Mind, Cognition, and Neuroscience

A Philosophical Introduction

Benjamin D. Young and
Carolyn Dicey Jennings

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

First published 2022
by Routledge
605 Third Avenue, New York, NY 10158

and by Routledge
2 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

Routledge is an imprint of the Taylor & Francis Group, an informa business

Library of Congress Cataloging-in-Publication Data
A catalog record for this title has been requested

ISBN: 978-1-138-39234-2 (hbk)
ISBN: 978-1-138-39236-6 (pbk)
ISBN: 978-1-003-24189-8 (ebk)

DOI: 10.4324/9781003241898

Typeset in Berling and Futura
by Apex CoVantage, LLC

CHAPTER 6

Introduction to Philosophy of Science

Carlos Mariscal

Chapter Overview

This chapter will be a brief survey of the concepts from the general philosophy of science for those interested in cognitive neuroscience. It covers several major topics in the philosophy of science: scientific explanation and underdetermination, reductionism and levels, 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 failed to discover explanations for phenomena and when established research did not withstand scrutiny, as well as the complex relationship between the study of the mind and the study of the brain and its parts.

These issues are common to many areas of science, but they can be particularly fraught in a field like cognitive neuroscience, as researchers from a wide variety of disciplinary backgrounds come together to develop a systematic understanding of the mind.

Key Terms

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

Levels: entities or processes that make up (or are made up of) other entities, such as psychological and biological properties; levels can be thought of as ontologically real ("levels of nature") or merely useful ("levels of description" or "levels of explanation").

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).

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.

DOI: 10.4324/9781003241898-7

Pragmatic explanation: a view of explanation that claims that explanations are acts of communication; successful explanations result in others' understanding.
Underdetermination: there is insufficient data to decide between interpretations (*weak or practical*); no amount of data would decide between various interpretations (*strong or logical*).
Ontological: a branch of metaphysics that focuses on the nature, structure, and categories of being and existence.
Methodology (methodological): the techniques followed in a particular discipline.
Scientific realism: the thesis that what science dictates (e.g., theory, entity, or relations) is approximately true.
Anti-realism (scientific): the thesis that what science dictates is not true. It is at worst a posit in need of wholesale rejection (eliminativism) and at best a useful fiction (instrumentalism).
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.
Fallibilism: the doctrine that all knowledge claims could, in principle, be mistaken (*weak*) or are probably false (*strong*).

1. INTRODUCTION

1.1. Historical Overview

The current debates in philosophy of science were largely set in the early-to-mid-1900s by a group of philosophers known as the logical positivists (also logical empiricists). Logical positivists criticized philosophical questions that could not be solved by empirical evidence, so claims about strangers' inner minds could not be a proper subject of science. This deeply affected the history of psychology in its turn toward behaviorism in the middle of the century and had knock-on consequences for the history of cognitive neuroscience (see Chapter 1). Many positivists had a scientific background, usually in physics, which was apparent in their positions. They believed that there was a strong relationship between the subjects of various sciences, advocating for a "unity of science" that was grounded in physical observations and logic (see Box 6.1). They also accepted that science explained phenomena by invoking laws of nature or other regularities (Hempel & Oppenheim, 1948).

Logical positivism was largely abandoned by philosophers of science who came to doubt their commitments. Positivists believed that science had a distinctive logic (Carnap, 1934/1937; Popper, 1935/1959), but that proved difficult to justify (Quine, 1951). Attention to the history of science showed that **theories** could change drastically, which led some to think that there were normal and revolutionary periods of science, questioning the long-held assumption that human knowledge was cumulative and sciences progressed (Kuhn, 1962). Taking this to the furthest extreme, some philosophers argued against the possibility of a distinctive scientific method, questioning whether science was any more rational than other human activities (Feyerabend, 1970). Attention to the diversity of approaches within the sciences further exacerbated this concern. It turns out that sciences other than physics operate quite differently, which led many philosophers to doubt that a unified view of science would be possible or desirable. This led to the splintering of philosophy of science into the philosophies of various sciences, such as philosophy of physics, philosophy of biology, and philosophy of neuroscience. In the Future Directions section (Section 3), we will discuss this trend and its implications for cognitive neuroscience.

In the rest of this chapter, we will discuss subsequent developments in the philosophy of science, focusing on the most relevant viable philosophical stances. Each of the topics we will address will shed light on the possibility of the unity and future direction of cognitive neuroscience. Understanding these background theories will be key to understanding the philosophical issues raised by the diverse and sometimes challenging research in cognitive neuroscience.

Box 6.1 Levels

Many researchers believe that minds are biological, biological organisms are chemical, and chemicals are made of physical particles. As such, philosophers sometimes describe the sciences as focusing on distinct "**levels**" of explanation, if not organization (Oppenheim & Putnam, 1958). In this view, we might distinguish the following levels: social groups, multicellular organisms, cells, molecules, atoms, and subatomic particles. Each level is composed of elements from lower levels, but not from higher levels. (One may notice the absence of "minds" from that list, which is to reflect the controversy over what minds are and how they relate to the other items. "Culture" is another, equally vexing omission.)

