CHAPTER 7

Metaphysical Issues of Relevance to Cognitive Neuroscience

Crystal A. L'Hote

Chapter Overview

Following an introduction to metaphysics, this chapter highlights basic metaphysical issues relevant to cognitive neuroscience. First, the relevance of metaphysics is evident whenever a cognitive entity and a neural entity are correlated, as when Kanwisher et al. (1997) correlated face recognition with activity in the fusiform face area (FFA). Here, a question arises: What accounts for the correlation? Some possible answers are that FFA activity and face recognition are identical, that FFA activity constitutes face recognition, or that FFA activity causes it. The correct answer inescapably depends on core metaphysical issues: identity, constitution, and causation. Second, metaphysical issues are raised by the fact that people can think thoughts with the same content despite vast differences in brain size, structure, and connectivity. How can physically different brains be cognitively the same? What accounts for this? Third, metaphysics is needed to explain how cognition figures into the causal nexus. After all, thoughts not only are caused but are also themselves causes of behavior. How do mental states and processes, as such, do this? To make both sense and science of cognition requires inquiry into the nature of causation. Fourth and finally, the chapter briefly introduces the free will debate before noting the equal relevance of cognitive neuroscience to metaphysics.

Key Terms

Metaphysics: reasoned inquiry into the ultimate, fundamental, or general nature of reality as a whole or of its parts.

Ontology: a branch of metaphysics that catalogs the entities and categories of entities that exist.

Causation: the relationship between cause and effect.

DOI: 10.4324/9781003241898-8

Constitution: in particular, material constitution, the relationship between an entity and the material or materials out of which it is made.
Composition: the relationship that exists between a whole and its parts.
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 referring to its object.
Content essentialism: the thesis that the content of a mental event, state, or process is essential to it.

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. However, the prefix "meta" has a more modest meaning in this context. The word metaphysics was coined by an early editor of the works of Aristotle (384–322 BC), who named the books following the treatise on physics, "te meta ta physika biblia," that is, the books *after* the book on physics. In those books, 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, while the subject matter of the sciences is individual beings and categories of beings, the subject matter of metaphysics is what all beings have in common: beings as such. Aristotle seeks generality, ultimacy, and fundamentality, aiming at the "first causes and principles of things," but he seeks them in this world.

Long before Aristotle, however, philosophers were pursuing answers to core metaphysical questions. What is the nature of reality as a whole? What is fundamental? Why? The earliest answers on record are strikingly scientific. Thales (624–546 BC) argues that water is the source of all things. Anaximenes (586–526 BC) contends that air is. According to Democritus (460–370 BC), the fundamental nature of reality is "atoms and void." Notably, these early thinkers also seek to understand the place of such intangibles as mind (νοῦς), love (φιλία), and soul (ψυχή) in the grand scheme of things. Empedocles (494–434 BC) construes love as a natural force that mixes the basic elements (water, air, earth, fire), while strife drives them apart. During this fascinating pre-Socratic period, to separate science from philosophy is difficult, if not impossible. The metaphysics is often scientific, and the science is metaphysical. Metaphysics and science are united in the pursuit of unified knowledge.

Over the last few millennia, metaphysics and science have also competed or simply parted company and gone their separate ways. Some ancient and medieval philosophers, for example, sought to affirm the existence of an essentially transcendent reality by using the power of reason alone, without the aid of the senses or empirical sciences. Centuries later, Kant (1724–1804) would deny the possibility of metaphysics so understood. Impressed by the relative success of the scientific revolution, Kant is critical of attempts to seek knowledge of a transcendent reality in itself and apart from any possible experience.

Without the aid of the senses, he contends, reason's attempts to determine whether the world has a beginning or whether there is a transcendent God, for instance, will end in puzzle and paradox. Such matters exceed reason's limits.

Kant had delivered a significant blow to the "high" metaphysics of the medievals. Metaphysics was slow to recover until the 20th century, when Quine (1951) championed a "low" metaphysics aimed at uncovering what exists within the world of possible experience rather than beyond it. In Quine's view, the task of metaphysics is to sort out the ontological commitments of our best scientific theories. If we assume that our best theories are true, what exists? For example, does our best science entail the existence of causes or merely successive events? So understood, metaphysics is not *a priori* inquiry into the nature or structure of reality as a whole. It is an empirically informed inquiry into the realities affirmed by our best science, an inquiry that clarifies, refines, supplements, and unifies.

