraone, & Seidman, 2009](#)). Individuals with schizophrenia also have difficulty with what is referred to as "**processing speed**" and are frequently slower than healthy individuals on almost all tasks. Importantly, these cognitive deficits are present prior to the onset of the illness ([Fusar-Poli et al., 2007](#)) and are also present, albeit in a milder form, in the first-degree relatives of people with schizophrenia ([Snitz, Macdonald, & Carter, 2006](#)). This suggests that cognitive impairments in schizophrenia reflect part of the risk for the development of psychosis, rather than being an outcome of developing psychosis. Further, people with schizophrenia who have more severe cognitive problems also tend to have more severe negative symptoms and more disorganized speech and behavior ([Barch, Carter, & Cohen, 2003](#); [Barch et al., 1999](#); [Dominguez Mde, Viechtbauer, Simons, van Os, & Krabbendam, 2009](#); [Ventura, Hellemann, Thames, Koellner, & Nuechterlein, 2009](#); [Ventura, Thames, Wood, Guzik, & Hellemann, 2010](#)). In addition, people with more cognitive problems have worse function in everyday life

([Bowie et al., 2008](#); [Bowie, Reichenberg, Patterson, Heaton, & Harvey, 2006](#); [Fett et al., 2011](#)).



Some with schizophrenia suffer from difficulty with social cognition. They may not be able to detect the meaning of facial expressions or other subtle cues that most other people rely on to navigate the social world. [Image: Ralph Buckley, <https://goo.gl/KuBzsD>, CC BY-SA 2.0, <https://goo.gl/i4GXf5>]

Some people with schizophrenia also show deficits in what is referred to as social cognition, though it is not clear whether such problems are separate from the cognitive problems described above or the result of them ([Hoe, Nakagami, Green, & Brekke, 2012](#); [Kerr & Neale, 1993](#); [van Hooren et al., 2008](#)). This includes problems with the recognition of emotional expressions on the faces of other individuals ([Kohler,](#)

[Walker, Martin, Healey, & Moberg, 2010](#)) and problems inferring the intentions of other people (theory of mind) ([Bora, Yucel, & Pantelis, 2009b](#)). Individuals with schizophrenia who have more problems with social cognition also tend to have more negative and disorganized symptoms ([Ventura, Wood, & Helleman, 2011](#)), as well as worse community function ([Fett et al., 2011](#)).

The advent of neuroimaging techniques such as structural and functional **magnetic resonance imaging** and **positron emission tomography** opened up the ability to try to understand the brain mechanisms of the symptoms of schizophrenia as well as the cognitive

impairments found in psychosis. For example, a number of studies have suggested that delusions in psychosis may be associated with problems in “salience” detection mechanisms supported by the ventral striatum ([Jensen & Kapur, 2009](#); [Jensen et al., 2008](#); [Kapur, 2003](#); [Kapur, Mizrahi, & Li, 2005](#); [Murray et al., 2008](#)) and the anterior prefrontal cortex ([Corlett et al., 2006](#); [Corlett, Honey, & Fletcher, 2007](#); [Corlett, Murray, et al., 2007a, 2007b](#)). These are regions of the brain that normally increase their activity when something important (aka “salient”) happens in the environment. If these brain regions misfire, it may lead individuals with psychosis to mistakenly attribute importance to irrelevant or unconnected events. Further, there is good evidence that problems in working memory and cognitive control in schizophrenia are related to problems in the function of a region of the brain called the dorsolateral prefrontal cortex (DLPFC) ([Minzenberg, Laird, Thelen, Carter, & Glahn, 2009](#); [Ragland et al., 2009](#)). These problems include changes in how the DLPFC works when people are doing working-memory or cognitive-control tasks, and problems with how this brain region is connected to other brain regions important for working memory and cognitive control, including the posterior parietal cortex (e.g., [Karlsgodt et al., 2008](#); [J. J. Kim et al., 2003](#); [Schlosser et al., 2003](#)), the anterior cingulate ([Repovs & Barch, 2012](#)), and temporal cortex (e.g., [Fletcher et al., 1995](#); [Meyer-Lindenberg et al., 2001](#)). In terms of understanding episodic memory problems in schizophrenia, many researchers have focused on medial temporal lobe deficits, with a specific focus on the hippocampus (e.g., [Heckers & Konradi, 2010](#)). This is because there is much data from humans and animals

showing that the hippocampus is important for the creation of new memories ([Squire, 1992](#)). However, it has become increasingly clear that problems with the DLPFC also make important contributions to episodic memory deficits in schizophrenia ([Ragland et al., 2009](#)), probably because this part of the brain is important for controlling our use of memory.

In addition to problems with regions such as the DLPFC and medial temporal lobes in schizophrenia described above, magnitude resonance neuroimaging studies have also identified changes in cellular architecture, white matter connectivity, and gray matter volume in a variety of regions that include the prefrontal and temporal cortices ([Bora et al., 2011](#)). People with schizophrenia also show reduced overall brain volume, and reductions in brain volume as people get older may be larger in those with schizophrenia than in healthy people ([Olabi et al., 2011](#)). Taking antipsychotic medications or taking drugs such as marijuana, alcohol, and tobacco may cause some of these structural changes. However, these structural changes are not completely explained by medications or substance use alone. Further, both functional and structural brain changes are seen, again to a milder degree, in the first-degree relatives of people with schizophrenia ([Boos, Aleman, Cahn, Pol, & Kahn, 2007](#); [Brans et al., 2008](#); [Fusar-Poli et al., 2007](#); [MacDonald, Thermenos, Barch, & Seidman, 2009](#)). This again suggests that that neural changes associated with schizophrenia are related to a genetic risk for this illness.

