Introduction

An exemplary thinker and a pivotal figure in the broader philosophical conversation of the so-called Age of Dialogue, Śāntarakṣita (c.725–788) provides an encyclopedic guide to the dominant Buddhist and Brahmanical currents of thought of his time. First associated with the great University of Nālandā, Śāntarakṣita would eventually journey to Tibet, at the invitation of King Khriṣrong lde btsan, to establish the first monastic university, Samye, and serve as its abbot. Well known for his synthesis of Yogācāra and Madhyamaka thought, Śāntarakṣita fuses together Nāgārjuna’s (fl. 150) anti-essentialism, and his notion that all phenomena lack inherent existence and are thus empty (śūnya), with Yogācāra phenomenology, and its stance on the ineliminability of consciousness from any account of ultimate reality. What is remarkable is that he accomplishes this task largely by adopting the epistemological tools of Dignāga (c.480–c.540) and Dharmakīrti (550–610/600–660), two thinkers closely affiliated with Yogācāra, without sharing their commitment to the irreducibility of the conventional account of mental properties.

The notion that mind and world arise in dependence upon a multitude of causes and conditions is central to Buddhist metaphysics and philosophy of mind, and much of Śāntarakṣita’s syncretic project is conceived as a defense of this view. The manifestation of both mental and physical phenomena is thus grounded in an overarching causal principle: dependent arising (pratītyasamutpāda). Conceived largely as an alternative to causal models that rely on the concept of potentiality, the principle of “dependent arising” depicts causation as nothing but an incessant flow or succession of events. With regard to entities defined primarily in terms of their capacity for agency and conscious
Consciousness and Causal Emergence: Śāntarakṣita

awareness, dependent arising underscores a dynamic picture of consciousness and cognition: each moment of awareness is thus part of a recursive, object-oriented, but also self-revealing causal stream. For the Buddhist, just as it is the case for the Pre-Socratic philosopher Heraclitus, the world is in perpetual flux; and, to draw another Western comparison, like Husserl, the Buddhist regards the phenomenal world (lokasaṃjñā) as an intentionally constituted stream of lived-experiential events.

Concern with finding an adequate causal explanation of phenomena is fundamental to Abhidharma, that system of thought and method of descriptive analysis that makes up one of the three collections of Buddhist canonical literature (the other two are the sūtra, or discourses of the Buddha, and the vinaya, or the monastic rules of conduct). It is also at the core of Śāntarakṣita’s syncretic project, which synthesizes a millennium of Buddhist thought in India by integrating Abhidharma thought, specifically as presented through the lens of Vasubandhu’s (c.400–480) Saṭrāntika perspective, within his Yogacāra-Madhyamaka system.

By advancing a conception of causation that includes consciousness and cognition as causally efficient categories, Abhidharma 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. But the project of identifying and mapping out these irreducible elements (e.g., sensations, volitions, magnitudes, arrays) may also be viewed as a naturalistic phenomenology. Indeed, given a notion of conscious cognition as pragmatically efficacious—a necessary consideration given assumptions about the efficacy of the Eightfold Noble Path, which forms the cornerstone of the Buddhism project—there are specific questions about the ways in which consciousness is also constitutive of the world of phenomena. It is largely in response to questions about the descriptive and epistemic features of consciousness that Śāntarakṣita conceives of his defense of the irreducibility of consciousness against the charge of physicalism, specifically as articulated by the Indian materialists, the Cārvākas.

The Indian materialists (also known by the more common Sanskrit term Lokāyata) occupy a unique position in the Indian philosophical tradition as the only philosophers to reject such cardinal notions as karmic retribution, rebirth, and liberation. Tracing their origins to the classical world, the Cārvākas do not cohere into a distinct school of thought until the eighth century CE—the date of their earliest recorded text, the Aphorisms of Bṛhaspati (a.k.a., Cārvāka-sūtra). Articulating a thoroughly material-ist conception of reality, with the four great elements (earth, water, fire, and air) as the only things that exist, the Cārvākas are also known for denying the reality of atoms and space on account of their imperceptibility, limiting all causality to material causality, and treating consciousness as an emergent property of elemental combination in a manner similar to the intoxicating power of fermented grains.

The key issue in the dispute between Śāntarakṣita and the Cārvākas concerns the relation between cognition and the body, and the role this relation plays in causal-explanatory accounts of consciousness and cognition. Here a number of questions arise. Does the central principle of reductionism apply to consciousness? Is there a causal
criterion for the presence of consciousness? If there is, can this causal criterion account for the specific features of consciousness, for example, its intentionality, phenomenality, and reflexive character (*svaṃvedana*)? In other words: can a causal account of phenomena be reconciled with the seeming irreducibility of consciousness? Sāntaraksīta’s answer to the challenge of Cārvāka physicalism, which he meticulously articulates in his masterful encyclopedic work, the *Compendium of Realities* (*Tattvasaṃgraha*),

displays many of the common features of classical Indian metaphysical debates on personal identity. The analysis that follows shows how arguments against reductive physicalism can benefit from an understanding of the structure of phenomenal consciousness that does not eschew causal-explanatory reasoning.