Critics of Quine's scientific naturalism (Box 7.1) find him to be "anti-metaphysical," a charge that reveals a competing view about the proper aims and methods of metaphysics. Recent decades have seen the return and rise of metaphysics in both its old and new forms as well as a rapid expansion of empirically informed approaches. Some philosophers aim to give an account of the whole of reality in terms of its fundamental structure, to "understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term," as Sellars (1960) famously put it. Others investigate specific domains in collaboration with mathematicians, physicists, biologists, and others to confront a range of questions fundamental to these particular areas of inquiry. What is the nature of probability? Time? What are the units of evolutionary selection? Here, the working relationship between metaphysics and the sciences is often a dialogical or dialectical back-and-forth. The metaphysics encourages higher levels of generality, broader consistencies, and deeper sense-making. The science discourages unmoored abstraction.

Box 7.1 Naturalism

Work at the intersection of metaphysics and cognitive neuroscience generally presupposes naturalism. Naturalism is the view that everything that exists is natural (not supernatural) and, so, can be studied by the natural sciences. Naturalism has different versions as an ontological or a methodological claim. The strongest version of ontological naturalism not only denies the existence of supernatural entities like ghosts and gods and immaterial souls but goes further to deny the existence of abstract entities like numbers and sets because these lack space-time dimensions. A weaker version of ontological naturalism will admit abstract entities and mental phenomena like thoughts and consciousness, as long as these can be shown to reduce to or depend on patently natural entities. As a strictly methodological thesis, naturalism is agnostic about the existence of supernatural, mental, and abstract entities. Still, according to the methodological naturalist, our investigation of reality ought to proceed as though everything is natural and can be known through science. In both versions of naturalism, metaphysics turns out to be importantly continuous with the natural sciences.

How do these developments bear on cognitive neuroscience? First, history is the beginning of an answer to skepticism about the relevance of metaphysics. Even a cursory survey reveals fruitful collaborations between metaphysics and the sciences and long periods during which even distinguishing them is a challenge. Second, history shows the different ways that metaphysics and cognitive neuroscience might collaborate. Metaphysics might describe the **ontology** of a special science like cognitive neuroscience or prescribe its ontology, or it might do both. Metaphysics might elicit foundational assumptions or provide them, or both. It might alter or be altered by cognitive neuroscience, or both. Last, to appreciate specific metaphysical issues in cognitive neuroscience requires familiarity with metaphysics, generally, and with the second-order debates that shape its collaborations. A historical definition of metaphysics thereby complements an ahistorical definition of metaphysics: metaphysics is reasoned inquiry into the ultimate, fundamental, or general nature of reality as a whole or of its various parts.

2. CONTEMPORARY ISSUES

An introductory course in metaphysics is likely to treat topics in philosophy of mind (see Chapter 5), including free will and determinism, mind and body, and the nature of selfhood and consciousness, all of which are metaphysical at root, in addition to topics that cut across domains of inquiry: existence and identity, objects and their properties, **causation**, possibility, space and time, and God. This section focuses on general metaphysical issues relevant to contemporary cognitive neuroscience before very briefly turning to the issues of free will and selfhood as they arise in neuroscientific and neurotechnological contexts.

2.1. Correlations

To establish correlations between cognition and neural activity is no easy task. Even in the wake of Kanwisher's (1997) pioneering work, disagreement persists about how well the ability to recognize faces correlates with fusiform face area activity (see 35 on the brain map). Some research suggests that other brain regions are face selective; other research suggests that the fusiform face area (FFA) activates in the presence of nonfaces (Slotnick, 2013). Here, we focus on questions that remain after correlations have been established: What accounts for them? What is the relationship between face recognition and FFA activity, for example, that explains their correlation? It is one thing to discover a correlation; it is another to explain it.

One possible explanation is that recognizing faces and FFA activity are the same thing, that they are identical. If true, this would (more than) explain their correlation. Another explanation is that, although face recognition and FFA activity are not strictly identical, FFA activity *constitutes* face recognition. This, too, would account for their correlation. A third possible explanation is that face recognition is caused by or somehow emerges from FFA activity. Which of these is the correct explanation inescapably depends on core issues in metaphysics: the meaning and nature of identity, **constitution**, and causation.