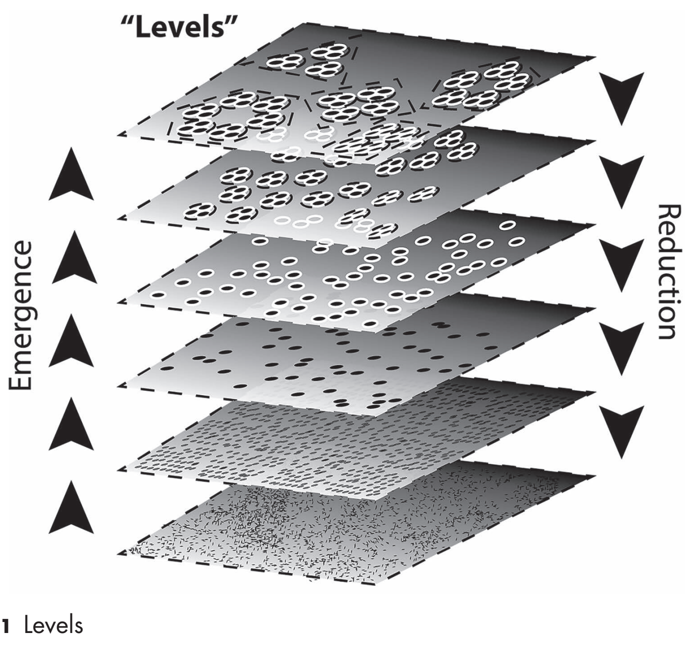

FIGURE 6.1 Levels

Logical positivists believed that science could be unified via *theoretical reduction* (see Section 1.2.2 on reductionism). In this view, "higher" sciences like biology or economics could conceivably be understood solely in terms of "lower" sciences, like physics and chemistry. There would be some sort of "bridge law" mapping thoughts onto neural patterns, for example. Sciences are discrete and autonomous in that they study a "level" of reality that exists independently of scientists, but each level can be reductively explained in terms of another, lower level.

While this view of levels of organization makes some intuitive sense, we should be careful to note that the levels analogy breaks down in several ways. First, scientists often use the tools and terms of nearby sciences to explain phenomena in their own field, so the levels are not exclusive. Second, many sciences study multiple "levels" of reality: biologists, for example, study everything from cells to ecosystems, from an instant to deep evolutionary time, and from observation to experiments. Third, aside from basic physics and chemistry, there are few bridge laws to be found in science, and many people have lost hope of discovering more. Finally, it is not obvious that nature divides into the same levels that our sciences do, as many levels interact with each other.

Instead of this "layer cake" model, many philosophers now prefer the view of sciences as investigating overlapping subject matter according to their own principles, theories, or methods. The sciences do not focus on discrete levels of reality (if there are such levels); they overlap. Because they intersect, there are no bridge laws to discover. That said, given that sciences presume different principles, theories, and methods between sciences that intersect, there is room for philosophers and scientists to develop and clarify incongruous theoretical concepts.

Cognitive science, neuroscience, psychology, and related disciplines may all treat different aspects of the same phenomenon or overlapping phenomena. Replacement of the "layer cake" model by a more complex Venn diagram of research approaches results in many open questions about how each approach affects the others. No approach is ultimately superior to the rest, so the search for unity or translation between the various philosophical assumptions of research programs is likely to be fruitful for decades to come.

1.2. Background Theories

1.2.1. Scientific Explanation

Scientific explanations show why the world behaves this way and not some other way. Philosophers have developed many accounts to understand the nature of scientific explanations, with the hope of understanding *how* explanations work, when they work, and how they might be distinct from everyday human explanations.

Recall that logical positivists accepted that explanations invoked laws of nature. This might be adequate for some explanations, especially in the field of physics, but laws and lawlike regularities are rare in most areas of science. There may be laws that govern action potentials (e.g., Lucas, 1909), but at other scales of cognitive neuroscience, laws are hard to come by and often riddled with exceptions. As a result, philosophers have sought alternative accounts of explanation. Two accounts are relevant to us here: **causal** and **pragmatic explanations**:

- **Causal explanations** are explanations that describe why a state is the way it is due to prior states. There are several varieties, of which two are relevant to our purpose: *manipulationist* views (Woodward, 2003), which accept that causation is a dependence relation such that manipulating a cause alters the effect, and *mechanism* accounts (Machamer et al., 2000), which invoke *entities* organized in such a way as to produce regularities. The differences may be subtle, but the former focus on systematic difference-making *patterns*, whereas the latter seek individual, discrete *mechanisms*. Scientific models may invoke both, as in gears turning a clock. Some explanations invoke many mechanisms for a single effect, as in psychiatric disorders, which may have many triggers. Conversely, a single mechanism can sometimes reliably cause multiple effects, such as neurotransmitters, which can serve many roles.

