What is the Philosophy of Consciousness?

Uriah Kriegel

Introduction to U. Kriegel (ed.), *Oxford Handbook of the Philosophy of Consciousness*

Consciousness is one of the most elusive phenomena of the natural world. But it is, after all, *part* of the natural world. It has presumably evolved at some point, as a result of certain natural processes taking place within the causally integrated spatiotemporal system we call Nature. What need is there, then, for a *philosophy* of consciousness? As a *natural* phenomenon, should it not submit to theoretical explanation by the *natural* sciences? There is no philosophy of owls; owls are natural phenomena, so the theory of owls is part of a natural science – zoology. We now also have a lively science *of consciousness*, conducted by cognitive psychologists and neuroscientists among others. What do philosophers have to contribute here?

One might answer: our empirical knowledge about consciousness is so fragmentary and incomplete, at this early stage of scientific inquiry, that philosophers may be called upon to offer more or less disciplined speculations about the part of the story about which we simply do not as yet have sufficient scientific knowledge. It would follow that, as scientific knowledge of consciousness grows, the need for a philosophy of consciousness will recede – until it will disappear and be entirely replaced by the science of consciousness.

The assumption behind the present volume is that philosophy has a more significant role to play in shaping our understanding of consciousness; that even a complete science of consciousness will involve certain lacunae calling for philosophical supplementation. The plausibility of this notion depends in part on what one calls "a science." To bracket verbal issues, we will concern ourselves here with three areas in which the science of consciousness *as pursued today* leaves certain questions unaddressed. The next three sections introduce and discuss these three areas; each serves as a motivating introduction to one of the volume's three parts.

1 Phenomenal Grounding: The Varieties of Conscious Experience

Science tries to explain. But to get busy explaining, we first need to be clear on what are the phenomena which need explaining. That is, we need to be clear on our *explananda* before launching our explanatory enterprise. A science of consciousness, in particular, needs a clear list of the kinds or types of conscious experience that may occur in the natural world, the scientific explanation of which would be welcome.

Early psychologists in the 1870s, alive to this dual task, distinguished between *descriptive psychology* and *explanatory psychology*. Descriptive psychology had the task of describing psychological phenomena – manifest mental occurrences – as well as providing an initial taxonomy of them; in this task *introspection* had a central role. Explanatory psychology, in contrast, had the task of providing causal explanations for the occurrence of the relevant phenomena, to record regularities in their occurrence, and to articulate the natural laws governing their (regular) occurrence; here introspection was comparatively subsidiary, with "physiology" (in essence: neuroscience) playing the more central role. Many psychological phenomena (e.g., visual experiences) are introspectible, but they occur as the result of subpersonal neurophysiological processes that are not themselves introspectible.

In this division of labor, descriptive psychology had the job of delineating the domain of phenomena to which explanatory psychology is answerable. However, the introspective method on which it relied soon fell into disrepute. The two main issues were (i) the persistent irresolvability of introspective disagreements and (ii) the apparent theory-ladenness of introspective judgments. Behaviorists were incensed about "psychology exceptionalism": the fact that psychology just did not look like a regular, bonafide natural science, in which objective standards of testing and replication could be discerned and certain methods were available for resolving disagreements (Watson 1913: 163). This story is familiar. What is little understood is the way the behaviorist critique effectively and decisively banished the research area of *descriptive* psychology from the science of the mind. Explanatory psychology, with its neurophysiological methods and goals, was little affected by the behaviorist turn. What needed radical reframing, for the behaviorist, was not our conception of *how to explain* in psychology, but our conception of *what needs explaining* in psychology. The new explananda were not subjectively experienced internal occurrences, but pieces of publicly observable behavior.

Crucially, when cognitivists mounted their critique of behaviorism, starting in the late 1950s, their critique was in truth very partial: they insisted that psychological *explanation* could advert to internal, genuinely psychological posits

("internal" as opposed to behavioral, "genuinely psychological" as opposed to merely neurological). But they left intact the behaviorist's conception of the *explananda* of psychology. Still today, what psychologists (or "cognitive scientists") try to explain are just bits of *behavior* (including verbal reports). Internal mental processes *are* invoked in post-behaviorist psychology, but they are treated as *explanatory posits* – theoretical entities posited in the context of trying to make sense of behavioral phenomena, not phenomena with which we have pre-scientific acquaintance and for which we would like to provide scientific explanation.