Consider, first, the view that the correlation is to be explained by the identity of face recognition and FFA activity. Why would one think that these correlated phenomena are identical, that they are one and the same? The bare fact that face recognition and FFA

activity are correlated is not a reason to prefer this explanation over the other two. So, what more must be shown to establish that correlated entities are identical?

Although little may seem more obvious than A = A, few philosophic issues are more difficult or consequential than identity. According to Leibniz (1646–1716), famously, to demonstrate that some X and some Y are genuinely identical, one must show that there are no property differences between X and Y. Other standards are more demanding, but the Leibnizian standard is demanding enough to cast doubt on the identity of face recognition and FFA activity. After all, each seems to have properties that the other lacks. Neural activity in the FFA has the property of being oxygenated, for instance, and electronically charged. It would be odd to ascribe these properties to the act or experience of recognizing a face. Equally, facial recognition seems to possess a unity that is not possessed by the distributed neural activity accompanying it (see Chapter 19). Finally, face recognition is essentially about faces; if it were not, it would not be face recognition. But the accompanying neural activity still would be the neural activity that it is, even if it were not about faces. These differences are reasons to deny the strict Leibnizian identity of face recognition and FFA activity.

Some metaphysicians employ a weaker standard of identity that hews more closely to ordinary linguistic usage. Most would say that a statue is the same object as its clay, for example, while knowing full well that the statue and its clay have different properties. Unlike the statue, for example, the clay could have remained as a slab and could again become a slab without thereby being destroyed. To recognize these differences between the statue and the clay but to assert, nonetheless, that they are identical is to defend the possibility of *contingent identities*, i.e., actual identities that might not have been. Is there reason to think that FFA activity and face recognition are weakly identical, that they are identical but might not have been identical? What is more: FFA activity and face recognition seem to differ not only in their possibilities but also in their actualities. Any actual differences between them (oxygenation, unity, aboutness) are reasons to deny even their weak and contingent identity.

What, then, accounts for the correlation of FFA activity and face recognition? Perhaps FFA activity *composes* or *constitutes* face recognition. Face recognition and FFA activity might be related as a bicycle is related to its various parts (**composition**) or as a coin is related to its silver (constitution). A bicycle is not identical to its replaceable parts, but neither can it exist without them. A coin is not identical to its silver, but it has no existence apart from it. As Baker (1997) sees it, constitution is importantly between identity and distinctness. If FFA activity composes or constitutes face recognition, as bicycle parts compose a bicycle or as silver constitutes a coin, this would account for their correlation.

Another possible explanation is that FFA activity *causes* face recognition. At a first approximation, causes are distinct from their effects and precede them in time. If Sam hits a baseball and causes the window to break, what causes the breaking is distinct from the breaking and also comes before it. However, the neural correlates of cognition do not precede cognition. They are simultaneous with it. FFA activity does not come *before* face recognition but might be said to come *below* it and, so, to "give rise to it." Now, if one's model is baseballs and broken windows, then such upward causation will seem impossible. But the baseball model of causation is only one among many. To wit, Aristotle recognized four types of causation, only one of which (efficient causation) conforms to the baseball model. Magnetism arguably causes events in a different no-touch fashion, and to subsume

events under physical laws is a standard mode of explaining them (see Chapter 6). Some philosophers further allow that an entity may be a *causa sui* (cause of itself). There is no absurdity, then, in FFA activity being the cause of the face recognition with which it is simultaneous. This, too, would explain the correlation.

Although this section has focused on the neural correlates of cognition, the same issues arise in relation to consciousness (see Part III). The search for the neural correlates of consciousness (NCC) might take the form of a search for correlates of being generally awake and responsive, a search for correlates of conscious sensations, perceptions, or thoughts, or a search for the correlates of some other form of consciousness (see Chapter 16). Whatever form the search takes, it is meant to end in a correlation of neural and conscious entities. In this case, too, the correlation will stand in need of explanation.