- A **pragmatic explanation** is one that serves the researcher's current explanatory goals, rather than some abstract ideal. As such, pragmatic explanations are tied to others' understanding. Pragmatists accept that explanations are acts of communication, which is a key insight into the social aspects of science. For example, "why is Broca's area involved in language?" is open to at least two interpretations: "why is it Broca's area that is involved in language (and not some other area)," and "why is Broca's area involved in language (and not some other behavior)." An explanation of the former might invoke all of the brain areas involved in language and how Broca's area fits in. An explanation of the latter might be a thorough analysis of all of the times Broca's area is used and what it may be used for. Scientific explanations depend on the nature of the problems involved and the *interests* of the researchers.

The philosophical debate over scientific explanation has gone on for decades, but two widely accepted conclusions are relevant: *multiple kinds of explanations are possible* and *explanations can be quite hard*. For example, we may accept the subjectivity of pragmatic explanations without thinking that they undermine the objectivity of science, because science also uses causal explanations. Similarly, it may be that talking on a cell phone increases the risk of brain cancer without us being able to articulate the precise mechanisms by which that happens. (In this latter case, we would have a successful causal-*manipulationist* explanation but not a complete causal-*mechanist* explanation.)

One issue with explanation comes with a special name: "**underdetermination**." We may think about it in terms of evidence. When the evidence *determines* its interpretation, we mean that there is only one reasonable interpretation of the evidence. We can also have multiple lines of evidence that all point to a single, unique interpretation. This would be termed "overdetermination." When fMRIs, autopsies, and surgeries all point to the same part of the brain being involved in, say, vision, we conclude that the area's role in vision is *over*determined by the evidence.

The opposite, and more relevant, phenomenon is underdetermination. In most cases in science, we have insufficient evidence for any particular interpretation. In controlled settings, a particular behavior may correspond with increased blood flow to a region of the brain, but it is unclear whether that region was involved in that behavior, whether some area connected to it was involved, or whether the individuals being studied are simply unusual.

