

8 Freedom from responsibility

Agent-neutral consequentialism and the bodhisattva ideal

Christian Coseru

Introduction

Is there such a thing as free will in Buddhism? Do moral and mental forms of cultivation at the heart of Buddhist practice imply some notion of agency and responsibility? And if they do, how are we to think of those individuals who embark on the path to liberation or enlightenment, considering that all Buddhists give universal scope to the no-self doctrine? Of course, Buddhism is not alone among the world's great philosophical traditions in providing ample testimony for the possibility of cultivating to a high degree such cardinal virtues as non-violence, wisdom, compassion, and a general spirit of tolerance. But it is unique among them in articulating a theory of action that, it seems, dispenses altogether with the notion of agent causation. Buddhists pursue what are unmistakably moral ends, but there is no stable self or agent who bears the accumulated responsibility for initiating those pursuits, and seemingly no normative framework against which some dispositions, thoughts, and actions are deemed felicitous, and thus worthy of cultivation, while others are not so deemed. It is not surprising, therefore, to find a near universal lack of agreement among contemporary interpreters about how best to capture the scope of Buddhist ethics using the vocabulary and theoretical frameworks of Western ethical discourse.

In seeking an answer to the questions above, the plan, then, is first to show that despite some straightforward metaphysical tenets, the conception of agency in Buddhism is less alien than it may seem at first blush—indeed, it is not unlike conceptions of moral agency that we find in Stoic thought, and more recently in Nietzsche (2006) and several strands of contemporary moral phenomenology; next, to argue for a solution to what is widely regarded as a clear conflict between traditional conceptions of moral agency and the agent-neutral metaphysical picture of causality that we glean from Abhidharma literature. Recent accounts (Flanagan 2002; Meyers 2014; Siderits 1987, 2008) seek 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. Specifically, my claim is that 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.¹ 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 Mahāyāna (later Buddhist) bodhisattva (one who has taken the altruistic vow) tirelessly, yet effortlessly, working to put an end to ultimately non-existent human suffering (on account of the non-existence 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.² 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 (cf. Abelson, Ch. 13, this volume) 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 responsibility and freedom to be intelligible, there needs to be a conception of first-personal agency in place. On the view I defend here, mental states are irreducibly first-personal: the idea of generic pain apart from individually realized sensations of burning, itching, or stinging is thus deeply incoherent.

In what follows, I argue that influential Mahāyāna ethicists, such as Śāntideva, who allow for moral rules to be proscribed under the expediency of a compassionate aim, seriously compromise the very notion of responsibility. Moral responsibility is intelligible only in relation to conceptions of freedom and human dignity that reflect a participation in, and sharing of, interpersonal relationships. As critics of hard determinism (the view that universal causal necessitation is incompatible with free will and moral agency) have argued, there is no threat to human agency so long as we understand that agency is essentially grounded in a range of participant reactive attitudes and feelings (e.g., resentment, gratitude, anger, etc.) that are impossible without the ascription of agency and moral responsibility (see Strawson 1973, p. 11, and discussion in Goodman 2009, pp. 147ff.). But bodhisattvas cannot be seen to harbor such participant reactive attitudes, at least not once they are sufficiently advanced on the path to understand that no beings exist whatsoever as ordinarily conceived. For compatibilists, thus, the extraordinarily demanding bodhisattva ideal—informed, as it is, by a steadfast commitment to forego the body, its enjoyments, and all virtue for the sake of accomplishing the welfare of all sentient beings—makes a compelling case for allowing special dispensation. On an agent-neutral consequentialist interpretation of the Mahāyāna ethical project, we must grant the Buddhist saint dispensation for the unfathomable and mysterious ways in which utterly impersonal psychophysical aggregates accomplish their aim, while the unenlightened must be content with merely following rules.

I do not dispute this claim. Nor do I disregard the importance of revising our traditional notions of agency and moral responsibility to accommodate new findings about the sociobiological roots of morality. Rather, I simply caution that such revisionary strategies, insofar as they seek to explain agency in event-causal terms, may well (if they have not done so already) set the stage for moral epiphenomenalism.³ Indeed, on the view I defend here, an effective compatibilist solution to the problem of reconciling freedom of the will and determinism depends on expanding, rather than eliminating, the complex register of factors that underpin the experiential aspects of our moral life. In short, although social and cognitive psychology has significantly augmented our *knowledge* of agency, there is a widespread sense that mapping out human action in impersonal terms—a project of significant affinity to Abhidharma—has advanced only marginally our *understanding* of agency, of *what it is like* to show responsiveness to norms, reasons, and principles.