Whether the correlation between neural and cognitive (and/or conscious) entities is best explained by identity, constitution, causation, or by some other relation is not settled by data or experiments alone. Explanation equally depends on the nature of the explananda: identity, constitution, and causation. Metaphysics, then, presents an invitation to make deeper sense of neuroscience. And there are scientific reasons to accept the invitation. If the metaphysical ideas implicit in a research program are inconsistent, seemingly scientific disagreements may be unnecessarily protracted. In such cases, scientific progress requires identification of the metaphysical assumptions that are causing the trouble, whether to clarify or to replace them.

For example, disagreements about the boundaries of Broca's area (see 28 on the brain map) may be reformulated as disagreements about the nature of identity and sameness. What counts as the same area? Why prefer a cytoarchitectural border to a myeloarchitectural one (Geyer et al., 2011)? At what point is a subject's brain area so dysfunctional that it no longer counts as an area or event of the same kind? At what point are parts so dysfunctional or nonfunctional that they no longer count as parts of the (same) whole? As Viola and Zanin (2017) show in the case of Broca's area, where neuroanatomists draw these lines presupposes some criterion of identity or sameness. Even brain parcellation, then, has interpretive and metaphysical dimensions.

2.2. Mental Sameness

Even if all brains were physically the same, sense would need to be made of correlations between cognitive and neural entities. By contrast, the following set of issues arises because physically different brains may be cognitively the same. Despite having brains that vary in size, structure, and connectivity, people manage to think about many of the same things: dogs, economies, Tokyo. Given the significant physical differences between brains, how is such sameness possible? What accounts for it?

The distinction between predicates and properties is important here. A predicate is a description like "(is) round." The corresponding property is roundness, the real and repeatable shape feature that satisfies the predicate. There may be no property answering to a predicate. For instance, no property answers to the predicate "is a unicorn" or "is a nonfrog." These predicates do not pick out features of reality. Properties, by contrast, are features of reality. They are difference makers. Because an object is round, it can roll. But there is nothing that nonfrogs, as such, can do; that horses are nonfrogs adds nothing to

their powers. Among the implicit aims of science and neuroscience is to sort the (mere) predicates from the (real) properties.

Mental properties are (real) properties characteristic of minds. Consciousness is a mental property. Another is intentionality: our thoughts, memories, hopes, and plans are about things such as dogs, economies, and Tokyo, and they are essentially about those things (see Chapter 13). When Raquel thinks about dogs, it is not an incidental feature of her thought that it is about dogs. It is defining. Change what Raquel's thought is about, and the thought ceases to be. In this respect, mental states are unlike ordinary physical objects like tables and rocks, which do not have intentional properties (see Chapter 5).

The thesis that a thought's **content** is essential to it is **content essentialism**. Accordingly, just as it is essential to Raquel's dog thought that it is about dogs, so is it essential to Sylvia's dog thought that it is about dogs. And so is it that all dog thoughts have the same essential aboutness property, no matter how different are the brains of their thinkers. What is the nature of this shared aboutness property? What is common to Raquel's, Sylvia's, and all other dog thoughts, in virtue of which they are all thoughts about dogs? What accounts for this?

Three categories of response are available. First, a researcher might search for the neural correlate of all dog thoughts or for a common pattern of neural activity that repeats when and only when thinkers think about dogs. However, brains are so physically different from one another that the likelihood of finding a unique pattern is slim. At least, any neural pattern that is common to all (and only) dog thoughts will be so abstract that its causal relevance will be dubitable.

Alternatively, having failed to identify a signature neural pattern, a researcher might simply give up on the idea that dog thoughts have any neural property in common. Consider an analogy, developed to a similar purpose by Wittgenstein (1953). Basketball and soccer share characteristics, as do chess and checkers, and tag and hide-and-seek. Although these different games resemble each other in various and overlapping ways, no single feature is common to all (and only) games except, perhaps, that we apply the predicate "game" to them. For related reasons, a researcher might deny the existence of a property that is common to all dog thoughts except, perhaps, that we apply the same predicate to them.

But this response is hardly satisfying. The question is not whether all dog thoughts have a property in common. Despite the differences between them, they are all about dogs. The question concerns this shared aboutness property. What accounts for it?

A third response appeals to the distinction between intrinsic and extrinsic properties. Loosely, intrinsic properties do not depend on anything besides the entity that has them. A coffee mug's intrinsic properties include its shape, size, and solidity. By contrast, extrinsic properties depend on external factors. A mug's extrinsic properties include being half-full or on the kitchen counter or the same color as the rug. By invoking the intrinsic-extrinsic distinction, a neuroscientist can deny that all dog thoughts share any (intrinsic) neural property without denying that they share any property. The property that all dog thoughts share might be extrinsic rather than neural.