In this respect, consciousness science differs from other scientific disciplines. Chemistry tries to explain chemical phenomena, zoology tries to explain zoological phenomena, and so on. But psychology does not try to explain psychological phenomena; it tries to explain *behavioral* phenomena. Furthermore, while the zoologist works with certain theoretical posits, such as DNA, with which we have no pre-scientific familiarity, she also recognizes manifest zoological phenomena, such as owls and wings, with which we do have pre-scientific familiarity; these she treats as observable explananda rather than as unobservable, purely theoretical explanatory posits. In contrast, the psychologist treats subjective experiences as theoretical posits, as though we have no pre-scientific familiarity with them. But in truth, it is evident that we very much *do* have pre-scientific and indeed pre-*theoretic* familiarity with experiential phenomena, such as feeling embarrassed at a party, smelling a camembert, suddenly understanding what the speaker meant in response to our question, and so on. If DNA did not help explain anything, we would never suspect its existence. But if the experience of sadness did not help explain anything, we would all still believe that there is such a thing as feeling sad.

In taking behavioral rather than psychological phenomena as its explananda, and in treating *all* psychological occurrences as mere theoretical posits, mainstream academic psychology involves a form of exceptionalism after all. No other scientific discipline banishes its own proprietary phenomena from its domain of explananda, substituting to them alternative explananda; and none proceeds as though its proprietary phenomena are purely theoretical posits even when we clearly have pre-scientific familiarity with them. Perhaps this exceptionalism is ultimately justified, say on methodological grounds. Still, none of this makes the experiential phenomena with which we have pre-scientific familiarity *go away*. Descriptive psychology may have been banished from scientific psychology, but the experiential phenomena themselves – the *subject matter* of descriptive psychology – cannot be banished by fiat.

This opens up a first area where the philosopher has a contribution to make to our understanding of consciousness. Wearing the moth-eaten cap of the

descriptive psychologist, the philosopher may hope to produce a respectable inventory of experiential phenomena, as well as to impose initial order in it (through taxonomy and topology).

Two central questions define the core of this endeavor. First: of all the possible forms of conscious experience, which ones have psychological reality in our conscious life? (Depending on one's ambition, the "our" can be interpreted in a variety of ways, from ranging only over normal human adults to ranging over every sentient being.) Second: of all the psychologically real phenomena of conscious life, which can be analyzed into combinations of other, more fundamental experiential phenomena, and which must be treated as fundamental, unanalyzable, primitive types of conscious experience? The types of experience which are not only psychologically real but also fundamental and unanalyzable – the "experiential primitives" in our stream of consciousness, if you will – are the holy grail of descriptive psychology. Collectively, they constitute the complete fund of experiential ingredients the combinations and recombinations of which yield every subject's total conscious experience at a time.

It is the fashion of the day to frame philosophical issues in terms of *grounding*, the metaphysical relation canonically picked out by the locution "in virtue of" in its philosophical use. Not to be deficient in this particular, we may frame the second issue just presented in terms of a distinction between (a) types of experience which are grounded in (combinations of) other types of experience, and (b) types of experience which are grounded in no others. The latter are the *ungrounded grounds* of the experiential realm, the "building blocks" of our stream of consciousness. Early introspectionist psychologists referred to these as the *elements of consciousness*, and had indulged in bold speculations about their number. (Edward Titchener, probably the most prominent of all introspectionist psychologists, proffered that there were around 42,415 "elements of consciousness"!) But although such questions are incredibly hard to settle, or even address in an epistemically responsible way, they answer to certain *facts of the matter* that empirical cognitive science does not, perhaps judiciously, venture to contend with.

