

Section III

Flanagan and Naturalized Buddhism



5 Consciousness, Naturalism, and Human Flourishing

Christian Coseru

Introduction

What does Buddhist metaphysics, epistemology, and ethics have to contribute that would be of interest to analytic philosophers? In his engaging and intellectually daring foray into cross-cultural philosophy, *The Bodhisattva's Brain: Buddhism Naturalized*, Owen Flanagan tackles this question head on. The answer – for philosophers who would care to listen – is that Buddhism offers a metaphysics anchored in such robust principles as impermanence, no-self, and the ubiquity of causation, an epistemology that is thoroughly empiricist, and a eudemonistic ethics that prizes compassion. Most importantly, argues Flanagan, is the claim “that there are logical connections between these three” (Flanagan 2011, 206), and the promise of reliable frameworks for exploring them. A philosopher working at the intersection of multiple spaces of meaning would find that these logical connections open up new possibilities for enhancing, refining, and expanding the range of philosophical arguments and possibilities, the ultimate and obvious aim of which is to make progress in solving enduring philosophical problems.

At first blush it may seem as though Flanagan's aim in *The Bodhisattva's Brain* is a modest one: Unpack for a broader, mostly philosophical, audience unacquainted with Buddhism what scholars have known all along – that Buddhism is host to a complex array of theories and practices of unique scope and enduring relevance that could be put in the service of addressing many of our most pressing existential and metaphysical concerns. Closer scrutiny, however, reveals it to be, in Flanagan's own words, “a work of advocacy for something that doesn't yet have any traction, at most a tenuous foothold” (Flanagan 2011, 4), but that he thinks ought to exist: namely, Buddhism *naturalized*. Given that most philosophers today take scientific naturalism to provide a robust basis for advancing empirical claims to knowledge – and, according to some (Kitcher 1992; Stroud 1996), the only such viable basis – if Buddhism can be shown to support such claims, then it stands a good chance of making a viable contribution to ongoing debates about causation, agency, and

human flourishing. As someone who has argued at length for the need to make Buddhist epistemology receptive to the findings of cognitive sciences (Coseru 2009, 2012, 2018), I am quite sympathetic to Flanagan's approach. But neither Buddhism nor the sciences of the mind speak with one voice, notwithstanding popular representations of Buddhism as a sort of mind science on the brink of revolutionizing Western conceptions of consciousness and cognition (Wallace 2003). The situation is further complicated by the fact that, as Flanagan himself has admitted elsewhere (Flanagan 2006), naturalism lacks a common core.

In what follows, I want to pursue the question of precisely what conception of naturalism, if any, is best suited to capture the scope of Buddhist Reductionism, and whether this conception can still accommodate the distinctive features of phenomenal consciousness (e.g., subjectivity, intentionality, first-person givenness, etc.). In the first section, I review dominant conceptions of naturalism, and their applicability to the Buddhist project. In the second section, I provide an example of problematic issues more stringent conceptions of naturalism under the guise of neurophysicalism confront, and evaluate Flanagan's response to these issues. In the third section, I consider briefly the reflexivity thesis (the thesis that consciousness consists in conscious mental states being implicitly self-aware), specifically as articulated by Dignaga, Dharmakirti, and their followers, and use this thesis to articulate a conception of minimal agency as mineness that, I argue, further challenges Flanagan's neurophysicalism stance and his compatibilist account of moral agency. I conclude, in the fourth section, by suggesting a way in which no-ownership conceptions of reflexive self-consciousness can help us both to get the structure of phenomenal consciousness right and to ground our conceptions of agency, intentionality, and moral responsibility.

Naturalism and Buddhist Reductionism

A term with multiple and imprecise meanings, 'naturalism' denotes a specific philosophical attitude and methodological approach that gained momentum at the beginning of the last century, when calls for philosophy to discard the supernatural and ally itself with science were first heard. As such, naturalism reflects a growing conviction, strengthened by advances in the empirical sciences, that reality is exhausted by nature. Given the rather imprecise meaning of 'naturalism' in contemporary philosophy, coming anywhere close to a unified view would be a daunting task. Undeterred by such a challenging task, Flanagan draws a sprawling, if eclectic, list of its varying and contested uses: Among them some read like rules of etiquette ("1. Philosophy should 'respect,' 'be informed by,' and 'wholeheartedly accept' the methods of science"), others like grown-up advice ("6. There is no room, or need, for the invocation of immaterial agents or forces or causes in describing or accounting for things"), and

yet others as enticements (“10. Naturalism is a form of non-reductive physicalism; there are genuine levels of nature above the elemental level”) (Flanagan 2006, 431f).

However capacious and enticing naturalism might be in its many (and often conflicting) guises, there are clear objections to its adoption as a methodology for philosophy, and Flanagan is careful to mention two of the most obvious: First, Wittgenstein’s (1922) claim that philosophy “aims at the logical clarification of thoughts” (*Tractatus* 4.112) and “is not one of the natural sciences” (*Tractatus* 4.111); and second, Bouwsma’s glib dismissal of naturalism for its belief (bordering on faith) in the “universal applicability of the scientific method” and its ignorance of the role that mathematics plays in experimental science (Bouwsma 1948). To these objections, we may add the observation of phenomenologists who, from Edmund Husserl to Dan Zahavi, have argued that what makes philosophy, especially after the phenomenological turn, immune to naturalization is that it conceives of itself as a form of transcendental inquiry that seeks to reflect on the conditions of possibility for experience and cognition (Zahavi 2013; Moran 2013). Of course, this conception of the task of philosophy, which goes back to Kant, does not rule out the possibility that empirical studies of consciousness might one day vindicate some version of naturalism fine-tuned to accommodate mental phenomena. Varela’s neurophenomenological project (1996) – first sketched in Laughlin, McManus, and d’Aquili (1992) – speaks to this vision of cognition as embodied, embedded, and enactive, and thus, as seemingly continuous with the environment of which it is a part (Lutz 2002; Lutz and Thompson 2003; Thompson 2007).

