Chapter 1: Introduction to the Textbook
Carolyn Dicey Jennings and Benjamin D. Young

Brief outline:

The chapter provides an overview of the structure and content of the textbook to help situate the reader. It begins by introducing this unique collaborative project, including a general introduction to the fields of philosophy, cognitive science, and neuroscience. It then segues into explaining the structural features of each chapter that provide uniformity across the textbook. The chapter concludes with an overview of the content provided in the textbook. Through a survey of the major themes and their interconnections the reader will be better able to appreciate the rich interplay between philosophy and cognitive neuroscience in each of the topical chapters.

Keywords

Cognitive psychology: a subfield of psychology devoted to studying thoughts, thinking, and the psychological processes responsible for our mental states.

Cognitive science: a broad multidisciplinary field that seeks to study cognition (understood as above) from a variety of perspectives using a multiplicity of research methods and experimental paradigms.

Neuroscience: narrowly understood, neuroscience studies the workings of the neural system. As an interdisciplinary field it utilizes research from biology, medicine, chemistry, computational science, math, physics, and psychology.

Cognitive neuroscience: a subfield of neuroscience that employs the methods and knowledge base of neuroscience to study cognition and mental phenomena. While there are still some labs, departments, and institutes purely devoted to cognitive science in the classical sense, cognitive science has transitioned into cognitive neuroscience in the last two decades.

Philosophy: perhaps the oldest academic discipline, philosophy is devoted to the study of a wide range of phenomena using a wide range of methods, including formal logic, ethics, and phenomenology. The subset of philosophy dedicated to the study of mind tends to be more theoretical and sometimes employs introspective methods.

Philosophy of cognitive neuroscience: a combination of philosophy of science and empirically informed philosophy, covering issues regarding both the practice of cognitive neuroscience and the evidence yielded by cognitive neuroscience.

Naturalism: a philosophical approach that seeks to explain all aspects of reality, especially phenomena studied by the sciences, purely in terms of naturally occurring objects, properties, and relations.
Chapter 2: Introduction to Cognitive Neuroscience
Adina L. Roskies, Dartmouth College

BRIEF OUTLINE:
1. Introduction:
   a. Overview: What is cognitive neuroscience?
   b. Foundational assumptions
      1. Materialism
      2. Information-processing model of mind
      3. Localization of function
      4. Biology constrains function
      5. Levels of processing/description
   c. Methods in cognitive neuroscience: a historical introduction
2. Contemporary issues
   a. Representation: What is the neural code?
   b. Transformations, processing
   c. Modularity vs. holism
   d. Context dependence
3. Future directions
   a. New technologies
   b. Focus on social and affective neuroscience
   c. Applied cognitive neuroscience: Social justice
   d. Philosophical foundations: Reductionism

BRIEF OUTLINE: BOX
This chapter introduces cognitive neuroscience. It begins with a brief historical overview, charting how different disciplines and methods combined to illuminate how the brain gives rise to mind and cognition. It then turns to the foundational assumptions of cognitive neuroscience and provides a historical view of key methods. Section 2 addresses a selection of contemporary issues in cognitive neuroscience, including the nature of neural representation, transformations or neural processing, the modularity debate, and the importance of context-dependence. Section 3 discusses future directions, including the development of new technologies, increasing focus on high-level and social phenomena, potential applications, such as implications for social justice, and foundational philosophical questions that still need to be addressed, such as the way in which high-level psychological constructs relate to lower-level biology.

KEY TERMS and DEFINITIONS:

Aphasia – loss of ability to understand or express speech, caused by brain damage.
Bayesianism – relating to or involving statistical methods that assign probabilities or distributions to variables based on priors and that apply Bayes' theorem to revise the probabilities and distributions after obtaining new evidence.
Chapter 3: Introduction to Molecular and Cellular Cognition
John Bickle and Ann Sophie Barwich

Since the 1990s, molecular and cellular cognition (MCC) has elucidated critical causal mechanisms behind higher-level processes of perception and cognition. New intervention techniques, including genetic tracers and fluorescent visualizations, facilitated targeted access and manipulation of molecular pathways to investigate their precise role in cognition and link to behavior. Insights from MCC transformed central ideas of cognition, especially memory. This Chapter introduces key developments in late 20th century neuroscience that explain how the discovery of the molecular mechanisms modified theoretical understanding of memory as a cognitive function. It further looks at contemporary innovations in optical imaging microscopy that further exemplify the central argument that technological innovations in the study of cellular processes constitute fundamental drivers also of high-level cognitive theories.

Keywords:
Florescence: physical property of absorbing light of shorter wavelength and emitting light of longer wavelength.
Genetics: field of biology studying inheritance of traits (and variations).
Long-term potentiation (LTP): lasting strengthening of synapses due to recent activity in connected neurons.
Memory: facility of organism to store, access, and use information about past experiences
Neuronal signaling pathway: series of biochemical reactions within and between neurons controlling cellular functions.
Receptor protein: typically embedded in a biological membrane, with a highly specific affinity for binding agents to initiate cellular activity.
Reductionism: philosophical program of explaining all complex phenomena in terms of fundamental or at least simpler phenomena.
Synaptic plasticity: capacity for synapses to strengthen and weaken over time in response to neuronal activities.

1. Introduction: Molecularizing the Mind

Neuroscientific research starts with the premise that cognitive processing results from cellular physiology, and that cellular processes result from molecular mechanisms. Intuitively, the distance between these mechanisms and cognition seems vast. Introductions to cognitive neuroscience seldom cover this gap. They usually start with some details of membrane and action potentials, and synaptic transmission, explaining briefly how neurons fire and communicate signals throughout the brain. The molecular machinery behind neuronal activity gets presented as a necessary evil to be endured en route to the “real” topics of the neural correlates of cognition. Examination of molecular and cellular mechanisms often stops there, such that molecular and cognitive accounts appear oddly disconnected.

A remarkably successful area in mainstream neurobiology has been bridging the gap between molecules and mind for over a quarter-century: “molecular and cellular cognition” (MCC). Knowing something about how MCC researchers design and execute experiments is important for all students of cognitive science, so it’s critical to demystify how that molecular machinery presents us with materialized cognition.
Chapter 4: Introduction to Experimental Methods in Cognitive Neuroscience
Kristina C. Backer

Topic Outline:

One of nature’s great mysteries is how the brain gives rise to our thoughts, our feelings, and our behavior. Although there are still many unknowns, the pieces of this puzzle are slowly coming together—thanks to classic methodologies and to technological advances that have given rise to modern brain imaging techniques. This chapter provides an introduction to how cognitive neuroscientists approach today’s cutting-edge brain research. We start with some general concepts and terminology that apply to all experiments. Then, we explore the various methodologies that are used to measure brain function and brain structure. We end the chapter with a discussion of the importance of conducting research with rigor and transparency.

Key Terms and Definitions:

General Experimental Terms:

Research question: a specific question that addresses a gap in knowledge about the phenomenon you would like to research.

Hypothesis: a statement about the expected results based on your observations and the scientific literature on the topic.

Independent variable: the variable that the researcher experimentally manipulates.

Dependent variable: a variable that the researcher experimentally measures.

Confounding variable (or confound): a variable that affects the dependent variable(s) but was not intentionally manipulated in the study.

Neuroanatomy Terms:

Gyrus (plural: gyri): the flat plateaus of cortical tissue observed when viewing the surface of the cerebral cortex.

Sulcus (plural: sulci): the crevices between the gyri of the cerebral cortex.

Gray matter: brain tissue that primarily contains the neuronal cell bodies.

White matter: brain tissue that primarily contains myelinated axons.

Ventricle: a cavity within the brain that produces and contains cerebrospinal fluid. The brain contains four ventricles, which are interconnected and also serve to cushion the brain.