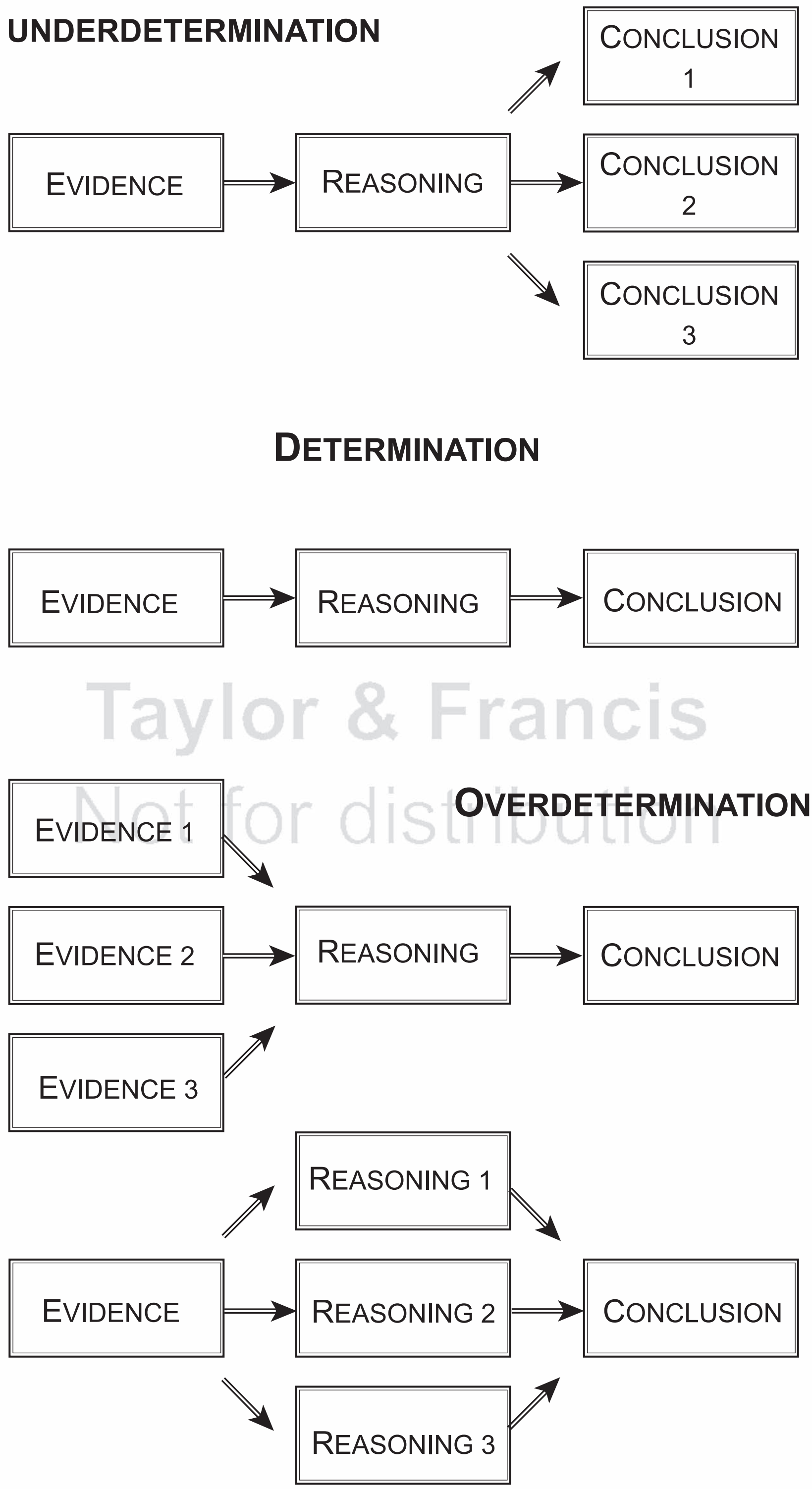

FIGURE 6.2 Underdetermination, determination, and overdetermination

Many areas of cognitive science are particularly vulnerable to underdetermination worries. When scientists work with humans, there are many interventions that are unethical or illegal to perform. Conversely, when scientists work with nonhumans, they are unable to communicate, and translation between animal models and humans is not always direct. If the question involves various sciences or disciplinary approaches, the underdetermination worries are often much harder, as reasonable experts may disagree about methods, data, and interpretation. As a result, there are often many viable explanations and scientists must choose between them, assessing them by some standard virtues of scientific theories: simplicity, precision, falsifiability, conservativeness, fruitfulness, reproducibility, generality, etc. [see Kuhn (1977) and more recent discussions such as Lacey (2004), Douglas (2013), or Keas (2018)].

1.2.2. Reductionism and Anti-reductionism

Many undergraduates have an intuitive tendency to seek a particular kind of explanation: **reductionism** (also discussed in Chapter 3). A quick, slogan definition of reductionism may be "smaller science is better." There are many varieties of reductionism, only some of which are about explanation (as with the previous section), while others are about scientific theories or the entities in them. It is important not to confuse notions of reductionism. A brief overview follows:

- **Theoretical reductionism**: the view that theoretical terms can be linked to lower-level terms, rendering one a special case of the other. If the theoretical term is "belief," for example, the lower level term might be the brain states involved in that belief. However, it is not enough to connect a single thought with a single brain state: theoretical reductionism requires all like thoughts be connected, in principle, to a set of like brain states. Philosophers in the mid-1900s were quite concerned with the issue of *theoretical reduction*, but there has been little success outside of physics (e.g., thermodynamics → statistical mechanics, Newtonian mechanics → Einsteinian relativity). Nevertheless, some projects in cognitive science may be viewed as attempts to bridge large theoretical gaps and may involve theoretical reduction (see Chapters 11 and 26).
- **Ontological reductionism**: the view that entities at a higher level simply *are* entities at a lower level. In this view, there is no such thing as a novel entity that arises from the combination of other properties. So, cognitive agents simply *are* collections of cells and organs. Ontological reductionism can be about kinds of objects or individual instances of those objects, what some philosophers might call "types" and "tokens." Tokens are widely accepted as easier to explain. For example, there may always be *some* brain state that corresponds with my thinking about waffles, without it being the same corresponding neurons firing each time for everybody who has ever thought about waffles.
- **Explanatory reductionism**: the view that explanations are always improved by appeal to lower-level phenomena. For example, in this view, facts about human psychology are supplemented by appeal to underlying neural correlates, which in turn are augmented by biological facts, chemical facts, and their underlying physical facts. For explanatory reductionists, we choose not to use quantum mechanics to explain thoughts about waffles merely due to our own cognitive limits.

- **Methodological reductionism**: the investigative approach to always provide a lower-level explanation if possible. Behaviorists were (at least) methodological reductionists, seeking to investigate behavior on its own rather than posit mental properties. Contemporary neuroscientists are also often methodological reductionists when they seek neural correlates to the mental or behavioral properties they study.

On the other hand, for each reductionist view, there are those who deny it. *Ontological anti-reductionists* believe that nonreducible entities exist at two or more levels. For example, the mind and the brain are distinct, but each is a distinct kind of entity. *Theoretical anti-reductionists* hold that there are theories at different levels that will never be successfully translated into one another: perhaps psychology and neuroscience require distinct theoretical posits, such that discoveries in one cannot be translated into the other without accepting those posits and related terms. *Explanatory anti-reductionists* accept that some explanations are superior without reference to lower-level empirical facts. Note that some areas of study have resisted explanatory reductionism, despite the best efforts of their practitioners. For example, there are currently no biological tests for any of the psychiatric disorders included in the widely used *Diagnostic and Statistical Manual of Mental Disorders* (Nesse & Stein, 2012) (see Chapter 29). Finally, *methodological anti-reductionists* accept that some scientific questions are better pursued without reference to their constituent parts. For example, linguists need few of the methods of biology.