**Causation, Emergence, and the Mind**

What does the notion that mind and world are part of a complex causal nexus mean for a conception of the kinds of things that can enter into causal relationships? And what kind of causal relationships can obtain between ontologically distinct domains? The canonical literature presents us with a standard formula for the dependently arising phenomenon of consciousness: “Dependent on the eye and forms, visual-consciousness arises. The meeting of the three is contact. With contact as condition there is feeling. What one feels, that one perceives. What one perceives, that one thinks about. What one thinks about, that one mentally proliferates.”

The principle of dependent arising, at least in this original formulation, does not specify how the causal relation works. Phenomena are grouped together depending either on whether they are in tight proximity of each other or on how clearly they display patterns of regularity. Common examples include breath and the living body, fire and smoke, and the tripartite relations that obtain between kinds of awareness (visual, auditory, introspective, etc.), their corresponding objects (visibles, sounds, thoughts, etc.), and their psychological basis (e.g., vision, hearing, thought, etc.). As for the constitutive components of mind and matter, these can be regarded as either mutu-

ally entailing or as emergent features of a more fundamental causal domain, the ultimate nature of which admits of neither exclusively phenomenal nor physical description.

The first option—mutual entailment—reflects the ontological position of one of the main Abhidharma schools, the Sarvāstivāda. Here the claim is that there are two kinds of existents: primary, such as partless atoms and partless phenomenal primitives that constitute the psychophysical domain, and secondary, specifically the composite entities (or J. L. Austin’s “middle size dry goods”) that populate our mundane existence. The primary existents comprise the building blocks of material form or *rūpa*, the first of the five irreducible bundles (*skandhas*), as well as the irreducible elements of the phenomenal domain, which together constitute the remaining four bundles of the psychophysical domain. The secondary existents are secondary because they can be dissolved or made to “disappear under analysis,” unlike primary existents, which are further irreducible. The chariot can be reduced to its component parts (wheels, felly,
Consciousness and Causal Emergence: Śāntarakṣita

axle, etc.), and these in turn can be reduced to their atomic totality, but the atoms themselves are impartite and thus further irreducible. Likewise, each experiential moment, which non-Buddhist philosophers mistakenly identify with the self, can be reduced to the categories of “sensation” (vedanā), “perception” (saṃjñā), “volition” (saṃskāra), or “consciousness” (vijñāna).

The second account—emergentism—finds its best articulation among the Indian physicalists, the Cārvākas, although it also informs the conception of personal identity of one particular school of thought self-identified as Buddhist Personalism (Pudgalavāda). Emergentism postulates that a system of sufficient organizational complexity has emergent causal powers that its components, taken individually, do not possess. In the language of Abhidharma reductionism, secondary existents, while reducible to their more basic constitutive elements, nonetheless have emergent causal powers that neither element in isolation can manifest: the chariot, when assembled and fastened to a horse, has locomotion (a property not available to any of its constitutive components taken in isolation). Likewise persons, which supervene on the aggregates, but are neither identical to nor different from them, have agency, a property that, at least according to the Pudgalavāda, does not belong to either of the constitutive elements of personhood. Emergentism is thus the view that at different levels of organization the same constitutive elements can display novel properties, properties that reflect not the intrinsic natures of the elements themselves but their causal powers.

**Physicalism and Its Discontents**

Śāntarakṣita opens chapter twenty-two of his *Compendium of Realities*, aptly titled the “Examination of Physicalism,” with a summary of the Cārvāka position on consciousness and causality (vv. 1856–1870), which serves as a point of departure for his criticism:

If there is no entity that has continuity of existence, then there can be no other world, because there is none belonging to the other world. The body, cognition, the sense faculties, and the rest being destroyed every moment, they do not endure in another world. There is nothing else that is admitted (by you, Buddhists). Hence, consciousness must be regarded as produced from and/or manifested by, certain material substances, just like intoxicants, the constituent elements of fermented grains and such. The names, “body,” “sense-faculty,” and so on apply to special combinations of earth and other material substances; there is nothing besides these. There can be no relation of cause and effect between the two streams of consciousness under dispute, because they subsist in different bodies, just like the cognition of the cow and that of the horse . . . It follows that consciousness proceeds, indeed, from the body which displays the activities of breathing in and out, and the rest of the five life-breaths, as it is said by Kambalāśvatara. It is sheer audacity to posit the existence of cognitive awareness in the foetus, etc., since no object is apprehended (at this stage of development), as the sense-faculties are yet to emerge.
Consciousness can have no other form than the apprehension of objects. It is for this reason that cognitive awareness is absent in someone who has fainted. There being no soul that may be the support of consciousness, the body is the only support for it. So, at the end, when the body has ceased to exist, what will support cognition? Furthermore, how could different minds be related to the same stream of consciousness—insofar as those minds are no different than those of an elephant, a horse, and other animals? For reasons concerning the basis for cognitive awareness, you (sc. Buddhists) either have to commit to the notion of a beginningless consciousness or accept physicalism.