Considering the centrality of Abhidharma reductionism, with its cardinal principles of momentariness, dependent arising, and no-self, it would seem that Buddhism is friendly to naturalism, at least in a prescientific sense that reflects commitment to empiricism. As I have argued elsewhere (Coseru 2012, 3f), epistemological inquiries in India never gave birth to the sort of anti-naturalism that is associated in the West with the legacies of Descartes and Kant. Nor did Indian epistemology introduce a distinction between causal questions (How are veridical states of cognitive awareness produced?) and questions of justification (What criteria ensure that we are justified in holding a particular belief?). This lacuna, rather than indicating a shortcoming, simply reflects the pragmatic orientation of epistemological inquiry in pre-modern India, where pragmatic rather than normative concerns drive most debates about knowledge, its mode of acquisition, and its function. With Dharmakīrti, an examination of the underlying process of cognition becomes instrumental in determining which epistemic practices are conducing to effective action, giving birth to a rich scholastic tradition of both empirical inquiry and debate that continues in parts of Asia to the present day.

Can the Abhidharma reductionist accounts of experience, then, be extended to accommodate the findings of cognitive science? And would

such an extension of its scope offer a viable way of integrating its methods, ideas, and arguments into contemporary philosophical discourse? Describing it as the “First Moral Psychology,” Flanagan finds in its techniques of moral and mental discipline a useful parallel to the Socratic directive that an unexamined life is not worth living. But he is quick to point out that, unlike Socrates and his followers in the West, Abhidharma is a normative rather than a descriptive project: One does not merely build a register of mental states as they become manifest in contemplative practice; rather, one learns to identify them as wholesome, unwholesome, or neutral. For the uninitiated, Abhidharma is likely to come across as metaphysics rather than cognitive psychology. In effect, given its focus on property particulars and their relations – the much disputed elements of existence and/or experience (*dharmas*) that are constitutive of all composite entities – Abhidharma is best understood as a trope theory.

Consider, for instance, Vasubandhu’s account of the operations of analytic insight in light of Buddhist reductionism:

One examines the body with regard to its proper and general characteristics, as well as sensation, mind and the other constitutive elements of existence. Their own nature is their proper characteristic. But the general characteristic is the impermanence of produced things, the fact that everything that is connected with the four afflictions is suffering, and the fact that all things are empty and not the self.

(Pradhan 1975, 206)

On the view articulated here, the practice of analytic insight that enables us to apprehend the specific characteristics of phenomena, also discloses their partite nature, and, given the principle of momentariness, also their impermanence. At first, it would seem that what motivates the Abhidharma project are metaphysical considerations about personal identity and causality (the first, concerned with establishing the no-self view, and the latter, with the idea that to exist is to have causal efficacy). But in advancing a conception of causation that includes consciousness and cognition as efficient causal categories, Abhidharma also presents us with a metaphysics of experience: The irreducible elements of existence and/or experience (*dharmas*) are not essences or substances, but activities, properties, and patterns of connectedness. The project of identifying and mapping out these irreducible elements (e.g., sensations, volitions, etc.), with a view of achieving specific ends (e.g., virtues such as mindfulness, compassion, and equanimity), is both descriptive and experiential. It is, thus, a kind of naturalized phenomenology (Roy et al. 1999), that is, as a method for bringing into focus, capturing, and categorizing variable mental operations and contents that are normally difficult to attend to, while also submitting to empirical scrutiny about their causal and

conditioning factors (unlike, say, the conception of phenomenology that Flanagan appeals to, which stands for a sort of introspective awareness).

That such an attentional skill is itself realized within a continuum of causally interconnected states does not mean that its phenomenal properties are attributable to some sort of persisting entity or to a formal and invariant structure of consciousness. Metaphysical explanations typically look for a substratum or process, a self or self-grasping tendency that can explain why experience has the features that it does. In rejecting such a permanent owner and/or locus of experience, however, Buddhism offers an opportunity to explore the structure of awareness and the problem of personal identity not only on metaphysical and empirical grounds, but also in terms of its descriptive and constitutive features: The question why self-awareness comes bound up with a sense of self (whether owned or merely occurrent) can thus be pursued independently of metaphysical concerns about what a self is and what are its fundamental attributes. It also allows for an analysis of the structure of attentive awareness without assuming that such structure reflects an external relation of ownership between consciousness and its self-specifying features (or their analogues in the dynamic structures of brain activity).¹

I think what Buddhism naturalized ends up looking like largely depends on whatever conception of naturalism is in play. A stripped-down, bare bones Buddhism without beliefs, set free of its ancillary ‘hocus pocus’ notions of rebirth, a karmic system, and ‘bodhisattvas flying on lotus leaves’ – the sort that Flanagan favors – most likely will appeal to those who reject outright the existence of ‘nonphysical states of mind’ (Flanagan 2011, 3). If we are to advance on behalf of the Buddhist any robust metaphysical claim – the story goes – nonphysical states of mind cannot be any part of it. The metaphysical claim in question is that of *physicalism* – essentially the view that everything that exists is physical or supervenes on the physical (Stoljar 2010). The problem with this claim, as critics have argued at length, is the very conception of the ‘physical,’ which, in the absence of a definition of the essential features that all physical things have, but which nonphysical things lack (a task well-nigh impossible prior to a complete investigation of all things physical), is too vague to serve as a foundation for a complete theory of what there is (Chomsky 2006; Dowell 2006). Briefly, defining ‘physical’ simply as that which is acknowledged by the science of physics, faces the well-known Hempel dilemma: If defined in terms of current physics, well, that is an incomplete science; and if defined in terms of a future, perhaps ideal physics, well, that is too vague to serve a useful explanatory function (Hempel 1969, 1980). Methodological and demonstrative definitions, likewise, run into similar difficulties. The first, which defines the ‘physical’ in terms of what is acknowledged by the basic methodology of physics, turns a metaphysical question into an epistemological one: It defines what there is in terms of how we discover basic facts about the world.