Method-Specific Terms:

Electroencephalography (EEG): a method to non-invasively record the brain’s neural activity via the electrical brain potentials, which are produced by tens of thousands of neurons and are detected on the scalp’s surface.
Chapter 5: Introduction to philosophy of mind
Joseph M. Vukov

Outline of topics covered

Our everyday discourse appeals to both mental and physical phenomena. On the mental side: I may say I have a desire for a second slice of cake; you may hope to secure a high-paying job; a friend may feel anxiety at their prospects on the academic job market. Desires, beliefs, hopes, anxieties: all are mental phenomena. We also regularly refer to physical phenomena: my physician may describe me as weighing 84 kg at my annual visit; I may describe my office as located close to the copy room; a neuroscientist may inform me I have elevated levels of serotonin. Weight, location, serotonin levels: all are physical phenomena.

Philosophy of mind is the discipline that explores the relationship between mental and physical phenomena. Of course, several chapters in this book explore that relationship. But most explore it by focusing on specific features of human psychology or cognitive processing. This chapter instead focuses on theories aiming to “carve nature at its joints” (Plato Phaedrus 265d-266a) by describing the relationship between mental and physical phenomena as it is most fundamentally. Theories in philosophy of mind are therefore meta-physical theories—they aim to chart the boundaries of what can be explained by the natural sciences, and so cannot themselves be proven or disproven exclusively by the deliverances of those disciplines. For example, philosophy of mind addresses questions such as: does the world at bedrock have both mental and physical properties, or only one of the two? Are brains and minds different? If so, how? What is the relationship between the feeling I have while eating that second piece of cake and the elevated dopamine levels that accompany this feeling? Is the relationship one of mere accompaniment, or one of identity, causation, emergence, or grounding? Or something else? Philosophical theories addressing these kinds of questions have motivated historical and contemporary reflection, and no chapter can cover them exhaustively. Instead, the chapter offers an overview of important historical and contemporary theories, issues they face, and some exciting new directions in research.

Key Terms
(in order of appearance in the text)

Monistic theories: theories that claim there is fundamentally one kind of entity or property: either mental or physical
Dualistic theories: theories that claim there are fundamentally two kinds of entities or properties: both mental and physical
Idealism: a monistic theory that claims fundamentally everything is mental
Physicalism: a monistic theory that claims fundamentally everything is physical
Reductive physicalism: a physicalist theory that affirms mental phenomena can be reduced to—that is, are straightforwardly derivable from—physical phenomena
Non-reductive physicalism: a physicalist theory that denies mental phenomena can be reduced to—that is, are straightforwardly derivable from—physical phenomena
Multiple-realization: this idea that psychological similarity need not entail physical similarity, and that mental phenomena can therefore be instantiated or realized by a variety of systems
Type physicalism: the idea every mental category corresponds to a physical category
Token physicalism: the idea that every mental instance corresponds to a physical instance
Supervenience: for two sets of characteristics, A and B, to say that A supervenes on B is to say things cannot differ in regards to A without also differing in regards to B
Functionalism: the idea that mental phenomena should be characterized in terms of the functional role they play
Property dualism: a dualistic theory that claims some entities have two different types of properties: mental and physical
Chapter 6: Introduction to Philosophy of Science
Carlos Mariscal

Key Terms and Definitions:

**Reductionism**: the view that higher-level entities, theories, or explanations (such as mental processes) are more fully or accurately described by appeal to the most basic science (i.e. physics).

**Levels of nature**: a realist interpretation of Levels of Organization such that they are not mere descriptions, but actually the way reality is organized.

**Levels of organization**: the entities or processes studied by scientists that make up (or are made up) by other entities that are also studied by science, such as groups, individuals, cells, molecules, etc.

**Scientific realism**: the thesis that what science dictates (e.g. theory, entity, or relations) is approximately true.

**Anti-realism**: the thesis that what science dictates is not approximately true.

**Instrumentalism**: a type of anti-realism that holds that scientific theories should be judged by their utility and not their truth-value; or that scientific theories are not attempts to describe reality beyond experience.

**Scientific explanation**: an account of why something is the case (rather than a mere description of the fact that it is).

**Causal explanation**: scientific explanations that advert to prior causes, usually by appeal to manipulations, mechanisms, models, or, more rarely, constant conjunctions and counterfactuals.

**Pragmatic explanation**: a view of explanation that claims explanations are acts of communication; successful explanations result in others’ understanding.

**Underdetermination**: (weak/practical) there is insufficient data to decide between interpretations; (strong/logical) no amount of data would decide between various interpretations.

**Theory**: a system of ideas that purports to explain a phenomenon.

**Ontological**: about the nature of being or existence rather than about how we know about it.

**Methodology**: the techniques followed in a particular discipline.

**Fallibilism**: the doctrine that all knowledge claims (weak) could, in principle, be mistaken; or (strong) are probably false.

Introduction

This chapter will be a brief survey of the concepts from general **philosophy of science** for those interested in cognitive science. It covers several major topics in the philosophy of science: **scientific explanation** and **underdetermination**, **reductionism** and **levels of nature**, and **scientific realism**. We will discuss the goals of science, the methods of science, and the most plausible interpretations of science. To demonstrate the importance of these topics, the chapter includes cases in which confusion over these issues has led scientists astray. These cases include instances in which cognitive neuroscience has failed to discover adequate explanations for phenomena, when previously established research did not withstand further scrutiny, and the increasingly complex and bewildering interrelationship between the study of the mind and the study of the brain and its parts.
Chapter 7: Metaphysical Issues of Relevance to Cognitive Neuroscience

C. L’Hote

Brief outline of topic covered:

Following an introduction to metaphysics, this chapter highlights three metaphysical issues that are relevant to cognitive neuroscience. The first issue arises anytime researchers establish correlations between cognitive and neural processes, for instance when Kanwisher (1997) established a correlation between face recognition and activity in the fusiform face area (FFA). Here, a question arises: Why are the cognitive and neural processes correlated? What explains their correlation? Possible explanations are that face recognition and FFA activity are actually identical, that recognition is constituted by FFA activity, or that face recognition is somehow caused by FFA activity. Which explanation is correct depends on core metaphysical issues: the meaning and nature of identity, constitution, and causation. A second metaphysical issue arises because brains that are physically different from one another may nonetheless be cognitively the same. Despite the many physical differences between our brains, my friend and I often have the same thoughts. What, then, is the nature of the property that our thoughts share? The answer depends on the nature of mental properties. A third metaphysical issue arises because cognition makes a difference. Thoughts are not only caused but are also themselves causes of behavior. How do thoughts do this? How does cognition figure into the causal nexus? Making both sense and science of cognition requires a closer look at the nature of causation.

Key terms:

Composition: is the relationship that exists between a whole and its parts. A bike is composed of its frame, wheels, shifters, and so on.
Constitution: in particular material constitution, is the relationship that exists between an entity and the material or materials out of which it is made. A statue is constituted by its clay.
Content Essentialism: is the thesis that the content of a mental event, state, or process is essential to it. A thought about dogs is essentially about dogs.
Ontology: is a branch of metaphysics that focuses on the nature, structure, and categories of being and existence.

1. Introduction to metaphysics

What, then, is metaphysics? To the uninitiated, the word may evoke the supernatural or paranormal. Visit a local bookstore and, to the chagrin of philosophers, a few metaphysics titles may be shelved in the new-age section. Undoubtedly, the word “metaphysics,” which seems to point to what exists “above” or “beyond” the material world, is partly responsible for the confusion. However, the prefix “meta” has a more modest meaning here. The word “metaphysics” was coined after an early editor of the works of Aristotle (384-322 BC) named the books following his treatise on physics, “te meta ta physika biblia,” i.e., the books after the books about physics. In this very work, Aristotle explicitly rejects the otherworldly metaphysics of his teacher Plato (427-347 BC), providing a conception of metaphysics that is influential to the present day. According to Aristotle, the subject matter of the sciences is individual beings and categories of beings, while the subject matter of metaphysics is what all beings have in common: being as such. In this way,
Chapter 8: Epistemic Issues Pertaining to Neuroscientific Methods
Nina Atansova

Box: Overview
The chapter provides an overview of the main epistemological issues in contemporary neuroscience: validation and integration of experimental knowledge. Cognitive neuroscience traditionally relied on animal models and computer simulations. However, powerful neuroimaging technologies have expanded its experimental and observational capacities enormously. Every technology comes with its own epistemological challenges, which the chapter introduces and then explains how these challenges can be addressed by following the best practices of experimental design.