Vision scientist David Marr famously proposed an anti-reductionist account of the mind (Marr, 1982). For the same phenomenon, one could investigate the goal or *function* of an activity, its *abstract* representation, or its neural realization or *implementation*. Perhaps due to the predominant computational and informational metaphors at the time, he termed these levels "computational," "algorithmic," and "implementation." Marr introduced his account after critiquing earlier approaches as either vague or mere promissory notes that never materialized (Bickle, 2015; Eliasmith & Kolbeck, 2015). In any case, Marr showed how a mature scientific study could operate at three distinct, compatible, and equally important levels of description.

1.2.3. Scientific Realism and Anti-realism

One question that may interest scientists is whether our models somehow reflect the structure of reality or merely organize our thoughts in ways that happen to be useful. Philosophers have long been interested in the reason for the success of science. There are two broad approaches: **scientific realism** and **scientific anti-realism**. Scientific realists accept that science is in the business of describing reality and is (sometimes) successful at doing so. Scientific anti-realists deny one or both of these claims.

As with the prior sections, there are several distinct kinds of scientific realism. Realists may accept that future science will describe reality more and more, even if current science does not (*progressivism*), they may accept the accuracy of some current scientific theories (*theory realism*), they may accept that scientific objects exist even if the theories describing them are false (*entity realism*), or they may accept that the mathematical and conceptual relationships discovered by science are true even if the entities and theories that discovered them turn out to be false (philosophers call this "*structural realism*," but perhaps we ought to call it "relations realism" so as not to confuse those who think about brain structures). The strongest argument for scientific realism has been the continued success of science. As philosopher Hilary Putnam put it, scientific realism "is the only philosophy that doesn't make the success of science a miracle" (1981, p. 1).

As an example of scientific realism, consider Alvarez and Squire (1994)'s theory of memory. Their model holds that the medial temporal lobe (MTL) acts as a way stop between the cortices for short-term memory. The model accounts for some observations that people with hippocampal damage tend to lose recent memories but retain older ones. What would be a scientific realist approach to this model? A *progressivist* may accept that it is a useful model and is perhaps on its way to an accurate description of how memory works. A *theory realist* approach would accept it as broadly true. *Entity* and "*relations*" *realists* may accept that the broad structures or relationships are real and possibly involved in memory, though there may be other structures and other pathways that are not included in this **theory**.

When faced with explanatory, predictive, and observable phenomena, the temptation toward realism may be quite strong. Nevertheless, many philosophers and scientists remain anti-realists. In fact, we all are scientific anti-realists about *something*: no current scientist accepts phrenology as accurately representing reality, for example. Scientific anti-realists also come in many varieties: some accept that what we can observe is real, but that we should withhold judgment of the reality of the unobservable, theoretical, or abstract entities we use to explain those observations (*constructive empiricist*). Others deny that science is in the business of seeking truth altogether (**instrumentalism**), while still others believe that scientific explanations only make sense with respect to local norms and cannot be applied more broadly (*constructivists*). The strongest argument for anti-realists comes from careful attention to past and alternative views of science. Most science throughout history has been wrong, even though the scientists involved were every bit as talented as current scientists. One may view this history as a reason to be wary that some of our current views may turn out to be false (**weak fallibilism**) or as a reason to be suspicious that most of them are fundamentally mistaken (**strong fallibilism**).

Let's think back to the Alvarez and Squire model and explore it through the lens of scientific anti-realism. *Instrumentalists* will think the model is important as long as it is useful, but they will remain silent on whether it reflects any deeper "truth." For them, the purpose of science ends when it can make reliable predictions and explanations. If the model were known to be false, but still more useful than the more accurate, but convoluted reality, we should still use the model. C*onstructive empiricists* would be entity realists about the observations that led to the model, while remaining *instrumentalists* about its theoretical aspects. On the other hand, *constructivists* will deny that this is the sole or even the most accurate description of reality—it is merely what fits the data that scientists are interested in explaining in ways scientists consider explanatory. *Fallibilists* may hesitantly accept this model (*weak*) or deny it (*strong*), while being continually aware that future science could show it to be entirely mistaken.

As with explanation and reduction, debates about scientific realism followed a similar pattern: views that were thought to be opposites were eventually accepted as able to coexist. Science can (and does) use multiple notions of explanation in different contexts, and scientists can be reductionists or realists about some scientific claims, but not others.

With these issues in philosophy of science under our belt, let us return to the question of how philosophers of science understand science and what they make of cognitive science.

1.3. Historical Development

Cognitive neuroscience uses the **theories**, principles, and methods of neuroscience, psychology, philosophy, and many other fields that may have wildly different theoretical commitments and interests. These range across many axes, including *time span*, from millisecond-scale recordings, such as EEGs, to long-term lesion studies; *spatial scale*, from microelectrode single-unit recordings to whole-brain imaging; and *substrate*, from computer models to behavior (see Chapters 2 and 4). For example, some researchers have attempted to map brain-neuron connections digitally, in the hopes of discovering secrets about consciousness and other mental activities (Markram, 2006). Even if such a project is viable, how it relates to other cognitive neuroscience projects, such as, say, single-unit recordings, is something left for future theorists to unify. This will not be an easy task, and similar tasks have eluded continued investigation.