Freedom, destiny, and the will

As with classical Western conceptions of causality and agency, the Indian philosophical context at the time of the Buddha presents us with a wide array of philosophical views: for (arguably) strict determinists like the Ājīvikas, who embrace a fatalist conception of human existence, all actions are predetermined by an external force of destiny (*niyati*), from which there is no escape; at the other end of the spectrum we find the Cārvākas, the Indian physicalists, for whom the most probable explanation for the existence of the universe is a series of random events. The Cārvākas reject both the law of karma and the concept of destiny because implicit in these notions is a view of existence as inherently purposeful (Bhattacharya 2011). The latter view is not unlike that of some contemporary libertarians who, drawing on the findings of quantum mechanics, argue that the statistical probabilities that characterize events at the subatomic level extend to the everyday realm of human experience: actions, though biologically and psychologically conditioned, and constrained by the norms of social conduct, are not strictly causally determined (Kane 1996, 1999; Wallace 2011).

It is worth noting from the outset that causal determinism was far less an issue of concern for the historical Buddha than the fatalism of the Ājīvikas. In a discourse on “The Fruits of the Ascetic Life” (*Sāmaññaphala Sutta*), the Buddha is particularly concerned to reject the view of Makkhali Gosāla, who, by removing all trace of effort from human action, renders the ethical life meaningless: neither defilement nor purification have any cause or condition, no action is voluntarily undertaken either by oneself or by another, and, generally speaking, humans lack power, energy, and steadfastness; it is simply their lot in life to experience pain and pleasure in a manner beyond their control (DN2.19–20, in Walshe 1987, pp. 94ff.; Meyers 2014, pp. 62ff.). In condemning this view, the Buddha not only emphasizes 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, 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, that one either strives to overcome some resistance or to reach the sort of reflective equilibrium that comes with understanding and insight (AN.6.38, in Bodhi 2012, p. 901).

Should this rejection of fatalism be taken to mean that the Buddha is championing freedom of the will? More importantly, is there a notion of personal autonomy at work in the Buddha's clear admonition to his followers to jettison the extremes of both determinism and indeterminism, and devote themselves instead to an ethical life in the pursuit of liberation? It would appear that the picture of the ordinary human condition, mired in ignorance and moved by short-term pragmatic goals, precludes such a notion of personal freedom. The evaluative attitude implicit in this complex analysis of cognitive and affective states, however, seems to suggest otherwise. Clearly, an expression of self-concern and concern about the consequences of one's actions is a ubiquitous feature of Buddhist teachings. Indeed, while the value placed on shame and apprehension suggests that the Buddha favors a conception of responsibility and moral self-regard for those pursuing the Eightfold Path, such a perspective is not ultimate (see Meyers' 2014 deft analysis of this view). Specifically, morally reactive attitudes, whether unwholesome (e.g., anger and hatred) or wholesome (shame and apprehension), are still impersonal mental factors. But, like the complex notion of disposition (*cetanā*), they are also self-referential mental states: "they presuppose the notion of oneself as a morally responsible agent" (Meyers 2014, p. 63). Is this conception of moral agency self-referentiality constitutive or is it merely an emergent feature of these mental states? Are our choices and the sense of control and ownership of action illusory or is agency built into the very fabric of lived experience?

Agency, choice, and self-referentiality are complex notions with a rich and contested history of interpretation. Some attention to Western theories of intentionality, the will, and motivation for action, then, is necessary if we are to make any progress in clarifying whether, and in what way, we can make sense of the Buddhist conception of selfless agency. This contrasting analysis must recognize that while Western and Buddhist ethical discourses are embedded in their own intellectual histories, there is no neutral stand from which to assess their merits and possible limitations. While contemporary philosophers do address meta-ethical questions, historically the paucity of inquiries into the nature of ethics in Buddhist philosophy makes the matter all the more complicated. By noting this absence (first pointed out in Siderits 1987), I do not mean to suggest that Buddhists either deny or doubt the possibility of moral knowledge. Nor do I mean to endorse the sort of moral skepticism championed by Mackie (1977), which says that judging a particular action morally permissible is simply a statement about one's participation in a specific way of life, Buddhist or non-Buddhist. I do, however, share the generally Kantian line of argumentation, which says that there are better and worse ends, and thus better and worse ways of achieving those ends depending on the criteria that we adopt (Kant 1993).