Box: Terms
Analogical models: models that are made up of objects that share a number of similarities with other, more complex, objects, which are themselves the targets of scientific inquiry. When such models are manipulated and additional characteristics of their objects are consequently identified, then the inference can be made that the target objects also possess the newly identified characteristics.

Animal models: experimental systems which include animals, or their parts or tissues, whose physiology or behavior are manipulated and studied with the ultimate goal to learn something about human physiology and behavior by extrapolation.

Convergent validity: the ability of a test to produce results that are reproducible through alternative experimental interventions.

Deduction: in logic, a mode of inference in which the conclusion can be derived from the premises with certainty.

Epistemology: the study of knowledge; a branch of philosophy.

Face validity: the ability of a model to superficially mimic its target.

Falsification: a logical procedure for disproving a scientific theory based on contradicting evidence.

Integration: the ability of multiple scientific disciplines to produce knowledge claims that together provide a better understanding of a studied phenomenon than any discipline by itself.

Operationalization: specifying observable or measurable effects to be produced under the assumption that a theoretical construct is real.

Predictive validity: the ability of an experimental intervention to reproduce expected results.
Chapter 10: Modularity
Aleksandra Mroczko-Wąsowicz

Brief outline – textbox

The internal structure of the mind has been studied in philosophy of mind, psychology, and cognitive science for more than fifty years. A significant contribution was provided during the 1960s and 1970s by the development of the information-processing theories of mental phenomena, couched in terms of the computational theory of mind, a view describing the mind as an information processing system (a computer). Computational approaches enjoyed mainstream status for many years. According to computationalism, cognition is computation, i.e., the transformation of sequences of symbols according to precise rules, the result or act of calculating amount or figuring an answer. Both artificial and biological systems can perform computations; well-known examples are the operation of a vending machine, the process of cell division, plant growth, or information processing in the nervous system. Intelligent behavior can be causally explained by computations performed by the agent's cognitive system (or brain) (e.g., Piccini 2012; Rescorla SEP). Modular models of the mind, emphasizing that mental phenomena arise from the operation of multiple distinct processes (specialized computational devices called modules), became influential as a general motivation for computationalism. In a modular computational system, the information to which certain processes have access is restricted. This makes information processing in the system efficient and fast. In contrast, in a totally integrated non-modular system, all computational processes have access to all the information in the system, which makes its information processing rather slow and ineffective. Although researchers agree that the brain contains different regions and the information-processing systems underlying perception differ from those underlying reasoning or motor control, there is little agreement as to what extent the mind exhibits a modular architecture, how to exactly define modularity, and functional isolation exhibited by modules. The fact that the assumed modularity of processing within the brain doesn’t translate to modularity of mental architecture suggests that more interdisciplinary collaboration between cognitive neuroscientists and philosophers is needed to progress on this topic, i.e., to provide decisive empirical evidence for one of theoretical modular approaches.

Key terms and definitions – sidebar textbox

**Information processing** – an approach in cognitive psychology based on the computer metaphor of the brain popular in 1950s in which cognition and behavior are described as a sequence of processing stages. The approach concerns the idea that our brains actively process the information they receive (e.g., from senses), record it, and store it, like a computer does.

**Module** – a cognitive system dedicated to performing a domain-specific information-processing task. Typically held to be informationally encapsulated, but not necessarily to have a fixed neural architecture.

**Informational encapsulation** – information processing in the module cannot be affected by information in the rest of the brain, e.g., the effects of a visual illusion, created by low-level sensory processes, persist despite high-level processing caused by conscious awareness of the illusion itself.

---

1 Faculty of Philosophy, University of Warsaw; National Science Centre, Poland (2019/35/B/HS1/04386)
Chapter 11 Mental Architecture - Computational Models of Mind
Othalia Larue, Jean-Nicolas Bourdon, Mylène Legault, Pierre Poirier

Brief outline of topics covered

A cognitive architecture is a complete computational or mechanistic model of cognition that aims at integrating all components responsible for the variety of cognitive abilities, and hence the intelligent behaviors of an agent, whether the agent is natural or artificial, human, or animal. The question of cognitive architectures occupies an important place in cognitive psychology, cognitive neuroscience, artificial intelligence (AI), and philosophy of mind. In these disciplines, the term "cognitive architecture" is centered around two main concerns: 1) the description, explanation, and justification of the fundamental structure of cognitive systems and 2) the design of complete cognitive organization (including sensory input and motor output) leading to the construction of artificial agents or computational models. The resulting research traditions have the common aim of developing theories that account for all cognitive mechanisms. Many computational models have been developed in the context of these two traditions. While each model comes with its own goals and background assumptions, agreements have emerged on fundamental elements. The future of cognitive architectures is likely headed toward the development of a common model of cognition and of artificial general intelligence, as well as an increase in the biological plausibility of cognitive architectures.

Key terms and definitions

Computational modeling: the study of complex systems (their properties, behaviors, interactions, etc.) through the use of computer programs.

Cognitive architecture: a theory of cognition and its computational instantiation that aims at integrating all components responsible for the variety of cognitive abilities.

Symbolic computationalism: adopts a mind as computer metaphor, where, like computers, minds are seen as symbol manipulators.

Connectionism: models cognition using artificial neural networks, which link artificial neurons through variably weighted connections; Information is encoded in the values of the connections, which are the system’s main parameters.

Dynamical systems: to account for cognition, this approach replaces representations and neural connections with relevant couplings of interacting variables, for instance sensory and motor variables, or organism and environmental variables.

Unified theories of cognition: research in cognitive modeling that aims at overcoming cognitive science’s “theoretical myopia” (Newell, 1990) by reorienting research from the development of models of single cognitive capacities (vision, memory, decision-making, etc.) to the development of integrated frameworks capable of accounting for the human mind as a whole.

Biological plausibility: the general goal according to which models of cognition should be compatible with biological sciences.
Chapter 12 - Language
David Pereplyotchik

Outline

The interaction between philosophy, psychology, and linguistics has been remarkably productive since the 1950s. Cognitive neuroscience is contributing novel results, making psycholinguistics one of the most fruitful interdisciplinary research areas within cognitive science.

This chapter surveys some of the field’s foundational principles, its most important findings, and some ongoing debates, with an eye toward ascertaining their bearing on philosophical questions. Psycholinguists routinely make use of philosophically contested notions such as mental representation, modularity, and innateness. So their empirical findings ought to constrain theorizing in philosophy of language and philosophy of mind, just as philosophical work should inform empirical practice.

The first section introduces some general facts about human languages. Section 2 then reviews what psycholinguists have discovered about how we acquire language as children and how we use it thereafter. The chapter concludes in Section 3 by showing how the findings from the previous section inform longstanding philosophical issues concerning innateness, mental representation, and ontology.

