

BUDDHIST PERSPECTIVES ON FREE WILL: AGENTLESS AGENCY?

Edited by

Rick Repetti

Routledge

Taylor & Francis Group

LONDON AND NEW YORK

Sarva mangalam!

Especially those trying to figure out whether or not anything is really *up to us*...

There is free action, there is retribution, but there is no agent that passes from one set of momentary elements into another one, except the lawful connection of those elements.

—*Paramārtha Śūnyatā Sūtra* (*Discourse on Ultimate Emptiness*), in S. Edelman,
Computing the Mind: How the Mind Really Works, p. 477

CONTENTS

Foreword by Daniel Cozort

Acknowledgments

Contributors

Preface

Introduction

- 1 Why the Buddha Did Not Discuss ‘the Problem of Free Will and Determinism’
Christopher W. Gowans

- 2 Why There Should Be a Buddhist Theory of Free Will
Rick Repetti

- 3 Uses of the Illusion of Agency: Why Some Buddhists Should Believe in Free Will
Charles Goodman

- 4 Just Another Word for Nothing Left to Lose: Freedom, Agency and Ethics for
Mādhyamikas
Jay Garfield

- 5 Negative Dialectics in Comparative Philosophy: The Case of Buddhist Free Will
Quietism

Owen Flanagan

6 Free Will and the Sense of Self

Galen Strawson

7 What Am I Doing?

Susan Blackmore

8 Freedom from Responsibility: Agent-Neutral Consequentialism and the Bodhisattva Ideal

Christian Coseru

9 Free Will, Liberation and Buddhist Philosophy

Marie Friquegnon

10 Buddhism and Free Will: Beyond the 'Free Will Problem'

B. Alan Wallace

11 Degrees of Freedom: The Buddha's Implied Views on the (Im)possibility of Free Will

Martin T. Adam

12 Buddhist Paleocompatibilism

Mark Siderits

- 13 Shifting Coalitions, Free Will, and the Responsibility of Persons
Ben Abelson
- 14 Psychological versus Metaphysical Agents: A Theravāda Buddhist View of Free Will and
Moral Responsibility
Peter Harvey
- 15 Emotions and Choice: Lessons from Tsongkhapa
Emily McRae
- 16 Grasping Snakes: Reflections on Free Will, *Samādhi*, and *Dharmas*
Karin Meyers
- 17 Agentless Agency: The Soft Compatibilist Argument from Buddhist Meditation, Mind-
Mastery, Evitabilism, and Mental Freedom
Rick Repetti

Index

FOREWORD

American politics and governance has in recent years been thrown into turmoil by the emergence of a strong Libertarian streak, exemplified by the Tea Party. Libertarians, as their name implies, believe that we are free. Libertarians tend to believe that we are autonomous beings whose free choices create our futures; we fully own ourselves; we owe nothing to others. They accordingly tend to advocate laissez-faire capitalism, protection of private property, and ethical egoism (that moral agents ought to seek their own self-interest). They believe that the poor are poor because of bad choices and that an unequal society may nonetheless be a just one.

Ayn Rand, a Russian expatriate whose best-selling novels include *Atlas Shrugged* (New York: Random House, 1957), is Libertarianism's patron saint. In her collection of essays, *The Virtue of Selfishness* (New York: New American Library, 1964), Rand argues against altruism because it is the enemy of self-fulfilment: "if a man accepts the ethics of altruism, his first concern is not how to live his life but how to sacrifice it" (p. 27). Among her many admirers is the current Speaker of the House of Representatives (and third in line for the Presidency), Paul Ryan, who required staffers to read *Atlas Shrugged*. Hard-line Libertarians such as Ron Paul and his son, Kentucky Senator Rand Paul (named after the novelist) have had little success in running for President, but only slightly more moderate Libertarians running under the Tea Party banner have had strong showings.

I mention all of this because, as this volume makes very clear, the Buddha was the original *anti*-Libertarian. Several essays collected here show that the Buddha refuted the underlying premise

of Libertarianism, namely, that we are autonomous moral agents. There are no autonomous moral agents in the first place because no agent is persistent; the Buddha's analysis of the person demonstrates (in the way now reflected in Western post-modern thought) that 'self' is something merely imputed to the changing components of body and mind. And second, there is no autonomous moral agent because everything and everyone is highly conditioned by many causes and conditions. Hence, the premises upon which Libertarianism in any of its forms is founded are false, even absurd.

The Buddha taught that we are conditioned beings, the product of biological imperatives, the history of our relationships, the norms and customs of our societies, and all of our experiences. I might think that I am making choices of my own free will. But what if someone knew me very, very well? What if, for instance, I had a twin who had always been my constant companion and with whom I had always freely shared my feelings and thoughts? Is there any doubt that he could almost always predict exactly how I would make choices and even what I would say in a particular situation? If I am in fact so completely determined, how could I in any meaningful sense call my choices 'free'? On the other hand, the Buddha taught that our fate is not inevitable; we can change the trajectory of our lives. His Eightfold Path offers that possibility. Does that mean that I do, in fact, have some kind of free will?

This is the first volume ever to collect the wide range and nuances of Buddhist thinking on the matter of free will. Here, readers will find many fascinating questions discussed. Certainly the Buddha was a determinist, but was he a 'hard' or 'soft' one? The former excludes the possibility of free will, whereas the latter admits at least the possibility of it. Does the answer to that

question turn on whether we refer to the Buddha of the Pāli scriptures or the Buddha of the Mahāyāna? Does the process of developing insight and virtue ‘soften’ determinism or does it replace one type of determinism (the force of countless lifetimes of delusion, ill-will, and greed) with another (the force of compassion)? Does the Buddhist assertion that there is no permanent, independent self, controller of mind and body, mean that there is no sense in talking about ‘free will’ in the first place? Nevertheless, since ‘free will’ is a concept almost universally accepted in the world, is it ‘conventionally true’ even if ultimately there is no agent whose will could be free? If there is no free will, is there no moral responsibility? Might revealing to people that they have no free will discourage them from pursuing a spiritual path?

This volume is also important because it brings Buddhist thought into dialogue with Western philosophers such as Harry Frankfurt, J.M. Fischer, Galen Strawson and Peter Strawson. It is also notable that many of the contributors, while specializing in Buddhism, have their training in Western philosophy. Although Owen Flanagan laments, “I am wary of asking Buddhists to talk about our problem of free will because it is a bad and idiosyncratic problem”, others have found it more fruitful to use Western categories to explore Buddhist thought.

Rick Repetti has been thinking about these issues for many years. I know, because he published several articles in the *Journal of Buddhist Ethics*, of which I am the editor, surveying all of the writing done by Buddhist scholars on the subject of free will over the past century. He is now drafting a monograph on the subject, to which the present volume will be the companion. He is almost ideally placed to bring together the many points of view represented here and has skillfully sequenced them so that each after the first may be seen as commenting on the one

prior. This book will provoke discussion on the subject of free will in relation to Buddhism for many years to come.

Daniel Cozort

Carlisle, PA

ACKNOWLEDGEMENTS

The inspiration, thinking, and writing that went into my efforts to put this book together owe in significant part not only to many of the contributors with whose writings I have been engaged over the years, many of them in this volume, but to many others, directly or indirectly, among the following, who contributed to my life-long philosophical curiosity, particularly about meditation, mental freedom, and free will. To my first meditation teachers, Hilda Charlton, Ram Dass, and Ma Jaya, and many other teachers and communities of practitioners. To my undergraduate mentors, Andrew Wengraf, who has read and commented on most of my philosophical writings since, and Eric Steinberg, for whom I first wrote on free will (in an independent study). To my dissertation advisor, Michael Levin, who supported converting many of my footnotes on Buddhism to text, as well as the other members of my committee, John Greenwood and Steven Grover, and my two external examiners, Harry Frankfurt and Galen Strawson. To the members of my Free Will Circle at the CUNY Graduate Center, especially Tziporah Kasachkoff and Gregg Caruso. To Daniel Cozort, the editor of the *Journal of Buddhist Ethics*, for patiently editing several previous versions of my thinking on this subject for publication in *JBE* and for encouraging me to parlay those efforts into an edited collection and/or monograph. To the Columbia Society for Comparative Philosophy and its members for feedback on an earlier version of my ideas on this topic, particularly Owen Flanagan, Christopher Gowans, and Joerg Tuske. To the NEH 2012 Summer Institute, *Investigating Consciousness: Buddhist and Contemporary Philosophical Perspectives*, both faculty and participants, who supported my thinking on this subject and where I developed or deepened relationships with our contributors Christian Coseru, Jay Garfield, Emily McCrae, and Mark Siderits. To the PSC-CUNY for grant

support on some earlier papers that fed into this research project. To my chairperson Michael Barnhart, for mentoring me, directing me to the *JBE* and to the NEH Summer Institute, for support with my teaching schedule, and for reading earlier versions of some of my writings here.

I thank the following publishers and journals for permission to reproduce (some revised) material from the following publications: Oneworld Publications, for Susan Blackmore's *Zen and the Art of Consciousness* (2011), pp. 135-149; Oxford University Press (www.oup.com), for Jay Garfield's "Just Another Word for Nothing Left to Lose: Freedom, Agency and Ethics for Mādhyamikas", in Matthew R. Dasti and Edwin F. Bryant, eds., *Free Will, Agency and Selfhood In Indian Philosophy* (2014), pp. 164-85, and for Galen Strawson's *Freedom and Belief* (1986), chapters 5 and 6; the *Journal of Buddhist Ethics*, for Emily McRae's "Emotions, Ethics, and Choice: Lessons from Tsongkhapa", vol. 19, pp. 344-369; the *Journal of Consciousness Studies*, for B. Alan Wallace's "A Buddhist View of Free Will: Beyond Determinism and Indeterminism" (2011), vol. 18, no. 3/4, pp. 21-33; the *Journal of the International Association of Buddhist Studies*, for Martin T. Adam's "No Self, No Free Will, No Problem: Implications of the *Anattalakkhana Sutta* for a Perennial Philosophical Issue" (2011), special volume, pp. 239-265; and *Philosophy East & West*, for Mark Siderits's "Buddhist Paleocompatibilism", vol. 63, no. 1, pp. 73-87.

I thank our publisher, Routledge / Francis & Taylor, their anonymous readers whose reports helped improve this collection, and the Asian Studies series editor Dorothea Schaeffer and especially her assistant Jillian Morrison (and, formerly, Sophie Iddamalghoda) for their patient support throughout the publication process. I also thank my family and friends for understanding

why I could not be present as much as I would prefer because, as my stepdaughter, Greta Triantis put it (using gestures to indicate ‘scare quotes’), I was ‘working’. Finally, my deepest gratitude is to Valerie Triantis, without whose ceaseless love, faith in me, and support of my work, this would not be possible. Of course, if some of the arguments in this collection are sound, then none of these individuals are real, nor do they ‘deserve’ any gratitude. But, as far as I can tell, that cannot be right, although ‘I’ could be wrong...

CONTRIBUTORS

Ben Abelson is a Visiting Assistant Professor at Mercy College in Dobbs Ferry, NY. A New York City native, he received his B.A. from Hunter College, where he taught as an adjunct for several years, and his Ph.D. in Philosophy (2015) from the City University of New York Graduate Center. His dissertation was an analysis of the concept of a person, in which he defines a person as a being that jointly possesses the capacities for self-consciousness and concern. Ben is currently co-editing two anthologies with Marie Friquegnon: *Great Disputes: Iconic Clashes in Philosophy* and *Window to Truth: Collected Papers of Raziel Abelson*. He has also been a vocalist for various local NYC hard rock and heavy metal bands, including Scribes of Fire and God Ox.

Martin T. Adam is an Associate Professor in the Religious Studies Program at the University of Victoria, Canada. He received his Ph.D. in Religious Studies (McGill University) and an MA in Philosophy (University of Waterloo). His recent interests lie principally in Buddhist meditation theory and practice, and in Buddhist ethics, about which he published well over a dozen articles and several book chapters. In addition, he has edited *Indian Buddhist Metaethics: Panel Proceedings from the Meeting of the IABS* (Atlanta, 2008) in the *Journal of the International Association of Buddhist Studies* (2010), 2011, special volume, and translated *Bhāvanākrāma (The Process of Meditation)* by Ācārya Kamalaśīla (International Buddhist Academy, 2012).

Susan Blackmore is a psychologist, lecturer and writer researching consciousness, memes, and anomalous experiences, and a Visiting Professor at the University of Plymouth. She is a TED

lecturer, blogs for the Guardian, and often appears on radio and television. *The Meme Machine* (1999) has been translated into 16 other languages; more recent books include *Conversations on Consciousness* (2005), *Zen and the Art of Consciousness* (2011), and a textbook *Consciousness: An Introduction*, 2nd ed. (2010).

Christian Coseru is an associate professor of philosophy at the College of Charleston. He works in philosophy of mind, phenomenology, and cross-cultural philosophy, especially Indian and Buddhist philosophy in dialogue with Western philosophy and cognitive science. He recently published *Perceiving Reality: Consciousness, Intentionality, and Cognition in Buddhist Philosophy* (Oxford 2012). In 2012 he co-directed (with Jay Garfield and Evan Thompson) an NEH Summer Institute, *Investigating Consciousness: Buddhist and Contemporary Philosophical Perspectives*. He is currently completing a book on the intersections between perceptual and affective consciousness, *Sense, Self-Awareness, and Sensibility*. He taught in the Centre for Asian Societies and Histories at the Australian National University, received his Ph.D. from the Australian National University in 2005, and holds a B.A. and M.A. in philosophy from the University of Bucharest. While at ANU, he worked on a proof of concept model for parsing Sanskrit based on the Interlingua System (funded by an ARC grant). He travels extensively for his research, and spent over four years in India studying Sanskrit and Indian Philosophy, affiliated with several research institutes, including the Asiatic Society in Calcutta, the Bhandarkar Oriental Research Institute and De Nobili College in Pune, and the Central Institute of Higher Tibetan Studies, Sarnath, Varanasi. He was a visiting scholar at Queens College, Cambridge University (2000) and Institut de Civilisation Indienne, Paris (2001).

Daniel Cozort is associate professor in the Department of Religion at Dickinson College in Carlisle, Pennsylvania, where he teaches comparative religion. A native of North Dakota, he earned degrees from Brown University and the University of Virginia. He has written two books on Buddhist *tantra*, a catalogue for an exhibition on tantric art, a video documentary on sand mandalas, a book on the schools of Indian Buddhism, another specifically on the Prasaṅgika Mādhyamika School, and several articles and book chapters on topics such as anger, suffering, Western Buddhist teachers, Buddhist ethics, and Tibetan Buddhism in America. He is the editor of the *Journal of Buddhist Ethics* and the *Oxford Handbook of Buddhist Ethics*.

Owen Flanagan is James B. Duke Professor of Philosophy and Co-Director of the Center for Comparative Philosophy at Duke University, Durham, NC. He is the author of *The Bodhisattva's Brain: Buddhism Naturalized* (MIT 2011) and *The Geography of Morals: Varieties of Moral Possibility* (Oxford 2016).

Marie Friquegnon is Professor of Philosophy and of Asian Studies at William Paterson University. She earned her Ph.D. in Philosophy at NYU. She is the author of *On Shantarakshita* and *Reflections on Childhood*, and editor of *A Short Introduction to the Philosophy of Shantarakshita*. She edited and co-translated with Khenchen Palden Sherab Rinpoche and Khentrul Tsewang Dongyal Rinpoche, Santaraksita's *Tattvasiddhi* and *Madhyamakalankara* with Rinpoches' commentaries (forthcoming, Coolgrove Press). She co-edited with Raziel Abelson and Michael Lockwood *The Philosophical Imagination*, and co-edited with Raziel Abelson *Clarity and Vision* and *Ethics for Moral Life*. She is currently at work on a translation and commentary of Santaraksita's *Madhyamakalankara Vritti*, a collection of essays, *Buddhism*

and Philosophy, and an intellectual autobiography *Finding Santaraksita*. She is editing with Benjamin Abelson *Window to Truth: Collected Papers of Raziel Abelson* and *Great Disputes: Iconic Clashes in Philosophy*. She has also published many articles and given many lectures. She is founder and actively engaged in the Society for the Study of Indian and Tibetan Buddhist Philosophy.