In support of this response, Putnam (1975) offers a well-known (and well-worn) thought experiment. Imagine that Oscar lives here on Earth, where H_2O fills the lakes, flows from the faucets, and so on. At the same time, Oscar's molecule-for-molecule

duplicate, Twin-Oscar, lives on a distant planet, Twin-Earth, which is an exact physical duplicate of Earth with just one exception: the clear and refreshing liquid that fills Twin-Earth lakes and flows from Twin-Earth faucets is not H_2O but XYZ. Because H_2O and XYZ taste the same, have the same boiling points, and are otherwise indistinguishable, Oscar and Twin-Oscar have the same experiences. While Oscar's thirst is quenched by the clear liquid that he calls "water," Twin-Oscar's thirst is quenched by the clear liquid that he calls "water." Still, Putnam argues, Oscar and Twin-Oscar mean different things by "water." Oscar refers to H_2O, and Twin-Oscar refers to XYZ. The meaning of their words depends on external factors, a lesson that extends to the content of their thoughts. While Oscar thinks about the liquid in his world (H_2O), Twin-Oscar thinks about the liquid in his world (XYZ). Although Oscar and Twin-Oscar are physically identical, their thoughts differ.

If this is right, then at least some mental properties are extrinsic. It is partly because we share a world with dogs in it, then, that our dog thoughts are about dogs, and partly because we share a world with faces in it that our FFA activity is about faces. More generally, it is partly because we share a world that physically different brains may be cognitively the same. Although mental sameness may sometimes or partly be accounted for by neural sameness, it also may sometimes or partly be accounted for by the world.

2.3. Mental Causation

Jess crosses the lawn because she believes it is the quickest path. Ivan waves because he recognizes the face of his friend. Thoughts are causes, whether of behavior or other thoughts. Such mental causation is commonplace, yet how do thoughts, inferences, and other forms of cognition, as such, fit into the causal nexus?

Ivan recognizes his friend's face and waves. Moreover, Ivan waves *because* he recognizes his friend and wants to greet him. Cognitive events, states, and processes do not simply cause behavior. They cause behavior in virtue of having the specific meaning or content that they do. Had Ivan believed that the figure before him was a stranger or a tree, he would not have waved. What Ivan believes seems to make a difference to what he does.

Yet, it seems that Ivan's behavior also could be explained without appealing to his beliefs. An alternative explanation might appeal to the prior physical state of Ivan's body-brain and explain his waving hand by appealing to the physical mechanisms that produce it, as one might explain the movement of a robot's arm. As Kim (1998, 2005) notes, the possibility of a complete physical explanation challenges the view that Ivan's belief makes a difference. If the physical facts about Ivan are sufficient to cause his behavior, it appears that no causal role remains for his belief. His belief seems to be an epiphenomenal extra. And if mental causes were something in addition to physical causes, then they would really be "something else" (Lowe, 1993).

That thoughts, beliefs, hopes, and desires cause behavior is not in question here. The present challenge is to understand how mental causation, as such, fits into the causal nexus, a challenge that is heightened by the extrinsic nature of some mental properties (see earlier discussion). Oscar believes that water is thirst quenching. Because he believes this, he turns on the faucet and fills a glass. But does it make a difference that his belief is about H_2O (and not XYZ)? On Twin-Earth, where XYZ flows from the

faucet, Twin-Oscar also turns on the faucet and fills his glass. The difference between the contents of their beliefs—Oscar's about H_2O and Twin-Oscar's about XYZ—does not seem to make any behavioral difference. The real causal work, it appears, is done by intrinsic physical properties. How, then, do mental causes figure in?

At the same time, there is a strong case to be made that Oscar and his twin do not behave the same way. Oscar not only fills his glass; he fills it with H_2O. By contrast, Twin-Oscar fills his glass with XYZ. Oscar showers with H_2O. By contrast, showering with H_2O is not something that Twin-Oscar can do. Arguably, then, Oscar and his twin behave differently in many instances. However, to establish that they behave differently is not yet to demonstrate that they behave differently *because* they have different beliefs. Both of these differences might be explained by a third variable: the difference between the liquids in their worlds. But behavioral differences at least make room for the possibility that beliefs, thoughts, hopes, and desires make a causal difference, a matter that ultimately depends on what it means for one thing to cause another (see previous discussion).