As I already noted, efforts to capture the scope of Buddhist ethics are mired in disagreements, mainly between those who favor a virtue ethical model, on

account of the presence of a rich catalogue of virtues and of practices conducive to their mastery (Keown 2001), and those who advance consequentialist interpretations, owing mainly to the identification of happiness and the elimination of suffering as key elements in a comprehensive list of factors that define wellbeing (Goodman 2009). Efforts to map out the theoretical structure of Buddhist ethics in sui generis terms—and thus to steer clear of both consequentialist and virtue ethical models—have so far gestured in the direction of moral phenomenology as the basis for the tradition’s normative claims: only the experience of enlightened beings can serve as a criterion for moral blame or praiseworthiness (see Garfield 2010, 2014). I have no intention to weigh in on this debate, relevant as it may be to the broader question of whether a specific conception of agency underwrites the Buddhist path all the way to awakening or only up to a point (namely, the moment immediately preceding it). Of course, this question assumes that we know what kind of agency, if any, enlightened beings exhibit, an assumption that lies at the heart of yet another controversy, about the very nature and possibility of enlightened agency (see Garfield 2006; Finnigan 2011).

Let us, then, briefly consider the concept of the will. An integral part of the Western vocabulary, captured by such ubiquitous statements as “Did you do that of your own free will?”—the idea of will occurs in the classical worlds of neither India nor Greece. For Aristotle, who provides much of the technical philosophical vocabulary for virtue ethics, voluntary action is conveyed by the less ambiguous concept of deliberative desire (*bouleutikê orexis*), which captures what it means for an action to be within our power. For an action to be deliberately and effectively undertaken, desire and reason must converge: it is only when reason is desiderative (*orekticos nous*) and desire is thoughtful or deliberative (*orexis dianoetike*) that we are in a position to make informed decisions (NE.VI.2, 1139b4, in Aristotle 1999, p. 87; Murphy 2001). When these two work in concert they give rise to the more capacious idea of moral purpose (*prohairesis*). In classical India, there are different avenues of volitional pursuit, typically classed alongside material (*artha*), affective (*kāma*), social (*dharma*), and ultimately emancipatory (*mokṣa*) ends. These ends are not necessarily continuous. The ethical life, centered on the first three aims, is almost entirely dispensed with in the generally spiritual and otherworldly quest of Upaniṣadic lore. If Aristotle heeds the Socratic dictum ‘knowledge is virtue,’ the Upaniṣadic sage Yājñavalkya is proleptically Humean in regarding morality as the play of emotions over reason. The early Buddhist conception of the ethical life, likewise, shares in this Upaniṣadic impulse to transcend all inclination, desire, and emotion in the pursuit of the higher, if still self-referentially constituted, goal of liberation (see Bilimoria *et al.* 2007, pp. 40ff.).

The idea of ‘will’ as a distinct faculty occurs for the first time in the writings of St. Augustine. In *On Free Will (De libero arbitrio)*, St. Augustine sets out to address the problem of theodicy by the introduction of a new faculty, free will (*liberum arbitrium*), which alone is responsible for moral acts being deemed praiseworthy or blameworthy.⁴ An omniscient and omnibenevolent God, as creator of the world, cannot be the cause for the primal fall from grace that

marks the human condition (according to the old Hebrew myth of the Fall from Eden). Hence, the invention of a new faculty, the will, capable of producing uncaused free action—that is, action that is not accountable in terms of natural events and processes (see Stump 2001). But he also uses the term *voluntas* ('the will'), which he adopts from Cicero and especially Seneca, who use it to refer to the Stoic manner of assenting to a given proposition. For the Stoics, this assent to propositions of the sort, "No man can compel you to receive what is false,"²⁵ has moral valence. Thus, to the extent that Augustine's use of '*voluntas*' captures the notion that we morally assent to various propositions, his understanding of agency is ultimately continuous with the Greek conception of moral purpose (*prohairesis*), especially as it finds articulation in Epictetus. With one caveat: the Stoic and Christian conceptions of human nature are radically different. For the Stoics, who take human nature to be ultimately pure, pursuing the moral life is basically living in accordance with right reason (not to be confused with the (Kantian) idea of a morality grounded in rationality). As Seneca so eloquently puts it in *Epistulae morales*, the pursuit of such moral ends is predicated on the notion that "conduct cannot be right unless the will to act is right" (1917–1925, XCV, pp. 56ff.). For Augustine, the 'right will to act' becomes 'free will,' not as a condition for the possibility of right attitudes of the mind (*habitus animi*), but as a necessary condition of the justice of divine retribution (see Rist 2001, pp. 34ff., for a detailed discussion).

Clearly, nothing resembling the Augustinian conception of a free will tied to divine justice is to be found in Indian philosophy. Karma, or the reward and punishment for action, is the closest we come to a conception of justice. But this is cosmic, rather than divine, justice: karmic consequences depend on the universality of the causal principle of dependent arising, not on uncaused divine judgment.