Jay Garfield is Doris Silbert Professor in the Humanities, Director of the Logic Program and of the Buddhist Studies Concentration at Smith College, Regular Visiting Professor of Philosophy and Buddhist Studies at the Harvard Divinity School, Professor of Philosophy at Melbourne University and Adjunct Professor of Philosophy at the Central University of Tibetan Studies. Garfield's most recent books include *Engaging Buddhism: Why it Matters to Philosophy* (Oxford 2015), *Moonpaths: Ethics and Emptiness* (with the Cowherds, Oxford 2015), *Madhyamaka and Yogācāra: Allies or Rivals?* (edited with Jan Westerhoff, Oxford 2015), *The Moon Points Back: Buddhism, Logic and Analytic Philosophy* (edited, with Yasuo Deguchi, Graham Priest and Koji Tanaka, Oxford 2015), *Indian Philosophy in English from Renaissance to Independence* (with Nalini Bhushan, Oxford 2011), *Contrary Thinking: Selected Papers of Daya Krishna* (edited with Nalini Bhushan and Daniel Raveh, Oxford 2011) and *Sweet Reason: A Field Guide to Modern Logic*, 2nd ed. (with James Henle and Thomas Tymoczko, Wiley-Blackwell 2011).

Charles Goodman earned his Ph.D. in Philosophy from the University of Michigan. He is an Associate Professor in the Philosophy Department and the Department of Asian and Asian-American Studies at Binghamton University. He is the author of several published articles on Buddhist philosophy, ethics, and applied ethics, and of *Consequences of Compassion: An*

Interpretation and Defense of Buddhist Ethics. His second book, *The Training Anthology of Śāntideva: A Translation of the Śikṣā-samuccaya*, is forthcoming from Oxford University Press.

Christopher W. Gowans is Professor of Philosophy at Fordham University in New York City. He is interested in the philosophical implications of moral diversity and conflict, contemporary normative moral philosophies, and ancient philosophical/spiritual traditions, especially Buddhism. His publications include *Philosophy of the Buddha* and *Buddhist Moral Philosophy: An Introduction*. For more information, go to www.christophergowans.com.

Peter Harvey gained his doctorate in Buddhist Studies at Lancaster University, under Ninian Smart, and is Emeritus Professor of Buddhist Studies at the University of Sunderland, UK. His research focuses on early Buddhist thought and practices, and Buddhist ethics. He edits *Buddhist Studies Review*, the journal of the UK Association for Buddhist Studies, having founded the Association in 1995 with Ian Harris, and from 2002 to 2011 ran an online MA Buddhist Studies program. He is author of *An Introduction to Buddhism: Teachings, History and Practices* (University of Cambridge Press, 1990, 2nd ed., 2013), *The Selfless Mind: Personality, Consciousness and Nirvana in Early Buddhism* (Curzon Press, 1995) and *An Introduction to Buddhist Ethics: Foundations, Values and Practices* (University of Cambridge Press, 2000). He has edited an anthology of Theravāda, Mahāyāna and Vajrayāna texts, *Common Buddhist Text: Guidance and Insight from the Buddha*, published for free distribution by Mahachulalongkornrajavidyalaya University in Thailand. He is a member of the Samatha Trust and teaches a free online course on Samatha meditation.

Emily McRae is Assistant Professor of Philosophy at the University of New Mexico. She has interests in moral psychology, Buddhist ethics, feminism, and social justice. She has published articles on these subjects in *American Philosophical Quarterly*, *History of Philosophy Quarterly*, *Philosophy East and West*, *Journal of Buddhist Ethics*, and *Sophia*.

Karin Meyers is Assistant Professor and Director of the MA in Buddhist Studies at Centre for Buddhist Studies at Rangjung Yeshe Institute and Kathmandu University. She is the author of another article and dissertation on the topic of Buddhist views of free will: “Free Persons, Empty Selves: Freedom and Agency in light of the Two Truths” (2014) and “Freedom and Self-Control: Freedom in South Asian Buddhism” (2010). Her other research interests include Buddhist theories of the path, Abhidharma and Yogācāra thought, and consciousness.

Rick Repetti received his Ph.D. at the CUNY Graduate Center in 2005 on the problem of free will. He has taught at Brooklyn College (CUNY), LaGuardia Community College (CUNY), and Vassar College, and is an Associate Professor of Philosophy at Kingsborough Community College (CUNY), where he teaches Asian Philosophy, Philosophy of Religion, Ancient Philosophy, and Global Ethics, and leads a weekly meditation group for the college community. He is the author of *The Counterfactual Theory of Free Will: A Genuinely Deterministic Form of Soft Determinism* and several articles on Buddhism, meditation, and free will, a yoga and meditation practitioner since 1973, and instructor in both since 2000.

Mark Siderits recently retired from the Philosophy Department of Seoul National University, where he taught analytic Asian philosophy. His research interests lie in the intersection between

classical Indian philosophy, on the one hand, and analytic metaphysics and philosophy of language, on the other. Among his more recent publications are *Buddhism As Philosophy* (Ashgate/Hackett 2007), *Personal Identity and Buddhist Philosophy: Empty Persons*, 2nd ed. (Ashgate, 2015) and, together with Shōryū Katsura, *Nāgārjuna's Middle Way: Mūlamadhyamakakārikā* (Wisdom 2013). He has also edited several collections of work on Indian/analytic philosophy. A collection of his papers on Buddhist philosophy is forthcoming from Oxford.

Galen Strawson is President's Chair in Philosophy, University of Texas at Austin. He taught at the University of Oxford from 1979-2000, where he was a Fellow of Jesus College. He was Professor of Philosophy at the University of Reading, UK, from 2001 to 2013, and Distinguished Professor of Philosophy at the City University of New York Graduate Center from 2004 to 2007. He has held visiting research positions at the Research School of Social Sciences, Australian National University, Canberra; the University of Copenhagen; Princeton University, where he was a Council of the Humanities Old Dominion Fellow; and the Ecole des hautes études en sciences sociales in Paris. He has taught as a Visiting Professor at NYU, Rutgers University, MIT and Princeton. He is the author of *Freedom and Belief*, *The Secret Connexion: Realism, Causation and David Hume*, *Mental Reality*, *Selves: An Essay in Revisionary Metaphysics*, *The Evident Connexion: Mind, Self and David Hume*, and *Locke on Personal Identity*. He is principal author of *Consciousness and its Place in Nature*. A selection of his philosophical papers, *Real Materialism and Other Essays*, was published in 2008 (Oxford).

B. Alan Wallace began his studies of Tibetan Buddhism, language, and culture in 1970 at the University of Göttingen and then continued his studies over the next fourteen years in India, Switzerland, and the U.S. During most of that time he trained as a Tibetan Buddhist monk, ordained by H.H. the Dalai Lama, for whom he has often served as interpreter. After graduating from Amherst College (1987), where he studied physics and philosophy of science, he earned his Ph.D. in religious studies at Stanford (1995). He taught for four years in the Department of Religious Studies at UCSB, and is now founder/president of the Santa Barbara Institute for Consciousness Studies (<http://sbinstitute.com>). He has edited, translated, authored, and contributed to over forty books on Tibetan Buddhism, medicine, language, and culture, and the interface between science and Buddhism, including *Meditations of a Buddhist Skeptic: A Manifesto for the Mind Sciences and Contemplative Practice*, *Mind in the Balance: Meditation in Science, Buddhism, and Christianity*, and *Hidden Dimensions: The Unification of Physics and Consciousness*.

PREFACE

The research project that led to this book began when, at age 15, during my first meditation I had a life-changing out-of-body experience, after which I took up a very serious meditation practice that led to many repeated mystical experiences, a number of which involved extremely complex, highly improbable series of vividly accurate precognitions, each equivalent to hitting the lottery 100 times in a row, some involving others who had the identical precognitive experiences as me. Such experiences challenge our ordinary conception of the relationship between free will, time, and causality: If I have free will, how can the future—which involves choices I, and others, have yet to make—cause my perception of it now? My attempts to understand these puzzles led to my dissertation on free will at the CUNY Graduate Center, but my meditation-based insights were either couched in Western terms or significantly relegated to the footnotes. Since then, most of my research and writing have been explicitly about Buddhism, meditation, free will, and mental freedom, as well as their interrelationships. After writing several articles on the subject, mostly for the *Journal of Buddhist Ethics*, *JBE*'s editor Daniel Cozort suggested I have enough material for a book, and this book project was born, along with a corresponding monograph, in progress. All of this, ironically, reveals ways in which conditions causally influence experiences and choices, but does not seem to determine them.

Though I argue for the philosophical plausibility of certain positions on the free will issue, my personal 'jury is still out' on free will, particularly given what originally prompted my interests and how little anyone has to say about those features of it. Protagoras said "Concerning the gods, I have no means of knowing whether they exist or not or of what sort they may be. Many things

prevent knowledge including the obscurity of the subject and the brevity of human life.” (Diels and Kranz 1985, 80b4) My thinking about free will is slightly more optimistic, but my arguments, both in the Introduction and in my own two articles here, are, so to speak, made *arguendo*.

The issue of *Buddhist* thought about free will—which some of our contributors think involves some kind of category error—further complicates matters, and is in its relative infancy. In the histories of both Western and Asian philosophy, it is often the case that the first few individuals to pose dilemmas, construct arguments, or level counterexamples significantly shape enduring, sometimes paradigm-forming trajectories of dialectic generated therefrom. The positive or negative value of these influences aside, this philosophical domain is in that fertile early period. Perhaps it is with this sensitivity to the potential hijacking of opinion that some of the writers here figuratively beat a loud drum, myself included, but that is because it matters to us. For this reason, I am privileged to contribute to, and more so to be entrusted editing, this collection. For the same reason, I am sometimes less neutral than an editor should be. Regardless of our arguments, however, it is the reader who ultimately must decide what makes the most sense.

Let’s talk first about why free will is an issue of general philosophical interest, that is, independently of such things as precognition and one-way temporal succession (time), then about why Buddhist thought on the subject may be enlightening. Prior to philosophical reflection, even if some elements of our lives seem beyond control, most of us exert significant control over other elements, such as being able to move our eyes, limbs, and, among others, vocal chords, *at will*, satisfying our desires, bringing about changes in the world. Studies suggest

that children age 4 consider free will *being able to do what you want*; around age 6 they understand free will as being able *not* to act on desires (Kushnir, Gopnik, Chemyak, Seiver, and Wellman 2015). As we mature, we come to think free will includes but surpasses these abilities, but specifying what that entails is the challenge of defining free will: What makes some choices and actions (if any) free? *What is* free will? If we cannot be clear on what free will *is*, then how can we know if we *have* any?

Though we know we are shaped by contingent constellations of experience, biology, and apparently random circumstances that influence our character, one key element in our thinking we have free will is the feeling that what we do is *up to us* somehow. Whether performing a mindless action, deliberating consciously about a major decision, or confronted with a dilemma pulling us in opposite directions, it seems, the action we consider free is one we intended: its occurrence didn't *happen* to us: we *caused* it; it would not have occurred had we not *made* it happen. Forces *influence* us, but we seem able to *resist* such influences, as if not subject to physical laws, though our bodies are. How could that be? If causal laws govern our neurons the way they govern electrons and galaxies, where whatever happens is the lawful consequence of prior conditions, then the sense that our doings are *up to us* seems illusory. Conversely, if our choices result from *purely random* processes (in our brains, circumstances, etc.), the neural equivalent of coin tosses, how can we think we *authored* them? Does the person suffering from epilepsy *author* a seizure? For us to truly *be* the way we experience ourselves, must we possess nonphysical minds, souls? Could we be free even if determinism is true and we are entirely law-governed physical beings, akin to complicated wasps, self-driving vehicles, or robots? Must determinism—the scientific view that all events (on our—non-quantum—level) are lawfully

necessitated by prior events—be false? If indeterministic chance infects choice, how can choice be *up to us*? This is the free will puzzle.

Since the latter quarter of the previous century, Buddhist and Western philosophers have begun to explicitly discuss *Buddhism and free will* together. Interestingly, however, this issue has *not* puzzled Buddhist philosophers since the very beginnings of Buddhism, roughly 2,500 years ago, perhaps because the Buddha explicitly rejected what may be understood as *inevitable causation*—by fate, chance, karma, gods, and matter. Could it be that the Buddha’s rejections of inevitabilism sufficed to render the issue closed? If not, why haven’t Buddhists thought about it until recently? Either way, can Buddhist philosophy—with its rich philosophical psychology, philosophies of mind and action, metaphysics, epistemology, phenomenology, logic, philosophy of language, and related tools for understanding *and transforming* the mind and attaining what they take to be the *maximum* of mental freedom, *nirvāṇa*—help us answer these questions? Do Buddhists think we have *something* like free will, as the Buddha’s rejection of inevitabilism seems to suggest? The answers to these questions are not black and white, as this collection is designed to reveal.

Supposing Buddhism has sophisticated explanations about all things necessary to explicate agency, the glaring problem in the core of Buddhist understanding is that the agent/self is a psychological fiction—indeed, the central illusion responsible for all our suffering. How could the agent/self have free will, if there *is no* agent/self? How can the things ‘I’ do be *up to* ‘me’ if ‘my’ sense of ‘myself’ is an illusion? Has the key question that drives this book—about what

Buddhists think about free will—been answered in the previous sentence? If so, the idea that our choices and actions are genuinely brought about by us—that we have free will—is false.

The reader might think the rest is redundant, but may be mistaken, for *at least* three reasons: the Buddha rejected inevitabilism (an opposite of free will), Buddhist philosophers disagree about whether we have free will, and there are many related issues that complicate their disagreement: Buddhist causation (dependent origination, similar to determinism), ultimate versus conventional truths, karma, moral responsibility, personhood, enlightened (agentless) agency, and, *inter alia*, Buddhist meditation and the highly-effective self-regulative powers it cultivates. In the Introduction, I offer initial indications of how some of these subjects connect, why they are philosophically interesting, the Buddhist traditions some contributors reflect in addressing them, points of critical contact with contemporary Western philosophy and science, and some (possibly biased) evaluative comments.

This collection is significantly but not entirely a menu of Buddhist positions on the free will issue, but also an attempt to engage the latest Western philosophical thinking about free will with Buddhist thinking about it, and vice versa. For that reason, some primarily Buddhist contributions lack the textual/historical grounding (in Pāli, Sanskrit, and other Buddhist language translations, terms, and citations) otherwise typical of Buddhist philosophical scholarship in English. Some Buddhist scholars are implicitly addressing the Western philosophical community, informing them of what their Buddhist tradition thinks about free will, or informing other members of their own community about the other tradition's thinking about free will and of how such thought squares with their own on this issue. Some are written in fairly strict accord

with publishing norms of Buddhist scholarship; I have inserted parentheticals or notes explaining most such terms or standards where those authors may have not.

Likewise, some of the Western contributions lack the sort of dialectical formality and traction typical of contemporary Western philosophical publications in which claims are tightly connected with previous arguments, publications, and related discourse conventions, for the simple reason that these philosophers are intending to present simplified, intuitive models of otherwise elaborate Western argument threads to Buddhists, without the usual terminological and logical hieroglyphics. As with some of the Buddhist terms, I have likewise inserted parentheticals or notes explaining some of the Western philosophical jargon. Inevitably, scholars on both sides will find simplifications and generalizations applicable to their own areas of expertise somewhat *wanting*, but putative remedies would swell the collection to unmanageable proportions and render much of the text impenetrable to all but a handful that specialize in Western *and* Buddhist philosophy. For similarly pragmatic reasons, the issues of compatibility between free will and theological philosophies, or with other Asian philosophies, etc., however interesting in their own right, are mostly excluded here.

Most articles here ought to be informative to scholars of both types and to the average intelligent reader or student interested in either free will or Buddhist philosophy. That these intended communications with different imagined audiences are at work, in varying degrees on both sides, ought to become clear when I overview the main focal points of each of the contributed articles in the Introduction. That this is also the first collection of articles on Buddhism and free will justifies presenting it in as broadly accessible a format as possible, for maximum inclusiveness.

This collection presents ten articles written exclusively for this collection, as well as seven writings from seminal thinkers that have already been published but which their authors have modified to better fit this collection, whose inclusion renders the collection comprehensive. To date, there are no other books on Buddhist theories of free will (apart from my own draft monograph in progress, intended to complement this collection). This collection aims to bring Buddhist philosophy more explicitly into the Western philosophical discussion of free will, both to render more perspicuous Buddhist ideas that might shed light on the Western philosophical debate, and to render more perspicuous the many possible positions on the free will debate that are available to Buddhist philosophy.