RISK FACTORS FOR DEVELOPING SCHIZOPHRENIA

It is clear that there are important genetic contributions to the likelihood that someone will develop schizophrenia, with consistent evidence from family, twin, and adoption studies. ([Sullivan, Kendler, & Neale, 2003](#)). However, there is no “schizophrenia gene” and it is likely that the genetic risk for schizophrenia reflects the summation of many different genes that each contribute something to the likelihood of developing psychosis ([Gottesman & Shields, 1967](#); [Owen, Craddock, & O'Donovan, 2010](#)). Further, schizophrenia is a very heterogeneous disorder, which means that two different people with “schizophrenia” may each have very different symptoms (e.g., one has hallucinations and delusions, the other has disorganized speech and negative symptoms). This makes it even more challenging to identify specific genes associated with risk for psychosis. Importantly, many studies also now suggest that at least some of the genes potentially associated with schizophrenia are also associated with other mental health conditions, including bipolar disorder, depression, and autism ([Gejman, Sanders, & Kendler, 2011](#); [Y. Kim, Zerwas, Trace, & Sullivan, 2011](#); [Owen et al., 2010](#); [Rutter, Kim-Cohen, & Maughan, 2006](#)).



There are a number of genetic and environmental risk factors associated with higher likelihood of developing schizophrenia including older fathers, complications during pregnancy/delivery, family history of schizophrenia, and growing up in an urban environment. [Image: CCO Public Domain]

There are also a number of environmental factors that are associated with an increased risk of developing schizophrenia. For example, problems during pregnancy such as increased stress, infection, malnutrition, and/or diabetes have been associated with increased risk of schizophrenia. In addition, complications that occur at the time of birth and which cause hypoxia (lack of oxygen) are also associated with an increased risk for developing schizophrenia

([M. Cannon, Jones, & Murray, 2002](#); [Miller et al., 2011](#)).

Children born to older fathers are also at a somewhat increased risk of developing schizophrenia. Further, using cannabis increases risk for developing psychosis, especially if you have other risk factors ([Casadio, Fernandes, Murray, & Di Forti, 2011](#); [Luzi, Morrison, Powell, di Forti, & Murray, 2008](#)). The likelihood of developing schizophrenia is also higher for kids who grow up in urban settings ([March et al., 2008](#)) and for some minority ethnic groups ([Bourque, van der Ven, & Malla, 2011](#)). Both of these factors may reflect higher social and environmental stress in these settings. Unfortunately, none of these risk factors is specific enough to be particularly useful in a clinical setting, and

most people with these “risk” factors do not develop schizophrenia. However, together they are beginning to give us clues as the **neurodevelopmental** factors that may lead someone to be at an increased risk for developing this disease.

An important research area on risk for psychosis has been work with individuals who may be at “clinical high risk.” These are individuals who are showing attenuated (milder) symptoms of psychosis that have developed recently and who are experiencing some distress or disability associated with these symptoms. When people with these types of symptoms are followed over time, about 35% of them develop a psychotic disorder ([T. D. Cannon et al., 2008](#)), most frequently schizophrenia ([Fusar-Poli, McGuire, & Borgwardt, 2012](#)). In order to identify these individuals, a new category of diagnosis, called “Attenuated Psychotic Syndrome,” was added to Section III (the section for disorders in need of further study) of the DSM-5 (see Table 1 for symptoms) ([APA, 2013](#)). However, adding this diagnostic category to the DSM-5 created a good deal of controversy ([Batstra & Frances, 2012](#); [Fusar-Poli & Yung, 2012](#)). Many scientists and clinicians have been worried that including “risk” states in the DSM-5 would create mental disorders where none exist, that these individuals are often already seeking treatment for other problems, and that it is not clear that we have good treatments to stop these individuals from developing to psychosis. However, the counterarguments have been that there is evidence that individuals with high-risk symptoms develop psychosis at a much higher rate than individuals with other types of psychiatric symptoms, and that the inclusion of Attenuated Psychotic Syndrome in Section III will spur

important research that might have clinical benefits. Further, there is some evidence that non-invasive treatments such as omega-3 fatty acids and intensive family intervention may help reduce the development of full-blown psychosis ([Preti & Cella, 2010](#)) in people who have high-risk symptoms.