Key terms and definitions:
Phonology: the study of phonemes, the basic sounds of a spoken language, corresponding (imperfectly) to the written letters of an alphabet—e.g., the /b/ and /p/ that distinguish “big” from “pig” in English.
Morphology: the study of morphemes, the smallest parts of words that carry a distinctive contribution to the word’s overall meaning. For instance, the word “farmers” is composed of three morphemes: “farm,” “-er,” and “-s.”
Syntax: the study of the syntactic categories to which words belong and the grammatical relations they can bear to one another in larger phrases and sentences, including syntactic category (e.g., noun, verb), syntactic structure (e.g., word arrangement), and syntactic rules (e.g., subjects precede verbs in English; “Sam runs”).
Semantics: the study of the meanings of words and compound phrases.
Prosody: the intonational structure of speech—e.g., the contrast between statements (“Pigs fly!”) and questions (“Pigs fly?”) when spoken aloud.
Mental lexicon: a cognitive structure that contains information about all of the linguistic properties of each word/morpheme that one knows—e.g., their phonological structures, syntactic categories, and semantic contents.
Competence: possession of a language; contrasted with linguistic performance—the use of a language.
Performance: the exercise of one’s linguistic competence in producing, understanding, or otherwise processing language.
Competence Grammar: a hypothesized cognitive structure, consisting of the linguistic rules acquired by language users and employed in real-time language processing (see Figure 2).
Universal Grammar (UG): a hypothesized innate cognitive system that guides the process of language acquisition.
Chapter 13 - Mental Content
Krzysztof Dolega, Tobias Schlicht

forthcoming in: Mind, Cognition, and Neuroscience: A Philosophical Introduction, ed. by B. Young, C. D. Jennings

Brief outline of the topic covered – 1-2 paragraph textbox

Mental representation is one of the key explanatory concepts in cognitive psychology, neuroscience, and philosophy of mind. The idea that mental states are contentful plays a crucial role in how we explain behavior and understand each other. Mental abilities like perception, imagination, memory, and action planning are commonly thought to require the capacity to represent external states of affairs. Consequently, many of the failures in exercising these abilities can be explained by the cognitive system misrepresenting germane aspects of the world.

This chapter provides an overview of the key debates concerning the nature of mental representation in philosophy of cognitive science. The chapter starts with a brief historical overview, followed by the main part aimed at elucidating three central issues about mental representation — their function, the nature of representational content, and the role that the notion plays in explanation. The last section of the chapter is devoted to connections with other areas of research and prospects for tackling the outlined issues.

Key terms and definitions – sidebar textbox

*Intentionality:* the feature of mental phenomena (states, processes, events) in virtue of which they are about or directed at some object, event, or state of affairs.

*Representational theory of mind:* the view according to which cognition consists in processes of formation, manipulation, and transformation of contentful representational vehicles.

*Representational vehicle:* the physical entity which serves as the physical medium into which the content of a representation is encoded.

*Content:* the meaning of a representation — what the representation is about, usually expressed in terms of truth or accuracy conditions describing when the representation is properly representing its object.

*Wide vs. narrow content:* content can be construed as being either narrow — depending only on features of the cognitive system, or wide — depending on environmental features as well.

*Personal vs. subpersonal:* distinction between different levels of explanation of cognitive phenomena. Explanations of behavior in terms of mental states and processes that are ascribed to whole persons are considered to be on a personal level, while those explanations of behavior and cognitive functions that are localized to parts of the cognitive system and can be subdivided into simpler elements that operate unconsciously are usually
Ch. 14 Concepts and non-conceptual content
Arnon Cahen

Brief Outline
This chapter concerns a distinction within the philosophy of mind and cognitive science between conceptual and nonconceptual ways in which a person may mentally represent the world as being. Correlatively, the distinction is between two kinds of representational mental states—states with conceptual or nonconceptual representational content. Briefly, the distinction is between ways of representing the world that depend upon the conceptual capacities of the creature having the mental state and ones that do not—between, for example, one’s thinking or judging that the Eiffel Tower is in Paris, and one’s experiencing the Eiffel Tower when standing before it. Among other things, the distinction has significant implications for epistemology—how perceptual experiences can inform empirical thoughts—as well as for our understanding of mental architecture—the informational and functional relations among our cognitive capacities, and in particular between our perceptual modalities and our more central, higher cognitive capacities. After introducing and looking more closely at the origins and nature of this distinction within philosophy of mind, we will address its relation to current debates in contemporary cognitive science, with special emphasis on the modularity and cognitive penetrability of perception. As the distinction has its origins in discussions of visual perception, we focus our exposition there. Yet, as we will see, the notion of nonconceptual content may be applicable to many other cognitive systems as well.

Keywords:
Propositional attitude: a mental state that involves having some attitude toward a proposition, usually denoted by an attitude followed by a ‘that’ clause (e.g., belief that p, desire that p, etc.).
Representationalism: the theory that mental states are constituted by contentful representations.
Semantic evaluability: being assessable for truth or accuracy.
Perspectival constraint: attribution of mental content is answerable to the subject’s perspective on what is being represented.
Conceptual content: content of a mental representation that depends on the subject’s concepts.
Nonconceptual content: content of a mental representation that does not depend on the subject’s concepts.
Experiential fineness-of-grain: the claim that our perceptual discriminations outstrip our conceptual repertoire.
Mental architecture: a specification of the relations between the different cognitive capacities that constitute the mind.
Chapter 15: Animal Cognition
Irina Mikhalevich

Keywords:
Anthropocentrism: the assumption that humans are the measure of all things.
Anthropomorphism: the attribution of human traits to nonhuman entities.
Associative learning: formation of associations between stimuli or a stimulus and behavior.
Chemotaxis: movement along chemical gradients
Cognitive model: an explanation of animal behavior describing cognitive mechanisms.
Phenomenological model: an explanation in terms of surface-level features (e.g., behavior).
Comparative cognition: the interdisciplinary study of the evolution, development, and function of cognitive processes and mechanisms in human and nonhuman animals.
Convergent evolution: the independent evolution of structural or functional similarities due to shared selection pressures.
Connectome: a complete diagram of an organism’s neural connections.
Deflationary explanation: in comparative psychology, an explanation in terms of simpler mechanisms or processes.
Encephalization quotient (EQ): ratio of brain and body size. Higher EQ is thought to correlate with higher general intelligence (see Box 1). 
Evolutionary constraint: a restriction on the course of evolution in a lineage (e.g., body plan constraint).
Homology: similarity due to shared ancestry.
Logical empiricist theory of meaning: mid-20th view in which terms that do not refer to observable phenomena are considered meaningless metaphysics.
Metacognition: knowledge of one’s own mental states.
Mindreading (theory of mind): ability to attribute mental states to another. (See discussion in Chapter 28).
Neuroethology: the study of animal brain activity under naturalistic conditions.
Phylogeny: history of evolutionary relatedness.
Principle of evolutionary continuity: the notion that complex traits are likely to be distributed to various degrees throughout the living world.

Chapter Outline

This chapter begins with an overview of the multidisciplinary science of comparative cognition and the current philosophical controversies in the field. These controversies include: (i) whether anthropomorphism is an error; (ii) whether simpler explanations of animal behavior are (ceteris paribus) superior to more complex ones; (iii) how we should think about the proper role of evolutionary considerations in the study of animal minds; and (iv) where the boundaries of cognition may lie. Section 2 provides a historical overview of the field that places these controversies in historical context by focusing on the emergence of two prominent subdisciplines of comparative cognition: comparative psychology and ethology. It begins with Darwin’s influence on early comparative psychology and ethology, proceeds through the rise and fall of behaviorism, and concludes with the cognitive revolution and the eventual, if partial, synthesis of comparative psychology and ethology into comparative cognition. Section 3 moves this discussion into the present. Here we critically assess the view that anthropomorphism is a uniquely dangerous error; consider the merits of simplicity preferences; examine the implications of insect, spider, and octopus intelligence for accounts of mind; and draw on evidence of apparently “intelligent” behavior in plants and protists to query how and whether to draw the line between the cognitive and the non-cognitive. The final section on future directions examines the contributions of animal cognition science to robotics and AI.

1. Introduction

Human minds are likely not the only kinds of mind that evolution has engineered. If we wish to understand the nature of mind broadly construed, our investigation must be comparative: it must sample across a wide swath of animal life in an effort to extract patterns and regularities governing the
Chapter 16: Kinds of Consciousness
Jacob Berger

Brief Outline
Consciousness is central to our lived experience. It is unsurprising, then, that the topic has captivated many students, neuroscientists, philosophers, and other theorists working in cognitive science. But consciousness may seem especially difficult to explain. This is in part because the term “consciousness” has been used in many different ways. The goal of this chapter is to explore several kinds of consciousness: what theorists have called “creature,” “phenomenal,” “access,” “state,” “transitive,” “introspective,” and “self” consciousness.