In my introduction—written from a perspective shaped by my Western analytic philosophical training, my personal interests in Buddhist philosophy, and my multiple-decades meditation practice—I identify most of the main issues and arguments in broad strokes, explaining why they were placed where they appear, and noting some problems that might face some of them. The general sequence of the articles was selected so as to loosely resemble a philosophical dialogue, such that each article somehow either supports the previous or subsequent article, or undermines it, but any of the articles may be read in any sequence. No other organizational structure (e.g., Part I, II, III: Skeptical Theories, Optimistic Theories, etc.; Theravāda, Yogācāra, Madhyamaka, etc.) would cohesively parse the articles into groupings that would make their division into parts fitting, as many of the articles defy those sorts of categorizations. After reading the Introduction, readers may be better equipped to select their own sequences. The collection may be read as an elaborate presentation of most of the extant competing Buddhist theories of free will —

optimistic, skeptical, and metaphilosophical. Together, the collection promises to touch on a multitude of dimensions relevant to the topic.

There are some overlapping threads across the articles, as to be expected, as each article was written to be read independently; thus, many will share foundational or introductory ideas and terms. However, some overlapping happens in some articles' main arguments or conclusions, which illustrates some convergence, if not consensus. Conversely, there is relatively equal divergence in other articles, illustrating that Buddhist philosophical thought on the topic is as complex and dialectically open as Western philosophical thought.

Some remarks about Buddhist texts, Sanskrit and Pāli, and 'Buddhism' are in order. The originally orally recorded writings of Buddhism were first written in Pāli, and most subsequent Buddhist writings were written in Sanskrit (although others were translated into Chinese, Japanese, Tibetan, etc.). If a contributor is quoting a Pāli text or enmeshed in that tradition, then that is the language they typically will use, and so on. Some Sanskrit words are more popular in English, so they are used more universally, such as *karma*, *nirvāṇa*, and *sutra* (instead of the Pāli: *kamma*, *nibbana*, and *sutta*), or sometimes vice versa, such as *vipassanā* and *sati* (instead of *vipaśyanā* and *smṛti*). Authoritative and canonical Buddhist texts are standardly referred to without publication information, the way Plato's *Republic* or Aristotle's *Nicomachean Ethics*, for example, might be, in their first occurrence, and if to be repeated are defined by their first letter or letters (e.g., "R" or "NE"), as well as regarding numeric references to original divisions of sections, verses, etc. Different contributors prefer slightly different annotation methods regarding punctuation with these abbreviated methods; I have done what I can to make them not

only internally consistent, but as consistent with the rest of the volume as possible while honoring their preferences.

Buddhism is not a singular thing, not unlike Western philosophy, theism, or empiricism. Flanagan, for instance, uses the term ‘the Buddhisms’ to acknowledge this while glossing over irrelevant differences. The context ought to make clear the sense intended. There are major divisions within Buddhism, though most forms accept certain core ideas, teachings, principles, figures, and texts. The original texts constitute the Pāli Canon, universally regarded as authoritative within Buddhism; most subsequent Sanskrit (and other language) texts are typically authoritative for later, Mahāyāna schools. When a contributor uses the term ‘Buddhist’ or its derivatives, ideally they are using it in the sense shared by most Buddhists, or, it should be obvious which tradition they are explicitly using as their reference for the term, without needlessly repeating some qualifier. While this may be a nuisance for learned Buddhist scholars, nonetheless they will know immediately if such references are insufficiently detailed, and they will likely also be able to make any necessary mental adjustments relative to the background tradition that informs that contributor. The real worry is that *novices* will erroneously think that *all* Buddhists adhere to a certain idea when only *some* do; for example, even the no-self doctrine is conceived differently by different traditions, and one early school—long without followers, but originally rather widespread—advocated a real-person doctrine (*Pudgalavāda*). But the same worry arises when Western philosophers speak about free will, fatalism, chance, in/determinism, in/compatibilism, and so on. That there are more such problems in collections that include multiple theoretical frameworks than in those restricted to a single tradition ought to be no

surprise, and to a certain extent cannot be avoided. If I have done my work adequately, the extent to which this is tedious is reasonably manageable.

References

Diels, H. and Kranz, W., 1985. *Die Fragmente der Vorsokratiker*. Zurich: Weidmann.

Edelman, S., 2008. *Computing the mind: how the mind really works*. New York: Oxford University Press.

Kushnir, T., Gopnik, A., Chemyak, N., Seiver, E., and Wellman, H.M., 2015. Developing intuitions about free will between ages four and six. *Cognition*, 138, pp. 79-101.

INTRODUCTION

Hermeneutical Koan—What Is the Sound of One Buddhist Theory of Free Will?

Some of the major reasons that motivate philosophers to think that the idea of free will is problematic, as well as why the issue has caught the attention of Buddhist philosophers and scholars, have been explained in some detail already in the Preface, so I will not repeat them here. Many of them will arise in connection with my discussion below of the articles in this collection. Instead, therefore, here I summarize the main arguments set forth by our contributors, I offer some observations about how they relate to the articles immediately preceding or following them, as well as strengths and weaknesses of some of them. As mentioned in the Preface, I have sequenced the articles to loosely resemble a philosophical dialogue, placing each where they afford either supporting or opposing views regarding the immediately adjacent articles.

The earlier articles question whether the very concept of a Buddhist ‘theory’ of free will (“FW”) is coherent, with mixed responses and qualifications, and some that presuppose the falsity of the notion of FW, then some that split Buddhist attitudes across two dimensions (ultimate and conventional, the former of which is negative about FW and the latter positive), culminating with articles that advocate increasingly FW-friendly views.

To ask “What is the sound of one Buddhist theory of free will?” is to implicitly acknowledge that the issue involves a conceptual knot, if not a non-starter. After all, Buddhism has remained ‘mostly’ silent for over two millennia about FW. More problematically, since

Buddhism rejects the reality of the self, isn't the notion of an autonomous agent incoherent?
If not, is, was, can there be, or ought there to be an implicit/explicit Buddhist theory of FW?

Christopher Gowans places the main issue of this text—the meta-question of how to think about the propriety of Buddhist theoretical considerations about FW—into the appropriately problematic but explanatory context of its historical absence: until recently, Buddhism mostly ignored FW, despite vast resources to address it. Gowans argues that the main reason is that Buddhist philosophical analysis is limited by soteriological parameters: whatever promotes enlightenment. Gowans concludes, however, that if Buddhism must pronounce on any theoretical position, it would only be justified as 'skillful means'—the doctrine whereby beliefs, speech, and actions otherwise known by the wise to be avoided may be justified *if they have soteriological utility*, a view shared by Goodman, below—but Buddhism nevertheless would remain silent on the metaphysics.

However, a Buddhist ethical theory might be soteriologically justified, yet Buddhism has none, except recently. Additionally, whatever justifies the many extant Buddhist (often metaphysical) theories of intentionality, phenomenology, and so on, arguably justifies FW theory. Arguably, these theories arose historically against competing views, and are thereby soteriologically warranted. If so, however, in its present historical encounter with West, Buddhism may acquire soteriological warrant for free will theorizing.

Contrary to this soteriologically restrictive view, as well as to some of the increasingly restrictive views to follow his, perhaps casting aside hermeneutical caution, Repetti offers three arguments in support of the idea that there ought to be a Buddhist FW theory. First, the Buddhist path to mental freedom prescribes meditation and related methods for cultivating

virtuoso-level, self-regulation-increasing abilities associated with (naturalistic) FW, e.g., mind-control, volitional/metavolitional regulation, reason-responsiveness, etc. Intuitively, if the virtuoso has greater FW-related skills than the average non-practitioner, the virtuoso has (naturalistic) FW, which increases, paradoxically, proportionate to the decrease in the self-sense, and peaks in total mental freedom, *nirvāṇa*, the cessation of the self-sense. Repetti argues that this skill undermines the most powerful FW skepticism, ‘hard incompatibilism’, a view Goodman advances (immediately below) to the effect that there is no autonomy *regardless* of whether we are determined, because either *we are*, and thus not responsible for our choices or *we’re not*, and thus our choices are random and thus not *up to us*. The meditation virtuoso, however, can escape from previous and present mental state conditioning, irrespective of its causal history. That virtuosos cultivate skills that theoretically defeat the most powerful FW skepticism justifies a Buddhist FW theory.

Second, proper understanding of agency, even if that understanding is negative, say, denying agency altogether, implies a (negative) FW theory, but proper understanding of agency is constitutive of enlightenment. And third, an abundance of Buddhist teachings are multi-leveled and progressive, prescribing a soteriological ‘path’ designed to take individuals from where they are, at a conventional level of understanding that presupposes agency, and to guide them to more subtle, ultimately impersonal levels of understanding. Relative to this progressive path, a theory of agency that capitalizes on the corresponding Buddhist distinction between relative, conventional truth and absolute, ultimate truth, serves a skillful-means-type therapeutic function for those at the conventional level, rather than simply deflating their pragmatically necessary belief in agency and evitabilism by assaulting non-Buddhists or initiates with an unremitting FW skepticism. In a somewhat striking analogy,

Repetti likens this unto the therapeutic use of methadone for heroin-addiction recovery, one point the next author might accept.

Taking a strongly opposing view except on the last point, one of the first Buddhist philosophers to argue forcefully that Buddhism flat-out rejects FW, Charles Goodman has argued that Buddhism is hard incompatibilist (it considers FW impossible whether determinism or indeterminism is true, though he thinks Buddhist causation is deterministic). He repeats that argument here, but now allows that the doctrine of skillful means might sanction Buddhist belief in FW, which Gowans suggested earlier. A step in the right direction, some would think, but Jay Garfield and Owen Flanagan, the next two contributors, would disagree.

Garfield would likely reject Goodman's view, that Buddhism is implicitly hard incompatibilist, for Garfield argues that Buddhism, especially, Madhyamaka (later, 'Middle Way') Buddhism, lacks a FW theory because it lacks a monotheistic theodicy, thus requires no conception of the agent operating outside the causal nexus, needed to relieve God from blame for evil, which could be threatened by a deterministic model of causation. For Garfield, absent the Western theodicy that generated the FW conception, the problem does not and cannot arise in Buddhism: there's no framework in which the discussion may legitimately arise, so no grounds even for a negative theory (like Goodman's).

However, contra Garfield, Mādhyamikas (followers of Madhyamaka) endorse the view that because there is no metaphysical foundation enabling the naïve realist's worldview to be reduced or eliminated, it makes as much sense to say there are tables as to say there are table-like phenomenological appearances that reduce to aggregates of atomistic psychophysical

tropes, which foundational binary is what many earlier Buddhists assert when they bifurcate reality or truth according to the understanding of the unwise and the wise, respectively. Thus, it arguably makes as much sense for Mādhyamikas to say people have FW, something Siderits, who we will get to further below, argues for, yet precisely from the earlier Buddhist reductionist foundationalist perspective Mādhyamikas reject.

Independently, but functioning like a tag team, Flanagan next argues against allowing any inroads into agency-free—in his view, therefore philosophically superior—Buddhism from Western FW conceptions tainted by their genesis within a monotheistic theodicy Buddhism lacks, echoing Garfield's argument that Buddhism has done fine over two millennia without any FW conception, doesn't need one, and is better off without one. What Flanagan adds is an independent argument from elimination in which he offers four hermeneutics for cross-cultural philosophical appropriation, so to speak, rejecting each as invalid for FW/Buddhism exchange. I'm not so sure his own writings elsewhere, about Buddhism naturalized, would make it through his argument from elimination, but I leave that to the reader to assess, as well as whether the other theorists to follow somehow violate his hermeneutic provisos and, if so, whether that invalidates them.

However, suppose *arguendo* that 'theory T' is the best theory on free will: the scientific community asserts T. Suppose Buddhism implies the opposite, $\sim T$, or worse: neither T nor $\sim T$. Is it cultural hubris to ask whether Buddhists ought to believe T? I don't think so. With exceptional exceptions, we should believe the truth. If T is the best we have, we/they ought to accept T. Do they? Can they? Are there resources in Buddhism that could support T? If T is true, and Buddhism implies $\sim T$, does one cease to be a Buddhist if she accepts T? Importantly, we're still disputing what 'FW' *means*; it's more like T1, T2, T3, etc. Buddhism

has interesting things to say about T1, T2, T3, etc., objections notwithstanding. We're dealing with different conceptual frameworks, but I doubt we're dealing with the sort of limitations Quine imposed on his hypothetical 'radical translator', such that he couldn't find out enough about the hypothetical natives' use of the term 'gavagai' to differentiate between their possible meanings of 'rabbit' or 'nondetached rabbit parts' (2013, pp. 23-72). After all, there seem to be what Quine would consider 'native bilinguals' contributing to this debate.

Considering Repetti's methadone/skillful means analogy above, the next two contributors may be said to prefer the metaphorical cold turkey of unremitting FW skepticism: they either outright deny FW or have doubts about its consistency with Buddhist philosophy. Galen Strawson (1986) was one of the first Western philosophers to link the Buddhist view of the unreality of the self with the unreality of FW. Strawson's FW skepticism rests implicitly on his earlier (1994) impossibility argument, which he takes to refute 'strong' FW ("SFW"), the belief that we are the ultimate originators of our choices/actions. Here Strawson focuses on determinism, and how even deterministic skeptics find determinism hard to assimilate into their daily lives. Unlike Peter Strawson, who argued (1962) that we cannot adopt the skeptical perspective in our daily lives because it's too alien to our interpersonal reactive attitudes (e.g., resentment), Galen Strawson thinks Buddhism represents a way of life that embodies that perspective. Reminiscent of Descartes suggesting we imagine an evil demon so as to sustain doubt against and thereby loosen up our habitual credence, Strawson proposes a thought experiment whereby we are to continuously attend to the impersonal deterministic causation of each thought, desire, action, to bring the resilience of our habitual agential thinking to light. When we see how we cannot maintain the perspective, Strawson advises us to take up the Buddhist practice of meditation, thought to reduce the gravitational pull of agential thinking.

Strawson rightly identifies a relationship between the Buddhist denial of the self and FW. However, Strawson's prescription may be premature: prognosis precedes prescription. Before we prescribe FW's post-mortem procedures, so to speak, FW must be dead. This reminds me of a scene from *Monty Python and the Holy Grail* (1975). A cart-man passes through a plagued medieval village collecting corpses while announcing, "bring out your dead!" He is about to prematurely collect someone's infirmed associate, who protests, "I'm not dead". The two negotiate a fee, the cart-man clubs the infirmed man, collects him on the cart, and continues on his way, "bring out your dead!" Oddly enough, this attitude has become quite fashionable among Western analytic philosophers of late (ed. Caruso 2013).

Coincidentally, Susan Blackmore suggests that meditation—precisely what was just prescribed to undo the FW belief—has contributed to her non-agential experience, implicitly confirming Galen Strawson's assertion, disconfirming Peter Strawson's, with her life depicted as a sort of *proof of concept* for the agentless view. Blackmore implies that the more she attends to her experience, the less she experiences agency. She draws the reader into an ongoing phenomenological analysis of 'doing' by attending to her ordinary activities and portrays how it is as difficult to locate any 'doing' as it is to locate the 'self'. Her essay is mostly narrative, a fresh alternative to 'arguments', but with many obvious implications for the arguments at issue here that seem to fall right out of her phenomenological scrutiny.

Or do they? This is a meta-question, at least for the non-Buddhist: meditative awareness resembles phenomenological reduction, as Coseru shows (2012), in which 'conceptual proliferation' is bracketed, but does meditation render putatively-existing agency invisible or, worse, disassemble it—a kind of psychic suicide practice? To return to *Monty Python*, does

the Buddhist version of the not-dead-yet person hit herself in the head with the metaphorical Zen bamboo staff, rendering her ‘self’ actually fit for the death-cart?

A Buddhist-doctrine-based objection (that might be an *ad hominem*) could be: that’s just clinging to ego. Perhaps, but the tables may be turned on this objection by noting that many meditation virtuosos attest, and there is much in Buddhism to support their claim, that meditation-theoretic self-regulative skills *increase* with practice, e.g., greater distance between impulse and action and, among others, ability to disapprove of, detach from, and thereby diminish the force of unwholesome volitional elements. These things arguably illustrate a psychologically functional, naturalistic ‘agent-lite’ or ‘self-lite’, not the *ātman* (immaterial self/soul) that the Buddha rejected. The agent/self-lite’s self-regulative abilities may be considered *weak* FW (“WFW”). In Buddhist terms, the virtuoso’s *conventional* self becomes *more* functional, not less. Arguably, then, like Blackmore, the virtuoso’s belief in *SFW* diminishes, but, unlike Blackmore, her *WFW* increases. Or so would argue Harvey, Meyers, and Repetti, further below.

True, SFW has bigger problems and presumably no place in Buddhism, but agency-lite threatens moral-responsibility-lite, a standard objection to WFW theories: they are too thin to support anything stronger than utilitarian or pragmatic justifications for moral responsibility, since, if there is no real self, nobody is truly responsible, thus nor does anyone *deserve* praise or blame. Christian Coseru, for instance, immediately below, and Ben Abelson, further down, find problems with WFW’s ability to carry the load imposed on it by the requirements of a genuinely justified—that is, desert-based, as opposed to a merely utilitarian—moral responsibility.