Part I of this volume is dedicated to philosophical discussions of some of the most central candidate ungrounded grounds of the sphere of experiential phenomena. It opens with two chapters on perceptual experience: a chapter by Pär Sundström on visual experience and a chapter by Casey O'callaghan on the varieties of nonvisual perceptual experience. There follow three chapters on three other types of phenomenologically lively kinds of experience: a chapter by Frédérique de Vignemont on bodily experience, a chapter by Julien Deonna and Fabrice Teroni on

emotional experience, and a chapter by Amy Kind on imaginative experience. Few philosophers would deny the very *existence* of these types of experience, but plenty of philosophers have maintained that these types of experience can be accounted for entirely in terms of the coming-together of other, more basic types of experience. We close Part I with four chapters on more controversial types of experience, ones the very existence of which has often been called into question: a chapter by Tim Bayne on the experience of thinking, a chapter by Myrto Mylopoulos and Josh Shepherd on the experience of agency, a chapter by Philippe Chuard on the experience of time, and a chapter by Farid Masrour on the experiential unity of consciousness. For each of these putative experiential phenomena, we can ask whether it is (a) an ungrounded primitive of conscious life (an "experiential primitive"), (b) a psychologically real phenomenon but one grounded in combinations of other, more fundamental phenomena (an "experiential derivative"), or (c) not a psychologically real phenomenon at all (a mere "experiential putative," if you will).

Consider the experience of conscious thinking. One possible view is that there is simply no such thing – thought processes are never *experienced*, though they may be systematically *accompanied* by certain experiences, say of auditory imagery. Another view is that we do experience some of our thoughts, but the relevant experience consists in some cocktail of perceptual, imagistic, and emotional feelings coming together just so. A third view is that the experience of conscious thought sometimes brings into the picture a new, sui generis kind of irreducibly cognitive feeling – a type of proprietary cognitive phenomenology. According to this third view, there is at least one aspect to the experience of conscious thought that constitutes an experiential primitive, an experiential ingredient that goes beyond any combination of ingredients found in other kinds of experience.

A similar dispute between three such positions – one broadly eliminativist, one broadly reductivist, and one broadly primitivist – can be framed for any of the putative types of conscious experience discussed in Part I of this volume. The collection of all phenomena for which the primitivist view is the correct one would give us the fundamental furniture of the experiential realm – the "elements of consciousness." Clearly, settling on the right collection or inventory of such elements is a foundational task of first importance in our understanding of consciousness. And it is the philosophy of consciousness that attends to this task. (To repeat, one could use the term "science" in such a way that descriptive psychology is part of the science of consciousness. My point is just that very few academics in science departments engage in descriptive psychology; it is typically academics in philosophy departments who do.)

2 Physical Grounding: Theories of Consciousness

When we speak of scientific research into consciousness, we speak in truth of research which targets in the first instance salient *correlates* of consciousness. Neuroscientific research is transparently framed in those terms, since its official aim is to identify the "neural correlates of consciousness." But as we will see in the next section, research into consciousness *in cognitive psychology* can similarly be framed as targeting, in the first instance, the *cognitive correlates* of consciousness. It is possible, of course, to hold that consciousness is in fact *nothing over and above* its neural correlates and/or *nothing over and above* its cognitive correlates. A certain kind of physicalist would hold the former and a certain kind of functionalist would hold the latter. But these additional claims – these nothing-over-and-above claims – are not themselves scientific claims, and the arguments bearing on their plausibility are not on their face empirical arguments based on laboratory results. On the contrary, nothing-over-and-above claims are paradigmatically philosophical claims and the considerations typically brought in their support are adduced rather from the armchair.

In other areas, what scientists provide us with makes it *straightforward* whether a nothing-over-and-above claim is appropriate. Once told that H_2O molecules are present wherever water is present, and that sufficiently large coalitions of them exhibit all the features and behaviors water does, it is straightforward that H_2O constitutes not only the "molecular *correlate* of water" but water itself. Nonetheless, it is still a non-empirical, *philosophical* claim that water is nothing but H_2O – albeit one that courts no controversy. In the case of consciousness, in contrast, the corresponding claim that consciousness is nothing but its neural (and/or cognitive) correlate does engender a kind of intellectual discomfort.