The basic distinctions among these kinds of consciousness are described in Section 1. Section 2 raises potential challenges for explaining these varieties of consciousness and describes a few current theories of them. Section 3 closes the chapter by exploring directions for future work in the cognitive science of consciousness. Along the way, some of the possible interrelationships among these kinds of consciousness are discussed.

Keywords:
Creature consciousness: an individual is creature conscious when the individual is awake and mentally responsive to stimuli—rather than, for example, asleep or anaesthetized.
Phenomenal consciousness: a mental state is phenomenally conscious when there is something that it is like for the individual to be in that state.
Access consciousness: a mental state is access conscious when the information contained in that state is available for use in thought and behavior; for example, the state’s information may be used as a premise in reasoning or for the rational control of action and speech.
State consciousness: a mental state is state conscious when an individual is subjectively aware of being in that state.
Transitive consciousness: an individual is transitively conscious (or, as some simply put it, aware) of something when the individual is mentally responsive to that thing, typically either by perceiving or having a thought about it.
Introspective consciousness: a conscious mental state (or, equivalently, an individual) is introspectively conscious when the individual is subjectively aware of being in that state in a deliberate and attentive way.
Self-consciousness: an individual is self-conscious when the individual is subjectively aware of itself.

1. Introduction

1.1 What are the various kinds of consciousness?

Perhaps nothing in the universe is more familiar to us than our consciousness, but perhaps nothing seems as hard to explain! A major obstacle to the study of consciousness has been the fact that people have used the term “consciousness” and related expressions such as “awareness” in many different ways. As a result, theorists over time have distinguished several kinds of consciousness, although there remains significant debate about how these kinds may relate to one another, or even whether they exist. This chapter thus begins by reviewing a few of the major types of consciousness discussed by cognitive scientists (for versions of these distinctions, see, e.g., Rosenthal 1986; Block 1995; Carruthers 2016).
Chapter 17: Philosophical Theories of Consciousness
William Lycan

TEXTBOX:

Issues in consciousness studies break down into three categories: specific phenomena that are of philosophical interest but are primarily empirical; more general phenomena that demand philosophical attention but are still scientifically tractable; and questions that seem in principle to resist scientific treatment and have afforded arguments against materialist theories of mind. This chapter addresses one issue of the second type, discussing philosophical theories of “state” consciousness, and then turns to four issues of the third type: the ontological status of sensory qualities, the intrinsically perspectival nature of the mental, the existence of intrinsically perspectival “what it’s like” properties as allegedly revealed by “knowledge” arguments, and Levine’s “explanatory gap” between neuroscience and what-it’s-like properties. Materialists have responded by defending a representational theory of sensory qualities and an account of “phenomenal concepts” that accommodates “what-it’s-like” properties but is entirely compatible with materialism.

SIDEBAR:

Key terms and definitions:
Perceptual consciousness: perceptual awareness of objects and states of affairs in the environment.
State consciousness: a mental state is state conscious when the individual is transitively conscious of being in that state in a subjective way.
Materialism: the view that minds and mental properties are constituted by physical matter and are entirely part of nature.
Sensory quality: the sort of qualitative feature of which a particular sense modality makes you aware, such as color or shape in your visual field or the pitch of a heard sound.
Phenomenal consciousness: a mental state is phenomenally conscious when it feels a certain way for the subject to be in that state, when there is something it is like for the subject to be in that state.
Phenomenal property: the introspectible property of a mental state described as “what it’s like” for the subject to be in that state.
Phenomenal concept: a concept deployed by a human being or other sentient creature to represent one of its own mental states in a distinctively internal and subjective way.

1. Introduction

The 20th century saw the end of German and British idealism, followed by a variety of philosophical theories of mind (for explanations of the labels, see Chapter 5): neutral monism (Russell 1956); substance dualisms, property dualisms, behaviorism (Ryle 1949), mind-brain identity theory (Smart 1959), and functionalism (Putnam 1975). Since the 1970s, functionalism has settled into a majority view. 21st-century philosophy of mind remains predominantly materialist, informed by cognitive science and neuroscience. But there is significant dissent based on phenomena that go under the broad heading of “consciousness” and that do not fit
Chapter 18: Neurobiological Theories of Consciousness
Myrto Mylopoulos

Brief outline of topics covered (1-2 paragraph textbox):
Whatever other properties conscious experiences might have, it is almost universally agreed that there is a tight relationship between consciousness and the brain (for exceptions, see Eccles 1994 and Dretske 1995). It is unsurprising, then, that neurobiological theories of consciousness, which seek to identify the precise nature of this relationship, are quite popular today. Such theories aim to understand consciousness by appeal to its underlying neural properties, processes, or mechanisms, which are in turn uncovered by way of standard techniques in neuroscience, including brain stimulation methods and neuroimaging techniques.

This chapter focuses on the neurobiological approach to consciousness. It starts in Section 1 by introducing the search for the so-called neural correlates of consciousness that began in the 1990s. In Section 2, we move beyond this approach and examine some influential theoretical frameworks that posit neurobiological mechanisms underpinning conscious experience. Such theories go beyond the neural correlates approach in that they do not merely propose answers to the what of the neural basis of consciousness, but the why and the how as well. We will focus on four specific theories, some of which we were introduced to in Chapters 16 and 17: the recurrent processing theory, the global neuronal workspace theory, the information integration theory, and higher order theory. In section 3, future directions of such research are discussed.

Key terms and definitions (sidebar textbox):
Neural correlates of consciousness (NCC): The minimal neural state that is sufficient for conscious experience.
Neural correlates of state consciousness (NCSC): The minimal neural state that is sufficient for a conscious state with a specific type of content (e.g., seeing faces, feeling pain).
Neural correlates of creature consciousness (NCCC): The minimal neural state that is sufficient for the conscious state of being awake and responsive to stimuli. May come in degrees.
Global neuronal workspace theory (GNWT): theory according to which conscious information is made accessible to specialized processors distributed throughout the brain by way of a “global workspace,” itself composed of widely distributed neurons with long-range axons that form reciprocal connections with the neurons local to these processors.
Recurrent processing theory: theory according to which conscious visual experience arises when information processed in higher visual areas, including parts of the parietal and temporal cortex, re-enters lower visual areas, including V1.
Integrated information theory (IIT): theory according to which consciousness is primarily a function of information integration in a system, and can be mathematically measured by a quantity labeled “Φ.”
Higher-order theory (HOT): theory according to which what it is for a psychological state to be conscious is for one to be aware of oneself as being in that state by way of a higher-order representation.

1. Introduction

What are the neural correlates of consciousness?
Chapter 19: The Unity of Consciousness
Rocco J. Gennaro

1. Introduction

1.1 Historical overview

The phrase “unity of consciousness” is ambiguous. First, the term “consciousness” can itself be ambiguous. Numerous notions of consciousness can be found in the philosophical literature (see Chapter 16) but the most often used is Nagel’s famous “what it is like” sense (Nagel 1974). When I am in a conscious mental state, there is “something it is like” for me to be in that state from the subjective or first-person point of view. When I smell a flower or have a conscious visual percept of a flower, there is something this perceptual experience seems like from my perspective. Second, and as we shall see in detail below, there is even greater ambiguity in the use of the term “unity” (Tye 2003; Bayne & Chalmers 2003). Perhaps most common is the general notion that, from the first-person point of view, we experience the world in an integrated way. In addition, there are multi-modal conscious experiences with content involving more than one sensory modality, such as experiencing the combination of auditory and visual perceptions at a concert. We might also at times refer to different aspects of the same object, such as the color, shape, and motion of a blue ball. These are all different forms of conscious unity, which will be explained in more detail below. Whatever its precise meaning, this crucial feature of consciousness played an important role in numerous Western philosophical debates, but especially in the philosophy of Kant (1781/1787). Kant argued that unified conscious experience must be the product of the (presupposed) synthesizing work of the mind, including the application of various concepts or “categories” (such as cause and substance). He famously called the activity of such synthesizing the “transcendental unity of apperception.” Although Kant had nothing to say about the mechanisms in neurophysiological terms, he had much of value to say about them in cognitive terms (Brook 1994, 2005; Kitcher 1990). The notion of the unity of consciousness has also played an important role in Indian philosophy going back centuries before Kant (Ganeri 2017; Siderits et al. 2011).