The more pressing question, however, is whether the idea of free will is compatible with determinism, given a conception of the world as causally ordered. Here, I want to draw on Stoic thought again as providing a better basis to conceptualize the Buddhist conception of moral agency. The Stoic emphasis on the causal antecedents of mental states does not mean that they are externally necessitated (Long 2002, p. 28). The occurrence of mental states may be causally governed, but their intelligibility is not: the latter requires assent, the only criterion of individuation that marks a mental state as mine, as occurring in my mental stream. It is the volition manifest in assenting, thus, that serves as the basis for Stoic conceptions of personal identity and moral agency. But assenting is not the same as judging. Evaluative judgments may intrude, but the entire scope of the Stoic life is to bracket them, to realize (with Epictetus) that "death is nothing terrible, else it would have seemed so even to Socrates; rather it is the idea that death is terrible that is terrible" (*Encheiridion*, 5, in Hard 2014, p. 288).

This Stoic conception of the moral purpose (*prohairesis*), indeed, is not unlike the role assigned to *cetanā* ('disposition' or 'volition')⁶ in the canonical Buddhist literature. Various rendered as 'will,' 'volition,' 'intention,' 'motivation,' 'conation,' 'drive,' 'stimulus,' 'determination,' 'effort,' 'choice,' and 'resolve,' *cetanā* is typically the sort of bodily, verbal, and mental activity one performs either on

one's own or conditioned by others (AN.II.158, in Bodhi 2012, pp. 563ff.). I can either voluntarily raise my arm or have it raised by another, as a referee would upon declaring the winner of a boxing match. Similarly, I can either think through an issue and volunteer an opinion or ponder a question and offer a response. Thus, I can say with certainty that my response to a question is causally determined by external factors. But this determinist picture of agency is too simplistic to capture the complexity of intersubjective relations (the sort of relations that, as already noted, Strawson has in mind when he suggests that participant reactive attitudes are indispensable to an account of moral agency and responsibility). My response might be solicited by a question, but that a string of sounds registers as a question requires a complex set of interpretive, evaluative, and analytic skills that can only be constituted as reasons. It is our responsiveness to reasons prompted by valuing judgments—of the sort that extol the cultivation of certain mental states as wholesome and the rejection of others as unwholesome—that serves as conduit for verbal and mental activity, even when caused by other things.

The question of freedom and determinism, thus, must consider not whether factors relevant to moral assessment can be causally assessed, but whether the agent-neutral framework of Buddhist Reductionism is compatible with a conception of responsibility-entailing moral agency.⁷

Agency, causation, and the moral domain

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? 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. The centerpiece of this metaphysical picture is the causal principle of dependent arising (*pratītya-samutpāda*) and a thoroughly reductionist account of persons, which takes volition to be but one of several contributing factors that shape human identity and agency.

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, whether “past, future, or present; internal or external; manifest or subtle . . . as it actually is . . . [as]: ‘This is not mine. This is not my self. This is not what I am’” (SN22, 48, in Bodhi 2000, p. 887). Rather, we are told, the arising of each element in the person series is only as conditioned by the presence of immediately preceding, causally efficacious elements:

[D]ependent 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.

(MN.I.111–112, in Ñāṇamoli and Bodhi 2001, p. 203)

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 Pāli canonical literature uses the locative absolute to capture the conditional nature of phenomena: 'when that, then this.' Hence, the central thesis (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.

(See, e.g., MN.II.32, in Ñāṇamoli and Bodhi 2001, p. 655)

It would appear thus that the conception of agency in Buddhism is not that of an autonomous, free willing agent or self, but of an embodied and self-referential bundle of aggregates. We can thus get on with the business of charting out the experiential domain using the 'when that, then this' formula: when there is touch, then there is feeling, when there is awareness, then there is grasping for objects. Can we go as far as to say 'when there is agency, there is moral responsibility'? The early Buddhist literature, as we have already noted, is unambiguous that initiative is essential to moral progress. But the philosophical innovations of later Mahāyāna Buddhism, specifically the doctrine of 'emptiness' (universal metaphysical insubstantiality) proposed by Nāgārjuna, complicate the ethical project. By making the agent-neutral metaphysical picture of selflessness indispensable to Mahāyāna Buddhist ethics, Śāntideva describes a way of living with the practical consequences of actions that effectively lack agency, and thus also lack the sort of intersubjective relation that entails moral responsibility.