How best to frame this discomfort is itself a topic of lively debate among philosophers; the opening chapter of this volume, by David Papineau, offers one potential diagnosis. According to Papineau, there is simply a specially resistant pre-doxastic "intuition of distinctness" that makes it hard for us to take seriously the proposition that consciousness might be nothing but its neural correlate. This is intended as an alternative to the most popular philosophical diagnosis of the challenge presented by consciousness, namely, the notion that even absolutely complete knowledge of a person's brain, or for that matter of the totality of physical reality, would be insufficient to derive any knowledge of what specific kind of conscious experience the relevant person is undergoing, and even of whether she is

undergoing any experience at all (more on this in §3). My own view may be put in terms of an "intuition of *categorical* distinctness": the reason we cannot see how the subjective quality of conscious experience could be generated, let alone *constituted*, by the silent transactions of so many brain cells lying motionless inside the darkness of the skull, is that the two appear to belong to two different ontological categories. All these brain cells ever do, after all, is increase and decrease the rates at which they transmit an electrical impulse to each other; the notion that the subjective quality of feeling sad, or smelling coffee, or getting angrier, *just is* such a change in cells' firing rates seems like a *category mistake*, something akin to saying that the number 5 is identical to that mango over there, or that justice is nothing but cement mixed with wood fiber.

However we ultimately choose to frame the relevant intellectual discomfort, its result is that there *is* a controversy over the non-empirical, philosophical nothing-over-and-above claim(s) associated with consciousness. Addressing and resolving this controversy is a second area where philosophical research is indispensable for a complete understanding of the nature of consciousness.

Accordingly, Part II of this volume is dedicated to discussions of some of the most prominent philosophical positions on the ultimate relationship between consciousness and its physical (notably neural) correlates. It opens with two chapters about the neural correlates of consciousness: one chapter by Jorge Morales and Hakwan Lau on the state of the art in this area of scientific inquiry, and another chapter by me about the different philosophical "explanations" of the very existence of a correlation between neural activity of consciousness. There follow six chapters about the most prominent such explanations, three physicalist and three anti-physicalist: a chapter by Brie Gertler on dualism, the view (roughly) that the experiential and the physical are mutually (metaphysically) independent, such that any links between consciousness and its neural correlate are at most causal and contingent, not constitutive and necessary; a chapter by Sam Coleman and Philip Goff on so-called neutral monism, the view (roughly) that consciousness and matter are ultimately two facets of a single underlying reality, but a reality somehow neither physical nor mental, or perhaps both physical and mental; a chapter by Michael Pelczar on idealism, according to which (roughly) consciousness is ontologically prior to matter (or else matter does not exist at all); a chapter by Liz Irvine and Mark Sprevak about eliminative physicalism, which denies the very existence of consciousness; a chapter by Frank Jackson on *a priori* physicalism, the idea that consciousness is not only grounded in its physical correlate, but is so in an epistemically transparent way, such that complete knowledge of a the physical facts regarding some person would make it possible to *derive* knowledge of the conscious