1.2 Historical development

There are numerous types of conscious unity that can be found in the literature, especially over the past thirty years or so. Tye (2003, 11-15), for example, distinguishes between object unity, neurophysiological unity, spatial unity, subject unity, introspective unity, and phenomenal unity. Phenomenal unity is the notion that he explains in terms of experiential contents being
Chapter 20: Attention
Carolyn Dicey Jennings

Brief outline of topic covered (1-2 paragraph textbox)

The main questions in philosophical research on attention concern its nature and impact. Regarding its nature, one might ask what sort of thing attention is; regarding its impact, one might ask what sort of thing attention does. While these questions have been asked by philosophers for thousands of years, they have had a resurgence in recent years due to advancements in the cognitive and neural sciences. This chapter will cover some historical context as prelude to a discussion of the contemporary debates, ending with issues that are as of yet still on the horizon. Topics covered include the divide between voluntary/involuntary, endogenous/exogenous, and top-down/bottom-up attention; the Posner cueing task; Treisman and the binding problem; Mack and Rock on inattentional blindness; Grassia, Campbell, and Dickie on attention, perception, and knowledge; the debate between Block, Dehaene, and others on consciousness and attention; Wu on attention and action; and future directions for attention research in predictive processing models, neuroethics, technology, and addiction.

Key terms and definitions (sidebar textbox)

Attention: the prioritization of some mental or neural phenomena over others, including sensation and behavioral response.
Endogenous attention: prioritization that is primarily due to an internal (to the mind or brain) source.
Exogenous attention: prioritization that is primarily due to an external (to the mind or brain) source.
Top-down attention: prioritization that benefits from prefrontal feedback.
Bottom-up attention: prioritization that is primarily driven by the sensory cortices.
Posner cueing task: a task that aims to measure behavioral differences between endogenous and exogenous attention through the difference in reaction times to central and peripheral cues.
Binding problem: determining the neural mechanism by which sensory features that are processed separately in the brain are bound into unified objects.
Inattentional blindness: the failure to notice a stimulus that is clearly visible in normal conditions due to an attention-involving task.
Spotlight metaphor: attention is likened to a spotlight being cast over a stage; intimately connected to the theater metaphor for conscious experience.

1. Introduction

(box with 3-5 paragraphs on neuroscientific findings relevant to topic)

For the past 40 years the cognitive and neural sciences have primarily examined attention through the Posner cueing task, dividing “exogenous” from “endogenous” attention (Posner 1980; see Image 1). Exogenous attention occurs when one cues a target in the location of the target. A box might appear around the very same location as a subsequent letter that the participant is then asked
Chapter 21 Memory  
Felipe De Brigard & Sarah Robins

[Brief outline of topic covered – 1-2 paragraph textbox.]

This chapter provides a discussion of memory in philosophy and neuroscience, with each section organized around three questions: 1) What is memory? 2) What is remembering? and 3) What are memories? Each section begins by exploring ways that memory has been taxonomized, both by scientists and philosophers of memory, emphasizing the kinds that have been particularly relevant for discussions at the intersection of these disciplines. The focus moves next to the mental act of remembering—the processes that are involved in our capacity to encode, store, and retrieve information—and the mechanisms that make remembering unique but also similar to other cognitive operations. Finally, each section draws attention to the individual mental states we call “memories,” whether or not they involve memory traces or representations, and what the nature of such entities may be. Throughout, we make note of the ways that philosophical and neuroscientific work on memory have been pursued independently and discuss the current trend toward interdisciplinary, collaborative work on these basic questions.

[Key terms and definitions – sidebar textbox]

**Episodic memory (EM):** memory for specific spatiotemporal events; it is said to be accompanied by autonoetic consciousness (i.e., awareness of oneself as having experienced the event in the past).

**Semantic Memory (SM):** memory for conceptual and factual information; it is said to be accompanied by noetic consciousness.

**Standard Model of Memory:** influential taxonomy of memory systems that divides long-term memory into declarative (explicit) and non-declarative (implicit).

**E-S-R model:** standard view of memory according to which information is first encoded, then stored, and then retrieved at recollection.

**Consolidation:** the stabilization of a memory trace that has been stored after encoding.

**Content (in)variantism:** the view according to which a memory’s content at encoding is identical to that memory’s content at retrieval.

**Engram/memory trace:** vehicles that enable the preservation of memories from encoding to retrieval.

**Representationalism:** in the context of this chapter, the view according to which memory is mediated by representations.

1. Introduction
Chapter 22: The Unconscious Mind
Alon Goldstein & Benjamin Young

Brief outline of topic:

Unconscious processes are mental states that occur in the absence of subjective awareness. We offer a focused historical survey of the robust debate about the nature of unconscious mental processing, from ancient and medieval theories that allow for bodily functions without subjective awareness to the 20th century acceptance of autonomous unconscious processing. The background introduction culminates with the rise of cognitive science in the latter half of the 20th century, as dual systems theories claimed that the mind had two forms of processing: an unconscious system that is associative, fast, and automatic and a conscious system that performs sequential, deliberative, and slow mental functions. More recently, the “cognitive unconscious” has been discovered to perform most if not all cognitive functions including higher-level cognitive processing, such as addition and subtraction. Despite a rich history of experimental methods and findings, recent criticisms are forcing scientists to reexamine these advancements. We are optimistic that the field will develop novel techniques that answer these criticisms, including measures for studying subjective reports in the absence of subjective awareness of the content of the unconscious state and the act of reporting.

Keywords:

**Conscious**: for the purposes of this chapter we follow experimentalists in defining a mental state as conscious if the subject is aware that they are undergoing an experience and can report on some of its content (see Chapter 16).

**Unconscious**: mental states (not mere neurologically mediated reflexive processes) that occur in the absence of subjective awareness, as defined above (i.e., reportability), sometimes referred to as subconscious.

**Preconscious**: historically, Freud considered these to be mental states that we are unaware of but are poised to become conscious. The term is currently used to describe a strong, durable activation spread across multiple regions in the brain that remains unreported but is poised to become conscious.

**Autonomous unconscious states**: mental states whose content occurs automatically without conscious mediation.

**Independent unconscious processing**: cognitive processes that can perform functions operating over informational states that can causally mediate behavior and influence other mental processes in the absence of subjective reportability or conscious mediation.

**Seriality/sequential processing**: mental processing constructed from several subprocesses, working one after the other, with each using its predecessors’ output and forwarding their outputs to their successors in an ordered manner.

**Implicit process**: a process that can only be inferred using indirect measures (i.e., not direct report).

**Automatic process**: a cognitive process that can efficiently perform a function in the absence of (or despite) volitional control, sometimes without awareness of the process occurring.

**Suppression**: rendering one stimulus unconscious by introducing other, internal or external, stimuli that draw attention away from it.
Chapter 23: Perception
Tony Cheng

Brief outline of topic covered – 1-2 paragraph textbox.

Humans and other animals perceive with many different sensory modalities, including olfaction, touch, audition, vision, echolocation, proprioception, gustation, and some other senses, depending on different criteria and definitions. Given its broad range, it is not possible to give a comprehensive overview of all the philosophical, psychological, and neuroscientific studies about perception in one chapter, so what will be offered below is quite selective. In the introduction we will discuss basic concepts such as figure-ground segregation and scene analysis. 1.1 provides a historical overview of psychological theories of perception since the late 19th century. 1.2 discusses the information-processing approach and hierarchical explanations. 1.3 looks into the historical development of philosophical studies of perception. 2.1 to 2.3 introduce contemporary issues, and 3.1 to 3.3 provide future directions.