TREATMENT OF SCHIZOPHRENIA

The currently available treatments for schizophrenia leave much to be desired, and the search for more effective treatments for both the psychotic symptoms of schizophrenia (e.g., hallucinations and delusions) as well as cognitive deficits and negative symptoms is a highly active area of research. The first line of treatment for schizophrenia and other psychotic disorders is the use of antipsychotic medications. There are two primary types of antipsychotic medications, referred to as “typical” and “atypical.” The fact that “typical” antipsychotics helped some symptoms of schizophrenia was discovered serendipitously more than 60 years ago ([Carpenter & Davis, 2012](#); [Lopez-Munoz et al., 2005](#)). These are drugs that all share a common feature of being a strong block of the D2 type **dopamine** receptor. Although these drugs can help reduce hallucinations, delusions, and disorganized speech, they do little to improve cognitive deficits or negative symptoms and can be associated with distressing motor side effects. The newer generation of antipsychotics is referred to as “atypical” antipsychotics. These drugs have more mixed mechanisms of action in terms of the receptor types that they influence, though most of them also influence D2 receptors. These newer antipsychotics are not necessarily more helpful for

schizophrenia but have fewer motor side effects. However, many of the atypical antipsychotics are associated with side effects referred to as the “metabolic syndrome,” which includes weight gain and increased risk for cardiovascular illness, Type-2 diabetes, and mortality ([Lieberman et al., 2005](#)).

The evidence that cognitive deficits also contribute to functional impairment in schizophrenia has led to an increased search for treatments that might enhance cognitive function in schizophrenia. Unfortunately, as of yet, there are no pharmacological treatments that work consistently to improve cognition in schizophrenia, though many new types of drugs are currently under exploration. However, there is a type of psychological intervention, referred to as cognitive remediation, which has shown some evidence of helping cognition and function in schizophrenia. In particular, a version of this treatment called Cognitive Enhancement Therapy (CET) has been shown to improve cognition, functional outcome, social cognition, and to protect against gray matter loss ([Eack et al., 2009](#); [Eack, Greenwald, Hogarty, & Keshavan, 2010](#); [Eack et al., 2010](#); [Eack, Pogue-Geile, Greenwald, Hogarty, & Keshavan, 2010](#); [Hogarty, Greenwald, & Eack, 2006](#)) in young individuals with schizophrenia. The development of new treatments such as Cognitive Enhancement Therapy provides some hope that we will be able to develop new and better approaches to improving the lives of individuals with this serious mental health condition and potentially even prevent it some day.

OUTSIDE RESOURCES

Book: *Ben Behind His Voices: One family's journal from the chaos of schizophrenia to hope* (2011). Randye Kaye. Rowman and Littlefield.

Book: *Conquering Schizophrenia: A father, his son, and a medical breakthrough* (1997). Peter Wyden. Knopf.

Book: *Henry's Demons: Living with schizophrenia, a father and son's story* (2011). Henry and Patrick Cockburn. Scribner Macmillan.

Book: *My Mother's Keeper: A daughter's memoir of growing up in the shadow of schizophrenia* (1997). Tara Elgin Holley. William Morrow Co.

Book: *Recovered, Not Cured: A journey through schizophrenia* (2005). Richard McLean. Allen and Unwin.

Book: *The Center Cannot Hold: My journey through madness* (2008). Elyn R. Saks. Hyperion.

Book: *The Quiet Room: A journal out of the torment of madness* (1996). Lori Schiller. Grand Central Publishing.

Book: *Welcome Silence: My triumph over schizophrenia* (2003). Carol North. CSS Publishing.

Web: National Alliance for the Mentally Ill. This is an excellent site for learning more about advocacy for individuals with major mental illnesses such as schizophrenia.

<http://www.nami.org/>

Web: National Institute of Mental Health. This website has information on NIMH-funded schizophrenia research.

<http://www.nimh.nih.gov/health/topics/schizophrenia/index.shtml>

Web: Schizophrenia Research Forum. This is an excellent website that contains a broad array of information about current research on schizophrenia.

<http://www.schizophreniaforum.org/>

DISCUSSION QUESTIONS

1. Describe the major differences between the major psychotic disorders.
2. How would one be able to tell when an individual is “delusional” versus having non-delusional beliefs that differ from the societal normal? How should cultural and sub-cultural variation been taken into account when assessing psychotic symptoms?
3. Why are cognitive impairments important to understanding schizophrenia?
4. Why has the inclusion of a new diagnosis (Attenuated Psychotic Syndrome) in Section III of the DSM-5 created controversy?
5. What are some of the factors associated with

increased risk for developing schizophrenia? If we know whether or not someone has these risk factors, how well can we tell whether they will develop schizophrenia?

6. What brain changes are most consistent in schizophrenia?
7. Do antipsychotic medications work well for all symptoms of schizophrenia? If not, which symptoms respond better to antipsychotic medications?
8. Are there any treatments besides antipsychotic medications that help any of the symptoms of schizophrenia? If so, what are they?

VOCABULARY

Alogia

A reduction in the amount of speech and/or increased pausing before the initiation of speech.

Anhedonia/amotivation

A reduction in the drive or ability to take the steps or engage in actions necessary to obtain the potentially positive outcome.

Catatonia

Behaviors that seem to reflect a reduction in responsiveness to the external environment. This can include holding unusual postures for long periods of time, failing to respond to verbal or motor prompts from another person, or excessive and seemingly purposeless motor activity.

Delusions

False beliefs that are often fixed, hard to change even in the presence of conflicting information, and often culturally influenced in their content.