The second, which singles out a representative sample of physical stuff ('matter' of various sorts and the middle-size dry goods that can be fashioned from it) as unequivocally representative of the 'physical' assumes that we can specify the conditions of similarity and difference for everything else. Lastly, demonstrative definitions too assume, rather arbitrarily, that if something is physical it is exclusively non-mental (Howell 2009, 87f, 2013, 19f).

The question, then, is this: Are there any alternative ways to advance the naturalism strategy that neither embrace physicalism wholesale nor reject the explanatory role of efficient causation in settling questions about the metaphysics of mind? As I have already noted, later Abhidharma, specifically Yogacara conceptions of the mental offer precisely such accounts, even as they bracket considerations about the ultimate irreducibility of the mental to a more basic substratum that serves merely as a repository for phenomenal qualities.

Neurophysicalism and the Selfless Mind

Like people seeking membership into a new and more progressive polity, ideas too, especially when hailing from our prescientific past, must undergo a process of naturalization to gain standing in our modern, scientifically grounded, republic of letters. Naturalism may cut a different profile in philosophy than it does in science; ultimately, however, it reflects a common ethos: the rejection of supernaturalism and of the whole repertory of principles, forces, or agencies whose actions are not amenable to efficient-causal explanation. Championed in America by such key figures in the pragmatist movement as John Dewey, Ernest Nagel, and Roy Sellars, naturalism was first conceived in transactional terms, as the transformation of nature in and through experience, a process that brings forth emergent properties like color and pain. Dewey, for instance, regards color neither as an exclusive property of the object, nor of light itself, nor of perception, but as a transactional phenomenon between all three (Dewey 1922/1988). In Europe, the situation is somewhat different: Influential members of the Vienna circle such as Neurath (1931) and Carnap (1932) conceived of naturalism largely as a thesis about the equivalence of statements about experience with statements about physics: "The sky is blue" is ultimately a statement about the psychophysics of light perception. Of course, the background philosophical assumptions that have kept these early conceptions of naturalism in place have now been largely abandoned. Physics and the natural sciences invoked at the beginning of the last century to support a conception of the nature of nature as material, and thus extra mental, have undergone a radical transformation. Quantum mechanics, as a set of mathematical principles for predicting the behavior of subatomic particles, may be a very efficient way of explaining what happens when phenomena at infinitesimal scales

are subjected to the instrumentation of science. But the idea that we can give an account of what the world is like at this scale *in and of itself* independently of any observation and measurement thereof is highly controversial (von Neumann 1955; Stapp 1993, 2007; Chalmers 1996).

Neurophysicalism, as the most recent incarnation of these early conceptions of naturalism, is predicated on the notion that the scientific method should be adopted in examining not only nature but human experience as well. A century later, we can state with confidence that the advice of these early champions of naturalism has been heeded. The claim that reality is exhausted by nature, however, remains problematic in light of ongoing debates about the meaning and extension of naturalism. This largely ‘semantic’ problem is further complicated by the varying degrees of commitment to naturalism: That is, those who operate with a rather unrestricted conception of nature embrace a less parsimonious ontology than stronger adherents, for whom naturalism serves as a platform for excluding most, if not all, of what belongs in the experiential domain. Among the latter, one encounters both eliminative physicalists, who seek to reduce all mental content to biological and neurobiological processes (Broad 1925; Sellars 1956; Quine 1960), and token-identity theorists who regard mental states and their neurobiological correlates as identical (Churchland 1986, 2013). Both groups are equally diverse, and count among their constituents both realist physicalists, who claim that consciousness is part of the physical world, and type-identity theorists, who think the subjective and physical domains in effect, coincide.

While the mind sciences do not rule out the possibility that mental states have nonphysical properties (indeed, the scope of cognitive science is precisely that of understanding the nature of such properties however they may be realized, rather than their reduction to more basic elements), the overwhelming evidence, argues Flanagan (appealing to an inference to the best explanation), is that “there are no such things.” As he further notes:

The reason has to do with mental causation. If mental events – for example, intentions to act – are, as they seem, causally efficacious, then the best explanation is that they are neural events. This is neurophysicalism, the thesis that mental events are brain events or, at least, bodily events, and that the subjective character of experience is explained by the way nervous systems are connected to the persons that house them.