Key terms and definitions – sidebar textbox

Sensation: uninterpreted sensory input.
Perception: sensations that have been processed and interpreted by perceptual systems to become exteroceptive, i.e., relating to stimuli that are external to the perceiver.
Perceptual experience: when perception is conscious, which might always or often be the case (see 3.1 below), this phrase is invoked to emphasize the conscious aspect of perception.
Information: the resolution of uncertainty, i.e., the amount of information negatively correlates with the level of uncertainty.
Representation: information-bearing states that have accuracy or truth conditions.

1. Introduction

(box with 3-5 paragraphs on neuroscientific findings relevant to topic)

We can perceive many types of things—physical objects (material bodies), shadows, rainbows, scenes, motions, colors, depths, sizes, sounds, odors, tastes, flavors, heat, speech, music, and so on. For the purpose of this chapter we will focus on just a subset, including many of those already listed.

Physical objects or material bodies are basic constituents of the world, at least according to commonsense: the world is populated by tables, chairs, trees, mountains, rivers, oceans, and people (or at least their living bodies). We perceive them with vision, audition, touch, and olfaction, etc.—again, at least according to commonsense. It is crucial for our perceptual systems to single out these objects from their background. “Figure-ground segregation” is one kind of perceptual grouping that is crucial for singling out objects in both sight and touch (e.g., Kimchi & Peterson 2008; Mac Cumhaill 2017). Similarly, “auditory scene analysis” and “olfactory scene analysis” play critical roles in auditory and olfactory perception (Bregman 1990; Ache, Hein, Bobkov, & Principe 2016). The empirical world makes sense to us due to such scene analyses and segregations in different sense modalities: the physical and chemical stimuli on the sensory receptors are ambiguous, and objects can be blurred or hidden. Moreover, objects can look different from different angles, and this can be true of some other modalities as well, such as audition and olfaction. Given that the empirical world is
Chapter 24: Mental Imagery
Amy Kind

Brief Outline (in textbox)

Mental imagery is typically defined as a perception-like representation that is brought about in the absence of an appropriate external stimulus. Philosophers have standardly taken our experience of mental imagery to be impoverished in comparison with perceptual experience, in some way decayed (Hobbes 1651/1968) or less vivid and forceful (Hume 1739/1985). But mental imagery has nonetheless been recognized as playing an important role in a variety of mental contexts—in memory, in both spatial and abstract reasoning, in problem-solving, and in future-planning, to name just a few.

In the neuroscientific literature, considerable attention has been devoted to determining whether and to what extent the same neural pathways responsible for perception are also responsible for mental imagery. In the philosophical literature, relevant issues concern the representational nature of mental imagery, whether and to what extent introspection can shed light on the representational nature of mental imagery, how mental imagery differs across individuals, and the connection between mental imagery and imagination. After providing a historical background to the treatment of mental imagery in the philosophical literature, this article focuses on late 20th and early 21st century discussion of this topic and then briefly explores some of the future directions for mental imagery research.

Key terms (in sidebar textbox):

- **Amodal perception**: when we visually represent parts of objects that are occluded by something else, such as when we see a complete dog through the slats of a fence (also known as amodal completion).
- **Aphantasia**: an inability (or greatly diminished ability) to voluntarily produce mental imagery that occurs in 2% to 12% of the population; some may be able to experience mental imagery involuntarily, as in dreams.
- **Descriptionalism**: a theory in which mental images represent the same way that linguistic or other symbolic representations do; such representations need not share any features in common with the state of affairs represented.
- **Pictorialism**: a theory in which mental images represent the same way that pictures do, by sharing at least one visible feature in common with what is represented.

Relevant Neuroscientific Findings (in textbox)

Though mental imagery is not confined to the visual domain, the vast majority of empirical work on mental imagery has focused on visual mental imagery. In the 1960s and 70s much of this work came from cognitive psychology; studies by cognitive psychologists such as Shepard and Meltzer (1971), Finke (1980), and Kosslyn (1973) showed functional similarities between images and percepts (see Finke 1985 for a survey). The existence of these functional similarities was taken to support the claim that the representational structures underlying mental imagery were of the same kind as those underlying perceptions.

In the 1980s neuroscientific studies were brought to bear on these questions. In a 1988 survey of relevant neuroscientific findings, Farah noted several different kinds of empirical studies on vision that support the claim that perceptual mechanisms are involved in mental imagery. First, regional cerebral blood flow techniques able to spatially localize brain activity revealed that subjects asked to perform
Chapter 26: Embodiment and Enactivism
Amanda Corris and Anthony Chemero

Brief outline of topics covered:
Typically we think of the brain as responsible for cognition. But the brain is, importantly, embedded in a body—a body that moves around and interacts with features of the environment. What role, then, does the body play in cognition? Some philosophers would argue that it has no significant role in determining how to think about cognitive processing. But others argue that the body is fundamental to cognition, since the body is deeply involved with cognitive processes such as acting and perceiving. Embodied and enactive cognition are theories that emphasize the importance of the body in discussing cognition. This chapter discusses how some philosophers and cognitive scientists have investigated bodily involvement in cognition.

We begin in Section 1 by providing an overview of the historical developments that influenced enactive and embodied cognitive science: namely, phenomenology, ecological psychology, and the phenomenological critique of artificial intelligence. We show how these developments were advanced with the introduction of dynamical systems theorizing in psychology and situatedness in robotics. Embodied, extended cognition and enactivism are then introduced as theories that emerge from this discussion. In Section 2, we consider some key debates in embodiment and enactivism, including debates over the role of representations and the status of dynamical modeling. We end in Section 3 by highlighting some of the innovative research developments that have come out of an embodied approach to cognition with a focus on insights regarding social cognition.

Key terms:
The lived body: in phenomenology, the body is considered as experiencing, perceiving, and acting; this conception is contrasted with the physical body as merely a material thing in the world of other material things.
Affordances: opportunities for action in the environment for a particular organism. These are determined by both the abilities of the organism and its physical environment.
Direct perception: the claim, by ecological psychologists, that perceptual contact with the environment is not mediated by mental representations.
Dynamical system: a set of variables that change together concurrently and continuously over time, explained using the tools of dynamical systems theory, a branch of calculus. Dynamical systems and dynamical systems theory are key explanatory tools in embodied cognitive science (see Chapter 11).
Subsumption architecture: a layered robot control system where higher control layers can be constructed from the activity of lower control layers, resulting in greater behavioral competence.
Autopoiesis: literally, “self-making” or “self-producing.” According to enactivists, living systems are autopoietic systems, and autopoiesis is necessary for cognition.
Adaptivity: an enactive concept that captures how organisms actively self-regulate in response to both internal and external perturbations in order to maintain their viability (that is, at a minimum, to stay alive).
Representationalism: the view that mental contents represent states of the world (see Chapter
Chapter 27: Emotion
Sarah Arnaud and Jesse Prinz

Brief outline of topic covered

Are emotions real? Are they as we take them to be in folk psychology? Are they natural kinds or socially constructed? Leading theories give different answers to these questions. These theories also commit to ideas about the potential neural correlates of emotions. One might, therefore, think that neuroscience can adjudicate. Alternatively, neuroscience might motivate a search for an alternative to leading theories. A third possibility is that neuroscience has nothing to bring to our understanding of emotions. In this chapter, we investigate the relation between our folk categories of emotions, such as anger, sadness, fear, etc. and their potential neural mechanisms. We show how current research echoes important issues that have long been debated in emotion research, and we present recent enactivist views as a promising area for future research on emotions.

Key terms and definitions – sidebar textbox

**Reductionism**: according to reductionism, emotions can be reduced to biological mechanisms. These are usually presumed to include brain mechanisms but may also include behavioral dispositions and facial expressions.

**Anti-reductionism**: anti-reductionism entails that emotions cannot be reduced to biological mechanisms. Their identity conditions lie at a different level of analysis, for example at the level of cognitive states that do not have fixed correlates in the brain or body.