2.4. Free Will and Selfhood

Philosophers have doubted the compatibility of determinism and free will for millennia. It was only because Epicurus's (341–270 BC) ontology included randomly swerving atoms that Lucretius (99–50 BC) saw any room for free will. Maimonides (1138–1204) defended free will by denying that God's infinite foreknowledge was determining. Finding no place for free will in the temporal realm, Kant infers that its source must exist in the noumenal realm. The view that there is no room for free will in a universe that is wholly determined, whether the determining causes are material or immaterial, is called *incompatibilism*. The incompatibilist assumption motivates more than the view that free will resists or eludes the causal realm (libertarianism). Coupled with the premise that there is no realm beyond the causal realm, the incompatibilist assumption also motivates the view that there is no such thing as free will, that free will is an illusion (hard determinism). According to hard determinism, when Oscar reaches for a glass of water or Ivan waves, they do so because the prior physical state of the universe and the laws of physics necessitated it. Because no alternative beliefs, desires, or actions were possible, their effective volitions were not free, however strong their impressions to the contrary.

Advances in neuroscience and neurotechnology have inspired contemporary free-will skeptics. Neuroscience does not provide any direct or additional evidence for physical determinism (Roskies, 2006, 2014; Roskies & Nahmias, 2016). Still, when researchers can predict above chance whether a person will press a button before they are aware they will press it (Libet et al., 1983; see also, Shurger et al., 2012), it may be more difficult to believe that the choice is freely made. When a functional magnetic resonance imaging (fMRI) scan makes visible the neural mechanisms of planning, making it plain that no magic is involved, it may seem less likely that those plans are freely made. Finally, when deep-brain stimulation (DBS) and other forms of brain manipulation modulate experience, behavior, and sense of agency, the view that we are ultimately controlled by whatever it is that controls our brains may seem more plausible. It is tempting to conclude that we lack free will if alternative possibilities are not genuinely open to us (Harris, 2012). It may seem that moral practices of praise and blame or reward and punishment

make little sense once we are knowledgeable about even probabilistic causes of volition (Greene & Cohen, 2004).

The metaphysical distinction between incompatibilism and compatibilism is critical to this debate. The compatibilist claims that determinism and free will are compatible and may also claim that free will requires determination (vs. randomness) or responsiveness of some kind. According to such a compatibilist, free will does not exist where causation is absent, but exists where the right kind of causation is present. A volition or action is free so long as it is appropriately caused, for example, by the agent's own beliefs and desires and is not forced or coerced. The compatibilist marks the difference between reaching for a glass of water to quench one's thirst and being forced to do so, between moving one's limbs and having one's limbs move (or seem to move) of their own accord. The compatibilist may distinguish between the ability to act on one's desires and the ability to act on desires that align with one's deep or higher self (Frankfurt, 1971). She may further distinguish between the ability to act on desires or reasons that one endorses on reflection and the ability to act on desires or reasons that one ought to endorse (Wolf, 1987). Although compatibilist analyses of free will differ in detail, none requires thinking or acting against or outside of the causal order.

In a compatibilist approach, free will, autonomy, or agency may be absent or diminished in particular cases, as may occur in the case of addiction. Degrees of freedom also may be increased. A compatibilist conception of freedom is thus implicit in the design of neuroprosthetics such as brain-computer interfaces (BCIs) and deep-brain stimulators (DBS), insofar as they provide a neurotechnological foundation for freedom. These technologies also invite questions about the nature of the self. The relevant debate on the matter concerns whether the self is to be understood as a single, unified, substantial, or persisting entity; whether it is better understood as a shifting bundle of experiences, perceptions, and memories; or whether the very idea of a self has outlived its usefulness (Parfit, 1987). As with the outright denial of free will, the outright denial of a unified self does not countenance different degrees of selfhood, and it may therefore render nonsensical attempts to restore lost unity and recuperate the self, whether through technological, narrative, or other techniques (see Chapter 19).