It is obvious from this outline that the Cārvākas, much like some contemporary physicalists, do not merely advance a new theory of causation. Rather, they put forward an austere ontological picture according to which every phenomenon reduces to the four basic material elements—earth, water, fire, and air—which alone are ultimately real. Emergent phenomena, like the fully formed body and discerning cognition, supervene on these basic elements as their functional properties. What the Cārvāka physicalists take issue with is the notion that consciousness originates outside the order of the causal domain, perhaps in a previous instance of cognitive awareness associated with a different body (as demanded by the view of karmic rebirth). But the causal closure of the physical domain at work in Cārvāka metaphysics does not preclude an event-causal explanation of consciousness itself. Indeed, the supervenience argument against the autonomy of cognition is meant to refute the existence of a distinct metaphysical realm of mental phenomena, not the efficacy of mental processes. On this emergentist picture, cognition is a graded phenomenon that tracks closely the development of the body: the absence of sensory organs in the embryonic stage, according to Cārvāka, precludes the attribution of any sort of perceptual awareness at that level of development. Nor can consciousness be present for a person in a state of coma.

For the Cārvāka, thus, consciousness is not merely a latent phenomenon, which becomes manifest as the body develops, but an expression of the body’s functional organization and responsiveness to a world of objects, situations, and things. Just like fermented grain yields a liquid with the capacity to intoxicate, so also consciousness must be regarded as nothing more than a product of the type of material organization that is constitutive of biological organisms. Indeed, the Cārvāka offers a rather stark picture of the human condition, as we see in one their earliest accounts:

We shall now examine the principles of reality. Earth, water, fire, and air are the principles (of reality), nothing else. Their combination is called the "body," "senses," and "objects". Consciousness [arises] out of these [elements], as the power to intoxicate [arises] out of fermenting ingredients. The human being is [nothing but] the body endowed with consciousness. [Cognition arises] from the body itself, because of the presence [of consciousness] when there is a body.

If consciousness is nothing but an emergent property of the body, then consciousness can be present neither without a faculty of apprehension, nor in the absence of suitable objects of apprehension. For the Cārvāka, then, seeking to circumvent these objections,
as the Buddhist does, by pointing to the momentary nature of the constitutive material elements further undermines the “continuity of consciousness” picture. The reason, claims the Cārvāka, lies with the idea of momentariness itself, which entails the serial dissolution of both the object and its apprehensibility.¹¹

It is interesting to note that the Abhidharma mereological reductionism invoked on behalf of the Cārvāka to bolster an argument against the autonomy of cognition works against both the existence of spatially extended objects and the epistemic reliability of cognitions that purport to establish the existence of such objects. It does not work, however, against cognition having a distinctive phenomenal character and epistemic status. For even assuming the Cārvāka view is correct, and only the four elements are real, the perception of objects as possessing various properties (motion, change, etc.), argues the Buddhist, cannot simply reflect the fundamental order of the causal domain. It is for this reason, then, that Śāntarakṣita’s (and his disciple, Kamalaśīla’s) critique of physicalism centers not on whether consciousness is amenable to causal explanation, but on whether an analysis of the structural features of the phenomenal domain mitigates a conception of consciousness as irreducibly qualitative, intentional, and first-personal.¹² Indeed, the physicalist must contend with the fact that grounding consciousness in the body not only raises important questions about the nature of the dependency relation, but also confronts well-known mereological issues. Assuming the arising of consciousness results from a change in the body, and taking the body to just be the four material elements so configured, as an emergent property of the body consciousness would violate the body’s physical identity. That is, the properties predicated of a body (extension, volume, weight, motion, etc.) must be so predicated from the start. Even assuming the body supervenes on the four elements, it is still the case that as a whole, the body’s emergent properties reflect the nature of the domain to which the elements belong. The Cārvāka, thus, is no better off at answering the causal question by replacing considerations about physical constitution with mereological considerations.

If the body is taken to be the cause of consciousness, as the Cārvāka claims, the question naturally arises: is it “cause” in the form of an aggregated whole, as a composite of various elements, or as an aggregate of atoms? And, given the requirement that the sensory systems are fully developed, does the body serve as a cause along with, or independently of, the senses? Finally, should the body (with its physical and functional properties) be regarded as a material cause or simply as a condition for the possibility of consciousness? Śāntarakṣita’s critique of physicalism is in large measure conceived as a series of answers to these pressing questions.

**Reasons, Causes, and the Principle of Invariable Concomitance**

Regardless of doctrinal considerations, Śāntarakṣita’s (and Kamalaśīla’s) arguments against the Cārvāka version of reductive physicalism offer an ingenious way of
conceptualizing the mind-body problem. In response to a largely emergentist picture, they offer a conception of mind and body as part of a complex causal chain of dependently arisen phenomena, complex enough, that is, to allow for both the multiple realizability of cognition, and for dissociating causes from other contributing factors. Śankarācārya’s commitment to the ontological difference between “cause” (kāraṇa) and “condition” or “conditioning factor” (pratyaya) plays an integral role in his critical assessment of the nature of causality: insofar 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.