(Flanagan 2011, 65ff)

Now, grounding the efficacy of mental events on that of neural events is just what the token-identity view consists in: A particular feeling and a given brain state are really the same thing so long as they are constituted by the same token-event. But the token-identity account fails in one

important respect: It does not and cannot explain how the phenomenal content of mental states is realized and its apparent capacity to impact physical events. Consider a typical example of intentional behavior: a cat pushing a water glass over the edge of the table, despite repeated warnings from its owner (such an incident can be easily witnessed courtesy of YouTube by simply searching for “cat knocks glass off table”). The events certainly admits of two different levels of description: one that takes into account goal-oriented behavior and the other that accounts for the brain processes underlying this instance of intentional behavior. On the token-identity view, the agent’s intention to do *x* is identical to a particular subset of neurons firing together. Since the dynamic brain event can be described intrinsically, without appeal to extraneous phenomena or events, it serves as the sole causal event. We are thus compelled to concede that descriptions in terms of dispositions and intentions are merely shorthand descriptions for brain processes. Thus we can understand this particular instance of intentional behavior as token-identical with the brain process, which alone is responsible for the causal interaction we witness at the macro-level: It’s not the cat, socialized into a world of domestic wares ripe for playful interaction, that pushes the glass, but its brain states. That leaves open the possibility of treating mental states as wholly epiphenomenal: They may well capture the seeming nature of experience, but they cannot exert any real influence on events occasioned by the only causes there are: brain processes.

One may wonder if neurophysicalism does indeed possess the explanatory resources necessary for making sense of phenomenal mental states, given that, as Flanagan acknowledges, “even the best contemporary scientific work does not yet reveal how even very simple conscious precepts, seeing a red patch, seeing a particular bent paper clip, are realized” (2011, 87). Not knowing how such basic percepts are realized and yet assuming that they are so realized takes a leap of faith. At this point we may wonder if neurophysicalism is indeed the best strategy of naturalization. Aren’t we better served by operating with a more capacious conception of nature, one that allows for non-supervenient mental causation, and for a theory of action that regards organisms as complex adaptive systems (where the mental is a self-organized structure with its own emergent dynamics that is not reflected in a change in the physical) to play a role?

The naturalism I propose here² as more suited to the task at hand is distinctly phenomenological. It argues that our conception of the mental must account for its phenomenal features in ways that capture their event-causal efficacy. These, in turn, become causally relevant in explaining how action is successfully accomplished with respect to criteria (e.g., deadlines, assent, opportunity) that are unintelligible in the third-personal language of neuroscience. Faced with such accounts of mental causation, critics typically invoke the causal closure of physical domain as evidence for the epiphenomenal character of mental states. But from

the point of view of phenomenological naturalism, epiphenomenalism seems odd and unattractive. Indeed, it is rather peculiar to think that cognitive events, which arise as a result of the tight causal coupling between perception, memory, reflection, and action, should themselves be causally inert. Furthermore, arguments that invoke the closure of the physical domain have the peculiar distinction of lacking empirical grounding: Such closure is often assumed by a priori postulation. Insofar as Flanagan favors the view that conscious mental states are identical with neural states, he too fails to provide such grounding: “the subjective character of experience is explained by the way nervous systems are connected to the persons that house them” (Flanagan 2011, 65f). That there is such connection, let alone how it is realized, is simply assumed without proof or argument.

The dominant direction of Abhidharma Reductionism does indeed point toward a naturalistic explanation of consciousness and agency. But limiting all causality to material causality, and treating consciousness as identical with brain states, forces Abhidharma Reductionism down too narrow a path for its rich accounts of consciousness and causality to have any explanatory purchase in charting the efficacy of the Eightfold Noble Path project. As such, neurophysicalism also sets the stage for moral epiphenomenalism. Yet, the causal closure of the physical domain on which neurophysicalism is predicated does not preclude an event-causal explanation of consciousness itself. Indeed, supervenience arguments against the autonomy of cognition are meant to refute the existence of a distinct metaphysical realm of mental phenomena, not the efficacy of mental processes. Lastly, as the late Jonathan Lowe argued at length, causation in the mental domain rests on principles of intelligibility (that is, on principles, which state that it is perfectly intelligible that intentions and motivations have a causal role in initiating behavior), rather than on principles of mechanism (that is, on principles, which explain how causality is actually realized) (Lowe 2008, 41).

Reasons, Causes, and the Character of Consciousness

The bodhisattva’s brush with MRI scanners is not the first time that Buddhism has faced the challenge of physicalism. First millennium Buddhist thinkers such as Dharmakīrti and Santaraksita had to wrestle with the Indian materialists, the Carvakas. This formidable and uncompromising group of thinkers, much vilified in a philosophical culture inimical to materialism, claimed that consciousness arises from the body in a manner similar to the way in which the power of intoxication emerges from fermented grains. Consciousness is a kind of high – a view that all respectable members of the Beat generation would have enthusiastically endorsed.

A brief sketch of this seminal philosophical debate should suffice to show why the framework of physicalism is problematic when introducing

Buddhist Reductionism to Western philosophical audiences. In response to the Carvaka's largely emergentist picture, Dharmakīrti and, following him, Santaraksita offer a conception of mental and physical elements as part of a complex causal chain of dependently arisen phenomena, complex enough, that is, to allow for the multiple realizability of cognition. An integral part of the Buddhist critique of physicalism is commitment to an ontological difference between 'cause' (*karana*) and 'condition' or 'conditioning factor' (*pratyaya*) in assessing the nature of causality: So long as something is deemed a cause it can only give rise to a specific type of effect. Conditions, on the other hand, can serve as a basis for the arising of multiple effects. A bulbil can only bloom into a specific type of water lily, but the same body of water can support any variety of water plants.

In claiming that the body alone is the cause of consciousness, the Carvaka physicalist assumes an unproblematic understanding of material causation. For the Carvaka philosopher Bṛhaspati (ca. sixth century CE), for instance, all manner of emergent properties can arise given the right combination of elements: While the elements themselves may not possess the qualities of the emergent phenomena, they serve as the latter's material support. The physicalist solution to the mind-body problem is thus simply a matter of either (i) treating mind and body as grounded in the same essential nature (*svabhava*) of the elemental domain or (ii) as a qualitative aspect (*guṇa*) of the body, or as an effect (*karya*) of the body (Prabhācandra 1990). If consciousness supervenes on the material elements, it does so because, as demanded by the supervenience thesis, the body's specific configuration and functionality constrains its psychology – a view not unlike the ancient Greek *harmonia* theory, which held the soul to be nothing but an attunement of the body's material elements (Castor 1997).