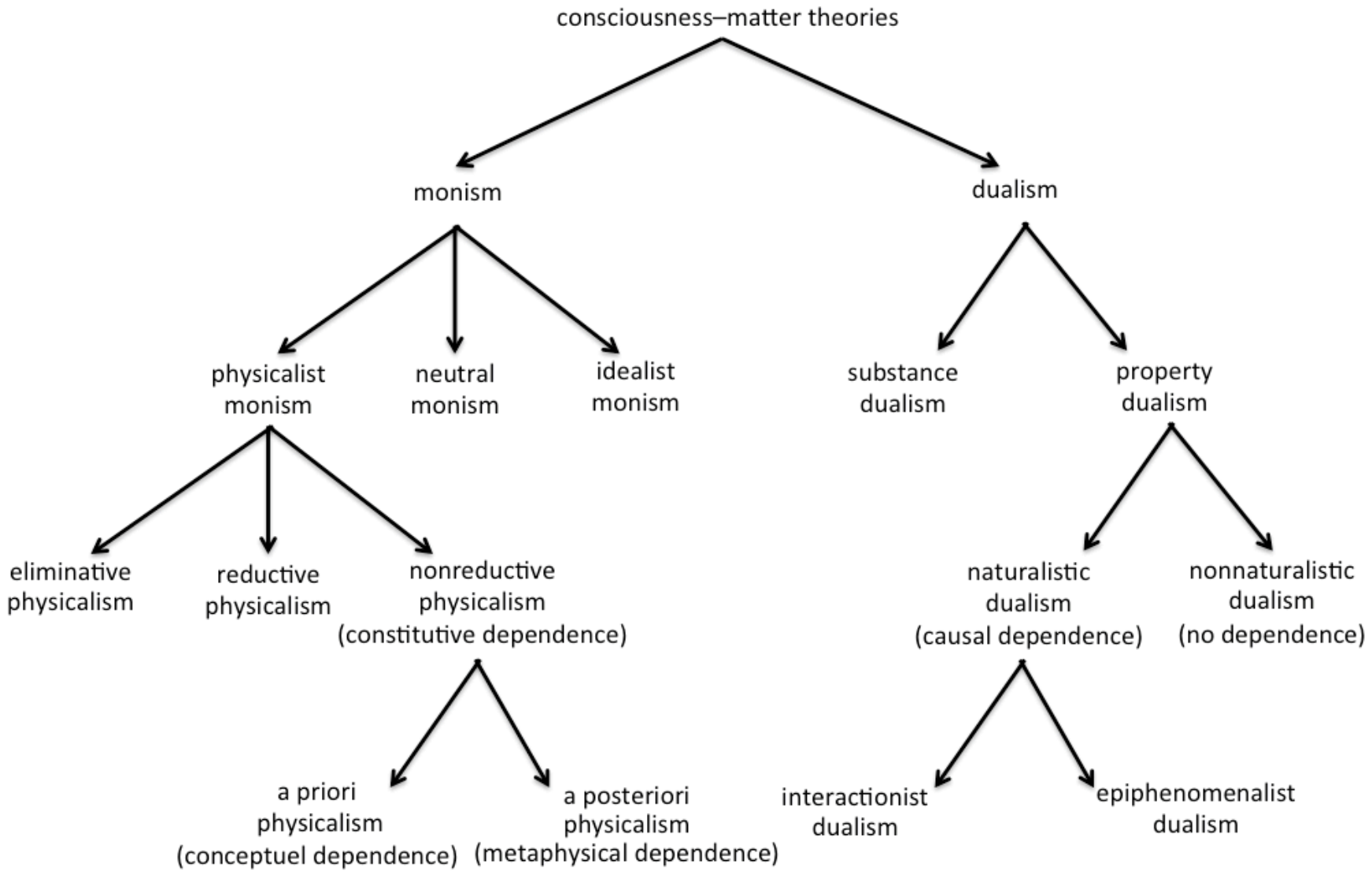
experience she is undergoing; and finally a chapter by Joe Levine on *a posteriori* physicalism, which holds that although experiential facts are grounded in physical facts, this link is epistemically opaque to us, in that complete knowledge of physical facts does not suffice to derive facts about conscious experience. (Note well: not all of these authors *endorse* the position on which they write!) The next three chapters of Part II explore three specific strategies often adopted by physicalists in trying to account for consciousness (but sometimes embraced without a physicalist agenda): a chapter by Adam Pautz about representationalist theories, according to which conscious mental activity is characterized essentially by its distinctive way of representing the world; a chapter by Josh Weisberg on higher-order representationalist theories, according to which conscious mental activity is characterized rather by the way it is targeted by representations *of it*; and a chapter by Tom McClelland on self-representationalist theories, according to which the crucial characteristic of conscious states is that, whatever else they represent, they always also represent *themselves*. The final chapter of Part II, by Daniel Stoljar, is dedicated to the so-called epistemic approach to the problem of consciousness, the view (roughly) that we simply lack knowledge of certain key empirical facts that would shed light on whether and how consciousness might "arise" from physical processes; in some versions, our ignorance of the relevant facts is principled and insurmountable, in others it is potentially merely temporary.

This selection of philosophical theories on the connection between consciousness and its physical correlate is perforce partial. There are many more options and sub-options for understanding the relationship between consciousness and matter. A more systematic and more nearly complete map of the available options is presented in Figure 1. It divides philosophical takes on the consciousness-matter relationship into monist and dualist. The dualist approaches it divides into substance dualism and property dualism, and property dualism it divides into naturalistic and nonnaturalistic versions: the former insists on a causal or nomic dependence of consciousness on physical facts, while the latter insists rather on the *absence* of any dependence of the sort. Naturalistic dualism is then divided into interactionist and epiphenomenal: the former grants, while the latter denies, causal efficacy to conscious experiences. In this organization of logical space, the four main dualist theories are substance dualism, nonnaturalistic property dualism, interactionist naturalistic property dualism, and epiphenomenalist naturalistic property dualism.