The crucial difference between the two accounts should now be obvious: in claiming that the body alone is the cause of consciousness, the physicalist assumes an unproblematic understanding of material causation. For the Cārvāka philosopher Bhāspatī (date unknown), 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 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 being grounded in the same essential nature (svabhāva) of the elemental domain, or (ii) as a qualitative aspect (guna) of the body, or (iii) as an effect (kārya) of the body.13 If consciousness supervenes on the material elements, it does so because, as the supervenience relation demands, 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.14

Can this causal model explain the specific characteristics of all emergent phenomena? In the case of the body, whose material properties are empirically discernible, 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 (asatkāryavāda). As a heterogeneous conception of causality, this principle ultimately entails the view that anything could come from anything. But—argues Śāntarakṣita, taking Dharmakīrti’s lead on this issue—causation follows the principle of homogeneity or “similar kind(s)” (sajāti). This alternative principle demands that phenomena arise not in an arbitrary manner, but thorough homogeneous causal chains. Like causes like: cows give birth to calves and fermented milk yields yoghurt. 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 Śāntarakṣita’s definition of consciousness as that which is opposite to insentience. However, 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 a hearing 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. This generative conception of cognition raises a different con-
cern: 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. So the Buddhist confronts a dilemma again.
Either consciousness has causal efficacy insofar as it constitutes a linchpin within the
psychophysical domain or its efficacy is such that it does not admit of causal description.
How is it, then, that the operations of mind can be realized within the body, but not be
reducible to its psychological states?

Regardless of the difficulties that causal explanations pose for understanding the
nature of consciousness, the critique of Čārvāka physicalism advanced by Śāntarakṣita
is instructive in the way it frames the principle of similar kinds. Simply put, the prin-
ciple 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. For phenomenal consciousness to be the effect of a body and its sensory
organs, its presence must be causally dependent on the latter. However, as Śāntarakṣita
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 con-
sciousness. 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 Śāntarakṣita's commentator,
Kamalaśīla, illustrate their distinctly phenomenological approach: severed limbs are
inert and defective sense organs lack cognitive function. Of course, this phenomeno-
logical orientation does not lack empirical grounding. A well-nourished body, argues
the Čārvāka, alters the quality of subjective experience in significant ways. True, recog-
nizes Kamalaśīla, but so does the sight of blood for hemophobic individuals. Similarly,
moods show little correlation with the body's physical strength and stamina. If the body,
as the Čārvāka 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 con-
scious 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.

If strict causal generation is untenable, perhaps a weaker conception of dependence
could be granted. By invoking the principle of invariable concomitance, the Čārvāka
presses the criticism further, offering a different example: just as light is invariably
concomitant with a light source (such as, for instance, a lamp), so too phenomenal
consciousness is invariably concomitant with the body. How light depends on and prop-
agates from a light source may not be obvious, but their invariable concomitance is not
only unmistakable but grounded in the essential nature of things: light sources illumi-
nate. Similarly, argues the Čārvāka, while the exact mode of generation is not apparent,
phenomenal consciousness is only observed when the body and its fully developed sensory systems are present. The model of generation at work in this case allows for heterogeneous causal factors to play a role. In the same way that fire is a cause for metals manifesting the property of fluidity, so also the fetus, as the body’s material cause, is the contributing cause for the manifestation of consciousness in the infant at birth. The Cārvāka is willing to grant that, once arisen, phenomenal consciousness may have properties that explain its subjective and temporal character, hence its diachronic continuity; what is untenable from a Cārvāka standpoint is the Buddhist notion that consciousness is beginningless.

Does, then, the relation between consciousness and causal explanation point to the autonomy of cognition or does it not? Cognitive awareness is obviously in some kind of dependency relation to the body, as demanded by the principle of dependent arising. For instance, visual awareness can only emerge in organisms that are sensitive to light. The Cārvāka does grant that cognitive awareness can have novel properties not observed in the material substratum (the body) that serves as its basis. But unlike the dualist picture the Buddhist paints, the Cārvāka contends that as an awareness of a certain type (visual, auditory, etc.) consciousness must be related to the body’s specific functionality in the respective cognitive domain. Given that consciousness takes the form of an apprehension of objects (that is, given its inherently intentional structure), and given that apprehension only occurs in specific modes of cognizing such as perceiving, imagining, or remembering, consciousness can be present neither when the sensory systems are not yet developed (as in the embryonic stage) nor when they are not responsive (as in a state of coma). 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 Cārvāka would simply not accept?

Causality and the Autonomy of Cognition

The Buddhist certainly faces a dilemma: the aggregated conception of personal identity seems to support the physicalist position that consciousness is an emergent property of certain types of material organization. Consciousness cannot be both part of this causal web and just an instance in a beginningless stream of conscious events. Can the Buddhist answer the physicalist challenge while retaining a causal-explanatory framework in accounting for the relation between cognition and the body?

Śāntarakṣita’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 instances of cognition occurring in the absence of sensory activity, Śāntarakṣita
argues for the self-intimating nature of cognitive awareness, specifically its self-reflexive character. This dimension of consciousness presumably is not affected by the temporary loss of a sense of self and one's surrounding in fainting. The postulation of a non-conceptual aspect of consciousness thus allows Śāntarakṣita to frame the question of the emergence of consciousness: where 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 (vv. 1922–1926):

What is the basis for asserting the absence of consciousness in sleep or 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 for the non-recollection of consciousness in such states, as is also the case with conscious experience in newly born infants.  