This complication is especially problematic when, in the *Bodhicaryāvatāra* (BCA), Śāntideva (1995), drawing on the principles of dependent arising and momentariness, claims that there is no continuity between agency and the experience of its consequences "in terms of a unity of the continuum of consciousness" (pp. 9, 72). Pressed with the objection that divorcing agency from the experience of moral responsibility makes the pursuit of virtues such as compassion for all sentient beings irrelevant, Śāntideva appeals to the 'two truths' (ultimate vs conventional) framework of Madhyamaka ('Middle Way') dialectic to make an even more radical claim: from the perspective of *ultimate* truth, the embodiment of such perfections as wisdom and compassion lacks the intentionality, aboutness or directness of ordinary (*conventionally* understood) mental states. In response to the crucially critical question: "for whom is there compassion if no being exists?" (BCA.9.75), Śāntideva contends that so long as the delusion—that there is a task to be done (e.g., bringing sentient beings to the realization of the truth of emptiness)—persists, the illusion of effort persists too (BCA.9.76).

Thus, when ethicists like Śāntideva reject even this minimal conception of agency as the ground for moral and mental cultivation, the Buddhist ethical project reaches an impasse. If there is no agent, and if actions are merely transient events arising within a continuum of causally interconnected states,

what explains the phenomenal character of experience? Touch, after all, is not generic contact, but an active and firm grip. Awareness is not bare wakefulness, but the sense of being present here and now. And compassion is no mere feeling for others, but empathetic self-disclosure in the presence of others. Furthermore, the capacity for self-regulation that grounds our moral sense presupposes that we are not merely self-aware but aware in a way that makes us implicitly responsive to action and their consequences. Even if we assume, as the evidence from cognitive neuroscience seems to suggest, that we are psychologically hardwired to attribute agency and hold others responsible for their actions, the question why such agency-attributing capacities should be accompanied by a moral sense remains to be explained (see Gray *et al.* 2007; Arico *et al.* 2011).

Whether the Buddhist no-self view is simply a theoretical construct, derived from metaphysical considerations about agency and causality, or a descriptive account grounded in the phenomenology of lived experience, matters to our conception of agency and moral responsibility. Compatibilists argue that reductionism about persons is not incompatible with the pursuit of an ethical life. Of course, the compatibilist must acknowledge that the conventional practice of morality (to which the Buddha offers precepts, inspiring tales, and rules of conduct) and Buddhist metaphysical doctrine are in conflict (this is precisely Siderits' 2008 view).

Can appealing to the 'two truths' framework of Buddhist philosophy solve this conflict? How is the moral life to be justified in terms that ultimately make no reference to anything experiential and intentional? If enlightened agency is no agency at all, what makes it desirable? The Buddha's concern to reject any conception of determinism that strips our efforts of causal efficacy becomes all the more pertinent. Indeed, if the Buddhist analysis of experience allows for persons to have the kinds of freedom necessary for the pursuit of moral ends, then the principle of dependent arising cannot function as a basis for strict causal determinism.

That agency and moral responsibility are deeply intertwined is obvious when we consider the relation between practical deliberation and theorizing about the nature of things: the latter looks for causal explanations of events, and ultimately finds them in impersonal elements and factors that are constitutive of the natural world. This naturalistic picture has no place for concepts like 'freedom' and 'responsibility.' But most, if not all, of our most pressing deliberations rest on practical reasoning of the sort that asks, "What should I do?"—and then looks for the most justifiable course of action. If such is the case, then holding myself responsible for actions that I undertake is integral for their success. That is, regardless of whether theoretical reason is able to demonstrate freedom or not, practical reason must assume that freedom is possible for the purpose of effective action. This Kantian perspective on human agency is motivated by the assumption that the kind of freedom we are supposed to consider (and criticize) is as described by libertarians or agent causal theorists. This conception of freedom gives agency its spontaneity within the logical space of reasons.

Do 'freedom' and 'responsibility' belong in a discourse about causation in the natural world? If such discourse does not eliminate consciousness from its ultimate

picture of what there is, then freedom and responsibility are no mere artifacts of practical reason, but epistemically objective features of lived experience. If, on the other hand, we find no room for practical concerns about how best to live in our ultimate ontology, then freedom and responsibility are confined exclusively to the domain of social convention. The Buddhist metaphysical picture of reality, as a product of theoretical reason, is devoid of any reference to selves and their concerns, or indeed to anything substantive. At least in principle, the no-self view would preclude any robust account of free will and responsibility.

Yet, Buddhist practice requires the observance of certain norms and the valuation of certain types of thought, speech, and action that are considered beneficial. Chief among these is the restraint of unmitigated willful thought, speech, and action. However, this valuation, and the psychological terms in which it is expressed, is at odds with an impersonal account of phenomena in causal terms. Proposals for some kind of Buddhist compatibilism to solve this conflict, as I have argued elsewhere (Coseru 2017), indirectly render agency in general, and moral agency in particular, epiphenomenal. The largely consequentialist framework of compatibilism, on my view, cannot give an adequate account of our moral institutions, and is generally indifferent to the concerns of practical reason.