Coseru thinks Buddhism's impersonal causal model compromises the notion of responsibility that requires a more robust FW that demands that a moral agent participate in interpersonal relationships and act in self-regulating ways relative to norms and reasons. Coseru argues that the compatibilist idea that we can dispense with SFW, in favor of a weaker notion of responsibility informed by cognitive science, compromises the notion of moral agency by eliminating responsibility from the purview of the truly enlightened, the Buddhist saint. Implicitly relying on his (2012) work with Buddhist phenomenology and the first-person (rather than the impersonal) perspective, Coseru offers interesting alternative ways of conceiving Buddhist ethics and agency, and explores an alternative view in which genuine responsibility is an ineliminable feature of moral agency. Whether and to what extent this line of thinking might be used to develop a Buddhism-friendly SFW I leave to the reader to discern.

Possibly indirectly supporting some of Coseru's ideas about enlightened agency, Marie Friquegnon argues that there are three distinct understandings of agency and freedom in various forms of Buddhism. First, all Buddhists understand agency as unconstrained by divine power or material causality, a point emphasized by subsequent contributors (Wallace and Repetti). Second, all Buddhists see unethical actions as the direct result of mental states (ultimately impersonally, deterministically) governed by anger/hatred, jealousy/attachment, and ignorance/fear. For instance, as Goodman (above) and Siderits (below) each emphasizes, Śāntideva asserts that we can no more blame someone—under the impersonally caused influence of such mental states—than we can blame fire for causing smoke or the liver for producing bile. And third, in all later (Mahāyāna) Buddhism this determinist attitude does not apply to 'selfless' actions flowing from enlightened nature, what Repetti describes as 'agentless agency'.

Friquegnon outlines the grounds for these divergent views of freedom, explores their implications within Buddhist thought, and articulates contributions they can make to the discipline of philosophy. We will see shortly some bifurcations and trifurcations that her distinctions have presaged.

Another one of the first luminaries to call attention to this aspect of Buddhist thinking about freedom was B. Alan Wallace. Wallace reviews how the Buddha rejected then-prevalent inevitabilist views about fatalism and chance that resemble determinism and indeterminism, respectively. Subsequent to this early Buddhist rejection of the subject, Wallace sees the Buddhist tradition taking a pragmatic turn, he explores ways we can acquire greater freedom to make choices conducive to well being, and highlights soteriological practices of Mahāyāna Buddhism that point toward mental freedom. One is the cultivation of mind-control, the ability to deliberately focus attention with continuity and clarity; another is the cultivation of insight into how attitudes shape experience, allowing for the possibility of altering not only the way we experience, but how we are influenced by memory.

Wallace's pragmatism rests on a liberating form of Mahāyāna metaphysics that emphasizes the 'pristine awareness' or 'substrate' dimension of consciousness transcending conceptualization and the causal nexus—and its determinism/indeterminism dichotomy. As alluring as Wallace's dichotomy-transcending perspective is, and despite that it arguably supports Repetti's 'soft compatibilism' (the opposite of hard incompatibilism), *infra*, Wallace's interpretation of the substrate consciousness is disputed even within Tibetan Buddhism, and his model resembles SFW (possibly stringer than Coseru's implied genuine agency) as a causality-

transcendent consciousness in which free actions may be initiated. Wallace's transcendental metaphysics aside, his pragmatic insight seems plausible: Buddhist soteriological practices *at least* generate a soteriologically pragmatic form of WFW, if not a SFW as well, regardless of the nature of causality.

These views are getting complex, further evidencing the difficulty of grouping them. Thus, it might help to consider a Western philosophical term that applies to some of them, a term no Buddhist FW theorists, apart from Repetti (2012, 2015) and Harvey (briefly, this volume), have entertained, 'semi-compatibilism'. In most Western usage, 'semi-compatibilist' denotes incompatibility between determinism and SFW, but compatibility between determinism and moral responsibility (or WFW in the moral-responsibility-entailing sense). Fischer (2006) argues that SFW presupposes ability to do otherwise *under identical causal circumstances*, implying indeterminism—obviously incompatible with determinism. However, Fischer argues that moral responsibility is compatible with determinism, as Frankfurt (1969) famously argued: an agent can voluntarily do X even if determined to, so long as the agent would have done X even if she could have done otherwise—and, Fischer adds, so long as she was 'moral-reasons-responsive' (able to respond to moral reasons for doing or not doing X), she is morally responsible for doing X. Prior to Frankfurt (and Fischer's emendations), most philosophers took responsibility and FW to be inseparable. Thus, some philosophers below are semi-compatibilists in this *narrow* (Fischer's) sense, though seemingly unaware of the term. However, some may be semi-compatible in a *broader* sense of thinking Buddhist (impersonal) causation—as opposed to *determinism*—is compatible with some sense of agency and/or moral responsibility. Most readings to follow are semi-compatibilist in both senses.

Developing his earlier view, Martin T. Adam concurs with an implicitly semi-compatibilist line of thought, based on an analysis of the no-self doctrine and of Frankfurt's view. Adam argues that the Buddha's views and those in the Pāli sutras (discourses of the Buddha) regarding the 'no-self' doctrine are patently incompatible with SFW, but compatible with the sort of WFW and moral responsibility connected with the possibility of *spiritual* freedom, a point most of the subsequent contributors would likely accept. Relying on distinctions between freedom of the person, of the will, and of action he attributes to Frankfurt, Adam argues that Buddhist freedom admits of degrees relative to the individual's spiritual development.

My only uncertainty here concerns Adam's attributing his triple distinction to Frankfurt, as it is not clear Frankfurt identifies a 'freedom of the person' in Adam's terms, and this might pose a problem for his overall argument. However, I will not develop that objection further here.

Mark Siderits is one of the earliest, seminal, and lasting voices in the contemporary dialectic on Buddhist views of FW. Siderits advances a uniquely semi-compatibilist view, in the broad sense, based on the Buddhist doctrine of 'two truths', the ultimate and conventional. Simplifying, the conventional is analogous to the naïve realist worldview and the ultimate is analogous to the scientific worldview, but in Buddhism the ultimate is what is real to the enlightened view. Conventionally, there are *whole* persons, but ultimately (within Abhidharma reductionism, which he focuses on) there are only deterministic atomistic psychophysical *parts* (tropes). Siderits argues that, ultimately, determinism applies and there are no person-level wholes that could have FW, but,

conventionally, there are persons who sometimes exhibit strong agency, SFW—unlike most other Buddhist compatibilists, who think conventional FW is WFW. To this he adds an argument for semantic insulation between the two discourses, violations of which threaten incoherence, from which it follows that impersonal-parts-determinism-discourse cannot have any consequences for whole-person-agency-discourse whatsoever, thus there can be no incompatibilism.

However, nowadays science-talk and person-talk increasingly mesh, particularly in the West, so it might be more parsimonious to say we have WFW in a way that is science-friendly, but it is questionable whether that can coherently be expressed on Siderits's semantic insulationism. Additionally, insofar as disagreement is impossible for two people (languages/discourses) who cannot communicate or compete for a zero-sum truth-value, it is not clear Siderits's 'paleocompatibilism' constitutes a genuine 'compatibilism', whereas the traditional in/compatibilism debate is about two who can communicate/compete.

Ben Abelson considers a different challenge to Siderits that resembles Coseru's concerns about tension between Buddhist metaphysics and ethics. On Siderits's reductionism, persons—ultimately unreal, but which reduce to impersonal psychophysical processes—have conventional existence because grouping some such parts as wholes rather than others has utility. Buddhist reductionists are committed to an 'Impersonal Description' (ID) thesis: a complete description of reality need not denote persons. Siderits defends against the charge that the ID thesis implies the 'extreme claim' that central features of our person-regarding practices cannot be

rationaly justified, such as interest in one's survival, concern for one's future, holding people responsible, and compensation for past burdens.

Abelson focuses on Siderits's defense against the responsibility part of the objection, which appeals to 'shifting coalitions' of self-revision processes in an individual which could provide enough agency to ground responsibility without violating elements of the no-self doctrine. Abelson argues that while the 'shifting coalitions' idea successfully disarms part of the objection, it cannot account for the robust responsibility Siderits wants, though it may ground a more modest—I think, semi-compatible—responsibility still stronger than the sort Siderits, like Coseru, dismisses as too weak.

Somewhat sharpening his earlier—implicitly *broadly* semi-compatibilist—view, Peter Harvey claims here the Buddhist FW problem concerns whether its impersonal conception of the person is compatible with moral-responsibility-entailing agency, the issue Coseru problematized. Restricting his analysis to Theravāda (the oldest surviving form of early Buddhism), Harvey concludes Theravāda is semi-compatibilist, a 'middle way' between seeing a person's actions as so impersonally conditioned as to lack the sort of proximal agency or WFW ordinarily understood as minimally required for moral responsibility and seeing the person as a strong self with SFW. Harvey adds that the Theravāda view identifies various factors that increase such agency, a point developed by the following contributors. Recalling *Monty Python*, Harvey could provide a different death-cart prognosis: "There's a cure for this man".

Also appealing to a Mahāyāna perspective focused on Buddhist practices designed to regulate something typically thought in the West to be inaccessible to the will—namely,

emotion—Emily McRae explores how we can exercise choice regarding emotional experiences and dispositions. Drawing on mind-training practices advocated by Tsongkhapa, McRae argues that Tsongkhapa’s analysis shows successful intervention in negative emotional experiences depends on four factors: intensity of the emotional experience, ability to pay attention to the workings of one’s mind and body, knowledge of intervention practices, and insight into the nature of emotions. Echoing Friquegnon’s observations about Śāntideva, McRae argues that this makes sense of Tsongkhapa’s seemingly contradictory claims that the meditator can and should control (and eventually abandon) her anger and desire to harm others, and that harm-doers are ‘servants to their afflictions’. McRae concludes with some (I think, semi-compatibilist) implications of Tsongkhapa’s account of choice in emotional life for the place of FW in Buddhism.

In a similar effort, Karin Meyers argues for a somewhat more nuanced implicitly semi-compatibilist view, an agency that is not so ‘lite’, grounded earlier in the Abhidharma, in Vasubandhu’s theories of karma, causation and liberation. Meyers argues that, though they differ from modern positions on FW (and the views of other Buddhists), Vasubandhu describes an understanding of the powers of mind, agency, and action—particularly those able to be cultivated to virtuoso levels of mind-control far beyond those typically associated even with SFW—that is nonetheless compatible with causation. On her analysis, everything is caused (perhaps not explicitly deterministically), and mental qualities explain FW and moral responsibility.

In the final selection, Repetti claims that Buddhists may accept semi-compatibilism, but they can also assert a broader ‘soft compatibilism’ based on the Buddha’s rejection of

inevitabilism (noted by Wallace and Harvey), and on a strong form of evitabilist agency noted by Meyers concerning the meditation virtuoso's ability to attain 'mind mastery'. Hard incompatibilism, recall, asserts that FW is incompatible with both determinism and indeterminism, but soft compatibilism is the equally comprehensive negation of that view: FW is compatible with both. The meditation virtuoso, as Repetti argued earlier, exhibits mental freedom regardless of the causal nature and genesis of her mental states. From a Buddhist perspective, to the extent ordinary folks unreflectively satisfy desires, they tend to decrease mental freedom, because doing as one pleases fortifies the chief culprit in our suffering, the false sense of self. Conversely, as we increase mental freedom from the ego-volitional complex, we increase self-regulative ability, subsequently exercising will *less* in the service of the ego-complex. Thus, the closer one gets to mental freedom, the greater one's self-regulative abilities.

But—and here's the paradoxical rub—as one attains the limit condition of mental freedom (*nirvāṇa*), one reaches maximum self-regulative ability, but there is no longer any sense of ego-volitional-self in need of regulation. Thus, the maximum of mental freedom and self-regulation (WFW) coincides with the absence of any sense of ego/self, a kind of *agentless agency* noted in connection with Friquegnon's third kind of freedom: a form of reason-responsiveness that is entirely dharmic (in accord with *Dharma*), or '*Dharma responsiveness*'. Reason-responsiveness is the central criterion in semi-compatibilist accounts; *Dharma-responsiveness* thus grounds a Buddhist form of semi-compatibilism. As Buddhist practitioners become increasingly dharmic (through soteriological practice), they not only increasingly approximate (or, on some views, instantiate) *nirvāṇa*, they increasingly approximate/instantiate soft compatibilist WFW,

agentless agency. Whether Repetti's account makes sense, I must leave to the reader to 'decide'.

References

- Caruso, G.D., ed., 2013. *Exploring the illusion of free will and moral responsibility*. Lanham, MD: Lexington Books.
- Coseru, C., 2012. *Perceiving reality: consciousness, intentionality, and cognition in Buddhist philosophy*. New York: Oxford University Press.
- Fischer, J.M., 2006. *My way: essays on moral responsibility*. New York: Oxford University Press.
- Frankfurt, H., 1969. Alternate possibilities and moral responsibility. *Journal of Philosophy*, 66(23), pp. 829-39.
- Monty Python and the Holy Grail*, 1975. [Film] Directed by T. Gilliam and T. Jones. UK: Python (Monty) Pictures.
- Quine, W.V.O., 2013. *Word and Object*. New ed. Cambridge, MA: MIT Press.
- Repetti, R., 2012. Buddhist reductionism and free will: paleo-compatibilism. *Journal of Buddhist Ethics*, 19, pp. 33-95.
- Repetti, R., 2015. Buddhist meditation and the possibility of free will. *Science, Religion & Culture*, 2(2), pp. 81-98.

Strawson, G., 1986. *Freedom and belief*. New York: Oxford University Press.

Strawson, G., 1994. The impossibility of moral responsibility. *Philosophical Studies*, 75(1-2), pp. 5-24.

Strawson, P.F., 1962. Freedom and resentment. *Proceedings of the British Academy*, 48, pp. 1-25.

1 Why the Buddha Did Not Discuss ‘the Problem of Free Will and Determinism’

Christopher W. Gowans

It is widely recognized that the problem of free will and determinism, a standard topic in contemporary Western philosophy, was not addressed in traditional Buddhist thought. In this chapter, I propose some reasons for this lacuna in the teaching of the Buddha as represented in the four primary *Nikāyas* (baskets or collections) of the Pāli Canon. The heart of the reason is the Buddha’s practical orientation: he said he would only discuss topics that directly pertain to overcoming suffering, and little in his teaching about this directly brought the problem into focus.

The Practical Nature of the Buddha’s Teaching

The place to begin is “The Shorter Discourse to Mālunkyāputta” (*Cūḷamālunka Sutta*) (*Majjhima Nikāya* (“MN”) I 426-32; references to MN are to Ñāṇamoli & Bodhi (1995)). In this well-known text, Mālunkyāputta observes with disapproval that the Buddha has not taken a position on ten ‘speculative views’. These pertain to whether the world is eternal or infinite, the relationship of body and soul, and the postmortem existence of the Buddha. Mālunkyāputta says that he will continue as a disciple of the Buddha if, and only if, the

Buddha takes a position on these views or at least declares that he does not know what to say.

In response, the Buddha says that he never claimed he would take a position on these views. By way of explanation, he offers a famous simile. Suppose a man wounded by an arrow covered with poison was brought to a doctor. The man then said he would not allow the doctor to treat the wound until he knew various facts about the attack, such as the name of the person who wounded him, his height, where he lived, the kind of bow and shaft he used, etc. It was observed that the man might well die before coming to know all these facts.

The immediate point of the simile is evident: Mālunkyāputta's requirement that he will follow the Buddha only if he first takes a position on the ten speculative views is similar to the wounded man's requirement that he will allow the doctor to treat him only if he first knows the facts about the attack. In both cases, the requirement is unreasonable because the knowledge demanded is not needed to address the practical issue at hand. In order to survive the attack, the wounded man needs proper medical care, and knowledge of the name of the person who attacked him and many other pieces of information about the attack are not relevant to this. Likewise, the Buddha says, in order to overcome suffering Mālunkyāputta needs to know the Four Noble Truths, and understanding the speculative views is not relevant to this. Specifically, the Buddha says that he has declared the nature of suffering, its origin, its cessation, and the way leading to its cessation, since understanding

these leads to enlightenment and *Nibbāna* (Pāli; Sanskrit: *nirvāṇa*). And he has not declared a position on the speculative views because understanding these does not have this benefit.