Sleep, swoon, and other seemingly unconscious mental 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 Śāntarakṣita, where the absence of consciousness is actually observed is that of the deceased. Furthermore, a distinction between unconscious and non-conceptual mental states is assumed, even if not explicitly spelled out. Indeed, the epistemological enterprise of Dignāga and Dharmakirti 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 sense modalities is also obvious, claims Śāntarakṣita, in the case of mind wandering.

Given a commitment to the idea that consciousness supervenes on the body and its functions, the Čārvāka may well reel at the thought that such supervenient consciousness not only could entirely lack conceptual content of any kind but even minimal subjectivity. Śāntarakṣita is aware, of course, that in rejecting this conception of consciousness and cognition, the Čārvākas are hardly alone. Nyāya and Mīmāṃsā philosophers were equally critical of the Buddhist position on direct, non-conceptual cognition, so Śāntarakṣita's arguments do double duty. Shoring up his defense of the autonomy of cognition against strict causal generation, Śāntarakṣita provides a descriptive account of affective states (a category of mental states that are generally overlooked in philosophical debates on the nature of consciousness), tracing their emergent features not to material processes but to such psychological factors as habit formation. Children, long before they have formed habits of attraction and revulsion, exhibit feelings of love and hate. Likewise, argues Śāntarakṣita, a grown man may associate his love for a woman with certain attributes like goodness and devotion, even when she lacks them.
Does appeal to dreams, infatuation, and mind wandering suffice to make the case against the physicalist claim that conscious cognition is an emergent property of the body? For an answer to this question we may turn to Kamalaśīla, whose extensive commentary on Śāntarakṣita’s “Examination of Physicalism” supplies much of the conceptual material for rethinking the relation between consciousness and causality. If, as the Cārvāka claims, consciousness arises from the body (in the same manner in which the power of intoxication emerges from fermented grains), then does it arise from the material elements alone, from their configuration as bodies, or from their atomic totality? Setting aside the question whether causality should be understood in terms of homogeneous or heterogeneous causal chains, Kamalaśīla points out the inadequacy of strict physical models of causal generation in accounting for mental phenomena:

It could be said, “All the atoms insofar as they occur in [its] proximity are the cause of cognition.” In that case a difference should be observable between the effect produced by a non-deficient cause and that produced by a deficient cause, as the two are different. Otherwise, a distinction in the [capacity of the] cause [to bring about different effects depending on its fitness] would be futile.24

The slightest variations in the causal chain are sufficient to cast doubt on a strictly physical model of causal generation. If a cause, which otherwise may appear perfect in the generation of some effect, fail to do so, specifically by not occasioning a difference in the mind and that which mental (mano-mati), then, it cannot be counted as such. We see here a clear acknowledgement that cognitive awareness depends upon the efficacy of all underlying causal factors (perceptual, volitional, dispositional, etc.), and the rec-ognition that, in turn, these factors reduce to their causal totality (kāranasāmagrī). On a strict account of causal generation, as the examples provided by the Cārvāka suggest, cognitive error would track closely deficient causation. But that does not always happen. One might perceive a sparkling lake where there is only a naturally occurring optical illusion. This perceptual illusion is not simply a case of misapprehension, for the illusion persists even after it has been disambiguated (that is, after one has come to apprehend the appearance of the lake as a mirage).

Śāntarakṣita’s critique of physicalism, thus, uses an error argument to target what in effect is a conception of brute emergence: that is, the notion that any thing may produce anything. The persistence of perceptual illusion even after disambiguation, and the possibility of effective action such disambiguation opens up (that is, not chasing after a mirage), works against the strict causal model of the physicalist, which reduces human agency to changes in the microphysical structure of each individual. When Śāntarakṣita claims that a trustworthy cognition must also be pragmatically efficacious,25 he advances a different naturalistic account of cognition than the one put forth by the Cārvāka, one that takes into account the intentional structure of awareness and its phenomenal character. Take the example of perception: the phenomenal character of perceptual awareness is not constrained by the apprehension of the particular as such; rather, the particular is only as perceived. This way of framing the character of cognitive awareness
allows for the perceiver's intentional stance to be disclosed. In the case of perceptual illusions such as mirages, it is not only the perceiver's vantage point but also the phenomenal character of the experience itself (illusory water cannot quench thirst) that ensures successful action.