3. FUTURE DIRECTIONS

The metaphysical issues introduced here are relevant to the sciences generally, and still other metaphysical issues are relevant to cognitive neuroscience. For example, metaphysical inquiry into the nature of space and time is relevant to work on spatial and temporal perception. Finally, this chapter has considered metaphysics insofar as it is relevant to cognitive neuroscience and has not considered cognitive neuroscience insofar as it is relevant to metaphysics. Among many exciting developments in recent work is a more dialogic and dialectical engagement between cognitive neuroscience and philosophy, as exemplified by this very volume.

Work on the nature of time, for example, has often involved exchanges between philosophers and physicists and may occasionally result in a historic debate, as with Bergson

and Einstein. Work on the nature of time is now more richly informed by cognitive science and its study of temporal experience. Paul (2016), for example, urges philosophers who study time to study the neuroscience of temporal perception and to take seriously the possibility that experiences of time are byproducts of cognitive processes. Metaphysics that does not take this research into account may be distorted by unduly favoring the view that time flows or passes.

Similarly, philosophic views about the nature of the self are increasingly informed by neuroscientific and neurotechnological work on agency and feelings of agency (see Chapter 19). A metaphysics of the self or person that does not take this research into account is unlikely to succeed in its aims. Just as philosophy illuminates possibilities that might otherwise go unnoticed, so too does cognitive neuroscience illuminate possibilities and complexities that philosophers might otherwise fail to consider.

In this same vein, a volume by Goldman and McLaughlin (2019) considers the bearing cognitive science might have on the practice and methods of metaphysics, more generally. In some cases, its contributors argue, cognitive science debunks metaphysics by showing that metaphysical beliefs were produced in a way that generally produces false beliefs. In other cases, cognitive science "unbunks" metaphysics by showing that metaphysical beliefs were produced in a way that generally produces true beliefs. In either case, as Paul also notes, the relationship between metaphysics and cognitive science is not one way. It is dialogical and dialectic.

Looking ahead: For some time, researchers have been refining algorithms to integrate data generated by various neuroimaging techniques (e.g., Laufs, 2012; De Martino et al., 2010). When subjects perform a cognitive task, techniques such as fMRI, electroencephalogram (EEG), magnetoencephalogram, and diffusion tensor imaging record different aspects of the accompanying neural activity. fMRI will pick up associated oxygen levels, for example, while EEG will pick up associated electrical activity (see Chapter 8). One result is that it is difficult to say whether all, some, or none of these recorded phenomena are the real correlates of cognition, a problem that the algorithms are meant to solve. It is likely that the focus on multimodal integration will renew interest in realism and instrumentalism (see Chapter 6). Realists will argue for the existence of a common neuronal source underlying the various recorded phenomena; instrumentalists will remain agnostic, proceeding as if the recorded phenomena have a common neuronal source only if doing so is useful for explanation and prediction.

Summary of Key Ideas

Metaphysics is key to making sense of the relationship between cognition and its correlates, to understanding what is common to mental states and processes with the same content, and to seeing how minds might figure into the causal order without thereby losing their characteristic properties. In these ways and others, metaphysics is relevant to cognitive neuroscience. Conversely, cognitive neuroscience is relevant to metaphysics.

Discussion Topics

- What is metaphysics? What is the relationship between metaphysics and the physical and natural sciences? What should it be? Why?
- What different metaphysical relationships might obtain between two correlated phenomena, for example, between a cognitive process and a correlated neural process? What are the differences between these relationships? What is the best way to determine which of these obtains in a given case? Why?
- If Ivan's behavior (waving) could be explained in wholly physical terms, does it follow that there is no room left for an explanation that appeals to his thoughts and desires? Can a single behavior have more than one correct explanation? Why or why not?

Further Reading

Allen, S. (2016). *A critical introduction to properties*. Bloomsbury Academic, an imprint of Bloomsbury Publishing Plc.

Antony, L. M. (2007). Everybody has got it: A defense of non-reductive materialism. In B. P. McLaughlin & J. D. Cohen (Eds.), *Contemporary debates in philosophy of mind*. Blackwell.

Baker, L. R. (2009). Non-reductive materialism. In B. McLaughlin, A. Beckermann, & S. Walter (Eds.), *The Oxford handbook of philosophy of mind*. Oxford University Press.

Goldman, A. I., & McLaughlin, B. P. (Eds.). (2019). *Metaphysics and cognitive science*. Oxford University Press.