Which action, whose responsibility?

Now that we have a clearer grasp of why understanding the nature of agency matters to morality, we can return to the metaphysical question of what personal agency entails. What does it mean to take ourselves as capable of choosing or directing our actions in a deliberate way? As it should be obvious, entertaining such a question and reflecting on what it means for the alternatives it presents (voluntary or intentional behavior vs behavior that is simply caused by a totality of causes and conditions) opens up the possibility that reason could serve as a causal motive for action. For someone like Kant, this deliberative process does not simply inform, but also enacts, the idea of spontaneity (1998, A533/B561, p. 533). In short, when we act, as opposed to merely being acted upon by causal factors beyond our control, we do so for reasons.

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. It is, after all, a demonstrable truth that how we choose to act (by deliberating about possible alternatives) makes a difference in how we actually act. Of course, choice means that the alternatives so entertained are equally attainable, and that deliberation is effective in charting the range of available possibilities.

The choice to act one way or another is also grounded in all sorts of practical considerations. Do we need a conception of free will or even an idea of freedom in a transcendental sense to ground our practical deliberations? Those who find the Kantian argument—about the independence of reason from the necessitation of impulses—compelling, take the view that we can only be free if we conceive of ourselves as such. But this way of framing the problem confronts us with yet

another metaphysical conundrum, in this case about whether conceivability entails possibility. It seems intuitively plausible that we can and indeed do conceive of ourselves as free agents, and thus that we do assume freedom for the possibility of action. And it is equally plausible that we can conceive of ourselves as lacking agency by externalizing the causes of our actions (Dennett 2004, p. 292 entertains this very possibility). This is precisely the strategy that informs Śāntideva's Madhyamaka ethics: pain, anger, and desire simply arise due to causes and conditions without there being someone for whom the pain is sharp or stingy, the anger righteous or impulsive, and the desire wholesome or unwholesome, respectively.

If conceiving of ourselves as free agents challenges the dominant picture of the universe as a causally closed physical system, conceiving of ourselves as lacking agency comes, it seems, at a significantly lesser cost: with human behavior explainable in terms of either external causes or internal, but subpersonal, cognitive processes, we can dispense with the notion of responsibility altogether. The utility calculus at the heart of agent-neutral consequentialism compels us to make the less expensive choice: sacrificing freedom also means the end of moral responsibility. It is no longer the individual but her brain or hormones that precipitate action. However, counterintuitive it may seem at first, the no-self picture is perfectly suited to accommodate this account of personal identity.

Also conceivable is that Buddhist ethical thinkers such as Śāntideva, are less concerned with the possibility of freedom in a causally ordered universe (such possibility is nonetheless taken to be the *modus operandi* of all enlightened beings), and more with minimizing suffering and/or maximizing happiness for all sentient beings. As there is no overarching normative framework and no need to demonstrate freedom, ethical conduct is simply a matter of pursuing certain pragmatic ends. Most importantly, the rules that regulate the pursuit of such ends vary depending on whether one is a novice bodhisattva or a realized Buddhist saint. And since Buddhist saints, unlike novice practitioners, are permitted to break moral rules in the service of carrying out compassionate actions, they are also free from the responsibility such actions entail for the unenlightened. In the end, it is precisely this freedom from responsibility, perhaps ironically, that makes the bodhisattva ideal the more costly alternative. No responsibility does not just mean no justification for action: it also means no agency, not even for the Buddha, who obviously could not have taught, let alone inspired, myriad generations to follow in his path.

Notes

- 1 See Caruso (2012), Smart (2006), and Wegner (2002) for various attempts to prove the illusory nature of experiences of mental causation. While not conclusive, Nahmias *et al.* (2004) review experimental data that seems to favor compatibilist over incompatibilist accounts of free will.
- 2 Despite the taboo on speaking about 'the self,' psychological and moral attitudes form an integral part of the Buddhist tradition (see Collins 1982, Ch. 6). Also, despite the dominance of the ultraminimalist account of agency developed in the Abhidharma,

there are good and compelling reasons to give ‘Buddhist personalists’ (pudgalavādins) credit for insisting that important features of personhood are ineliminable (see Carpenter 2015; Priestley 1999).