Can this causal model explain the specific characteristics of all emergent phenomena? In the case of the body, whose material properties are empirically discernable, it certainly offers a plausible account. But is it as effective in explaining the specific characteristics of consciousness? Here, we see called into question the principle, which Sāṃkhya philosophers too confront, that postulates the non-identity of the effect and the cause, (*asatkaryavada*). As a heterogeneous conception of causality, this principle ultimately entails the view that anything could come from anything. But – argues Santaraksita, taking Dharmakīrti's lead on this issue – causation follows the principle of homogeneity or "similar kind(s)" (*sajati*), which demands that phenomena arise not in an arbitrary manner, but through homogeneous causal chains: like causes like, cows give birth to calves, and fermented milk yields yoghurt (Shastri 1968, 449). Atypical cases, such as the caterpillar's metamorphosis into a butterfly, are just the exceptions that test the rule.

Following the principle of similar kinds, consciousness could not come from something non-conscious, a principle that also serves as the basis

for Santaraksita's definition of consciousness as that which is opposite to insentience (Shastri 1968, 478). But if consciousness is neither identical with, nor simply an emergent property of insentient matter, and yet it is only observed when the body is present, what precisely accounts for its manifestation? The dependent arising model of causation traces the arising of consciousness to the presence of corresponding sensory systems and their objects. Depending on a system of vision and a visual object, there is visual consciousness. Depending on an auditory system and sound there is auditory consciousness. In the case of introspective awareness, it is the constitutive elements of the mental domain (thoughts and desires) themselves that provide the causal link.

If an intentional mental state can serve as the basis for the arising of a subsequent moment of cognitive awareness, then consciousness is a *causa sui* and no longer fits the explanatory model of dependent arising. Here the Buddhist confronts a dilemma. Either consciousness is causally efficacious within the psychophysical domain or its efficacy is such that it does not admit of material explanation. How is it then that the operations of mind can be realized within the body, but not be reducible to its bodily (read *neural*) states?

Regardless of the difficulties that causal explanation poses for understanding the nature of consciousness, the critique of physicalism advanced by Dharmakirti and Santaraksita is instructive in the way it frames the principle of similar kinds. Simply put, the principle states that a causal relation cannot be established between two things, if changes in one do not result in changes in the other. For something to count as the effect of a cause it must be brought about by changes in the immediately preceding instance in the causal chain (Shastri 1968, 449). For phenomenal consciousness to be the effect of a body and its sensory organs, its presence must be causally dependent on the latter. But, as Santaraksita argues, experience suggests otherwise: Loss of hearing, sight, and other kinds of sensory and motor impairment do not diminish the self-reflexive character of phenomenal consciousness. So phenomenal consciousness is dependent neither on the body and the sensory systems working together, nor on each of them taken individually. The examples adduced in support of this view by Santaraksita's commentator, Kamalasila, illustrate their distinctly phenomenological approach: Severed limbs are inert and defective sense organs lack cognitive function (Shastri 1968, 452). But this phenomenological orientation does not lack empirical grounding, a view the Buddhist endorses. A well-nourished body, for instance, alters the quality of subjectivity in significant ways, claims the physicalist. True, recognizes Kamalasila, but so does the sight of blood for hemophobic individuals. Similarly, moods and desires show little correlation with the body's physical strength and stamina. If the body, as the Carvaka claims, is the material cause of consciousness, then a strong, vigorous body ought to result in a greater degree of clarity, wisdom, and understanding.

The opposite, however, is what is observed to be the case: The clarity and stability of a conscious feeling, say of love, correlates more strongly with a happy and peaceful mind than with a vigorous body. It would be a mistake to call youthful infatuation love, just as it would be a mistake to call physical stamina wisdom.

The Buddhist does not deny that cognitive awareness is in some kind of dependency relation to the body (such relation is demanded by the causal principle of dependent arising). For instance, visual awareness can only emerge in organisms that are sensitive to light. Like some contemporary emergentists, the Carvaka physicalist too grants that cognitive awareness can have novel properties not observed in the material substratum (the body) that serves as its basis. But the physicalist also claims that consciousness can be present neither when the cognitive systems are not yet developed (in the embryonic stage), nor when they are not responsive (e.g., in a state of being comatose). Is there a causal criterion for the presence of consciousness? And, more importantly, can the Buddhist answer the challenge of physicalism without appealing to the kind of evidence (e.g., the remembrance of past lives) the physicalist cannot accept?

Santaraksita's response, it seems, signals an important difference between the operations of causality in the physical domain and the limits models of material causation face when extended to consciousness and cognition. Noting the case of dreams, which are obvious cases of cognition occurring in the absence of sensory activity, Santaraksita argues for the self-intimating nature of cognitive awareness, specifically its self-reflexive character, a dimension of consciousness that presumably is not affected by the temporary interruption of awareness of oneself and of one's surrounding caused by fainting. The postulation of a non-conceptual aspect of consciousness thus allows Santaraksita to frame the question of the emergence of consciousness: Whereas the physicalist denies the presence of consciousness in the fetus on empirical grounds, the Buddhist posits it as simply a case of minimal conscious awareness:

What is the basis for asserting the absence of consciousness in sleep, swoon, and other similar states? If it is argued that "Such ascertainment comes from the absence of consciousness," then, we ask: how is such absence (of consciousness) cognized? If it is claimed, "no consciousness is cognized in that instance (of sleep or swoon)," then, that is a proof for the existence of consciousness in those states. It may be further argued, "If consciousness is present during such states, why is there no recollection of such states upon awakening?" This reasoning is not an effective refutation of our view. It is lack of vividness and other factors that account of the non-recollection of (consciousness in such states), as is also the case with conscious experience in newly born infants.