There are several texts in which the Buddha declines to answer questions that are put to him. The point of these texts varies. Often the point is that the questions have a false presupposition, and sometimes it is that any answer would be misunderstood by the questioner. There are other texts in which the ten speculative views, among others, are discussed. Once again, the point of these texts varies. In one prominent case it is to warn against the dangers of contact and craving (see *Dīgha Nikāya* (“DN”) I 41-45; references to DN are to Walshe (1987)). These diverse purposes often complement one another and need not be seen as signs of conflict.

The Mālunkyāputta discourse has been interpreted in different ways. But the most straightforward reading is that the Buddha’s sole concern is to enable us to overcome suffering and so he will only discuss matters that pertain to this preeminently important practical concern. This does not mean he will never discuss any ‘speculative views’. It means that he would do so only if this would help us to overcome suffering. After all, a dominant theme in the *Nikāyas* is the importance of having ‘right’ rather than ‘wrong’ views (views that reduce or promote suffering, respectively). Nonetheless, one message of this text is clear: don’t be distracted by matters irrelevant to the urgent practical task at hand.

Another important feature of the Mālunkyāputta discourse is that the Buddha’s teaching is compared to the medical practice of a doctor (see Gowans 2010). Just as a doctor heals physical ills of the body, so, too, the Buddha heals the broader ills encompassed by the term ‘suffering’ (*dukkha*). This medical analogy is reinforced by other texts in which *Nibbāna* is compared to a state of health and the Buddha is compared to a doctor (MN I 510-12, II 260). Texts such as these are in the background of Buddhaghosa’s (1999) use of a medical simile to interpret the Four Noble Truths:

“the truth of suffering is like a disease; the truth of the origin is like the cause of the disease, the truth of cessation is like the cure of the disease, and the truth of the path is like the medicine” (*Visuddhimagga* XVI 87).

The implication of the medical analogy is that the Buddha’s teaching, the *Dhamma* (Pāli; Sanskrit: *Dharma*), should be regarded as a kind of practical or craft knowledge similar to medicine. From this perspective, his teaching has two important features: it has a specific practical goal (overcoming suffering), and it is based on purported knowledge that enables us to achieve this goal. The teaching is first and foremost knowledge of how to do something: it is a program of spiritual exercises directed to the attainment of enlightenment. But the program is based on an understanding of the nature of the world, especially human nature. The Mālunkyāputta text tells us that the Buddha was only interested in discussing issues that were pertinent to his teaching so understood.

My suggestion is that these considerations provide a helpful framework for thinking about the Buddha's teaching in relationship to the contemporary problem of freedom and determinism. In Western philosophy, this problem usually arises because it is thought that there is at least *prima facie* reason to believe that human beings have free will, that everything that happens is causally determined, and that these contentions conflict with one another. Beliefs about free will are usually supported by intuitions people have about making choices and assumptions they make about the presuppositions of moral responsibility. Beliefs about determinism are usually maintained by reference to the scientific understanding of the world (at least above the quantum level) and sometimes by appeal to theological beliefs about the power of God. Determinism is thought to imply that whatever happens must happen in the way it does because it is causally determined by scientific laws or God. The concern about conflict stems from the thought that an exercise of free will, a free choice, is at least in some respect not causally determined in this sense.

Is there any reason to think that the Buddha would or should have shared these or similar concerns? It is worth reminding ourselves that the Buddha's enlightenment was not based primarily on the development of a philosophical theory. In a canonical account of his enlightenment, the Buddha first attains the four *jhānas* (states of deep meditative attainment), thereby passing through but stilling "sustained thought", and then reaching "neither-pain-nor-pleasure and purity of mindfulness due to equanimity" (MN I 21-24; cf. MN I 247-49). In this purified state, he achieves three forms of "true knowledge": that of his past lives, the operation of karma and rebirth, and the Four Noble Truths. It was through a form of meditation, then, that the Buddha saw the way to overcome suffering. His

enlightenment was based more on a kind of observation than on rational analysis. And he said that his teaching was “unattainable by mere reasoning” (MN I 167).

Philosophy enters, to the extent that it does, in the Buddha’s articulation of what he observed and especially his willingness to discuss challenges to the coherence of his teaching and its consistency with commonly held beliefs or prominent philosophical views of other traditions. This willingness is constrained by the pragmatic orientation of the Mālunkyāputta text, but the import of the discussions in the *Nikāyas* is a substantial account of human beings, our susceptibility to suffering and our ability to overcome it through enlightenment.

Dependent Origination

The heart of this account is a causal understanding of human life. The doctrines of karma and rebirth as well as the Four Noble Truths are crucially based on observations of causal regularities. The Buddha called these phenomena ‘dependent origination’ (*paṭicca samuppāda*: Pāli; Sanskrit: *pratītyasamutpāda*). In a widely quoted brief formulation of dependent origination, he says:

“When this exists, that comes to be; with the arising of this, that arises. When this does not exist, that does not come to be; with the cessation of this, that ceases.” (MN

II 32; cf. *Samyutta Nikāya* (“SN”) II 28, 65, 70, 78, and 95-6; references to SN are to Bodhi (2000))

This idea is the beginning, and in many ways the end, of the Buddha’s teaching: “one who sees dependent origination sees the *Dhamma*; one who sees the *Dhamma* sees dependent origination” (MN I 190-91).

Dependent origination is sometimes interpreted as a form of causal determinism. If this were correct, the Buddha would have been committed to one important part of what gives rise to the contemporary problem of free will and determinism. However, this interpretation attributes far more to the Buddha than we are encouraged to suppose in the *Nikāyas*. The brief formulation is a summary of a longer statement that elaborates eleven factors that explain the arising and cessation of suffering. The arising part of the statement says:

“That is, with ignorance as condition, formations [come to be]; with formations as condition, consciousness; with consciousness as condition, mentality-materiality; with mentality-materiality as condition, the sixfold base; with the sixfold base as condition, contact; with contact as condition, feeling; with feeling as condition, craving; with craving as condition, clinging; with clinging as condition, being; with being as condition, birth; with birth as condition, ageing and death, sorrow, lamentation, pain, grief, and despair come to be. Such is the origin of this whole mass of suffering.” (MN III 63-4)

This statement is immediately followed by the cessation part. This begins by stating, “but with the remainderless fading away and cessation of ignorance comes the cessation of formations”, and it concludes by saying “such is the cessation of this whole mass of suffering” (MN III 64). Since the Second Noble Truth says that craving is the source of suffering, and the Third Noble Truth refers to the cessation of suffering, it is evident that these longer formulations of dependent origination are expansions of the basic ideas in standard formulations of the Second and Third Noble Truths (see *Aṅguttara Nikāya* (“AN”) I 176-7 for this connection; references to AN are to Bodhi (2012)).

The longer formulation has been a source of perplexity both because of what is and is not included on the list and because of its apparently rigid linear sequence. A traditional interpretation based on a literal reading understands it as referring to a succession of three lives. However, though the longer formulation appears frequently in the texts (see especially “The Book of Causation” in SN), perhaps we should not be so preoccupied with a literal reading of it. There are variations of the longer formula that refer to other factors (for example, see SN II 29-32), and in general in the Buddha’s teaching many items beyond the eleven factors in the longer formulation are referred to as pertaining to the arising and cessation of suffering.

In any case, these interpretive questions should not obscure the fact that the brief formulation is a summary device for referring to the Buddha’s rather extensive analysis of why suffering arises and how we can overcome it. It is important to recognize that the

Buddha's teaching concerning dependent origination has two parts. The first part says: these are the causal conditions of suffering. Many of these conditions involve deeply ingrained habits that govern the lives of most ordinary adult human beings (ignorance, formations, craving, clinging, etc.). But the second part says in effect: you can use knowledge of these conditions as part of a training program to change your habits so as to reduce and eventually eliminate suffering. This is the heart of the Buddha's practical teaching: knowledge of human nature that is put to practical use.

The doctrine of karma, the other side of the Buddha's enlightenment experience, also conforms to this model. In the cycle of rebirth, morally good actions bring about greater well being and morally bad actions bring about lesser well being. But we can make use of these causal conditions to improve our lives: by living better morally we can increase our well being in the future. Dependent origination is central to both *saṃsāra* (the delusional reincarnational cycle) and liberation from *saṃsāra*.

In order to understand the Buddha's teaching in its proper context, as a form of practical knowledge, it will be helpful to consider a comparison from agriculture (a frequent source of similes in the Buddha's thought, e.g. SN I 227). Suppose a group of farmers often complained that their crops were doing quite badly. In response, a wise farmer made a series of careful observations about the circumstances in which crops did and did not do well. He then reported back by saying that poor crops are the result of planting seeds at the wrong time, inadequate nutrition in the soil, improper water, heat and light, damage from animals, insects and disease, and harvesting at the wrong time. Conversely, good crops will

result from reversing these conditions: planting at the right time, providing adequate nutrition in the soil, etc. The wise farmer claims to provide some valuable practical knowledge, and the main test of whether or not he has done so is basically practical—whether or not reversing the conditions brings about good crops.

Someone might say that the wise farmer's claims presuppose the philosophical theory of causal determinism, noting that he had summed up his account with the brief formulation of dependent origination, and then demand that he defend this philosophical theory. In response, the farmer might well say, I am showing you how to have good crops, and I don't need to defend that theory in order for you to grow better crops. Just focus on improving your soil, planting at the right time, etc. If this improves your crops, that is all that matters. For the purpose of improving our crops, we don't need to resolve philosophical questions about causality, for example, about whether Hume or Kant or someone else had it right. Surely this would be a proper response.

My suggestion is that we regard the Buddha's teaching about dependent origination in a similar way. He had observed some causal regularities, both about the relationship between the moral quality of our actions and our future well being (karma), and about the steps needed to achieve enlightenment and overcome suffering (the Four Noble Truths, especially the Eightfold Path). It is important that he understood the regularities correctly, and the Buddha was quite prepared to discuss this. But all that really mattered is that the teaching was effective: that we could improve our well being by being more virtuous, and

that we could overcome suffering by following the Eightfold Path (and related practical advice).

If some Mālunkyāputta figure demanded that the Buddha take a position on whether or not dependent origination was a form of causal determinism, he might well have said that we do not need to resolve this question in order to follow and benefit from his teaching. Just as the wise farmer's advice could be put to good use without resolving these philosophical perplexities, so too the Buddha's teaching could have valuable results without engaging these philosophical concerns. This does not mean that dependent origination, as the Buddha understood it, does not raise philosophical questions. It and the wise farmer's teaching might well do so. However, understood as forms of practical or craft knowledge, pursuing such philosophical questions is a distraction from the main point: it will not help, but will simply delay, overcoming suffering and improving our crops.

It might seem that the brief formula of dependent origination and other texts cannot be read otherwise than as implying strict causal determinism. But they have been read in other ways. Moreover, if we think about the observations that are likely to have been behind these texts we can see that this reading is not required. In contemporary psychology, experimental work establishes statistical correlations between observed factors that are expressed in terms of a correlation coefficient (a number between -1 and 1, indicating negative and positive correlations). Additional evidence is often thought to support judgments about direction of causality. It is reasonable to think that the texts about what contributes to suffering and overcoming suffering were based on observations by the

Buddha of what he took to be significant correlations indicating causality similar to those in contemporary psychology. For example, the Buddha saw that craving increases suffering, and that reducing craving decreases suffering.

In fact, judging from the texts we have, with their multiple and partially overlapping lists, he discovered a complex array of relevant phenomena. For example, he said that there are five hindrances to enlightenment (sensual desire, ill will, sloth and torpor, restlessness and remorse, and doubt) and seven factors of enlightenment (mindfulness, investigation-of-states, energy, rapture, tranquility, concentration and equanimity) (MN I 60-62, III 85-8). Such phenomena are often presented in a standard order in the texts, and it is sometimes suggested that there is a linear causal sequence, with one factor causing another, which then causes another, etc. (as in the longer formulation of dependent origination). But such ordering is more likely to reflect rhetorical strategies of presentation of relevant factors in an oral teaching than literal reports of observations of strict causal sequences.

None of this shows that the Buddha rejected the philosophical theory of determinism. Complex statistical correlations by themselves neither undermine nor support philosophical positions about determinism. But recognizing that observation of such correlations must have played an important role in the development of the Buddha's teaching helps us to see why the brief formulation of dependent origination need not be read as an endorsement of determinism. It is a quick summary device reminding us that there are causes of suffering and corresponding ways of intervening to overcome suffering (the arising and cessation parts of the formula), as the Buddha has explained in detail in

various places. The Buddha's teaching was primarily a practical instruction, based on observation, about the origin of suffering and how to overcome it. "Take these steps," the Buddha was saying, "and they will enable you to attain enlightenment and end suffering". This did not require a philosophical theory of causal determinism any more than the wise farmer's advice on growing good crops required such a theory.

'It Is Up To You'

Several passages sometimes thought to be in the neighborhood of the problem of free will and determinism offer support for this interpretation. For example, the Buddha rejects the view of Makkhali Gosāla, that the defilement and purification of beings have "no cause or condition" because, according to this view:

"There is no power, no energy, no manly strength, no manly endurance. All beings, all living things, all creatures, all souls are without mastery, power and energy; moulded by destiny, circumstance, and nature, they experience pleasure and pain in the six classes." (MN I 407; see also DN I 53-4)

The point here is that what we do and subsequently experience as pleasant or painful (on account of karma) are not the result of destiny, circumstance and nature. Our own mastery, power, energy, etc. also play a role. Elsewhere the Buddha denies that everything we experience is caused by what we have done in the past or by God's creative activity (AN I

173-5). The problem with these views, the Buddha says, is that those who hold them “have no desire [to do] what should be done and [to avoid doing] what should not be done, nor do they make an effort in this respect” (AN I 174-5). If we think that the quality of our future experiences depends entirely on what we have already done, or on God, then we will have no desire now to act so as to affect these experiences, as the doctrine of karma explains.

The thought that animates these passages is that the belief that it is not now up to me (because of destiny, past actions or God) how to act so as to have more or less positive future experiences would undermine my ability and motivation (power, energy, desire, etc.) to act well so as to increase pleasure and decrease pain. This pragmatic point does not mean that there are no other reasons for rejecting these views. But it makes evident the Buddha’s concern to address issues that could destroy our confidence that we can change our lives. If it were suggested that the Buddha’s own understanding of dependent origination would have a similar effect (as is sometimes said about theories of determinism), then it is likely that he would have said the same thing: the arising part of dependent origination should not be understood as undermining our ability and motivation to act on the cessation part.

Recall that the second part of the longer formulation of dependent origination has the implicit message, to each of us: it is up to us to overcome suffering by undoing these conditions, for example, by following the Eightfold Path. Many expressions of the Buddha’s teaching about karma and attaining enlightenment make this kind of point explicitly. Of course, that your actions are *up to you* is often taken to be a central intuition in support of

free will, and it is frequently said in contemporary discussions that the Buddha implicitly accepted some notion of free will. But what exactly do people mean by saying this?

The Buddha thought that human beings were capable of reflecting on and understanding his teaching, accepting it on this basis, forming intentions to carry it out, and acting on these intentions. This was the basic presupposition of the second part of the longer formulation of dependent origination. When it is said that the Buddha accepted a notion of free will, perhaps this is all that is meant. Roughly speaking, this is the kind of thing many compatibilists say when they endorse a notion of free will: human beings ordinarily can reflect, form intentions on this basis, and then act on these intentions, etc. If this were the only sense in which it was said that the Buddha accepted free will, there would be no objection to the claim.

However, to say this is not to ascribe to him a libertarian conception of free will according to which exercises of free will are in some sense uncaused. Nothing the Buddha says directly endorses such a conception. This does not mean that the Buddha was committed to a form of compatibilism. Rather, he saw no reason to think that the two sides of his teaching (why we suffer and what we can do about it) might be incompatible. None of the elements that give rise to the contemporary free will problem was clearly in view. So, no solution to the problem, such as compatibilism, was required.

Libertarians argue that the compatibilist understanding of free will is inadequate because it cannot make sense of the 'it is up to you' intuition (among other reasons). If it were thought

that this is correct, and that this intuition requires uncaused actions, then a question might arise about how the Buddha's assumption that it is up to us to follow the Eightfold Path comports with the arising part of dependent origination. But this was not recognized in the *Nikāyas*: libertarian free will was no more on the scene than was the Newtonian determinism that concerned Kant. And the pragmatic orientation of the Mālunkyāputta text suggests that the Buddha would have resisted being drawn into any such debate. Even if there were an important philosophical issue on the horizon, it is not clear that resolving it would facilitate overcoming suffering any more than doing so would help us grow better crops.