Diagnostic criteria

The specific criteria used to determine whether an individual has a specific type of psychiatric disorder. Commonly used diagnostic criteria are included in the Diagnostic and Statistical Manual of Mental Disorder, 5th Edition (DSM-5) and the International Classification of Disorders, Version 9 (ICD-9).

Disorganized behavior

Behavior or dress that is outside the norm for almost all subcultures. This would include odd dress, odd makeup (e.g., lipstick outlining a mouth for 1 inch), or unusual rituals (e.g., repetitive hand gestures).

Disorganized speech

Speech that is difficult to follow, either because answers do not clearly follow questions or because one sentence does not logically follow from another.

Dopamine

A neurotransmitter in the brain that is thought to play an important role in regulating the function of other neurotransmitters.

Episodic memory

The ability to learn and retrieve new information or episodes in one's life.

Flat affect

A reduction in the display of emotions through facial expressions, gestures, and speech intonation.

Functional capacity

The ability to engage in self-care (cook, clean, bathe), work, attend school, and/or engage in social relationships.

Hallucinations

Perceptual experiences that occur even when there is no stimulus in the outside world generating the experiences. They can be auditory, visual, olfactory (smell), gustatory (taste), or somatic (touch).

Magnetic resonance imaging

A set of techniques that uses strong magnets to measure either the structure of the brain (e.g., gray matter and white matter) or how the brain functions when a person performs cognitive tasks (e.g., working memory or episodic memory) or other types of tasks.

Neurodevelopmental

Processes that influence how the brain develops either in utero or as the child is growing up.

Positron emission tomography

A technique that uses radio-labelled ligands to measure the distribution of different neurotransmitter receptors in the brain or to measure how much of a certain type of neurotransmitter is released when a person is given a specific type of drug or does a particularly

cognitive task.

Processing speed

The speed with which an individual can perceive auditory or visual information and respond to it.

Psychopathology

Illnesses or disorders that involve psychological or psychiatric symptoms.

Working memory

The ability to maintain information over a short period of time, such as 30 seconds or less.

REFERENCES

- APA. (2013). *Diagnostic and statistical manual of mental disorders, Fifth Edition* (5th ed.). Washington, DC: American Psychiatric Association.
- Barch, D. M., & Ceaser, A. E. (2012). Cognition in schizophrenia: Core psychological and neural mechanisms. *Trends in Cognitive Science*, 16, 27–34.
- Barch, D. M., & Dowd, E. C. (2010). Goal representations and motivational drive in schizophrenia: The role of prefrontal-striatal interactions. *Schizophrenia Bulletin*, 36(5), 919–934. doi: sbq068 [pii] 10.1093/schbul/sbq068
- Barch, D. M., Carter, C. S., & Cohen, J. D. (2003). Context processing deficit in schizophrenia: Diagnostic specificity, 4-week course, and relationships to clinical symptoms. *Journal of Abnormal Psychology*, 112, 132–143.
- Barch, D. M., Carter, C. S., Macdonald, A., Sabb, F.

- W., Noll, D. C., & Cohen, J. D. (1999). Prefrontal cortex and context processing in medication-naïve first-episode patients with schizophrenia. *Schizophrenia Research*, 36(1–3), 217–218.
- Batstra, L., & Frances, A. (2012). Diagnostic inflation: Causes and a suggested cure. *The Journal of Nervous and Mental Disease*, 200(6), 474–479. doi: 10.1097/NMD.0b013e318257c4a2
 - Boos, H. B., Aleman, A., Cahn, W., Pol, H. H., & Kahn, R. S. (2007). Brain volumes in relatives of patients with schizophrenia: A meta-analysis. *Archives of General Psychiatry*, 64(3), 297–304.
 - Bora, E., Fornito, A., Radua, J., Walterfang, M., Seal, M., Wood, S. J., . . . Pantelis, C. (2011). Neuroanatomical abnormalities in schizophrenia: A multimodal voxelwise meta-analysis and meta-regression analysis. *Schizophrenia Research*, 127(1–3), 46–57. doi: 10.1016/j.schres.2010.12.020
 - Bora, E., Yucel, M., & Pantelis, C. (2009a). Cognitive functioning in schizophrenia, schizoaffective disorder and affective psychoses: Meta-analytic study. *The British Journal of Psychiatry: The Journal of Mental Science*, 195(6), 475–482. doi: 10.1192/bjp.bp.108.055731
 - Bora, E., Yucel, M., & Pantelis, C. (2009b). Theory of mind impairment in schizophrenia: Meta-analysis. *Schizophrenia Research*, 109(1–3), 1–9. doi: 10.1016/j.schres.2008.12.020