While cognitive neuroscience is often called an interdisciplinary science, it is probably more accurate to label it a "multidisciplinary" science. Unlike interdisciplinary fields, which focus on questions at the intersection of disciplines, and transdisciplinary fields, which do not quite fit into any traditional demarcations, multidisciplinary fields use the **theories**, tools, techniques, and terminology of several individual sciences. Cognitive neuroscience is not alone in this feature: planetary science, climate science, and astrobiology are also fields that attempt to integrate the findings of distinct fields into a single, global understanding. In those fields, it is important to translate between quite different approaches without making assumptions that are taken for granted in one field but unknown or disregarded in others.

How do we understand a field in which prominent scientists disagree about how to interpret its major, or even foundational, results? What if they disagreed as to the nature of the field itself? At first, it may seem as if the answer to this question could be answered in the abstract, but there are certainly social factors to consider. The role of social factors in scientific understanding divides the philosophical community. An optimistic view is to wait and see: perhaps future data will result in scientists accepting a unified, theoretical interpretation. Because scientists are engaging in experiments and discovering facts about the world, this view holds, we should just trust that the truth will eventually win out and scientists will come to an agreement on how to understand the mind. An opposing view accepts that "success" in science depends largely on contingent social factors. When those factors go away, so, too, will any semblance of unification under a common purpose. In this view, when universities stop hiring and funding sources dry up, cognitive neuroscientists will find other ways to describe their research and the field will be relegated to yet another chapter in the history of the study of the mind. For advocates of this perspective, there are a multitude of ways of seeking, interpreting, and integrating data.

One hopeful sign for the field is that it has reached several sociological milestones that are characteristic of developed sciences: proprietary (and conceptually unified) scientific journals, conferences, textbooks, and even several university programs and departments (Berkeley, University of California, San Diego, Johns Hopkins, etc.). Furthermore, as cognate fields have developed, cognitive neuroscience has not been discarded, but has rather changed and adapted. From the perspective of philosophy of science, cognitive neuroscience is on a positive trajectory, although still negotiating its disciplinary boundaries, approaches, and philosophical commitments.

2. CONTEMPORARY ISSUES

Along with the conceptual challenge of integration that faces cognitive neuroscience, there are practical problems about what standards of evidence we ought to use in accepting interpretations of research. In recent years, scientists in general, and cognitive scientists in particular, have become cognizant of certain issues facing science: inherent biases as to what gets published, interpretive errors, and the difficulty of replicability. These issues were shocking to many practicing scientists, but they have long been discussed by philosophers:

- A series of researchers have pointed to author, reviewer, and journal bias against the publication of negative findings or replications (Greenwald, 1975; Coursol & Wagner, 1986; Neuliep & Crandall, 1993a; Madden et al., 1995; Callaham et al., 1998). When these studies *are* published, they are usually not labeled as such (Neuliep & Crandall, 1993b), suggesting a pervasive bias in favor of novel, positive results.
- In 2005, John P.A. Ioannidis published the paper, "Why Most Published Research Findings are False," in which he argued that certain research practices serve to undermine individual scientific results, and not correcting those undermines the entire practice. For example, scientists regularly interpret many small studies as adding up to large, robust experiments, or they collect data until statistical significance is reached. So even the studies that make it to publication may face systemic issues.
- In 2009, Bennett et al. sparked a firestorm for a study involving the neuroimaging of a dead salmon. They were making a statistical and **methodological** point: if one uses statistical corrections improperly, they may result in false positives. It was not an idle point; several studies, they thought, were engaged in such statistical bad practice.
- Five years later, Sorge et al. (2014) published a study that suggested rodents had increased fear reactions when exposed to androgens, suggesting that prior studies that did not report the gender of the researchers conducting the experiments may have been studying rats in varying and unaccounted for conditions. That science may have ignored what turned out to be a significant and major variable rippled across the scientific community, causing many reflective scientists to worry about what unexamined factors they might still be overlooking.
- The next year, the Open Science Consortium (led by Brian Nozek) attempted to reproduce 100 well-known psychological studies, but it managed to only reproduce between 30 and 40 of them, depending on the criteria used for reproduction. Of the studies that were successfully reproduced, the effect size found averaged about half of what was initially claimed in the original studies. This

was a major point of worry and has since been performed (with similar results) in fields such as economics and cancer medicine (Camerer et al., 2016; Begley & Ellis, 2012).

- The next year, the magazine *Nature* carried out a survey that found that 70% of researchers had failed an attempted replication and that 52% believed there was a crisis; that number was largely consistent across fields (Baker, 2016).