- 3 The problem with event-causal theories of action is their failure to capture agency altogether, instead reducing it merely to things that happen to us. On this model, there are pushes and pulls but no one does anything ever. This is the so-called problem of the ‘disappearing agent’ (see Mele 2003, Ch. 10; Lowe 2008, pp. 159ff.; Steward 2013).
- 4 As Garfield (2014, p. 166, n.1) notes, both compatibilists and libertarians claim Augustine as their source, and both readings are possible.
- 5 As Epictetus (Hard 2014) notes in *Discourses* 1.17, moral agency is grounded in our capacity to assent:

Can any man hinder you from assenting to the truth? No man can. Can any man compel you to receive what is false? No man can. You see that in this matter you have the faculty of the will free from hindrance, free from compulsion, unimpeded.
- 6 Among the most common translations are ‘will’ (Rhys Davids 1898), ‘volition or conation’ (Aung and Rhys Davids 1979), ‘choice’ (Keown 2001), ‘volition’ (Guenther 1976), and ‘intention’ (Gombrich 1988; Heim 2014). Garfield (2014) thinks all arguments in favor of a conception of ‘the will’ in Buddhism are bad arguments because they rely on tendentious translations of *cetanā* as ‘choice’ rather than ‘intent’ or ‘volition.’ See also Repetti (2010) for an analytic review of Western discussions of free will in Buddhism that focuses on interpretations of the principle of dependent arising, and its possible interpretation as endorsing either a soft or hard determinism.
- 7 As Siderits (2008, p. 30) notes, since classical Indian philosophers did not directly address this problem, we cannot go to the historical record in search of an answer. Instead, the question should be framed in terms of what Buddhists ‘should say, given their other commitments.’

References

- Arico, A., Fiala, B., Goldberg, R.F., and Nichols, S., 2011. The folk psychology of consciousness. *Mind & Language*, 26(3), pp. 327–352.
- Aristotle, 1999. *Nicomachean ethics*. 2nd edn. Translated by T. Irwin. Indianapolis, IN: Hackett.
- Aung, S.Z. and Rhys Davids, C.A.F., trans., 1979. *Compendium of philosophy (Abhidhammatthasaṅgaha)*. London: Pali Text Society.
- Bhattacharya, R., 2011. *Studies in the Cārvāka/Lokāyata*. London: Anthem Press.
- Bodhi, B., trans. and ed., 2000. *The connected discourses of the Buddha: a new translation of the Saṃyutta Nikāya*, vols. I and II. Boston: Wisdom Publications.
- Bodhi, B., trans. and ed., 2012. *The numerical discourses of the Buddha: a translation of the Aṅguttara Nikāya*. Boston: Wisdom Publications.
- Bilimoria, P., Prabhu, J., and Sharma, R., eds. 2007. General introduction: thinking ethics, the West and India. In: P. Bilimoria, J. Prabhu, and R. Sharma, eds. *Indian ethics: classical traditions and contemporary challenges*, Vol I. Hampshire: Ashgate. pp. 1–32.
- Carpenter, A., 2015. Persons keeping their karma together: the reasons for the *Pudgalavāda* in early Buddhism. In: K. Tanaka, Y. Deguchi, J.L. Garfield, and G. Priest, eds. *The moon points back*. New York: Oxford University Press. pp. 1–44
- Caruso, G.D., 2012. *Free will and consciousness: a determinist account of the illusion of free will*. Lanham: Lexington Books.
- Collins, S., 1982. *Selfless persons: imagery and thought in Theravāda Buddhism*. Cambridge: Cambridge University Press.