(Shastri 1968, 461)

Sleep, swoon, and other seemingly unconscious states may differ in terms of their etiology, but what they all have in common, as states of a living organism, is the capacity of awakening. The only case, argues Santaraksita where the absence of consciousness is actually observed is that of the deceased. But Santaraksita also appeals to an important distinction between conceptual and non-conceptual mental states. Indeed, the philosophical enterprise of Dignaga and Dharmakirti (as informed by Yogacara phenomenology) is grounded on the possibility of direct, non-mediated access to the givenness of experience itself. Cognition, in its non-conceptual mode, can be self-intimating without being contentful or self-grasping. The autonomy of cognition from its different perceptual modalities is also obvious, as the claim goes, in the case of mind wandering.

Does appeal to dreams, infatuation, and mind wandering suffice to make the case against the physicalist claim that cognition is an emergent property of the body? As I have argued elsewhere (Coseru 2017), slight variations in the causal chain of interdependently arisen phenomenal cast doubt on a strict model of causal generation. If a cause, which otherwise may appear perfect in the generation of some effect, fails to do so, specifically by not occasioning a difference in the mind and that which is mental, then it cannot be counted as such. Santaraksita thus appeals to an error argument to target strict causal generation. What Santaraksita and his followers argue against, then, is this notion that each mental state is instantiated by a suitably relevant combination of physical elements and processes. The persistence of perceptual illusion even after disambiguation, and the possibilities of effective action such disambiguation opens up (not chasing after a mirage), work against the strict causal model of the physicalist.

Conclusion: No Flourishing Without Agency

My aim here has been not to argue for a *sui generis* Buddhist naturalism drawn on the Abhidharma reductionist model. Rather, in taking up the question of whether there are better and worse strategies of naturalization, first, I put forward a more inclusive conception of naturalism that allows for conscious mental states to be part of the chain of dependently arisen phenomena, without rendering them epiphenomenal. I call this view *phenomenological naturalism*: The notion that cognitive awareness is to be conceived not as an internal state of mind realized by brain processes locked into linear causal chains of sensory input, subpersonal processing, and behavioral output, but rather as a *structure of comportment, an intentional orientation and attunement to a world of actions, objects, and meaning*. On this trope-theoretical phenomenological naturalism, we can make sense of how Buddhists can be reductionists about pots and selves without being eliminativist about consciousness. Against the

claim that any strategy of naturalization whatsoever must necessarily be compatible with neurophysicalism – the interpretation of naturalism that Flanagan favors – I am arguing for an inclusive conception of nature that allows for a minimal conception of agency as self-reflexivity (the awareness by which an events is given as both self-intimating and intentional). According to this conception of agency, event-causal explanation (that is, causal explanation of an event in terms of its first-person givenness) must play a role in understanding the order of the causal domain.

Phenomenological naturalism is in keeping with Buddhist diagnosis of the human condition, and the path taken to remedy it. The Buddha, after all, emphasizes not only the reality of karmic action but also the efficacy of individual effort. To those who claim that nothing is done either by oneself or another (e.g., the Ajivika fatalists), the Buddha responds by pointing out the inconsistency of such statements: Taking a first step in articulating any view whatsoever shows that there is an element of initiative. The hard naturalist, thus, faces a moral dilemma: What would become of the Buddhist account of how human flourishing is achieved if this practical philosophy of enlightenment is thoroughly naturalized? That is, if morally reactive attitudes such as anger and hatred or, alternatively, shame and apprehension, are impersonal mental factors, how are they supposed to serve as a basis for moral agency?

It is widely acknowledged that there is a clear conflict between traditional conceptions of moral agency and the agent-neutral metaphysical picture of causality that we glean from Abhidharma literature. Flanagan, much like Siderits (1987, 2008), seeks to resolve this conflict by arguing that the two pictures are compatible because the discourse of ‘persons’ and the discourse of ‘causes’ belong in two distinct and incommensurable domains. But these compatibilist solutions compromise the traditional notion of moral responsibility and render ethical conduct indistinguishable from merely pragmatic acts. The main thrust of the compatibilist move is against the notion of agent causation itself, which social and cognitive psychology has presumably rendered incoherent. It is only to the extent that we dispense with such incoherent concepts – as compatibilist interpreters of Buddhist action theory argue – that some notion of moral agency and responsibility can be salvaged.

Despite the dominant and paradoxical image of the selfless Mahayana bodhisattva tirelessly, yet effortlessly, working to put an end to *ultimately* nonexistent human suffering (on account of the nonexistence of sentient beings as *conventionally* established), support for a robust notion of phenomenal agency can be found in nearly all major schools of Buddhist thought. Indeed, the Eightfold Path program, much like the promulgation of monastic rules of conduct (the *Vinaya*), comes in recognition of the complex range of personal and subpersonal factors that are constitutive of human agency. Because mental states such as greed, hatred, and delusion or, alternatively, loving kindness, compassion, and sympathetic

joy, can only be made sense of with reference to the person whose states they are, they are irreducibly phenomenal: They only exist first-personally. The impersonal description thesis at the heart of Abhidharma reductionism may allow for the analysis of mental states in terms of their constitutive factors, but for these states to be analyzable at all, and for the attribution of moral agency and responsibility to be intelligible, there needs to be a conception of first-personal agency in place.³

Does the Buddhist conception of agency demand a radical reassessment of our understanding of voluntary action and of the causal and motivational factors that inform, condition, and sanction our valuing judgments? In order to answer this question, we must consider the defining experience that transforms Siddhartha Gautama from a human being caught in the causal web into the Buddha, an enlightened being. This transformative experience becomes at once the source of the Buddhist metaphysical picture of reality and the culmination of all human aspiration for genuine freedom. Firmly situated within this causal web, yet unattached to its emerging phenomena, the Buddha can thus declare that we ought to regard any form of sensation, attention, and consciousness as first-personally *experienced*, but not first-personally *owned*: “This is not mine. This is not my self. This is not what I am” (Bodhi 2000, 887).