Has this Cārvāka's account of brute emergence been, then, satisfactorily challenged? Before we attempt an answer to this question, let us revisit once more Śāntarakṣita's contention that the senses are rendered ineffective by an impaired consciousness. Regardless of whether we take him to be arguing from a Sautrāntika or Yogācāra position,26 it is clear that even when he appears to reject the notion that the intentional object is causally related to the experience of a unique particular, Śāntarakṣita is in fact making a statement about the importance of phenomenological considerations (specifically, concerning the structure of awareness) in settling epistemological disputes. Indeed, his account of the efficacy of cognition hinges on whether conscious experience ultimately stands not merely for how things seem, but also for how they are. Clearly, Śāntarakṣita here takes the lead of the Buddhist epistemologists Dignāga and Dharmakīrti, who defend a reflexivist dual-aspect theory of consciousness. As Dignāga writes in his Collection on the Sources of Knowledge (Pramāṇasamuccaya):

Every cognition is produced with a twofold appearance, namely that of itself (svabhāsa) and that of the object (viṣayabhāsa).27 The objective aspect (grāhyākāra) or object-appearance (viṣayabhāsa) captures the intentional aspect of cognition or its object-directedness, while the subjective aspect stands for cognition's own self-intimating aspect, which marks it as happening for someone, as occurring in a given mental stream. Thus, regardless of whether an actual object is present or not, cognition arises having this dual aspect.

By adopting this dual-aspect theory of phenomenal consciousness, Śāntarakṣita does show commitment to an internalist epistemology. Unlike the physicalist, who operates with an externalist epistemology, Śāntarakṣita is committed to the view that cognitions do not effectively differ from their sources. Instead of understanding, say, empirical awareness in terms of a relation between perception and its object, Śāntarakṣita, much like Dignāga and Dharmakīrti, offers us an account of epistemically warranted cognitions as constituted by their content only insofar as this content is produced in the right way: that is, by the object as perceived. Incidentally, it is because illusory cognitions are not so constituted, rather than any strict causal process, that they lack both pragmatic efficacy and explanatory force. This view is motivated, at least in part, by the radical and essential separation of direct modes of apprehension (e.g., perception), which take the form of the givenness of experience itself, from the inferential processes they thus support.

If one does not factor in the dual-aspect theory of mental states in mapping out the relation between consciousness and causality, then one cannot understand why causal explanation retains an element of ontological subjectivity: whether causation is taken to be a necessary relation between events or merely a concept for observed regularities between events, it is undeniable that dependency relations are both grounded in, and reflective of, specific modes of causal reasoning. Thus, whether one rejects the Cārvāka notion that there are strict dependency relations between events, or takes the Buddhist
view that dependent arising provides an effective way to accommodate consciousness as a factor in the causal chain, a causal explanatory account must be epistemically warranted. The justification, then, for taking reflexivity (svasaṃvedana, svasaṃvitti) as a condition for the possibility of warranted cognition may indeed stem from a commitment to the self-luminosity theory of mental states, the theory that the mind has some inherently revealing capacity. What is important about this theory, however, is that it rules out the possibility of treating human cognition as an involuntary, purely mechanistic, or behavioral process. Walking is a complex motor skill, and once learned may appear involuntary. The same can be said about language and basic perceptual function. But what ensures their epistemic efficacy is the ongoing active presence of an implicit awareness that attends to the object at hand.

The acquisition of practical skills, like any other kind of moral and mental cultivation, is only possible if cognitive awareness is taken to have a specific kind of structure, specifically one that can accommodate the intentional and self-presentational aspects of each mind moment. The mind’s attentive capacity (manaskāra), thus, makes a certain dimension of human cognition not merely the effect of causal chains in the physical domain but a causal factor in its own right in the domain of cognition. This attentive or self-aware aspect of cognition can remain constant with respect to a given object of experience (say, a column of smoke), despite it being prompted by a deficient cause (in this case, a dust column). On account of this apparent variance, consciousness is neither entirely grounded in, nor explainable in terms of, physical elements and processes. The Buddhist does not deny that cognitive states are made manifest (abhivyajyate) when the body is present, only that their manifestation is not to be understood in strictly physical terms.

These considerations about why causal accounts of generation for material bodies are inadequate in explaining how conscious awareness comes to have the properties that it does bring us to the final question we asked at the beginning of this essay: Can we reconcile a causal account of phenomena with the seeming irreducibility of consciousness? It should in principle be possible to venture a positive answer on behalf of Śāntarakṣita if we abandon the notion that aggregation in the physical domain offers the only viable conception of causality. In lieu of the conventional translation of upādāna-kāraṇa as “material cause,” the more suitable notion of “appropriating” or “acquiring cause” could be adopted instead. The case could then be made that the elements, in their particular dynamic grouping as bodies, “acquire” the emergent property of sentience or awareness only in specific circumstances: that is, only when their combination serves as a basis for agency and attentiveness (that is, only when configured as living organisms). Kamalāśīla’s gloss on the relationship between cause and effect (kārya-kāraṇa), it seems, suggests precisely this reading of causation in the psychophysical domain: something that causes emergent properties such as consciousness to become manifest.

This understanding of the relationship between the autonomy of cognitive awareness and material causation differs in important ways from that of the Cārvāka. Not only is a minimal sense of non-reflective self-awareness or attentiveness covariant with the body and its sensory systems, it also assists in the emergence of cognition in its manifold
aspects. The character of cognition, as Śāntarakṣīta crucially reminds us, is contrary to insentient objects: “Cognitive awareness arises as something that is excluded from all insentient objects. This reflexive awareness of that cognition is none other than its non-insentience.”30

In articulating this conception of the phenomenal character of cognition, Śāntarakṣīta simply follows Dharmakīrti’s critique of the physicalist claim that consciousness arises from the material elements or from a transformation thereof.31 Rather, consciousness is merely covariant with, and only with, a specific type of material organization typical of sentient organisms.