- Bourque, F., van der Ven, E., & Malla, A. (2011). A meta-analysis of the risk for psychotic disorders among first- and second-generation immigrants. *Psychological Medicine*, 41(5), 897–910. doi: 10.1017/S0033291710001406
- Bowie, C. R., Leung, W. W., Reichenberg, A., McClure, M. M., Patterson, T. L., Heaton, R. K., & Harvey, P. D. (2008). Predicting schizophrenia patients' real-world behavior with specific neuropsychological and functional capacity measures. *Biological Psychiatry*, 63(5), 505–511. doi: 10.1016/j.biopsych.2007.05.022
- Bowie, C. R., Reichenberg, A., Patterson, T. L., Heaton, R. K., & Harvey, P. D. (2006). Determinants of real-world functional performance in schizophrenia subjects: Correlations with cognition, functional capacity, and symptoms. *The American Journal of Psychiatry*, 163(3), 418–425. doi: 10.1176/appi.ajp.163.3.418
- Brans, R. G., van Haren, N. E., van Baal, G. C., Schnack, H. G., Kahn, R. S., & Hulshoff Pol, H. E. (2008). Heritability of changes in brain volume over time in twin pairs discordant for schizophrenia. *Archives of General Psychiatry*, 65(11), 1259–1268. doi: 10.1001/archpsyc.65.11.1259
- Cannon, M., Jones, P. B., & Murray, R. M. (2002). Obstetric complications and schizophrenia: Historical and meta-analytic review. *The American Journal of Psychiatry*, 159(7), 1080–1092.
- Cannon, T. D., Cadenhead, K., Cornblatt, B.,

Woods, S. W., Addington, J., Walker, E., . . . Heinssen, R. (2008). Prediction of psychosis in youth at high clinical risk: A multisite longitudinal study in North America. *Archives of General Psychiatry*, 65(1), 28–37.

- Carpenter, W. T., Jr., & Davis, J. M. (2012). Another view of the history of antipsychotic drug discovery and development. *Molecular Psychiatry*, 17(12), 1168–1173. doi: 10.1038/mp.2012.121
- Casadio, P., Fernandes, C., Murray, R. M., & Di Forti, M. (2011). Cannabis use in young people: The risk for schizophrenia. *Neuroscience & Biobehavioral Reviews*. doi: S0149-7634(11)00073-X [pii] 10.1016/j.neubiorev.2011.04.007
- Cohen, A. S., & Minor, K. S. (2010). Emotional experience in patients with schizophrenia revisited: Meta-analysis of laboratory studies. *Schizophrenia Bulletin*, 36(1), 143–150. doi: 10.1093/schbul/sbn061
- Corlett, P. R., Honey, G. D., & Fletcher, P. C. (2007). From prediction error to psychosis: Ketamine as a pharmacological model of delusions. *Journal of Psychopharmacology*, 21(3), 238–252. doi: 21/3/238 [pii] 10.1177/0269881107077716
- Corlett, P. R., Honey, G. D., Aitken, M. R., Dickinson, A., Shanks, D. R., Absalom, A. R., . . . Fletcher, P. C. (2006). Frontal responses during learning predict vulnerability to the psychotogenic effects of ketamine: Linking

cognition, brain activity, and psychosis. *Archives of General Psychiatry*, 63(6), 611–621. doi: 63/6/611 [pii] 10.1001/archpsyc.63.6.611

- Corlett, P. R., Murray, G. K., Honey, G. D., Aitken, M. R., Shanks, D. R., Robbins, T. W., . . . Fletcher, P. C. (2007a). Disrupted prediction-error signal in psychosis: Evidence for an associative account of delusions. *Brain: A Journal of Neurology*, 130(Pt 9), 2387–2400. doi: 10.1093/brain/awm173
- Corlett, P. R., Murray, G. K., Honey, G. D., Aitken, M. R., Shanks, D. R., Robbins, T. W., . . . Fletcher, P. C. (2007b). Disrupted prediction-error signal in psychosis: Evidence for an associative account of delusions. *Brain*, 130(Pt 9), 2387–2400. doi: awm173 [pii] 10.1093/brain/awm173
- Dominguez Mde, G., Viechtbauer, W., Simons, C. J., van Os, J., & Krabbendam, L. (2009). Are psychotic psychopathology and neurocognition orthogonal? A systematic review of their associations. *Psychological Bulletin*, 135(1), 157–171. doi: 10.1037/a0014415
- Eack, S. M., Greenwald, D. P., Hogarty, S. S., & Keshavan, M. S. (2010). One-year durability of the effects of cognitive enhancement therapy on functional outcome in early schizophrenia. *Schizophrenia Research*, 120(1–3), 210–216. doi: S0920-9964(10)01222-3 [pii] 10.1016/j.schres.2010.03.042
- Eack, S. M., Greenwald, D. P., Hogarty, S. S., Cooley, S. J., DiBarry, A. L., Montrose, D. M., & Keshavan, M. S. (2009). Cognitive enhancement

therapy for early-course schizophrenia: effects of a two-year randomized controlled trial. *Psychiatr Serv*, 60(11), 1468–1476. doi: 60/11/1468 [pii] 10.1176/appi.ps.60.11.1468