This picture of causality, however, does not entail strict determinism. The enlightened being’s actions are not so much causally grounded as conditioned by an ongoing series of enabling factors. Unlike the typical – ‘if, then’ – formula of Western forms of sentential logic, the Buddhist canonical literature (in both Pali and Sanskrit) uses the locative absolute to capture the conditional nature of phenomena: ‘when that, then this.’ Hence, the central thesis (grounded in the principle of dependent arising) that all Buddhists endorse is: “When this is present, that comes to be; from the arising of this, that arises. When this is absent, that does not come to be. On the cessation of this, that ceases” (Ñāṇamoli and Bodhi 2001, 655).

The conception of agency in Buddhism, thus, is not that of an autonomous, free-willing agent or self, but of an embodied, self-referential, and self-specifying bundle of aggregates. We can thus get on with the business of charting out the experiential domain using the ‘when that, then this’ formula with enough confidence that when we act, as opposed to merely being acted upon by causal factors beyond our control, we do so for reasons that reflect both self-concern and concern about the consequences of our actions. We may dispute the libertarian conception of an unconditioned spontaneity. But reflection compels us to acknowledge its epistemic and phenomenological salience in differentiating between voluntary and involuntary actions, and the moral import of those actions that we voluntarily undertake. Of course, choice means that the alternatives so entertained are equally attainable, and that deliberation is effective in charting the range of available possibilities.

Notes

1. I develop this point at length in Coseru (2019).
2. Developed at length in Coseru (2012).
3. For an extensive treatment of the agent-neutral consequentialist framework of Buddhist ethical theory, see Coseru (2016).

References

- Bodhi, Bhikkhu, trans. And ed. 2000. *The Connected Discourses of the Buddha: A New Translation of the Saṃyutta Nikāya*, vols. I–II. Boston: Wisdom Publications.
- Bouwsma, O. K. 1948. “Naturalism.” *Journal of Philosophy* 45 (1): 12–22.
- Broad, Charlie Dunbar. 1925. *The Mind and Its Place in Nature*. London: Routledge & Kegan.
- Carnap, Rudolf. 1932/1933. “Psychology in Physical Language.” In *Logical Positivism*, edited by A. J. Ayer, 165–198. New York: The Free Press.
- Castor, Victor. 1997. “Epiphenomenalisms, Ancient and Modern.” *Philosophical Review* 106 (3): 309–363.
- Chalmers, David. 1996. *The Conscious Mind*. New York: Oxford University Press.
- Chomsky, Noam. 2006. *Language and Mind*. 3rd ed. Cambridge: Cambridge University Press.
- Churchland, Patricia. 1986. *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press.
- Churchland, Patricia. 2013. *Touching a Nerve: The Self as Brain*. New York: W.W. Norton & Company.
- Coseru, Christian. 2009. “Naturalism and Intentionality: A Buddhist Epistemological Approach.” *Asian Philosophy* 19 (3): 239–264.
- Coseru, Christian. 2012. *Perceiving Reality: Consciousness, Intentionality, and Cognition in Buddhist Philosophy*. New York: Oxford University Press.
- Coseru, Christian. 2016. “Freedom from Responsibility: Agent-Neutral Consequentialism and the Bodhisattva Ideal.” In *Buddhist Perspectives on Free Will: Agentless Agency?* Edited by Rick Repetti, 92–105. New York: Routledge.
- Coseru, Christian. 2017. “Consciousness and Causal Emergence: Santaraksita Against Physicalism.” In *The Oxford Handbook of Indian Philosophy*, edited by Jonardon Ganeri, 360–378. New York: Oxford University Press.
- Coseru, Christian. 2018. “Are Reasons Causally Relevant for Action? Dharmakīrti and the Embodied Cognition Paradigm.” In *Buddhist Philosophy: A Comparative Approach*, edited by Steven Emmanuel, 109–122. Malden, MA: Wiley-Blackwell.
- Coseru, Christian. 2019. “Presence of Mind: Consciousness and the Sense of Self.” In *Problem of the Self: Consciousness, Subjectivity, and the Other*, edited by Manidipa Sen, 46–64. New Delhi: Aakar Books.
- Dewey, John. 1922/1988. “Human Nature and Conduct.” In *John Dewey: The Middle Works*, edited by Jo Ann Boydston, Vol. 14, 3–227. Carbondale, IL: University of Southern Illinois Press.
- Dowell, Janice. 2006. “The Physical: Empirical, Not Metaphysical.” *Philosophical Studies* 131 (1): 25–60.