The No-Self Teaching

It might be thought that a problem with this interpretation arises from the Buddha's teaching about ignorance and wisdom. Recall that ignorance and the cessation of ignorance are at the beginning of the longer statements of dependent origination. For the Buddha, the heart of our ignorance—more accurately, delusion—is the mistaken belief that we are selves, and the key to enlightenment is the realization that we are not selves. It might be argued that the no-self teaching is a philosophical theory and that it puts into question the Buddha's assumption that 'it is up to you' to undertake the Eightfold Path. Here, it may be objected, is an important disanalogy with the wise farmer: the causal regularities in farming concerned plants, not human beings, and so while there was no conflict with the

farmer's advice, there was conflict with the implicit 'it is up to you' assumption of the Buddha's teaching.

The Mālunkyāputta text precludes discussing philosophical questions only if doing so is not relevant to overcoming suffering. The no-self teaching and the considerations put forward in support of it surely amount to at least a proto-philosophical position. So this is an important challenge. Karin Meyers (2014) has argued that the no-self teaching is incompatible with the implicit 'it is up to you' message in the Buddha's various statements to his followers because the message presupposes an 'autonomous agent' that the no-self teaching denies. She argues that the problem can be resolved by appealing to the distinction between ultimate and conventional truth: there are no selves in the former, but there are autonomous agents in the latter. However, though this may well be an apt resolution for later Buddhist traditions, in the Buddha's teaching in the *Nikāyas* the distinction between ultimate and conventional truth had not yet been explicitly articulated. How might the Buddha have considered this issue?

The Buddha's basic position is that what we call the 'self' is nothing but an ensemble of ever-changing and causally dependent processes that he categorized as the 'five aggregates' (material form, feeling, perception, formations and consciousness). There is no distinct and unchanging self over and above these, or somehow at the center holding all of them together. The discovery that this was so, that no such self could be observed, was important, he believed, because it undermined the primacy of the distinction between what is and is not 'me and mine' in practical deliberation. In particular, the Buddha discovered

that living our lives from the perspective of the dichotomy between what is and is not 'me and mine' gives rise to greed and aversion, which in turn gives rise to suffering. The realization that there is no self is crucial to dismantling this conditioning: when what is 'me and mine' loses its practical salience, greed and aversion, and consequently suffering, decline. This analysis conforms to the basic arising and cessation model of dependent origination discussed earlier.

In the Buddha's teaching, the processes that are mistakenly interpreted as selves are not randomly distributed across the universe. That they are organized in some significant way is implicit in the idea that they are causally conditioned (especially as exemplified in karma). From these patterns of organization arises the mistaken thought that "I am a self", that these particular processes (and not those other ones) belong to 'me' or 'myself'. When the Buddha addressed the not-yet-enlightened and assumed that they could understand and accept his teaching, form intentions to carry it out, and act on those intentions, he would have supposed that all this could be explained in terms of processes of aggregates. Of course, he would have recognized and taken into account that, insofar as his followers were still deluded, they would have tacitly assumed that it was their selves as agents who were thinking, intending, acting, etc. This was part of their delusion, and perhaps in striving to follow the Buddha's teaching they sometimes experienced some cognitive dissonance concerning this issue. But for the most part confidence (*saddhā*) in this teaching concerning the self evidently sustained the conviction among his disciples that with enlightenment they would come to observe and understand the explanation in terms of processes of

aggregates. We see here one of the roots of the distinction between ultimate and conventional truth.

Those who are naturally drawn to philosophical inquiry might well think that this account raises philosophical questions. However, in a similar vein, the Buddha's teaching about rebirth might raise questions about body and soul, and his teaching about escaping the cycle of rebirth might raise questions about the postmortem existence of an enlightened person. Yet these are precisely the questions the Buddha declined to answer in response to Mālunkyāputta's challenge. His pragmatic orientation is a guide to what he might have said had his followers expressed perplexity about how, if there is no self, no agent, they could act on his teaching. The Buddha might have said that they should focus their attention, less on developing a philosophically cogent explanation of the nature of action, and more on acting selflessly, on living in such a way that the distinction between what is and is not 'me and mine' is not salient in deciding how to act—that is, on living with compassion and loving-kindness for all beings.

From this perspective, the emphasis of the no-self teaching was as much ethical as it was metaphysical. It was first and foremost a practical teaching about living without greed and aversion so as to overcome suffering. Though it depended on an understanding of persons, this understanding was based more on meditative awareness than on metaphysical analysis. The wisdom of the Buddha, from this perspective, was primarily a practical rather than a theoretical wisdom, more a *knowing how* than a *knowing that*. This is why, on this

reading, the Buddha was not concerned with free will and how it might be reconciled with no-self, dependent origination, or other aspects of his teaching.

References

Primary Sources

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.

Buddhaghosa, B., 1999 (1975). *The path of purification (Visuddhimagga)*. Translated by B. Ñāṇamoli. Seattle, WA: Buddhist Publication Society Pariyatta Editions.

Ñāṇamoli, B. and Bodhi, B., trans. and eds, 1995. *The middle length discourses of the Buddha: a new translation of the Majjhima Nikāya*. Boston: Wisdom Publications.

Walshe, M., trans., 1987. *The long discourses of the Buddha: a translation of the Dīgha Nikāya*. Boston: Wisdom Publications.

Secondary Sources

Gowans, C.W., 2010. Medical analogies in Buddhist and Hellenistic thought: tranquility and anger. In: C. Carlisle and J. Ganeri, eds. *Philosophy as therapeia, Royal Institute of Philosophy Supplement 66*. Cambridge: Cambridge University Press. pp. 11-33.

Meyers, K., 2014. Free persons, empty selves: freedom and agency in light of the two truths. In: M.R. Dasti and E.F. Bryant, eds. *Free will, agency, and selfhood in Indian philosophy*. Oxford: Oxford University Press. pp. 41-67.

2 Why There Should Be a Buddhist Theory of Free Will

Rick Repetti

As reflected in some of the other contributions to this volume, some Buddhist scholars think Buddhism rejects free will, or they deny there is, was, or should be a Buddhist free will theory or even any Buddhist inquiry into free will. I disagree. I argue *for* a certain Buddhist theory of free will in chapter 17 of this volume that could but won't be repeated here: here I argue *why* there should be such a theory (see ch. 17).

The Buddhist path contains methods for cultivating virtuoso degrees of abilities that exhibit self-regulative agency—e.g., volitional/metavolitional regulation, reason-responsiveness—touted by contemporary philosophers sympathetic to naturalistic or ‘compatibilist’ free will, that is, the sort of free will that is compatible with scientific views of causation or nature. For Buddhists, false understanding of self-agency is considered the central cause of *dukkha* (suffering, existential unsatisfactoriness), and correct understanding is its antidote. The elimination of *dukkha*, coextensive with enlightenment, is the principal aspiration of Buddhist soteriology. Thus, whether free will is part of correct understanding of self-agency is a valid question; its answer, positive or negative, arguably constitutes a ‘Buddhist free will theory’, just as the answer to the question whether the self is real constitutes ‘the Buddhist theory of the self’.

The unanalyzed self-sense is not that of a passive witness,¹ but of one that wants, deliberates, resists, wills, chooses, acts, and whose actions are *up to* it. This reified understanding of self-agency, Buddhists claim, perpetuates *dukkha*. All Buddhist teachings and practices seek to disentangle this self-conception, sorting its illusory and non-illusory elements. It is not choice, action, etc. that disappear upon enlightenment, but the *illusory misconception* that one is the ontologically independent *ātman* (immaterial, changeless soul/self) that (virtually telekinetically) generates them.

Whether, to what extent, and how elements/abilities of agency are self-regulating, though the ‘agent’ is not considered metaphysically substantive, are questions the answers to which would provide a theoretical understanding of self-agency and of whether such elements/abilities constitute moral-responsibility-entailing free will, and thus a Buddhist free will theory that may be described as involving ‘agentless agency’. Such a theory is soteriologically warranted, among other reasons, but even if for no other reason than that individuals presumably would want to understand how it is that they can alter their lives in certain ways to attain enlightenment if there is no such thing as the individual agent. Such a theory also promises to illustrate the explanatory purchase of Buddhist psychology, metaphysics, and ethics regarding one of the most—if not *the* most—intractable of problems in analytic Western philosophy.

Āryas (advanced Buddhist practitioners) are meditation virtuosos, and they have cultivated great self-control, an element of free will that—at least theoretically—doesn’t require a substantive agent-self. According to Buddhist thought, the average non-practitioner is so heavily conditioned by the three poisons (greed, hatred, and delusion) as to be virtually determined—whether

determinism is universally applicable or not—insofar as they lack the sort of proximal self-regulative abilities the virtuoso exhibits. The virtuoso exercises self-regulation dharmically (in accord with *Dharma*, Buddhist teachings), reducing *dukkha* and cultivating enlightenment, not only not feeding ego-volitions, but diminishing them.

The most powerful free will skeptical argument in Western philosophy, ‘hard incompatibilism’, denies autonomy *regardless* of whether we are determined: if choices are determined, we’re not ultimately responsible for them; if they’re random, we cannot claim to author them. An example is Strawson’s ‘impossibility argument’: we choose as a function of our present mental state; we cannot be the cause of our first mental state; so, it’s impossible to be ultimately responsible for mental states we’re in when we choose, or for such choices (Strawson 1994).

Most Buddhists would accept most of Strawson’s claims, given their doctrine of dependent origination: whatever arises does so in dependence on everything else that has arisen or is simultaneously arising. That choices are conditioned by mental states is axiomatic, as is the idea that conditioning extends backwards indefinitely through previous mental states, choices, and actions; indeed, Buddhist philosophy considers this temporal sequence beginningless. Add to this the tautology that self-creation *ex nihilo* is impossible, and it seems to follow that nobody can be ultimately responsible for any mental state, choice, or action.

The central insight of Buddhism is that total mental freedom is possible in *nirvāṇa*, the antidote to *dukkha*. The meditation virtuoso thus can escape from the influences of previous and current mental states—irrespective of deterministic or random causes—contradicting Strawson’s

impossibility argument. Non-Buddhists might deny anyone attains *nirvāṇa*, but evidence supports some Buddhist meditation claims; to mention just one relevant study, Zen practitioners (trained to be ceaselessly responsive to the present) exposed to repetitive stimuli show no reduction in responsiveness, whereas control groups ignore repetitive stimuli after brief exposure (Kasulis 1985). Hundreds of studies conducted on Buddhist meditators similarly confirm their claims (Lutz, Dunne and Davidson 2007; Lutz and Thompson 2003; Thompson 2015). These studies indirectly render plausible the claim that the *arhat*—the enlightened being, the limit case meditation virtuoso—is free of mental state influences, with minor exceptions (Harvey 2007), and to the extent they approximate *nirvāṇa*, so are the not-yet-enlightened *āryas*. Buddhist metacognitive training thus cultivates a skill that theoretically defeats the most powerful free will skepticism in analytic Western philosophy: this alone justifies a Buddhist free will theory, but there are other justifications.

Gowans (this volume) rejects the idea of a Buddhist theory of free will, with one pragmatic exception—shared by Goodman (this volume), but for different reasons—because Buddhism restricts inquiry to the soteriologically relevant, and Gowans assumes an understanding of free will is not generally soteriologically relevant. But Westerners are not committed to the principle that only what is soteriologically relevant is warranted. Non-Buddhists—and Buddhists alike—clearly may benefit from a Buddhist theoretical understanding of free will that promises to be enlightening. Moreover, it is possible that by seeing the complexity of Buddhism Westerners might be drawn to it. Many who came to debate the Buddha were drawn into the *Dharma* (the Buddhist teaching) upon hearing how it handled philosophical puzzles. If Buddhism is

enlightening in general, then anything that brings listeners toward the *Dharma* is soteriologically relevant and thus justified.

Garfield and Flanagan (both this volume) cite theodicy in the Western genesis of the free will belief as a basis for rejecting its relevance to godless Buddhism, adding that the free will meme survives—implicitly, as a doxastic appendage—in the West, independent of the decline of its genesis in monotheism. If they are right, perhaps this Western meme may be rehabilitated by a Buddhist free will theory that denies the ultimate reality and causal autonomy of the self, but advocates that we may control our volitions, cultivating an agency that leads to its own transcendence in total mental freedom, *nirvāṇa*. What could be wrong with that, from a Buddhist perspective?

Regardless of its genesis, however, which is technically irrelevant to the validity of a concept, most of us have the free will meme, and even if we come to doubt it upon reflection (something that also arguably has a genesis only in Western thought), we tend to live as if it is true, with few exceptions, including Blackmore (this volume), who claims she doesn't experience free will, and Harris (2013), who claims that even the belief that we *experience* free will is illusory, because phenomenological examination shows that choice is mysterious. Blackmore and Harris are both long-term Buddhist meditation practitioners, so perhaps the genetic invalidation of ideas (based on their origin) actually does apply to them, in reverse.

For one key Buddhist meditation instruction—clearly based on traditional Buddhist assumptions—is to pay nonjudgmental attention to the *impersonal* arising of mental states. Thus,

what Buddhist meditation practitioners consider *discovered* or *verified* in phenomenological experience actually may be entirely generated by Buddhist-theory-laden conceptual expectations. Thus, unless one is an arguably Buddhism-biased meditator, our pre-theoretical agential phenomenology arguably universally depicts choices and actions as *up to us*: even Buddhism agrees that desires and thoughts that typically arise spontaneously *may* be regulated. In the two traditions' mutual search for truth, no holds are barred—in either direction.

Most of us *experience* some sort of free will: our beliefs and desires inform our choices and volitions, and these—we all seem to think we experience—*cause* our behavior. However, many branches of contemporary science reveal unconscious biases, errors, and illusions distorting our experience. Anyone thus informed must navigate repetitive cognitive dissonance between these findings and our humanistic self-conceptions. Anxiety accompanies existential uncertainty about the pre-theoretical narrative that no longer seems to sufficiently explain our place in the vast universe, our communities, bodies, brains, and minds.

These narrative voids in many Westerners' lives may be filled with Buddhist philosophy, which does not require belief in implausible myths about divine creation and human origins *ex nihilo*, anachronistic moral rules, crumbling folk psychology, soul, or self, but nonetheless presents a long-flourishing way of life guided by a comprehensive way of understanding one's place in a meaningful world that is comfortable with such seemingly depersonalized implications.

Buddhism offers hope to anyone in an *existential doxastic impasse*, unable to maintain outmoded moral, religious, and other humanistic beliefs, but reluctant to embrace what they suspect is the narrative bankruptcy of value-free scientific inquiry. Buddhism offers an equally sobering but

simultaneously coherent existential, psychological, philosophical, ethical, and soteriological narrative.

Compare the strategy of trying to forge a truce between religion and science, exemplified in Gould's 'non-overlapping magisteria' ("NOMA") thesis: science reigns over empirical fact, religion over values and meaning (2011). But most major religions—with the possible exception of Buddhism—assert claims about the history of the world and humanity that contradict facts in astrophysics, geology, and genetics, among others (ed. Caruso 2014). Most versions of Buddhism are immune to this criticism, and some versions of Buddhism (e.g., emerging Western versions) may count more as philosophies than religions, but most have far fewer claims that contradict scientific facts, and the few that threaten to contradict science have yet to be refuted, and are not core tenets. Buddhism is committed to empirically validated truth, dating to injunctions from the Buddha not to accept anything on authority, but on investigation.

Two major exceptions are karma and reincarnation, though some Buddhists do not take these literally. Flanagan suggests a naturalized Buddhism, 'Buddhism without hocus pocus' (Flanagan 2011), which would eliminate supernatural and related metaphysical elements but which would likely appear to traditional Buddhists as 'Buddhism without Buddhism'. Nonetheless, Buddhism is a humanistic and spiritual philosophy that satisfies many human needs typically satisfied by religion, and much if not all of its supernaturalism is plausibly optional. Indeed, revisionary Western forms of Buddhism that take Flanagan's naturalism for granted are emerging, despite traditionalists' objections (eds. Purser, Forbes and Burke 2016).

Buddhism *naturalized* instantiates a valid NOMA case. Buddhism *as is* almost qualifies for the status NOMA erroneously affords all religions. Buddhism is probably attractive to Westerners today because it mostly qualifies for NOMA status: without conflicting with science, it answers to deep needs for humanistic self-understanding; it is mostly consistent with contemporary moral thinking, unlike most religions; and it arguably supports a discounted but palatable view of agency or free will that coheres with that moral framework.