Kane, R. (2005). *A contemporary introduction to free will*. Oxford University Press.

Kanwisher, N. (2017). The quest for the FFA and where it led. *The Journal of Neuroscience*, *37*(5), 1056–1061.

Kind, A. (2015). *Persons and personal identity*. Polity.

Mumford, S., & Anjum, R. L. (2013). *Causation: A very short introduction*. Oxford University Press.

Roskies, A. (2014). Can neuroscience resolve issues about free will? In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 4: Free will and moral responsibility* (pp. 103–126). Massachusetts Institute of Technology Press.

REFERENCES

Baker, L. R. (1997). Why constitution is not identity. *The Journal of Philosophy*, *94*(12), 599.

De Martino, F., Valente, G., de Borst, A. W., Esposito, F., Roebroeck, A., Goebel, R., & Formisano, E. (2010). Multimodal imaging: An evaluation of univariate and multivariate methods for simultaneous EEG/fMRI. *Magnetic Resonance Imaging*, *28*(8), 1104–1112.

Frankfurt, H. (1971). Free will and the concept of a person. *The Journal of Philosophy*, *68*(1), 5–20.

Geyer, S., Weiss, M., Reimann, K., Lohmann, G., & Turner, R. (2011). Microstructural parcellation of the human cerebral cortex—from Brodmann's post-mortem map to *in vivo* mapping with high-field magnetic resonance imaging. *Frontiers in Human Neuroscience*, *5*.

Greene, J., & Cohen, J. (2004). For the law, neuroscience changes nothing and everything. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *356*(1451), 1775–1785.

Harris, S. (2012). *Free will*. Free Press, USA.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, *17*(11), 4302–4311.

Kim, J. (1998). *Mind in a physical world*. Massachusetts Institute of Technology Press.

Kim, J. (2005). *Physicalism, or something near enough*. Princeton University Press.

Laufs, H. (2012). A personalized history of EEG—fMRI integration. *NeuroImage*, *62*(2), 1056–1067.

Libet, B., Gleason, C., Wright, E., & Pearl, D. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): The unconscious initiation of a freely voluntary act. *Brain*, *106*, 623–642.

Lowe, E. J. (1993). The causal autonomy of the mental. *Mind*, *102*(408), 629–644.

Parfit, D. (1987). Divided minds and the nature of persons. In C. Blakemore & S. Greenfield (Eds.), *Mindwaves*. Blackwell.

Paul, L. A. (2016). Experience, metaphysics, and cognitive science. In J. Systma & W. Buckwalter (Eds.), *A companion to experimental philosophy*. John Wiley & Sons, Inc.

Putnam, H. (1975). The meaning of "meaning". *Minnesota Studies in the Philosophy of Science*, *7*, 131–193.

Quine, W. V. (1951). Main trends in recent philosophy: Two dogmas of empiricism. *The Philosophical Review*, *60*(1), 20.

Roskies, A. (2006). Neuroscientific challenges to free will and responsibility. *Trends in Cognitive Sciences*, *10*(9), 419–423.

Roskies, A. (2014). Can neuroscience resolve issues about free will? In W. Sinnott-Armstrong (Ed.), *Moral psychology, Vol. 4: Free will and moral responsibility* (pp. 103–126). Massachusetts Institute of Technology Press.

Roskies, A., & Nahmias, E. (2016). "Local determination," even if we could find it, does not challenge free will: Commentary on Marcelo Fischborn. *Philosophical Psychology*, *30*(1–2), 185–197.

Shurger, A., Sitt, J., & Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. *Proceedings of the National Academy of Sciences*, *109*(42), E2904–E2913.

Slotnick, S. (2013). *Controversies in cognitive neuroscience*. Palgrave Macmillan.

Viola, M., & Zanin, E. (2017). The standard ontological framework of cognitive neuroscience: Some lessons from Broca's area. *Philosophical Psychology*, *30*(7), 945–969.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe & R. Rhees, Eds., G. E. M. Anscombe, Trans.). Blackwell, UK.

Wolf, S. (1987). Sanity and the metaphysics of responsibility. In F. D. Schoeman (Ed.), *Responsibility, character, and the emotions: New essays in moral psychology* (pp. 46–62). Cambridge University Press, USA.