- Flanagan, Owen. 2006. "Varieties of Naturalism." In *The Oxford Handbook of Religion and Science*, edited by Philip Clayton and Zachery Simpson, 430–452. New York: Oxford University Press.
- Flanagan, Owen. 2011. *The Bodhisattva's Brain: Buddhism Naturalized*. Cambridge, MA: MIT Press.
- Hempel, Carl. 1969. "Reduction: Ontological and Linguistic Facets." In *Essays in Honor of Ernest Nagel*, edited by Sidney Morgenbesser, Patrick Suppes, and Morton Gabriel White, 179–199. New York: St Martin's Press.
- Hempel, Carl. 1980. "Comments on Goodman's Ways of Worldmaking." *Synthese* 45: 139–199.
- Howell, Robert. 2009. "Emergentism and Supervenience Physicalism." *Australasian Journal of Philosophy* 87: 83–98.
- Howell, Robert. 2013. *Consciousness and the Limits of Objectivity: The Case for Subjective Physicalism*. New York: Oxford University Press.
- Kitcher, Philip. 1992. "The Naturalists Return." *Philosophical Review* 101 (1): 53–114.
- Laughlin, Charles, John McManus, and Eugene d'Aquili. 1992. *Brain, Symbol and Experience: Toward a Neurophenomenology of Human Consciousness*. New York: Columbia University Press.
- Lowe, Edward Jonathan. 2008. *Personal Agency: The Metaphysics of Mind and Action*. Oxford: Oxford University Press.
- Lutz, Antoine. 2002. "Toward a Neurophenomenology as an Account of Generative Passages: A First Empirical Case Study." *Phenomenology and the Cognitive Sciences* 1: 133–167.
- Lutz, Antoine and Evan Thompson. 2003. "Neurophenomenology: Integrating Subjective Experience and Brain Dynamics in the Neuroscience of Consciousness." *Journal of Consciousness Studies* 10 (9–10): 31–52.
- Moran, Dermot. 2013. "'Let's Look at It Objectively': Why Phenomenology Cannot Be Naturalized." *Royal Institute of Philosophy Supplement* 72: 89–115.
- Ñāṇamoli, Bhikkhu and Bhikkhu Bodhi, trans. And eds. 2001. *The Middle Length Discourses of the Buddha: A New Translation of the Majjhima Nikāya*. 2nd ed. Boston: Wisdom Publications.
- Neurath, Otto. 1931/1983. "Physicalism: The Philosophy of the Vienna Circle." In *Otto Neurath: Philosophical Papers 1913–1946*, edited by Robert Cohen and Marie Neurath, 48–51. Dordrecht: Reidel.
- Prabhācandra. 1990. *Prameya-kamala-mārtaṇḍa*. Edited by Nyaya Shastri Mahendrakumar. New Delhi: Sri Satguru Publications.
- Pradhan, P. 1975. *Commentary on the Compendium of Superior Knowledge (Abhidharmakośabhāṣyam)*, revised 2nd ed., edited by A. Haldar. Patna: Kashi Prasad Jayaswal Research Institute.
- Quine, Willard Van Orman. 1960. *Word and Object*. Cambridge, MA: MIT Press.
- Roy, Jean-Michel, Jean Petitot, Bernard Pachoud, and Francisco Varela. 1999. "Beyond the Gap: An Introduction to Naturalizing Phenomenology." In *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, edited by Jean Petitot, Francisco Varela, Bernard Pachoud, and Jean-Michel Roy, 1–82. Stanford, CA: Stanford University Press.
- Sellars, Wilfrid. 1956. "Empiricism and the Philosophy of Mind." In *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*:

- Minnesota Studies in the Philosophy of Science*, edited by Herbert Feigl and Michael Scriven, Vol. 1, 253–329. Minneapolis: University of Minnesota Press.
- Shastri, Swami Dwarikadas, ed. 1968. *Tattvasamgraha of Āchārya Shāntaraksita with the Commentary “Pañjikā” of Shrī Kamalaśīla*, Vols. 1–2. Vārāṇasi: Bauddha Bhāratī Series.
- Siderits, Mark. 1987. “Beyond Compatibilism: A Buddhist Approach to Freedom and Determinism.” *American Philosophical Quarterly* 24 (2): 149–159.
- Siderits, Mark. 2008. “Paleo-Compatibilism and Buddhist Reductionism.” *Sophia* 47 (1): 29–42.
- Stapp, Henry. 1993. “A Quantum Theory of the Mind-Brain Interface.” In *Mind, Matter, and Quantum Mechanics*, edited by Henry Stapp, 145–171. Berlin: Springer.
- Stapp, Henry. 2007. *Mindful Universe*. Berlin: Springer.
- Stoljar, Daniel. 2010. *Physicalism*. New York: Routledge.
- Stroud, Barry. 1996. “The Charm of Naturalism.” *Proceedings and Addresses of the American Philosophical Association* 70 (2): 43–55.
- Thompson, Evan. 2007. “Neurophenomenology and Contemplative Experience.” In *The Oxford Handbook of Religion and Science*, edited by Philip Clayton, 226–235. New York: Oxford University Press.
- Varela, Francisco. 1996. “Neurophenomenology: A Methodological Remedy for the Hard Problem.” *Avant: Trends in Interdisciplinary Studies* 1 (1): 31–75.
- Von Neumann, John. 1955. *Mathematical Foundations of Quantum Mechanics*. Translated Robert T. Beyer. Princeton, NJ: Princeton University Press.
- Wallace, Alan. 2003. *Buddhism and Science: Breaking New Ground*. New York: Columbia University Press.
- Wittgenstein, Ludwig. 1922. *Tractatus Logico-Philosophicus*. New York: Harcourt, Brace & Company.
- Zahavi, Dan. 2013. “Naturalized Phenomenology: A Desideratum or a Category Mistake?” *Royal Institute of Philosophy Supplement* 72: 23–42.