- Eack, S. M., Hogarty, G. E., Cho, R. Y., Prasad, K. M., Greenwald, D. P., Hogarty, S. S., & Keshavan, M. S. (2010). Neuroprotective effects of cognitive enhancement therapy against gray matter loss in early schizophrenia: Results from a 2-year randomized controlled trial. *Archives of General Psychiatry*, 67(7), 674–682. doi: 2010.63 [pii] 10.1001/archgenpsychiatry.2010.63
- Eack, S. M., Pogue-Geile, M. F., Greenwald, D. P., Hogarty, S. S., & Keshavan, M. S. (2010). Mechanisms of functional improvement in a 2-year trial of cognitive enhancement therapy for early schizophrenia. *Psychological Medicine*, 1–9. doi: S0033291710001765 [pii] 10.1017/S0033291710001765
- Fett, A. K., Viechtbauer, W., Dominguez, M. D., Penn, D. L., van Os, J., & Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: A meta-analysis. *Neuroscience and Biobehavioral Reviews*, 35(3), 573–588. doi: 10.1016/j.neubiorev.2010.07.001
- Fioravanti, M., Carlone, O., Vitale, B., Cinti, M. E., & Clare, L. (2005). A meta-analysis of cognitive deficits in adults with a diagnosis of schizophrenia. *Neuropsychology Review*, 15(2), 73–95. doi: 10.1007/s11065-005-6254-9

- Fletcher, P. C., Frith, C. D., Grasby, P. M., Shallice, T., Frackowiak, R. S. J., & Dolan, R. J. (1995). Brain systems for encoding and retrieval of auditory-verbal memory: An in vivo study in humans. *Brain*, 118, 401–416.
- Forbes, N. F., Carrick, L. A., McIntosh, A. M., & Lawrie, S. M. (2009). Working memory in schizophrenia: A meta-analysis. *Psychological Medicine*, 39(6), 889–905. doi: 10.1017/S0033291708004558
- Fusar-Poli, P., & Yung, A. R. (2012). Should attenuated psychosis syndrome be included in DSM-5? *Lancet*, 379(9816), 591–592. doi: 10.1016/S0140-6736(11)61507-9
- Fusar-Poli, P., McGuire, P., & Borgwardt, S. (2012). Mapping prodromal psychosis: A critical review of neuroimaging studies. *European Psychiatry: The Journal of the Association of European Psychiatrists*, 27(3), 181–191. doi: 10.1016/j.eurpsy.2011.06.006
- Fusar-Poli, P., Perez, J., Broome, M., Borgwardt, S., Placentino, A., Caverzasi, E., . . . McGuire, P. (2007). Neurofunctional correlates of vulnerability to psychosis: A systematic review and meta-analysis. *Neuroscience and Biobehavioral Reviews*, 31(4), 465–484.
- Gejman, P. V., Sanders, A. R., & Kendler, K. S. (2011). Genetics of schizophrenia: New findings and challenges. *Annual Review of Genomics and Human Genetics*. doi: 10.1146/annurev-genom-082410-101459

- Gottesman, I. I., & Shields, J. (1967). A polygenic theory of schizophrenia. *Proceedings of the National Academy of Sciences of the United States of America*, 58(1), 199–205.
- Green, M. F. (2006). Cognitive impairment and functional outcome in schizophrenia and bipolar disorder. *The Journal of Clinical Psychiatry*, 67 Suppl 9, 3–8; discussion 36–42.
- Heckers, S., & Konradi, C. (2010). Hippocampal pathology in schizophrenia. *Current Topics in Behavioral Neurosciences*, 4, 529–553.
- Hoe, M., Nakagami, E., Green, M. F., & Brekke, J. S. (2012). The causal relationships between neurocognition, social cognition, and functional outcome over time in schizophrenia: A latent difference score approach. *Psychological Medicine*, 1–13. doi: 10.1017/S0033291712000578
- Hogarty, G. E., Greenwald, D. P., & Eack, S. M. (2006). Durability and mechanism of effects of cognitive enhancement therapy. *Psychiatric Services*, 57(12), 1751–1757. doi: 57/12/1751 [pii] 10.1176/appi.ps.57.12.1751
- Jensen, J., & Kapur, S. (2009). Salience and psychosis: Moving from theory to practise. *Psychological Medicine*, 39(2), 197–198. doi: 10.1017/S0033291708003899
- Jensen, J., Willeit, M., Zipursky, R. B., Savina, I., Smith, A. J., Menon, M., . . . Kapur, S. (2008). The formation of abnormal associations in schizophrenia: Neural and behavioral evidence. *Neuropsychopharmacology: Official*

Publication of the American College of Neuropsychopharmacology, 33(3), 473–479. doi: 10.1038/sj.npp.1301437

- Kapur, S. (2003). Psychosis as a state of aberrant salience: A framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry*, 160(1), 13–23.
- Kapur, S., Mizrahi, R., & Li, M. (2005). From dopamine to salience to psychosis—linking biology, pharmacology and phenomenology of psychosis. *Schizophrenia Research*, 79(1), 59–68. doi: 10.1016/j.schres.2005.01.003
- Karlsgodt, K. H., van Erp, T. G., Poldrack, R. A., Bearden, C. E., Nuechterlein, K. H., & Cannon, T. D. (2008). Diffusion tensor imaging of the superior longitudinal fasciculus and working memory in recent-onset schizophrenia. *Biological Psychiatry*, 63(5), 512–518.
- Kerr, S. L., & Neale, J. M. (1993). Emotion perception in schizophrenia: Specific deficit or further evidence of generalized poor performance? *Journal of Abnormal Psychology*, 102(2), 312–318.
- Kim, J. J., Kwon, J. S., Park, H. J., Youn, T., Kang, D. H., Kim, M. S., . . . Lee, M. C. (2003). Functional disconnection between the prefrontal and parietal cortices during working memory processing in schizophrenia: A [15O]H₂O PET study. *American Journal of Psychiatry*, 160, 919–923.
- Kim, Y., Zerwas, S., Trace, S. E., & Sullivan, P. F.