**Conclusion**

The “Examination of Physicalism” does not contain Śāntarakṣīta’s full account of the nature of mind as worked out in his Yogācāra-Madhyamaka synthesis,32 but does serve as a focal point for thinking about the relation between consciousness and causality. Theories of causation form an integral part of much philosophical speculation in classical India, because they are motivated by concerns with demonstrating the possibility of freedom, which is essential to overcoming the limitations of the human condition. Insofar as they eschew such concerns, the Cārvākas also discount the importance of efficient causation, focusing instead on material causes and conditions, as their emergentist account of consciousness demonstrates. In that regard, they both align with, and face the same challenges as, modern physicalists. Indeed, from a modern standpoint, it may be objected that consciousness is a subjective phenomenon and thus not amenable to scientific (hence, efficient causal) explanation. In reducing the analysis of consciousness and cognition to transactions in the physical domain, the Cārvāka, like present-day physicalists, is thus committing a category mistake. Perhaps even Śāntarakṣīta, insofar as he advances the claim that mental events can be accounted for in efficient-causal terms, is guilty of the very charge he levels against the physicalist. But ontological reductionism is not the same as epistemological reductionism. The Buddhists may well admit that aggregated entities reduce to their ontological primitives, which alone are real. But causally describable series of events are not incompatible with treating some basic events as irreducibly mental, so Buddhist reductionism does not necessarily entail physicalism.

Now, on a distinctly Kantian line of argumentation about the irreducibility of normative relations (relations that obtain in the logical space of reason), mental events that are constitutive of personal identity will not admit of efficient-causal description. Against such a normative framework, Śāntarakṣīta’s appeal to causal explanation as a criterion for the efficacy of epistemic practices would arguably render his account indistinguishable from that of the physicalist.33 The efficacy of reasons lies primarily in the relevance of their content rather than the fact that they fall under some efficient-causal description. The point of this Kantian take on the Buddhist epistemological project is that reasons engage our conceptual capacities in ways that cannot be accounted for
in terms of the dynamics of causally efficacious particulars. However, if Śāntarakṣita’s causal reductionism allows for material causation to play only an assisting role in mediating cognition, then the epistemological naturalism he proposes as a viable alternative to physicalism is less problematic, especially when considered in light of embodied accounts of cognition.

Śāntarakṣita’s response to the problem of how to think of consciousness using the language of efficient-causal explanation is framed by two sorts of considerations: first, about the basis or support of consciousness (that is, about what sorts of factors might be responsible for the arising of different aspects of consciousness given a metaphysical commitment to momentariness), and second, about the structure and character of consciousness (which reflects a commitment to the reflexivity thesis). Consciousness is thus constitutive of a constant and continuous stream of discrete cognitive events, not independently of, but rather alongside, various conditioning and dispositional factors. For Śāntarakṣita, thus, the problem is not how cognition can arise from purely causal interactions in the psychophysical domain. Rather, the problem is how this domain, which also includes irreducible phenomenal properties, conditions the arising of specific cognitive events.

If there is one lesson from contemporary philosophy of mind that can prove useful in negotiating Śāntarakṣita’s contribution to understanding the relation between consciousness and causality, it concerns a specific case of the so-called fallacy of ambiguity. As John Searle has convincingly argued, “there is a systematic ambiguity in the use of the word ‘empirical,’ an ambiguity between an ontological sense and an epistemic sense.” Sometimes “empirical” stands for contingent states of affairs, and sometimes for a method that can be used to establish something as factual. In particular, facts about one’s subjective experience are not empirically accessible in the way that facts about the external world are. The basis for the epistemic subjective-objective distinction, thus, is an ontological distinction in modes of existence. Even for Śāntarakṣita, whose ontological perspective with regard to the nature of reality most closely aligns with the Sautrāntika, entities like mountains and atoms are experience-independent (even as they are treated as conventional designations). On the other hand, there is a class of phenomena that are experience-dependent: pains and pleasures manifest only as phenomenal qualities of awareness. The experience-dependent status of these qualitative states, however, does not render them any less real. The pain, whatever its causal status with regard to prior events (say, a wound), has a distinctive qualitative aspect.

Is there a clear boundary between experience-dependent and experience-independent phenomena? Even in the prescientific world of first-millennium Indian Buddhism one can appeal to specific physical and psychophysical laws to negotiate this difference. On the one hand, we are physiologically limited to perceiving only a narrow range of light frequencies. On the other hand, the composition and temperature of fluid mediums such as air and water cause light to bend (heat waves at ground level cause straight lines to appear curved). Not only is the analysis of cognitive error central to the Buddhist epistemo-logical project, but also accounts of its etiology clearly distinguish between internal and external factors in an effort to mark the boundary between experience-dependent and experience-independent phenomena.
The capacity to unambiguously apprehend phenomena not only as they seem, but as they presumably are, suggests that we can have an epistemologically objective account of subjectivity. In short, ontological subjectivity is no bar to epistemic objectivity. Consciousness does not begin when attention is directed to a given object or mental state: that is, conscious awareness does not become manifest by being reflected upon (which is to say that consciousness has an observer-independent status). On the other hand, causality is an observer-relative phenomenon: the very notion of “event” presupposes an observer. Events thus stand in a particular kind of relation to their antecedents only to the extent that there is a conception of causality in place. That causality should be an observer-relative phenomenon does not mean, however, that it is not a real feature of our ontology. Rather, its observer-relative status simply suggests that it contains an element of ontological subjectivity.