**Strong Realism**: realism about emotions implies that our current folk categories of emotions exist, and their existence is independent of our categorization of them. Weak realism would deny the latter claim.

**Nominalism**: emotions exist but depend in some way on our categorization of them. Each instance of an emotion, say an episode of fear, qualifies as the emotion it is because of how it is classified. Categorization confers identity by labeling. This is a form of weak realism.

**Social Constructionism**: emotions exist, but they depend on sociocultural forces. Categorization and norms play a role in shaping emotions, but once an emotion has been shaped in this way, it can also occur on occasions where it does not happen to be classified. Categorization confers identity by shaping. This is a form of weak realism.

**Anti-realism**: according to anti-realism, our folk categories of emotions do not refer to any genuine, tangible, existing items or episodes in the world. Nothing like them has any chance to be scientifically studied or discovered. They are, at worst, posits in need of wholesale rejection (eliminativism) and, at best, useful fictions (instrumentalism).

**Revisionism**: revisionism implies that our folk categories of emotions do not exist as such, but something relatively similar does have scientific validity. Revisionism says our folk categories will need to be revised, depending on future scientific discoveries about emotions.

1. Introduction
Chapter 28: Social Cognition and Theory of Mind
Evan Westra

Brief outline of topic covered:

“Social cognition” refers to the psychological capacities that humans and other animals use to reason about other agents and navigate complex social environments. This chapter focuses on the dominant approach to social cognition in contemporary cognitive science, which is centered around a capacity known as **theory of mind** or **mindreading**. Subjects covered include the false-belief task, the social brain network, mirror neurons, major accounts of theory of mind, objections to the theory-of-mind framework, mindreading in non-human animals, and new research on the ways that theory of mind is affected by social context.

**Theory of mind** (also called “mindreading” or “mentalizing”) is the ability to predict and interpret the observable behavior of other agents in terms of their underlying psychological states, such as their beliefs, desires, emotions, and intentions.

**Module:** modules are a class of cognitive mechanism associated with a cluster of properties: **domain-specificity**; automaticity; fast processing, limited central accessibility; informational encapsulation; “shallow” outputs; a fixed neural architecture; characteristic and specific breakdown patterns; and innateness (Fodor 1983). There is disagreement about which of these features is of primary importance, with some authors emphasizing domain-specificity (Carruthers 2006) and others emphasizing informational encapsulation (Coltheart 1999).

**Nativism:** the view that innate, genetically based mechanisms, concepts, or processes are required in order to explain the presence of a particular cognitive ability.

**Domain-general:** domain-general mechanisms operate similarly across a broad range of inputs, such as associative learning and statistical inference. Associative learning, for example, functions similarly whether one is learning about inanimate objects, speech sounds, or the behaviors of other agents.

**Domain-specific:** domain-specific mechanisms are specialized cognitive systems that only process certain kinds of functionally defined inputs. For example, some neuroscientists have argued that the fusiform gyrus contains a specialized face-detection mechanism known as the fusiform face area (Kanwisher & Yovel 2006).

**Holism**, in this context, refers to the epistemic connectedness that holds within a person’s total set (or “web”) of beliefs. A belief-formation system is holistic when all of a person’s beliefs are in principle relevant to the formation and confirmation of any given belief. Philosophers of cognitive science developed these ideas via an analogy with epistemic holism in science, where seemingly remote disciplines (e.g., archaeology and astronomy) can end up informing one another in surprising ways (Antony 2003; Fodor 1983, 107–108).

**Scripts** or **schemata** are long-term memory structures that help us to efficiently categorize, respond to, and store information about different situations, events, or individuals. Because
Chapter 29: Neuroscience and Psychopathologies
Alex Pereira, Gemma Lucy Smart, & Dominic Murphy

Key terms and definitions (sidebar textbox)

Mechanistic explanation: explanation of phenomena by showing how they are produced by the interaction of parts of physical systems, rather than by departures from natural laws.

The Medical Model: mental illnesses are expressions of failures of normal function in cognitive and neurological systems.

Mental illness: a pattern of abnormal and distressing thought or behavior that impairs everyday life. Naturalistic analyses of mental disorder assume that these patterns are produced by failures of internal systems or by imbalance between those systems and the environment.

RDoC (Research Domain of Criteria) framework: an initiative by the US National Institute of Mental Health intended to redirect research towards projects that treat mental illness as disorders in brain circuits.

Feighner criteria: a set of proposals offered by a group of psychiatrists in 1972. The aim was to develop new diagnostic criteria for several important disorders, using a medical rather than a psychoanalytic approach in order to achieve greater validity and reliability.

Diagnostic and Statistical Manual of Mental Disorders (DSM): a compendium of psychiatric diagnostic criteria, listing signs and symptoms allowing for the diagnosis of disorders recognized by the American Psychiatric Association. The DSM was first published in 1952 but has been heavily revised and reformed.

Brief Outline of topics covered

This chapter looks at the foundations of modern psychiatry, with its stress on neurological malfunction, and asks about its strengths and limitations. We start by tracing some of the historical development of the ideas that have found their way into modern psychiatry from their roots in nineteenth century medicine and neuroscience. Turning to the present day, we briefly look at competing conceptions of mental illness, before we discuss the philosophy of science that form some of the theoretical foundations of modern psychiatry. We suggest that modern psychiatry makes a number of commitments that are quite familiar to philosophers of psychology, employing mechanistic explanations to address disorders that are seen as natural kinds. We end with an example, panic disorder, that seems to fit this picture, and one, addiction, that seems not to.

1. Introduction

Reductive, physical explanations for mental disturbances have existed in the West since classical antiquity, and these have often involved speculative theories about the brain (Scull 2015; Shorter 1997). The first wave of avowedly neurological psychiatry arrived in the late nineteenth century (Maudsley 1867) and included many neuroscientists throughout Europe. Freud’s first publications, for example, were in neuroscience (Glymour 1992). The great success story for nineteenth century biological psychiatry was the explanation of General Paresis of the Insane (GPI), which came to be understood as neurosyphilis (see box). GPI, with
Chapter 30: Neuroethics
Katrina Sifferd & Joshua VanArsdall

Neuroethics is the body of work exploring the ethical, legal, and social implications of neuroscience. This work can be separated into two rough categories. The “neuroscience of ethics” concerns a neuroscientific understanding of the brain processes that underpin moral judgment and behavior. The “ethics of neuroscience” refers to the potential impact advances in neuroscience may have on the ethical principles that should guide brain research, treatment of brain disease, and cognitive enhancement. The Contemporary Issues section of this chapter will consist of two sub-sections corresponding to these categories: one on what neuroscience can tell us about human ethics and the other on the ethics of neuroscientific applications.

Key Terms and Definitions

**Neuroethics**: a field of study at the union of neuroscience and ethics, divisible into two branches: the neuroscience of ethics and the ethics of neuroscience. The first branch explores moral judgment and action via an understanding of the brain, while the second focuses on the ethical implementation of neuroscience.

**Brain monitoring**: techniques used to observe the activity, structure, and function of the brain.

**Brain intervention or manipulation**: techniques used to modulate brain function, typically using chemical or electrical interventions.

**Brain privacy**: the idea that we have a right to keep the brain states and processes underpinning mentality private.

**Cognitive enhancement**: techniques used to improve brain function above a person’s typical level or above the standard for the average person.

**Moral responsibility**: whether agents can be considered morally praiseworthy or blameworthy for their actions.

**Moral judgment**: a decision regarding the moral nature of something; generally, whether it is good or bad, right or wrong.

**Dual process theory**: the idea that moral judgment can be separated into *controlled processes* (possibly associated with consequentialist or utilitarian judgments) and *automatic processes* (possibly associated with deontological judgments).

**Consequentialist moral theory**: the view that the morally right action is the action that produces the best results.

**Deontological moral theory**: the view that morally right action is action in keeping with certain rules or principles, and that some choices are morally forbidden regardless of their effects.

**Neuroplasticity**: the understanding that the nervous system is responsive to its environment; brains are literally shaped as the result of environmental experience.