(2011). Schizophrenia genetics: Where next? *Schizophrenia Bulletin*, 37(3), 456–463. doi: sbr031 [pii] 10.1093/schbul/sbr031

- Kohler, C. G., Walker, J. B., Martin, E. A., Healey, K. M., & Moberg, P. J. (2010). Facial emotion perception in schizophrenia: A meta-analytic review. *Schizophrenia Bulletin*, 36(5), 1009–1019. doi: 10.1093/schbul/sbn192
- Kring, A. M., & Moran, E. K. (2008). Emotional response deficits in schizophrenia: Insights from affective science. *Schizophrenia Bulletin*, 34(5), 819–834.
- Lieberman, J. A., Stroup, T. S., McEvoy, J. P., Swartz, M. S., Rosenheck, R. A., Perkins, D. O., . . . Hsiao, J. K. (2005). Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. *The New England Journal of Medicine*, 353(12), 1209–1223. doi: 10.1056/NEJMoa051688
- Llerena, K., Strauss, G. P., & Cohen, A. S. (2012). Looking at the other side of the coin: A meta-analysis of self-reported emotional arousal in people with schizophrenia. *Schizophrenia Research*, 142(1–3), 65–70. doi: 10.1016/j.schres.2012.09.005
- Lopez-Munoz, F., Alamo, C., Cuenca, E., Shen, W. W., Clervoy, P., & Rubio, G. (2005). History of the discovery and clinical introduction of chlorpromazine. *Annals of Clinical Psychiatry: Official Journal of the American Academy of Clinical Psychiatrists*, 17(3), 113–135.
- Luzi, S., Morrison, P. D., Powell, J., di Forti, M., &

Murray, R. M. (2008). What is the mechanism whereby cannabis use increases risk of psychosis? *Neurotoxicity Research*, 14(2–3), 105–112. doi: 10.1007/BF03033802

- MacDonald, A. W., III, Thermenos, H. W., Barch, D. M., & Seidman, L. J. (2009). Imaging genetic liability to schizophrenia: Systematic review of fMRI studies of patients' nonpsychotic relatives. *Schizophrenia Bulletin*, 35(6), 1142–1162.
- Mamah, D., Mbwai, A., Mutiso, V., Barch, D. M., Constantino, J. N., Nsofor, T., . . . Ndeti, D. M. (2012). A survey of psychosis risk symptoms in Kenya. *Comprehensive Psychiatry*, 53(5), 516–524. doi: 10.1016/j.comppsy.2011.08.003
- March, D., Hatch, S. L., Morgan, C., Kirkbride, J. B., Bresnahan, M., Fearon, P., & Susser, E. (2008). Psychosis and place. *Epidemiologic Reviews*, 30, 84–100. doi: 10.1093/epirev/mxn006
- Mesholam-Gately, R. I., Giuliano, A. J., Goff, K. P., Faraone, S. V., & Seidman, L. J. (2009). Neurocognition in first-episode schizophrenia: A meta-analytic review. *Neuropsychology*, 23(3), 315–336. doi: 10.1037/a0014708
- Meyer-Lindenberg, A., Poline, J., Kohn, P. D., Holt, J. L., Egan, M. F., Weinberger, D. R., & Berman, K. F. (2001). Evidence for abnormal cortical functional connectivity during working memory in schizophrenia. *American Journal of Psychiatry*, 158, 1809–1817.
- Miller, B., Messias, E., Miettinen, J., Alaraisanen, A., Jarvelin, M. R., Koponen, H., . . . Kirkpatrick, B.

- (2011). Meta-analysis of paternal age and schizophrenia risk in male versus female offspring. *Schizophrenia Bulletin*, 37(5), 1039–1047. doi: 10.1093/schbul/sbq011
- Minzenberg, M. J., Laird, A. R., Thelen, S., Carter, C. S., & Glahn, D. C. (2009). Meta-analysis of 41 functional neuroimaging studies of executive function in schizophrenia. *Archives of General Psychiatry*, 66(8), 811–822. doi: 10.1001/archgenpsychiatry.2009.91
 - Murray, G. K., Corlett, P. R., Clark, L., Pessiglione, M., Blackwell, A. D., Honey, G., . . . Fletcher, P. C. (2008). Substantia nigra/ventral tegmental reward prediction error disruption in psychosis. *Molecular Psychiatry*, 13(3), 267–276.
 - Ndeti, D. M., Muriungi, S. K., Owoso, A., Mutiso, V. N., Mbwai, A. W., Khasakhala, L. I., . . . Mamah, D. (2012). Prevalence and characteristics of psychotic-like experiences in Kenyan youth. *Psychiatry Research*, 196(2–3), 235–242. doi: 10.1016/j.psychres.2011.12.053
 - Nuechterlein, K. H., Subotnik, K. L., Green, M. F., Ventura, J., Asarnow, R. F., Gitlin, M. J., . . . Mintz, J. (2011). Neurocognitive predictors of work outcome in recent-onset schizophrenia. *Schizophrenia Bulletin*, 37 Suppl 2, S33–40. doi: 10.1093/schbul/sbr084
 - Olabi, B., Ellison-Wright, I., McIntosh, A. M., Wood, S. J., Bullmore, E., & Lawrie, S. M. (2011). Are there progressive brain changes in schizophrenia? A meta-analysis of structural

magnetic resonance imaging studies. *Biological Psychiatry*, 70(1), 88–96. doi: 10.1016/j.biopsych.2011.01.032