In sum, there are systemic problems with science publications. Such problems may undermine strong interpretations of the accuracy or utility of research explanations. If published studies are systematically more positive than general research suggests, that might make us doubt the results of studies. This problem is compounded in multidisciplinary fields like cognitive neuroscience, in which expertise in one area does not necessarily imply expertise in related areas.

One may wish to avoid relying on statistical claims like these by appealing to the reality of observed mechanisms, in line with the causal-mechanisms approach. But even observed mechanisms may be challenged by systemic biases in research programs or publications.

Science is in a crisis due to its overvaluation of novel and positive studies and the improper use of statistics, not to mention the always present fear that scientific **theories** are misunderstanding or missing key variables that would challenge established dogma if revealed. A community-wide reckoning has been ongoing for many years, but the nature of the problem makes it particularly hard to address. Systemic problems are ones that need to be addressed by a community-wide change in value structures and procedures, but culture changes slowly and unevenly. We will eventually require more mathematical rigor and a community-wide conversation about the philosophical goals of science, but to successfully address these problems for science may take more than a generation, if it is even possible to do so.

3. FUTURE DIRECTIONS

All science involves philosophical commitments. Scientists must be aware of these commitments or risk making simple conceptual mistakes. Scientists cannot avoid philosophy, as every scientific discovery, creation, or **theory** raises some basic philosophical questions:

- How do we know it?
- How does it fit into what we believe about the world?
- What ought we to do about it?

Take, for example, the growing field of cognitive genomics, an interdisciplinary study of the links between genetic and epigenetic markers and cognitive abilities. Such a field is at the intersection of other, more developed fields with specific paradigms, tools, methods, jargon, standards, and approaches. In an ideal case, these alternative approaches would coincide in a single explanation that appeals to experts from both backgrounds. As an example, consider Huntington's disease, in which a single coding section of the genome (a CAG repeat) is linked causally to a degenerative cognitive disease impacting adults. Other cases, such as intelligence, are fraught with confounding variables, such as the many implicated genetic loci that account for a tiny percentage of the known variability, cultural biases in what counts as intelligence, and methodological problems in measuring it. In such a case, there are philosophical problems with even concluding whether intelligence is a single phenomenon, whether it is stable or innate, and how the complex network of genes, development, and environment interact to produce it. Last, even if these issues are addressed adequately, there is still the ethical question of what ought we to do about it? Should we conduct such research at all and at what priority, given its terrible history? If we do conduct the research, what ought we to do with the results, knowing that even the choice of messenger and the message are loaded with moral import, much of which we can be sure will be (mis) used for problematic purposes.

This means that cognitive neuroscientists will always require philosophical thinking. The debates discussed in this chapter will come up again and again in different guises for decades to come. Hopefully, these will work hand in hand with scientific research, rather than as after-the-fact critiques. Understanding the multifaceted nature of cognition is one of the most important projects in all of science, but it is also one of the most difficult. Cognitive neuroscience involves a number of inferences: from intentional mental states to nonintentional physical states, from theory to practice, from one instance to many, etc. Each of these requires a theoretical choice that must be justified philosophically or run the risk of not being useful.

Summary of Key Ideas

In this chapter, we covered several major areas in the philosophy of science: explanation, reduction, and realism. With respect to explanation, many cognitive scientists believe that we should focus on causal-mechanistic explanations (see Chapter 8), which researchers hope will avoid the worries of underdetermination. One should remember that scientific explanations also have pragmatic elements, at least some of which may not be causal in nature. Next, students of cognitive neuroscience are often ontological, explanatory, and methodological reductionists, although the field has a long history of well-accepted, anti-reductionist approaches. Last, with respect to scientific realism, we explored a variety of ways to be realists: about progress, theories, entities, or relations. Many scientists are realists about their theories and the entities or relations discovered by other scientists. There are also anti-realists, who believe that either science is not in the business of seeking truth (instrumentalism), truth does not often result from scientific practice (constructivism), or past errors should make us wary of current confidence (fallibilism). Cognitive neuroscientists may fit anywhere across these broad conceptual landscapes or in multiple areas, depending on the issue in question. The chapter closed with a discussion of whether cognitive science is one science or many and how it may proceed given conceptual and practical challenges.

Discussion Topics

- How do we distinguish between manipulationist and mechanistic explanations? Can you think of cases that would fit one but not the other?
- Can you think of an example of underdetermination from your own studies? What possible evidence might determine the best interpretation?
- What does it mean to be an anti-realist about a theory? How does this differ from being an anti-realist about a relationship or an entity?
- Does it make sense to divide the world in terms of "levels" of, say, biology, psychology, etc.? If so, is each increasing level increasing in the same way? If not, what would that mean about the world?
- Can you imagine a situation in which very different standards of evidence and disciplinary approaches produce a complete theory of the world? How about the reverse: a situation in which a unified approach produces an inconsistent patchwork of results?

Further Reading

Barker, G., & Kitcher, P. (2013). *Philosophy of science: A new introduction*. Oxford University Press.

Bechtel, W. (2013). *Philosophy of science: An overview for cognitive science*. Psychology Press.