Loosely analogous to NOMA's bifurcation is the 'two truths' doctrine in Buddhism that Siderits (this volume) applies to free will. Simplifying, the two truths parallel the distinction between pragmatic and scientific levels of description in Western philosophy. (The *Stanford Encyclopedia of Philosophy* has an excellent entry on the "Two Truths".) *Conventionally* (pragmatically), there are tables and people, though *ultimately* only atoms configured certain ways, which we conveniently designate 'tables' and 'people'. Most forms of Buddhism treat the conventional similarly, but Madhyamaka (later) Buddhists deny, and Theravāda (earlier) Buddhists assert, that the ultimate is metaphysically substantive. Theravāda denies ultimate reality to partite wholes it views as reducing to mere aggregations of momentary atomistic psychophysical tropes it sees as ultimately real. On both Buddhist views, what appear to constitute independently existing macro-level wholes are viewed as constructed projections of our—dualistic and therefore illusory—conceptualizations, lacking metaphysically substantive or objectively real essences but, for Madhyamaka, everything lacks substance, all the way up and all the way down, so to speak: as Friquegnon notes (this volume), on this view *nirvāṇa* and *saṃsāra* (the world of cyclical reincarnational wandering in primal confusion, as perceived by the unenlightened) are one, simply viewed differently.

On analysis, almost everything Buddhism claims is only conventionally true, which implies not ultimately true, except perhaps the respective understandings about the ultimate. A Buddhist NOMA issue is arguably whether conventional Buddhist truth (e.g., Buddhist religion) is non-overlapping with ultimate Buddhist truth (akin to science). Applying this reasoning to free will, it is probably ultimately illusory (in agreement with many scientists and philosophers) but conventionally real (in agreement with the folk), a position made explicit by Siderits (this volume), who insists the two domains are non-overlapping, *semantically insulated* from each other—both ‘true’ under different conceptions of what it means to be true, or true in different discourse domains—yet, importantly, conventional truth may be thought true because it (mostly) *supervenies* on ultimate truth.

Any ways the two domains seem non-insulated are, on analysis, increasingly sophisticated conventional discourses, say, analogous to fiction or implicit meta-language, all ultimately false. Only statements describing ultimate reality are strictly true, and these reflect the insubstantiality of everything (except, for Theravāda, psychophysical atomistic tropes). Enlightenment consists in the embodied/integrated realization of metaphysical insubstantiality or ‘emptiness’. Buddhism aims at this negative realization—everything is ‘empty’ of ‘self-nature’ or metaphysical substance—and prescribes a way of life devoted to its attainment, conceived as life affirming and maximally liberating.

There is an ancient contemplative saying, “Before enlightenment, chop wood and carry water; after enlightenment, chop wood and carry water” (Roth and Montgomery 2007, p. 141),

suggesting enlightenment involves the same world, mental freedom being the difference. It's not that post-enlightenment there's no wood, water, chopping: these remain, but one doesn't misapprehend their insubstantiality. This applies to free will. It's not that post-enlightenment there's no considering consequences, choosing, acting: these remain, but one doesn't misapprehend their insubstantiality. There is effective agency, but no substantive agent—'agentless agency'. Or so I argue, despite being unenlightened and thus barely 'pointing at' this putatively nondualistic reality.

Though this is a possibly misleading comparison, Buddhism may be likened—for anyone 'recovering' who turns to Buddhism—to methadone, which helps transition from heroin addiction, insofar as Buddhism provides enough doxastic scaffolding to support a transformative process from faith in the (post-monotheistic-religious) conventional, humanistic, folk view of self and world to one consistent with ultimate reality. The two-tiered and multi-dimensional elements of Buddhist thought facilitate transition from simpler to more complex levels of understanding, enabling beginners to latch onto initially palatable doctrines that upon further inquiry reveal deeper insights that require (typically meditative) work to assimilate. Perhaps this is why Buddhism is essentially a (progressively developmental) 'path'.

For example, beliefs in *reincarnation* and *karma* fill a theodicy void in recovering Western theists who might still subscribe to a 'just world' thesis, for they assuage fear of death and existential injustice, respectively, and present a morally superior worldview. A Newtonian karmic law guaranteeing a just world and an infinite number of incarnations until enlightenment is more compassionate than eternal damnation for one bad life. From an impartial vantage,

Buddhism has the best theodicy among all the world's religions, however oxymoronic (if one defines 'theodicy' to require a deity). We can call Buddhist godless theodicy 'soteriodicy', its soteriological theory of cosmic justice.

That's the methadone part of the analogy. After 'drying out' on the metaphorical methadone of simplistic belief in karma and reincarnation, according to maturity and readiness, budding Buddhist converts come to understand these doctrines more subtly (say, there's no self that reincarnates or bears karma, or, subsequently, if there's no reincarnation, it's *like* reincarnation moment to moment anyway). Admittedly, this enables a bait-and-switch strategy, but that millions of Westerners are drawn to Buddhism and that Buddhism has spread globally without force, adapting to different cultures throughout history, is likely significantly explained by this progressive soteriodicy.

Buddhism may play a transformative role in the lives of many such Westerners and possibly a greater role if it had a well-worked out explanation for the dynamics of agency that captures something of our folk psychological sense that our lives are significantly up to us, particularly in light of its pragmatic arts of self-regulation, while it eases us through its hierarchy of levels of understanding about the ultimately agentless nature of agency, the latter of which converges upon the sort of deflationary pictures of ourselves being delivered to us increasingly from the latest scientific research. Buddhism already does this levels-of-complexity transition when it comes to its theory of the self: ultimately there is no metaphysically substantive thing that is a self, but conventionally talk of the self is pragmatic. As Siderits argues (this volume), ultimately there is no free will because there is no self, but conventionally we have free will because we are

persons (who act voluntarily, etc.). Coseru (this volume) sketches what might, in my view, make for an intermediate level Buddhist understanding of our essentially subjective/phenomenological first-personal agency. This two-tiered hierarchy comports with the need to balance our everyday conceptions of ourselves as social agents with our understanding of the impersonal ultimate reality revealed by science. If Flanagan and Garfield are right that the Western free will meme is a post-theodicy doxastic appendage, then Buddhism's complex, tiered theoretical understanding of agency may serve as its therapy.

A Buddhist free will theory may also be supported by Buddhist ethics. According to Goodman (2009), the Buddhist canon implies negative consequentialism: the reduction of suffering is its prime directive. Whether he is right, Buddhism is *consistent with* negative consequentialism. Buddhist ethics thus may be taken to indirectly support negative consequentialist arguments to the effect that doing x is morally acceptable if it leads to the reduction of suffering. Thus, if a Buddhist free will theory promises to reduce suffering, then it is dharmic. The Buddhist path is classically likened unto a raft constructed on the teachings of the *Dharma*, used to cross over *samsāra*, the sea of delusional/reincarnational wandering, and reach the shore of *nirvāṇa* (Repetti 2010a). Constructing a Buddhist free will theory promises to help the aspirant understand the use of the paddle.

My earlier argument (Repetti 2010a), that Buddhist meditation increases both a form of free will and its more wholesome exercise, confirms the paddle-use justification for a Buddhist free will theory. Buddhist meditation cultivates the ability to detach from first-order intentions (base-level cravings, impulses, desires, wants, etc.) by the repeated discipline of simply observing them,

which generates a causal/functional experiential distance between first-order intentions and awareness of them, enabling the practitioner to take those intentions *virtually off-line* from their typical, habituated, online dynamics of leading mindlessly into action, and thus to make wiser, more dharmic choices, approving some and disapproving others, forming dharmic second-order intentions (meta-intentions), to borrow Frankfurt's first-/second-order distinction—not to advocate a Frankfurt-style Buddhist theory of free will, contra Flanagan (this volume) and Tuske (2013).

Whether the process is deterministic or ultimately insubstantial is irrelevant to the fact that practitioners cultivate powerful self-regulation skills enabling them to reduce mindlessly acting on first-order intentions otherwise leading to further addictive tendencies and suffering, as Meyers, Wallace, and Harvey (all this volume) would agree—skills most consider forms of autonomy. Buddhist meditation therefore reduces suffering and increases self-regulation. Thus, Buddhist meditation supports a conventional conception of autonomy as self-regulation that is ultimately agentless, though the term 'autonomy' need not imply a real self. Some heating systems have thermostats; they're self-regulating, though lacking a self. The usual counter is that thermostats don't program themselves, but Buddhism suggests, as Wallace and McRae (both this volume) argue and I argued (2010a, 2010b), that we can reprogram ourselves to increase our freedom. I'm noting that doing so increases free will.

Because much recent interest in Buddhism is driven by interest in meditation, an analysis of the dynamics of meditation in a Buddhist free will theory promises to support that interest, and attract more individuals toward the *Dharma* by virtue of its explanatory elegance, promising to

reduce suffering, satisfying Buddhism's negative consequentialism. Buddhist free will theory promises to help Western meditators understand how their discipline enables them to increase psychological/functional (conventional) agency on the path to (ultimate) agentlessness—to freedom *from* the adharmic, ego-driven elements of the will.²

Many Western intellectuals are drawn to Buddhism because it espouses a philosophically sophisticated deflationary understanding and doctrine of the illusory nature of the self, together with a similarly sophisticated psychology, metaphysics, and ethics that provide its broader framework.³ Buddhism exemplifies a long-established, socially/culturally stable set of lived traditions in which that understanding is normalized. The worry among many Westerners who resist scientific findings is that our humanistic self-conception and corresponding way of life cannot withstand these sobering truths.⁴ But Buddhism has thrived for millennia after embracing sobering truths about the self, the world, the gods, etc. The implication is that Buddhism facilitates wholesome assimilation of enlightening truths, even about our dearest illusion—the self. Unless we are enlightened, each of us is the main protagonist in the narrative of our lives, the center of our perspectival universe (Coseru 2012).

The notion of free will—autonomy, self-governance—rests intuitively on the notion of the self. Since the self and free will are interdependent, as Goodman (2002) argues when he says because there is no self there cannot be free will, and since Buddhism has done such a great job transitioning folks from belief in the self to belief that the self-illusion is the chief cause of suffering, Buddhism can play the same role with belief in free will. This transition might be smoother than with the self, because the process is supported by a set of mindfulness-cultivating

and mindlessness-weakening practices that *increase* self-regulation: practitioners experience/cultivate *more effective agency* as practice deepens (Aronson 2004), to the point where the sense of *metaphysically autonomous* agency is predicted to shed spontaneously with the shedding of the sense of a substantive self, at a natural, wholesome pace.

An analogy might help make the point. In psychotherapy, a wholesome interpretation of ‘defense mechanisms’ as ‘coping strategies’ may be preferable, insofar as one may use them when one feels the need to, if understood as crutches the client will drop when healthier/stronger. The Buddhist path may be understood not only as a raft, but a crutch—training wheels for cycling toward enlightenment. If a Buddhist free will theory promises to function as *upāya*, soteriologically instrumental, it is dharmic.

Technically, there is no singular Buddhism, only different forms of Buddhism, so the sort of free will theory that comports with one form of Buddhism may not comport with another, as with even some core tenets. But each form of Buddhism might formulate its own version, and each might be justified by variations on these arguments.

Another obstacle is that Buddhism is essentially a collection of strategies for attaining freedom *from* the will’s typically mindless, adharmic expression, so Buddhism seems *prima facie* diametrically opposed to the idea of ‘free’ will, on grounds seized upon by philosophers quick to oppose the idea outright. For example, one everyday pre-philosophical free will conception involves being able to do as you please, to act on desires, to spontaneously express in action

whatever you feel like doing, etc., but, from a Buddhist perspective, if that's free will, Buddhism is out to eliminate it because feeding ego-desires causes suffering.

Another free will conception, often embraced by religious believers and others as a knee-jerk reaction to the idea that we have no free will, is a 'libertarian' view that we can act against all causal influences, 'contracausally' injecting free will from the nonphysical mind/soul into the physical world, like mini-god/soul-egos. Ideas like these account for the greatest resistance to the idea that there should be a Buddhist free will theory. But similar ideas were rejected by Buddhist metaphysics in its encounters with the Orthodox Indian *ātman* concept thousands of years ago (Federman 2010)—another reason to think *there was* an implicit Buddhist free will theory, however negative, from Buddhism's inception.

Considerations such as these lead to metaphilosophical questions about the various concepts of 'free will', what each really means and entails, whether any are true, and so on. I have articulated my views on that in detail elsewhere (Repetti 2010b), so I will not do so here. But, based on a fairly representative sampling of the state of dialectical affairs in the free will literature (ed. Caruso 2013), the most brilliant thinkers who specialize in the subject are in respectful disagreement on almost all such matters, so the issues are mostly open regarding the oldest, most discussed issue in Western philosophy. Few contemporary Western philosophers embrace *ātman*-like models of self-agency early Buddhists rejected, and few would endorse the pre-philosophical view that free will just consists in choice making.

Rather, arguments for the psychophysical models of mind and action in early and subsequent Buddhist texts are complex, compatible with, and/or presage those in contemporary Western philosophy, neuroscience, cognitive science, and the social sciences (Coseru 2012, Garfield 2015). This is one of the main reasons researchers and scholars in these fields have found Buddhism interesting. Thus, not only is no bar against Buddhism entering the Western dialectic, Buddhism has much to offer it; Garfield makes this argument convincingly (2015), despite his exception regarding free will (this volume). And while there is no already-explicit free will theory in Buddhism, the elements to be deployed in its construction abound in Buddhist psychology, action theory, metaphysics, and ethics. As I argue in chapter 17 (this volume), the Buddha explicitly rejected several contraries of free will: the notions of inevitable causation by matter, gods, fate, karma, and chance. By double negation, this implies a favorable attitude toward free will.

A talking point frequently repeated by critics and defenders of free will in the West is that the only sort of free will that matters is one that grounds moral responsibility. Why this line has been pressed is too complex to explore here at length, but one reason is that it emerged during the heyday of logical empiricism, which inclined philosophers to censor purely metaphysical issues as if they violated empirical meaning and thus standards of dialectical legitimacy, rendering free will suspicious if not wedded to something with pragmatic value, such as moral responsibility. Accordingly, there can be no moral responsibility without free will. Now, free will defenders worry that if eliminative forces from the sciences undermine free will, morals collapse with it. Buddhism, however, has flourished in the absence of an explicit free will theory, and most of its psychology and metaphysics suggest we ultimately lack free will, but Buddhism has managed to

remain a richly moral tradition. So, Buddhism illustrates that morals can survive eliminative attitudes towards belief in the self and free will. Therefore, a Buddhist free will theory promises to bring morals to the table even though its views on agency may be closer to being eliminativist than they are to being nonreductively substantive; they may be seen, as Harvey argues (this volume), as somewhere in the ('semi-compatibilist') middle between these two extremes.

A parallel to whether there is or should be a Buddhist free will theory concerns Buddhist ethical theory. Despite how *richly ethical* Buddhist doctrine is, there is no explicit *theory* of ethics in historical Buddhism, no metaethics, and no prescriptive *theory*, though many have recently attempted to discern whether such theories may be inferred from Buddhist prescriptions.

Possibly violating his own hermeneutical criteria (Flanagan, this volume), Flanagan (2011) has argued that Buddhist ethics is a kind of aretaic ethics, Goodman (2009) has argued that it is a negative form of consequentialism, and Garfield (2015) suggests a narrative model. We need not debate that, but simply note it as an important parallel to our question about a Buddhist free will theory. The point is that there has been no theoretical component to Buddhist ethics throughout its history, but there is enough in Buddhist ethics to render the project of adducing a Buddhist theory relevant, interesting, and worthy of pursuit. By analogy, it would be interesting and worthy of pursuit to find out what Buddhism can contribute to our understanding of free will.

For all these reasons, there can and should be a Buddhist theory of free will (or theories, given different forms of Buddhism). Some Buddhists disagree, but they seem to conflate a *negative* theory with *no* theory, though it is standard in Buddhism to treat the negative theory about the self as a theory, rather than as no theory; meanwhile, several other Buddhists are proposing such

theories (largely filling this volume). Indeed, their disagreement counts as evidence that there is room in Buddhism for the variety of free will theories analogous to those in Western philosophy, since a similar variety is emerging among Buddhists attempting a Buddhist position on free will. The remainder of this collection contains a number of them. I argue in chapter 17 (this volume) that the Buddha seemed to imply one.

Notes

¹ However, in meditation one *trains* in passive witnessing, construable as practicing enlightenment.

² See Aronson (2004) on the difference between the psychological/functional self, which becomes increasingly effective and well integrated along the Buddhist path, and the metaphysical self, which is increasingly seen as an illusion as one advances along that path.

³ See Garfield (2015) for an excellent argument in support of much of what is philosophically attractive in Buddhism for Western philosophers.

⁴ Some, like Peter Strawson (1962), deny we can lose our humanistic self-conception. Galen Strawson (this volume) argues we can, and Blackmore (this volume) appears to have done so, with some success.