- Owen, M. J., Craddock, N., & O'Donovan, M. C. (2010). Suggestion of roles for both common and rare risk variants in genome-wide studies of schizophrenia. *Archives of General Psychiatry*, 67(7), 667–673. doi: 10.1001/archgenpsychiatry.2010.69
- Preti, A., & Cella, M. (2010). Randomized-controlled trials in people at ultra high risk of psychosis: a review of treatment effectiveness. *Schizophrenia Research*, 123(1), 30–36. doi: 10.1016/j.schres.2010.07.026
- Ragland, J. D., Laird, A. R., Ranganath, C., Blumenfeld, R. S., Gonzales, S. M., & Glahn, D. C. (2009). Prefrontal activation deficits during episodic memory in schizophrenia. *American Journal of Psychiatry*, 166(8), 863–874.
- Repovs, G., & Barch, D. M. (2012). Working memory related brain network connectivity in individuals with schizophrenia and their siblings. *Frontiers in Human Neuroscience*, 6, 137. doi: 10.3389/fnhum.2012.00137
- Rutter, M., Kim-Cohen, J., & Maughan, B. (2006). Continuities and discontinuities in psychopathology between childhood and adult life. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 47(3–4), 276–295. doi: 10.1111/j.1469-7610.2006.01614.x
- Schlosser, R., Gesierich, T., Kaufmann, B., Vucurevic, G., Hunsche, S., Gawehn, J., & Stoeter,

- P. (2003). Altered effective connectivity during working memory performance in schizophrenia: A study with fMRI and structural equation modeling. *Neuroimage*, 19(3), 751–763.
- Snitz, B. E., Macdonald, A. W., 3rd, & Carter, C. S. (2006). Cognitive deficits in unaffected first-degree relatives of schizophrenia patients: A meta-analytic review of putative endophenotypes. *Schizophrenia Bulletin*, 32(1), 179–194.
 - Squire, L.R. (1992). Memory and the hippocampus: A synthesis from findings with rats, monkeys, and humans. *Psychological Review*, 99, 195–231.
 - Sullivan, P. F., Kendler, K. S., & Neale, M. C. (2003). Schizophrenia as a complex trait: Evidence from a meta-analysis of twin studies. *Archives of General Psychiatry*, 60(12), 1187–1192. doi: 10.1001/archpsyc.60.12.1187
 - Ventura, J., Helleman, G. S., Thames, A. D., Koellner, V., & Nuechterlein, K. H. (2009). Symptoms as mediators of the relationship between neurocognition and functional outcome in schizophrenia: a meta-analysis. *Schizophrenia Research*, 113(2–3), 189–199. doi: 10.1016/j.schres.2009.03.035
 - Ventura, J., Thames, A. D., Wood, R. C., Guzik, L. H., & Helleman, G. S. (2010). Disorganization and reality distortion in schizophrenia: a meta-analysis of the relationship between positive symptoms and neurocognitive

deficits. *Schizophrenia Research*, 121(1–3), 1–14.
doi: 10.1016/j.schres.2010.05.033

- Ventura, J., Wood, R. C., & Helleman, G. S. (2011). Symptom domains and neurocognitive functioning can help differentiate social cognitive processes in schizophrenia: A meta-analysis. *Schizophrenia Bulletin*. doi: 10.1093/schbul/sbr067
- Verdoux, H., & van Os, J. (2002). Psychotic symptoms in non-clinical populations and the continuum of psychosis. *Schizophrenia Research*, 54(1–2), 59–65.
- van Hooren, S., Versmissen, D., Janssen, I., Myin-Germeys, I., a Campo, J., Mengelers, R., . . . Krabbendam, L. (2008). Social cognition and neurocognition as independent domains in psychosis. *Schizophrenia Research*, 103(1–3), 257–265. doi: 10.1016/j.schres.2008.02.022

AUTHORS



Deanna M. Barch

Deanna M. Barch, the Gregory B. Couch Professor of Psychology, Psychiatry and Radiology at Washington University in Saint Louis, received the APA's Distinguished Scientific Award for Early Career Contributions to Psychology in Psychopathology and the Joseph Zubin Memorial Fund Award, and is a Fellow of the Association for Psychological Science.

CREATIVE COMMONS LICENSE



SCHIZOPHRENIA SPECTRUM

DISORDERS by [Deanna M. Barch](#) is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](#). Permissions beyond

the scope of this license may be available in our [Licensing Agreement](#).

LINK TO THE 1ST EDITION, MUCH THANKS TO THE NOBA PROJECT

The first edition of this work can be found at <https://nobaproject.com/textbooks/emily-marler-understanding-biological-behavior-first-edition>.