Monist theories are divided in Figure 1 into three groups: physicalist monism, neutral monism, and idealist monism. All three hold that reality is "at bottom" unified (i.e., unified at the fundamental level, the level of "ontological

bedrock"). But, roughly speaking, one view is that reality is at bottom physical, another is that it is at bottom mental, and a third that it is at bottom something else, something which is neither physical nor mental but is both proto-physical and proto-mental. Ignoring important subdivisions of idealist and neutral-monist views, Figure 1 goes on to divide physicalist views into three further groups: eliminative physicalism, reductive physicalism, and nonredudctive physicalism. The first claims that there are no experiential properties or types; the second that experiential properties or types are identical to physical ones; the third that experiential properties/types are not identical to physical properties/types, but the former are nonetheless *constitutively dependent* (and not merely causally or nomically dependent) upon the latter. This constitutive dependence was traditionally framed in terms of supervenience, a relationship whereby one family of properties (the supervenient ones) cannot vary without corresponding variation in another family of properties (the subvenient ones, the "supervenience base"). Traditionally, a further important distinction between a priori and a posteriori physicalism is then framed as a distinction between two different potential modal forces with which the experiential may supervene on the physical: a priori physicalism holds that the experiential supervenes on the physical with *conceptual* necessity, a posteriori physicalism that it supervenes with merely *metaphysical* necessity. (Incidentally, naturalistic dualism has sometimes been characterized as holding that the experiential supervenes on the physical with merely *nomic* necessity on the physical, and experiential nonnaturalistic dualism can be seen as rejecting *any* supervenience of the experiential on physical.)

It is worth noting that in recent philosophy supervenience relations have come to be seen as mere symptoms of the aforementioned more robust underlying metaphysical relation of *grounding*. Accordingly there emerged in recent philosophy of mind an interest in "grounding physicalism" as a potential position, either on the mind-body problem or on the associated "consciousness-matter problem" on which we have focused here. It is not always clear in the relevant literature how grounding physicalism is related to the more traditional versions of physicalism, but my own sense is that it is best viewed as a potentially better and more accurate reformulation of nonreductive physicalism. For grounding relations are not typically eliminative about the grounded, so grounding physicalism is not a version of eliminative physicalism. Meanwhile, reductive physicalism is naturally seen as requiring the identification of consciousness with something physical, whereas an important feature of grounding is that it differs from identity in being asymmetric.


consciousness–matter theories
monism
dualism
physicalist monism
neutral monism
idealist monism
substance dualism
property dualism
eliminative physicalism
reductive physicalism
nonreductive physicalism (constitutive dependence)
naturalistic dualism (causal dependence)
nonnaturalistic dualism (no dependence)
a priori physicalism (conceptuel dependence)
a posteriori physicalism (metaphysical dependence)
interactionist dualism
epiphenomenalist dualism

(This difference is moreover crucial to grounding physicalism's ability to capture the ontological priority of the physical over the mental without subsuming or absorbing the mental in the physical.) So I think grounding is best seen as that which metaphysical or conceptual supervenience is the symptom of, and accordingly grounding physicalism is best seen as what nonreductive physicalism was always supposed to be. Now, if we do frame nonreductive physicalism in terms of grounding rather than supervenience, then the distinction between a priori and a posteriori physicalism becomes a distinction between two kinds of grounding, or two different ways in which the experiential can be grounded in the physical. The idea is that there are distinctions between grounding relations that mirror the distinction between conceptual and metaphysical supervenience. Consider: the fact that Jimmy is a bachelor is grounded in one way in the fact that Jimmy is an unmarried man, but grounded in another way in the fact that the person who originated in the relevant sperm and the relevant egg is a bachelor. (I am indulging here, merely for the sake of illustration, the essentiality-of-origin thesis.) The first kind of grounding has conceptual (hence a priori) supervenience for a symptom, the second has metaphysical (hence a posteriori) supervenience for a symptom.

Figure 1 thus distinguishes, alongside the aforementioned four types of dualism, four central forms of physicalism: eliminative physicalism, reductive physicalism, a priori nonreductive physicalism, and a posteriori nonreductive physicalism. In addition it recognizes a fifth and sixth types of monism: idealism and neutral monism. Then there is the epistemic approach, which allows that we may simply not know, and may even be constitutionally incapable of knowing, which of these positions is correct. With this plethora of options, certain decisions had to be made to keep Part II of this volume manageable – decisions which were guided not only by intellectual but also by practical and sociological considerations.