Okasha, S. (2016). *Philosophy of science: Very short introduction*. Oxford University Press.

Staley, K. W. (2014). *An introduction to the philosophy of science*. Cambridge University Press.

REFERENCES

Alvarez, P., & Squire, L. R. (1994). Memory consolidation and the medial temporal lobe: A simple network model. *Proceedings of the National Academy of Sciences*, *91*(15), 7041–7045.

Baker, M. (2016). Is there a reproducibility crisis? A nature survey lifts the lid on how researchers view the crisis rocking science and what they think will help. *Nature*, *533*(7604), 452–455.

Begley, C. G., & Ellis, L. M. (2012). Drug development: Raise standards for preclinical cancer research. *Nature*, *483*(7391), 531.

Bennett, C. M., Wolford, G. L., & Miller, M. B. (2009). The principled control of false positives in neuroimaging. *Social Cognitive and Affective Neuroscience*, *4*(4), 417–422.

Bickle, J. (2015). Marr and reductionism. *Topics in Cognitive Science*, *7*(2), 299–311.

Callaham, M. L., Wears, R. L., Weber, E. J., Barton, C., & Young, G. (1998). Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. *Journal of the American Medical Association*, *280*(3), 254–257.

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., & Heikensten, E. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, *351*(6280), 1433–1436.

Carnap, R. (1937). *Logical syntax of language*. Kegan Paul Trench, Trubner & Co. Logische Syntax der Sprache (A. Smeaton, Trans.), (1934), Julius Springer Verlag, Vienna.

Coursol, A., & Wagner, E. E. (1986). Effect of positive findings on submission and acceptance rates: A note on meta-analysis bias. *Professional Psychology: Research and Practice*, *17*(2), 136–137. https://doi.org/10.1037/0735-7028.17.2.136

Douglas, H. (2013). The value of cognitive values. *Philosophy of Science*, *80*(5), 796–806.

Eliasmith, C., & Kolbeck, C. (2015). Marr's attacks: On reductionism and vagueness. *Topics in Cognitive Science*, *7*(2), 323–335.

Feyerabend, P. K. (1970). *Against method: Outline of an anarchistic theory of knowledge*. University of Minnesota Press, Minneapolis.

Greenwald, A. G. (1975). Consequences of prejudice against the null hypothesis. *Psychological Bulletin*, *82*(1), 1.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, *15*(2), 135–175.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, *2*(8), e124.

Keas, M. N. (2018). Systematizing the theoretical virtues. *Synthese*, *195*(6), 2761–2793.

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press.

Kuhn, T. S. (1977). Objectivity, value judgment, and theory choice. *Arguing About Science*, 74–86.

Lacey, H. (2004). *Is science value free? Values and scientific understanding*. Psychology Press.

Lucas, K. (1909). The "all or none" contraction of the amphibian skeletal muscle fibre. *The Journal of Physiology*, *38*(2–3), 113.

Machamer, P. K., Darden, L., & Craver, C. F. (2000). [MDC], Thinking about mechanisms. *Philosophy of Science*, *67*, 1–25.

Madden, C. S., Easley, R. W., & Dunn, M. G. (1995). How journal editors view replication research. *Journal of Advertising*, *24*(4), 77–87.

Markram, H. (2006). The blue brain project. *Nature Reviews Neuroscience*, *7*(2), 153–160.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Henry Holt and Co. Inc., USA

Nesse, R. M., & Stein, D. J. (2012). Towards a genuinely medical model for psychiatric nosology. *BMC Medicine*, *10*(1). doi:10.1186/1741-7015-10-5

Neuliep, J. W., & Crandall, R. (1993a). Reviewer bias against replication research. *Journal of Social Behavior and Personality*, *8*(6), 21.

Neuliep, J. W., & Crandall, R. (1993b). Everyone was wrong: There are lots of replications out there. *Journal of Social Behavior and Personality*, *8*(6), 1.

Oppenheim, P., & Putnam, H. (1958). Unity of science as a working hypothesis. *Minnesota Studies in the Philosophy of Science*, *2*.

Popper, K. (1959). *The logic of scientific discovery* (K. Popper, Trans.). Logik der Forschung (1935), Hutchinson, London.

Putnam, H. (1981). *Reason, truth and history* (Vol. 3). Cambridge University Press.

Quine, V. W. O. (1951). Two dogmas of empiricism. *Philosophical Review*, 60, 20–43.

Sorge, R. E., Martin, L. J., Isbester, K. A., Sotocinal, S. G., Rosen, S., Tuttle, A. H., Wieskopf, J. S., Acland, E. L., Dokova, A., Kadoura, B., Leger, P., Mapplebeck, J. C. S., McPhail, M., Delaney, A., Wigerblad, G., Schumann, A. P., Quinn, T., Frasnelli, J., Svensson, C. I., Sternberg, W. F., & Mogil, J. S. (2014). Olfactory exposure to males, including men, causes stress and related analgesia in rodents. *Nature Methods*, *11*(6), 629–632.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press, USA.