One of the central principles of Buddhist philosophy of mind is that causal explanation contains an element of ontological subjectivity. It is this conception of causal explanation that allows Śāntarakṣita to claim that not only cognition, but even such apparently involuntary and causally governed phenomena as breathing, are contingent on the presence of consciousness.

Notes

1. Consisting of an entire chapter, the “Examination of Physicalism” (Lokāyataparīkṣa), illustrates Śāntarakṣita’s critical and philosophically insightful engagement with the Cārvākas. Along with Kamalaśīla’s extensive Commentary (Pañjikā), this work serves as a locus classicus for framing the mind-body problem in Indian philosophy. The citations and summaries follow the Sanskrit edition of the text in Swami Dwarikadas Shastri, ed., Tattvasaṃgraha of Āchārya Śāntarakṣita with the Commentary “Pañjikā” of Shri Kamalaśīla, vol. 2 (Vārānasi: Baudha Bhārati Series, 1968), hereinafter cited as Commentary on Realities. Ganganatha Jha’s translation of the entire work, albeit dated, is still very useful in navigating the intricacies of this complex encyclopedic work. See Ganganatha Jha, Tattvasaṃgraha of Śāntarakṣita With the Commentary of Kamalaśīla, vols. 1–2 (Baroda: Gaekwad Oriental Series; reprint, Delhi: Motilal Banarsidass, 1986).


legitimate position within the Buddhist tradition. For the Pudgalavādins, personhood is an emergent and unity-conferring principle, neither identical to, nor different from, the five aggregates, and reflects the specific ways that human individuals differ from other sorts of causal continua.

6. For an excellent discussion of emergentism in Indian philosophy, which identifies some of the specific ways in which both Buddhist and Nyāya philosophers in particular adapt the notion that thinking can be an emergent phenomenon and yet enjoy a certain degree of autonomy, see Jonardon Ganeri, “Emergentism, Ancient and Modern,” *Mind* 120, no. 479 (2011): 671–703, and the “Reasons, Causes, and the Principle of Invariable Concomitance” section of the present essay.

7. *Compendium of Realities*, pp. 444–446. For a discussion of these opening stanzas (includ-ing the occurrence of some of these verses in Kumārila’s seminal work, the *Ślokavārttika*), in the context of broader engagements with the Cārvāka in Indian philosophy, see also R. Bhattacharya, *Studies in the Cārvāka/Lokāyata* (London: Anthem Press, 2011), chap. 11.

8. Stoljar defines physicalism as the thesis “that everything is physical” with the caveat that “most physicalists do not mean that absolutely everything is physical or that every par-ticular is physical or even that every property is physical.” Rather, physicalism should be taken to mean simply that “every instantiated property is necessitated by some physical property”; Daniel Stoljar, *Physicalism* (New York: Routledge, 2010), 5.


11. *Compendium of Realities*, p. 446.
22. *Compendium of Realities*, p. 461
26. It is possible to read Sāntarakṣita as endorsing both the Yogācāra view of the luminos-ity of the mind (*prabhāsvaraṇa cittam*)—with the epistemic consequence that refl exive awareness (*svasaṇvedana*) is taken as the only warranted type of cognition and the
Sautrāntika notion that the appearance of an object in cognition corresponds to stable particulars.


32. The general scholarly consensus is that Śāntarakṣita’s own philosophical views ultimately align with the Madhyamaka. As James Blumenthal convincingly argued, his entire philosophical synthesis is encapsulated in a single stanzas (v. 92) of his most mature work, *The Ornament of the Middle Way*: “By relying on Consciousness Only, know that external entities do not exist. And by relying on the Madhyamaka, know that no self at all exist, even in that consciousness.” See Blumenthal, *The Ornament of the Middle Way: A Study of the Madhyamaka Thought of Śāntarakṣita* (Ithaca, NY: Snow Lion), 22.


34. Embodied cognition refers to a complex research program in cognitive science and philosophy of mind that emphasizes the role that the body and the environment play in the development of cognitive process. For a detailed discussion of downward causation, emergentism, and the supervenience thesis in embodied cognitive science, see Evan Thompson, *Mind in Life: Biology, Phenomenology, and the Sciences of Mind* (Cambridge, MA: Harvard University Press, 2007), Appendix 2.

35. John Searle, *Seeing Things as They Are* (New York: Oxford University Press, 2015), 74. Searle traces this ambiguity to the British empiricists, specifically to Hume’s *Dialogues*, and argues that only recognition of the intentional structure of awareness can help negotiate the epistemic subjective-objective distinction.

**Further Reading**