3 Psychological Grounding: Consciousness and Neighboring Phenomena

At least historically, the problem of whether there is more to consciousness than its physical correlate has been the most central in the philosophy of consciousness. Indeed, this problem is often referred to in the philosophical literature simply as "the problem of consciousness." As noted, there are different ways of understanding the source of the problem. But a recurring theme in many presentations is that although the problem concerns in the first instance the relationship between consciousness and its *physical realizer*, it ultimately passes through worries concerning the relationship between consciousness and its *functional or psychological role* (see Levine 1983: 357, Chalmers 1995: 64).

It is sometimes said that there is nothing more to physical matter than *structure and dynamics*, whereas there *is* something more to consciousness than structure and dynamics, namely, *intrinsic feel*. The point can be put more informally and impressionistically as follows: there is no gap between what a physical phenomenon *does* and what it *is*, whereas there is more to what consciousness *is* than what it *does* (there is also how it *feels*). A physical phenomenon, system, mechanism, process, or property *is* what it does – its essence or nature is to do just that which it does. But although a conscious experience does various things – it performs many psychological functions, i.e. plays a certain psychological role – its nature is not exhaustively captured by a complete enumeration of all it does. There appears to be a "para-functional residue" in consciousness; the term "phenomenal consciousness" can be seen as designating precisely the aspect of consciousness that appears to go beyond the totality of its psychological functions (Block 1995).

To bring the point out, let us compare consciousness and memory. Both of them *do* something: they play a causal or functional role in the subject's overall psychological economy. Crushing many subtleties, let us say that what memory does is store information. What is important to note is that storing information is so essential to memory that it is natural to say that memory just is an *information-storer*. Now, the complete psychological profile of memory is surely more nuanced than this; but whatever it is, once we know exactly what memory does, we can identify memory with that-which-does-*that*. Any system or mechanism that did the same thing – that played the relevant psychological role – would ipso facto count as a memory system. This is why we have no problem attributing memory to such inanimate objects as desktop computers from the seventies. (To say this is not take a stand on whether computers could ever be conscious; it is only to take a stand on whether desktop computers from the seventies were conscious – the stand that they were not!)

Consciousness has its own functional or psychological profile, which seems considerably more multifaceted than memory's. Some attempt to unify that profile around the notion of *cognitive access*: consciousness makes information more available to a wide variety of downstream mechanisms, modules, and subsystems. Again, a full specification of the functional or psychological role of consciousness would surely be more nuanced than this, but whatever it turns out to be, it is what I referred to in §2 as the "cognitive correlate of consciousness": it is the specific, singular contribution consciousness makes to cognition. What is important for our purposes is that it seems perfectly possible, indeed rather *easy*, to imagine a system or mechanism that does exactly *that* (whatever consciousness does within our cognitive system) and yet does not amount to consciousness at all. If one day we

isolate the complete psychological role of consciousness, construct a system made of silicon chips that perform the exact same cognitive functions, and embed that system in a computer, we will not immediately take this to *settle the question* of whether that computer is conscious. There seems to be an *open question* left over: Does the computer's performance of these functions involve a subjective quality, a phenomenal character, an experiential feel? The very fact that this question makes sense, even where the psychological role of consciousness has been fully replicated, shows that there is more to our concept of consciousness than the idea of *that-which-does-such-and-such*. In other words, our concept of consciousness allows for a potential daylight between what consciousness is and what consciousness does – a daylight that our concept of memory, for instance, does not tolerate.

A philosopher could of course maintain that although this is true of our naïve, prescientific concept of consciousness, a more sophisticated scientific understanding of consciousness will reduce consciousness to its cognitive correlate, that is, will construe consciousness as exhausted by the sum total of its psychological functions. To this, another philosopher will respond that such a reconceptualization of consciousness is so radical, leaving out the very essence of consciousness (namely, its subjective feel), that it amounts to effectively denying the existence of consciousness as we ordinarily think of it. The first philosopher may then come back with the complaint that unless we accept her radical reconceptualization of consciousness, we will end up with a mysterian, nonnaturalist picture of the world overly welcoming of scientifically intractable, vaguely supernatural phenomena. Here we approach the great looming fear of the philosophy of consciousness: that all said and done, the only *stable* positions on the nature of consciousness are eliminative physicalism and nonnaturalistic dualism – and that the tremendous industry of fashioning clever intermediate positions is but an expression of a futile hope to have the cake and eat it too.