- Coseru, C., 2017. Breaking Good: Moral Agency, Neuroethics, and the Spontaneity of Compassion. In: J. Davis, ed., *A Mirror is for Reflection: Understanding Buddhist Ethics*. New York: Oxford University Press.
- Dennett, D., 2004. *Elbow room: the varieties of free will worth wanting*. Cambridge, MA: MIT Press.
- Finnigan, B., 2011. How can a Buddha come to act? The possibility of a Buddhist account of ethical agency. *Philosophy East and West*, 61(1), pp. 134–160.
- Flanagan, O., 2002. *The problem of the soul: two visions of mind and how to reconcile them*. New York: Basic Books.
- Garfield, J.L., 2006. Why did Bodhidharma go to the East? Buddhism's struggle with the mind in the world. *Sophia*, 45(2), pp. 61–80.
- Garfield, J.L., 2010. What it is like to be a bodhisattva? Moral phenomenology in Śāntideva's *Bodhicaryāvatāra*. *Journal of the International Association of Buddhist Studies*, 33(1–2), pp. 333–358.
- Garfield, J.L., 2014. Just another word for nothing left to lose: freedom, agency and ethics for Mādhyamikas. In: M. Dasti and E.F. Bryant, eds. *Free will, agency, and selfhood in Indian philosophy*. New York: Oxford University Press, pp. 164–185.
- Gombrich, R., 1988. *Theravāda Buddhism: a social history from ancient Benares to modern Colombo*. London: Routledge.
- Goodman, C., 2009. *Consequences of compassion: an interpretation and defense of Buddhist ethics*. New York: Oxford University Press.
- Gray, H.M., Gray, K., and Wegner, D.M., 2007. Dimensions of mind perception. *Science*, 315, p. 619.
- Guenther, H., 1976. *Philosophy and psychology in the Abhidharma*. Berkeley, CA: Shambhala.
- Hard, R., trans., 2014. *Epictetus, Discourses, Fragments, Handbook*. Oxford: Oxford University Press.
- Heim, M., 2014. *The forerunner of all things: Buddhaghosa on mind, intention, and agency*. New York: Oxford University Press.
- Kane, R., 1996. *The significance of free will*. Oxford: Oxford University Press.
- Kane, R., 1999. Responsibility, luck, and chance: reflections on free will and indeterminism. *Journal of Philosophy*, 96, pp. 217–240.
- Kant, I., 1993. *Grounding for the metaphysics of morals*. Translated by J.W. Ellington. Indianapolis, IN: Hackett.
- Kant, I., 1998. *Critique of pure reason*. Translated by P. Guyer and A.W. Wood. Cambridge: Cambridge University Press.
- Keown, D., 2001. *The nature of Buddhist ethics*. Basingstoke: Macmillan.
- Long, A.A., 2002. *A Stoic and Socratic guide to life*. London: Clarendon Press.
- Lowe, E.J., 2008. *Personal agency: the metaphysics of mind and action*. Oxford: Oxford University Press.
- Mackie, J.L., 1977. *Ethics: inventing right and wrong*. New York: Penguin.
- Mele, A.R., 2003. *Motivation and agency*. Oxford: Oxford University Press.
- Meyers, K., 2014. Free persons, empty selves: freedom and agency in light of the two truths. In: M. Dasti and E.F. Bryant, eds. *Free will, agency, and selfhood in Indian philosophy*. New York: Oxford University Press, pp. 41–67.
- Murphy, J.B., 2001. Practical reason and moral psychology in Aristotle and Kant. In: E.F. Paul, F.D. Miller, and J. Paul, eds. *Moral knowledge*. Cambridge: Cambridge University Press, pp. 257–299.
- Nahmias, E., Morris, S., Nadelhoffer, T., and Turner, J., 2004. The phenomenology of free will. *Journal of Consciousness Studies*, 11(7–8), pp. 162–179.

- Ñāṇamoli, B. and Bodhi, B., trans. and eds., 2001. *The middle length discourses of the Buddha: a new translation of the Majjhima Nikāya*. 2nd edn. Boston: Wisdom Publications.
- Nietzsche, F., 2006. *On the genealogy of morality*. Translated by C. Diethe. Cambridge: Cambridge University Press.
- Priestley, L., 1999. *Pudgalavāda Buddhism: the reality of the indeterminate self*. Toronto: Centre for South Asian Studies, University of Toronto.
- Repetti, R., 2010. Earlier Buddhist theories of free will: compatibilism. *Journal of Buddhist Ethics*, 17, pp. 279–310.
- Rist, J., 2001. Faith and reason. In: E. Stump and N. Kretzmann, eds. *The Cambridge companion to Augustine*. Cambridge: Cambridge University Press, pp. 26–39.
- Śāntideva, 1995. *The Bodhicaryāvatāra*. Translated by L. Crosby and A. Skilton. New York: Oxford University Press.
- Seneca, 1917–1925. *Moral epistles*, Vol. I. Translated by R.M. Gummere. Cambridge, MA: Harvard University Press.
- Siderits, M., 1987. Beyond compatibilism: a Buddhist approach to freedom and determinism. *American Philosophical Quarterly*, 24, pp. 149–159.
- Siderits, M., 2008. Paleo-compatibilism and Buddhist reductionism. *Sophia*, 47, pp. 29–42.
- Smart, J.J.C., 2006. Metaphysical illusions. *Australasian Journal of Philosophy*, 84 (2), pp. 167–175.
- Steward, H., 2013. Processes, continuants and individuals. *Mind*, 122 (487), pp. 781–812.
- Strawson, P.F., 1973. *Freedom and resentment and other essays*. London: Methuen.
- Stump, E., 2001. Augustine on free will. In: E. Stump and N. Kretzmann, eds. *The Cambridge companion to Augustine*. Cambridge: Cambridge University Press, pp. 124–147.
- Wallace, B.A., 2011. A Buddhist view of free will: beyond determinism and indeterminism. *Journal of Consciousness Studies*, 18(3–4), pp. 217–233.
- Walshe, M., trans., 1987. *The long discourse of the Buddha: a translation of the Dīgha Nikāya*. Boston: Wisdom Publication.
- Wegner, D., 2002. *The illusion of conscious will*. Cambridge, MA: MIT Press.