The point I want to make is that this is a paradigmatically philosophical debate, not a scientific one. At the end of inquiry, if you please, the cognitive psychology of consciousness will identify the complete cognitive correlate of consciousness. But the question of the ultimate relationship between consciousness and its cognitive correlate cannot be settled by cognitive psychology itself – just as the question of the ultimate relationship between consciousness and its neural correlate cannot be settled by neuroscience (and indeed, just as the question of the ultimate relationship between voting patterns and the behavior sub-atomic particles cannot be settled by particle physics). The claim that consciousness is *nothing but* its cognitive correlate, and the claim that it is *not* nothing but its cognitive correlate, are

philosophical claims. Here too, then, we find a distinctive role for philosophy to play in our overall understanding of the nature of consciousness.

Part III of this volume is dedicated in large part to studies of the relationship between consciousness and some of the psychological phenomena or mechanisms with which it is functionally bound. It includes chapters on consciousness' relationship with attention (Chris Mole), memory (Chris Hill), action guidance (Benjamin Kozuch), intentionality (Angela Mendelovici and David Bourget), perceptual knowledge (Berit Brogaard and Eli Chudnoff), the self (Dan Zahavi), introspection and self-knowledge (Maja Spener), and even moral knowledge (Josh Shepherd and Neil Levy); the volume closes with a chapter (by Mark Rowlands) on consciousness' embodiment.

There are surely other psychological functions of relevance to consciousness' functional role. But this partial list captures some of the most salient ones: conscious awareness tends to display a foreground/background attentional structure; consciously perceived stimuli tend to stay in short-term memory longer and more vividly than stimuli perceived subliminally or in blindsight; conscious experiences seem to guide certain of our on-the-fly actions; they tend to be intentionally directed at objects and features in the external world; they tend to induce and justify perceptual beliefs about the external world, as well as introspective beliefs about oneself and one's "inner world." So much is beyond serious doubt. The question is whether these intimate relations consciousness bears to attention, memory, action, intentionality, perceptual judgment, introspection, and the self are *causal and contingent* or on the contrary *constitutive and necessary*. Causal and contingent relations are to be studied by cognitive science; the philosopher's distinctive contribution is to examine the prospects for constitutive, essential connections between consciousness and these neighboring psychological phenomena.

Again, we may also put the question in terms of grounding: Is consciousness grounded in some of its neighboring psychological phenomena, such that once a cognitive system performs the relevant psychological functions, it is *guaranteed* to also exhibit (phenomenal) consciousness? Or does consciousness ultimately outstrip any collection of psychological functions and is thus a sui generis mental phenomenon? Grounding being a metaphysical relation, such questions of grounding pertain to the metaphysics of consciousness, not to the empirical science of consciousness. There is no laboratory instrument that can detect grounding relations in nature. The question is not *empirically* tractable.

Conclusion

Given a certain conscious experience – say, feeling grief-stricken about a relative's death, or nervous about an upcoming appointment – there are at least three kinds of question about relations of grounding or constitutive dependence that the experience bears to (increasingly removed) other phenomena: (1) Is the experience grounded in/constitutively dependent upon other types of conscious experience? (2) Is it grounded in/constitutively dependent upon psychological phenomena other than conscious experience? (3) Is it grounded in/constitutively dependent upon certain physical (or otherwise non-mental) processes or states? These questions define a sphere of inquiry where philosophical reflection on consciousness seems indispensable. This volume attempts to cover some of the most prominent issues that arise concerning each of these distinctively philosophical questions.

References

- Block, N.J. (1995), 'On a Confusion About the Function of Consciousness', in *Behavioral and Brain Sciences* 18: 227-247.
- Chalmers, D.J. (1995), 'The Puzzle of Conscious Experience, in *Scientific American* (December): 62-68.
- Levine, J. (1983), 'Materialism and Qualia: The Explanatory Gap', in *Pacific Philosophical Quarterly* 64: 354-361.
- Watson, J.B. (1913), 'Psychology as the Behaviorist Views It', in *Psychological Review* 20: 158-177.