References

Aronson, H.B., 2004. *Buddhist practice on western ground: reconciling Eastern ideals and Western psychology*. Boston, MA: Shambhala.

Caruso, G.D., ed., 2013. *Exploring the illusion of free will and moral responsibility*. Lanham, MD: Lexington Books.

Caruso, G.D., ed., 2014. *Science and religion: 5 questions*. Copenhagen: Automatic Press.

Coseru, C., 2012. *Perceiving reality: consciousness, intentionality, and cognition in Buddhist philosophy*. New York: Oxford University Press.

Federman, A., 2010. What kind of free will did the Buddha teach? *Philosophy East and West*, 60(1), pp. 1-19.

Flanagan, O., 2011. *The bodhisattva's brain: Buddhism naturalized*. Cambridge, MA: MIT Press.

Garfield, J., 2015. *Engaging Buddhism: why it matters to philosophy*. New York: Oxford University Press.

Goodman, C., 2002. Resentment and morality: Buddhism on moral responsibility. *American Philosophical Quarterly*, 39(4), pp. 359–372.

Goodman, C., 2009. *Consequences of compassion: an interpretation and defense of Buddhist ethics*. New York: Oxford University Press.

Gould, S.J., 2011. *Rocks of ages: science and religion in the fullness of life*. New York: Random House.

Harris, S., 2013. *Free will*. New York: Free Press.

Harvey, P., 2007. ‘Freedom of the will’ in the light of Theravāda Buddhist teachings. *Journal of Buddhist Ethics*, 14, pp. 271-291.

Kasulis, T.P., 1985. *Zen action, Zen person*. Honolulu, Hawaii: University of Hawaii Press.

Lutz, A. and Thompson, E., 2003. Neurophenomenology: integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies*, 10(9-10), pp. 31-52.

Lutz, A., Dunne, J.D., and Davidson, R., 2007. Meditation and the neuroscience of consciousness: an introduction. In: P.D. Zelazo, M. Moscovitch, and E. Thompson, eds. *The Cambridge handbook of consciousness*. Cambridge: Cambridge University Press. pp. 499-554.

Purser, R., Forbes, D., and Burke, A., eds., 2016. *Handbook of mindfulness: culture, context and social engagement*. New York: Springer.

Repetti, R., 2010a. Meditation and mental freedom: a Buddhist theory of free will. *Journal of Buddhist Ethics*, 17, pp. 165-212.

Repetti, R., 2010b. *The counterfactual theory of free will: a genuinely deterministic form of soft determinism*. Saarbrücken, Germany: LAP Lambert Academic Publishing.

Roth, R. and Montgomery, R., 2007. *The sacred light of healing: teachings and meditations on divine oneness*. New York: iUniverse Inc.

Strawson, G., 1994. The impossibility of moral responsibility. *Philosophical Studies*, 75(1-2), pp. 5-24.

Strawson, P.F., 1962. Freedom and resentment. *Proceedings of the British Academy*, 48, pp. 1-25.

Thompson, E., 2015. *Waking, dreaming, being: self and consciousness in neuroscience, meditation, and philosophy*. New York: Columbia University Press.

Tuske, J., 2013. The non-self theory and problems in the philosophy of mind. In: Steven

Emmanuel, ed. *Companion to Buddhist philosophy*. Wiley-Blackwell. pp. 419-428.

3 Uses of the Illusion of Agency: Why Some Buddhists Should Believe in Free Will

Charles Goodman

What if the free will issue was never really about determinism? What if the only reason determinism ever *seemed* relevant to the question of whether our actions are truly our own is that causal chains stretching far back into our past constitute a vivid way of showing that our actions belong to the universe? What if the absence of free will is just a special case of the frightening, liberating truth that *nothing* is our own, because there is no self? That might explain why the free will issue seems intractable: despite the enormous attention lavished on it, philosophers haven't fully understood it. It might help explain recent surprising empirical results suggesting that belief in free will and in determinism are not opposed, but represent two largely statistically independent factors (Nadelhoffer and Tocchetto 2013, p. 126). It might explain why critiques of free will feel threatening: they have the potential to undermine the one thing we are most invested in, our false belief in a real self.

Since what is most philosophically distinctive about Buddhism is its rejection of that false belief, Buddhists should be committed to the nonexistence of free will. Many scholars appear reluctant to accept this conclusion, though, perhaps because they overestimate how destructive and revisionary it would be. I have proposed that the most defensible Buddhist view of free will would be a form of 'hard incompatibilism' (Goodman 2009), the view that free will is incompatible with both determinism and indeterminism (Pereboom 2007). Hard incompatibilists

do not claim that we cannot change, or that our actions make no difference, or that making an effort to improve is pointless. Instead, they reject the existence of basic desert and regard interpersonal reactive attitudes such as anger and resentment as cognitively inappropriate.

We can be said to deserve many things: praise, blame, rewards, punishments, and so on. The specific form of desert that Pereboom says we should reject, considered in relation to blame or credit, is basic in the sense that the agent, to be morally responsible, would deserve the blame or credit just because she has performed the action (assuming she understands its moral status), and not by virtue of consequentialist or contractualist considerations (2007, p. 86).

Thus, Pereboom does not want to issue a blanket ban on blaming or praising, or to say that no one deserves anything. Practices of holding people responsible can have socially beneficial consequences that can create derivative, institutional, non-basic forms of desert and responsibility. What Pereboom denies is that our actions could give rise to reasons to treat us well or badly independently of the consequences resulting from such treatment. The issue about praise and blame is, in its direct practical implications, almost trivial; the rubber meets the road when we consider the justification of punishment and systems of criminal justice.

Note that much blame consists of accurate descriptions of a person's character. So, for instance, in a world without basic desert or deep moral responsibility, someone might persistently neglect his responsibilities due to laziness. To tell that person "You are lazy!" would be an accurate description, and 'appropriate' as such. But no amount of laziness could cause that person to deserve, in a basic sense, the distress it might cause him to be told so. We should tell him,

truthfully, he is lazy *only if that would be helpful*, and never simply to hurt his feelings, no matter how much havoc his laziness has caused.

Hard incompatibilism includes a claim I summarize by saying that anger, resentment, and other interpersonal reactive attitudes are cognitively inappropriate. When I say that a particular emotion is cognitively inappropriate, I mean it is partly constituted by a judgment that is false and unwarranted. Thus, a hard incompatibilist could say that anger at person X for doing action A involves construing X as being, in some deep and ultimate sense, the source of A; but we are never the sources of our actions in this sense.

It's hard to deny that Buddhists disapprove of anger and resentment; but few Buddhist texts consider the free will problem in any recognizable form—few, but not none. Many scholars agree that the problem is raised explicitly by Śāntideva in ch. VI of his *Bodhicaryāvatāra* (“*BCA*”; all references to *BCA* are to the Crosby and Skilton (1995) translation). (Goodman 2009, ch. 8, examines this in detail.) Here's one crucial verse from Śāntideva's discussion of the issue:

“In this way everything is dependent upon something else. Even that thing upon which each is dependent is not independent. Since, like a magical display, phenomena do not initiate activity, at what does one get angry like this?” (*BCA* VI.31; Crosby and Skilton 1995, p. 52)

This verse looks like an argument for incompatibilism. But we must ask: What is being claimed to be incompatible with what? On my interpretation, the core of this argument is the claim that the availability of an impersonal, causal level of description of actions is incompatible with the cognitive appropriateness of anger.

We might try to formalize part of the argument in *BCA* VI as follows:

1. Anger towards entity X is cognitively appropriate only if X is the independent, autonomous initiator of its actions.
2. That entity X is the independent, autonomous initiator of its actions is incompatible with X's agency being constituted by impersonal causal processes.
3. The agency of every sentient being is constituted by impersonal causal processes.
4. Therefore, anger towards sentient beings is never cognitively appropriate.

From this perspective, it is irrelevant whether the impersonal, causal processes appealed to in describing how X's agency is constituted are deterministic. What matters is that there is no basic, irreducible self. Since there is no such self, we are not responsible in any way that would generate basic desert or sustain the appropriateness of anger.

Even some readers who find these ideas somewhat compelling may be reluctant to publicly proclaim such a message. Far too many people find themselves leading chaotic, violent and miserable lives due to their systematically poor choices. And the nonexistence of free will seems to be the last thing they should hear. Dennett tells us that if one is confronted with someone who admits he is a ‘despicable villain’, one should urge him to change his ways and reform his life; but,

“[i]f, on the other hand, one was interested in compounding his misery (at whatever cost to society), one could urge on him a vision of his own utter degradation and helplessness, and foster in him an attitude of apathy and fatalism, thereby achieving (perhaps) an almost complete dissociation of his reflection from his deeds and projects and encouraging in him a cynical tolerance of his own worst side” (Dennett 1984, p. 167).

This concern is intuitively plausible, and psychological evidence supports it. For example, Baumeister, Masicampo and DeWall set out to test a hypothesis they expressed, in part, as follows:

“Feelings of responsibility and accountability may make people feel that they ought to behave in socially desirable ways, such as performing prosocial acts of helping and restraining antisocial impulses to aggress against others. The deterministic belief essentially says that the person could not act otherwise, which resembles a standard form of excuse (‘I couldn’t help it’) and thus might encourage people to act in short-sighted, impulsive, selfish ways.” (2009, p. 261)

These authors' experimental results supported their hypothesis. Subjects who read statements supportive of determinism subsequently expressed, in a survey, diminished willingness to help others (p. 262). In a second experiment, the authors found that people who held long-term views about the absence of free will volunteered fewer hours when asked to help someone in need (p. 265). Finally, they found that people who had read statements supportive of determinism were more prone to aggression, in the form of putting extra hot salsa in the food of people who had expressed a dislike of spicy tastes (p. 266). The authors' interpretation was that those who had been primed to take a deterministic stance that negated free will would be less likely to exert effort to inhibit their anger.

Against these findings we may set the well-established fact that our intuitive mechanisms for attributing responsibility are skewed in a highly morally problematic way. Psychologists have known for decades that we suffer from a pervasive cognitive distortion called 'the fundamental attribution error' (Ross 1977). It involves a strong tendency to attribute others' mistakes and morally problematic choices to their character, even when these are due to unfortunate circumstances. We cut ourselves far more slack: if I make a mistake, I will look for features of the situation that might causally explain my poor decisions, rather than attributing them to deep-seated character patterns.

Suppose we acknowledge the following. First, anger towards others can be very destructive to our relationships, our happiness, and our spiritual path. Second, given the psychology of us ordinary people, holding ourselves responsible for our own choices can be crucially important in

motivating ourselves to do better. Third, we are biased in our own favor, too often regarding our mistakes as caused by circumstances while being harshly judgmental towards others' mistakes. For someone with that kind of personality, the most spiritually beneficial attitude to cultivate about the issue of free will would correct the mistakes we are most prone to make, and so would need to be asymmetric along two dimensions: self versus other, and good versus bad actions.

Remarkably, if we read Śāntideva's *BCA* in a philosophically naïve way, the picture we get has this double asymmetry. On the resulting view, I should hold myself fully and robustly responsible for my destructive actions: "Therefore, this is just what I deserve, I who have caused distress to other beings" (*BCA* VI.42; Crosby and Skilton 1995, p. 53; see *BCA* II *passim*, especially II.29 and II.54). But I should take no responsibility for my own good actions and virtues: "As a blind man might find a jewel in heaps of rubbish, so too this Awakening Mind has somehow appeared in me" (*BCA* III.27; Crosby and Skilton 1995, p. 22).

Meanwhile, I should not hold others responsible for their wrongdoing, seeing this as non-culpable because it results from a combination of causes and conditions (the main thrust of *BCA* VI.22-33; Crosby and Skilton 1995, pp. 52-53) and was brought about by the madness of emotional reactivity: "When, driven insane by their defilements, they resort to killing themselves, how is it that not only have you no pity but you become angry?" (*BCA* VI.38; Crosby and Skilton 1995, p. 53) Yet I should rejoice in the virtuous character traits and actions of others (e.g. *BCA* III.1-3; Crosby and Skilton 1995, p. 20) and express this delight in conversation: "One should speak of others' virtues in their absence, and repeat them with pleasure" (*BCA* V.76; Crosby and Skilton 1995, p. 40). In short: "One should acknowledge

oneself as having faults and others as oceans of virtues” (*BCA* VIII.113; Crosby and Skilton 1995, p. 98).

Considered as philosophy, a view with this structure is extremely implausible. It runs aground on the very consideration Śāntideva uses to refute our innate inclination to selfishness: the undeniable fact that others and I are fundamentally alike (see *BCA* VIII.94-96; Crosby and Skilton 1995, p. 96). But as medicine for self-cherishing, such a view could have important advantages. The doubly asymmetric view is distorted, but directly opposed to the distortions of our asymmetrical biases. Cultivating the doubly asymmetric view could therefore potentially correct these serious flaws in our makeup.

But if our innate attitudes about responsibility and the doubly asymmetric view bend in opposite directions, what kind of mental attitude would count as straight? For Buddhists, the undistorted condition, free from error, is a state of wakeful presence. In such a state, there is no sense of being a person who acts; rather, actions emerge through spontaneous responsiveness and are thus appropriate to the situation and expressive of compassion for all who suffer, without any feeling of agency or choice, and entirely free from the illusion of free will.

This state of wakeful presence also involves certain emotional attitudes towards others. These are summed up beautifully by the *Vimalakīrti-nirdeśa-sūtra*:

“Like Mount Sumeru, you are unmoved by honor or scorn.

You love moral beings and immoral beings equally.

Poised in equanimity, your mind is like the sky.

Who would not honor such a precious jewel of a being?" (Thurman 2008, p. 14)

In this sky-like equanimity, there is no room for resentment (BCA VI.64; Crosby and Skilton 1995, p. 55). Instead of resentment, an awake person would experience boundless and equal compassion for all sentient beings, regardless of their past actions.

To be in this state of wakeful presence does not require endorsing hard determinism; indeed, many Buddhists would say the fully awake do not have philosophical views. But to recommend this state, to hold it up as an ideal, is to endorse a version of hard determinism that commends the abandonment of resentment, and of the interpersonal reactive attitudes in general.

It is, of course, perilous for someone like me to offer any description of an awakened state. But I think the claims I am making about that state are well motivated textually and philosophically. Frankly, would any Buddhist really assert that a buddha would spend her time consumed by guilt over her own past actions, or seething with resentment at another's?

As regards their standing within a Buddhist understanding of free will, the different components of the doubly asymmetric view are not on a par. When that view tells us to hold ourselves deeply responsible for our mistakes and others for their virtues, it is giving advice we will eventually need to transcend. But when the doubly asymmetric view tells us to let go of pride that flows from taking responsibility for our accomplishments and virtues, to view the mistakes of others

with compassion, and to abandon resentment by recognizing the lack of ultimate responsibility, it is giving advice that describes aspects of the goal of the Buddhist path.

This account reveals the insights and the limitations of Daniel Breyer's 'Buddhist Perspectivalism', which is asymmetric along one of the two dimensions I have discussed: he recommends that each practitioner regard herself as fully responsible for her choices, and others as not responsible (2013, p. 377). How could we justify this seemingly irrational view? If we assume Buddhist anti-realism, Breyer suggests, this will be easy: since there is no way the world is, all views are justified by their role in enabling successful practice, and for the reasons discussed above, an asymmetric view of free will can enable successful practice.

If we assume instead Buddhist realism, our task will be harder. Breyer attempts, with limited success, to refute the arguments I (and others) have offered for the claim that Buddhist realists are committed to determinism.¹ Breyer suggests they should instead hold that the existence of morally significant freedom is highly improbable, but physically and epistemically possible (p. 375). Then, since an asymmetric view leads to successful practice, we are warranted both in regarding others as not responsible, guided by the probabilities, and in holding ourselves responsible, guided by the bare possibility that we really are.

Breyer's article is an important contribution, and its ideas helped me to formulate the doubly asymmetric view. But Breyer misses the reasons for thinking that any asymmetric view of free will can have only a limited and provisional role in Buddhism. For practitioners in certain stages of spiritual development, the attitudes recommended by Buddhist Perspectivalism could lead to

successful practice. But at a higher level, an even more successful form of practice can result from the full abandonment of the ascription of responsibility and the illusion of agency. Thus, any asymmetric view, including Breyer's, should be seen as a mere *upāya* (skillful means), whereas Buddhists can assert without qualification the nonexistence of basic desert and ultimate responsibility.

As I have explicated hard incompatibilism, it commits us to asserting the inappropriateness of all instances of anger and resentment. One way to cast doubt on this view would be to argue that the appropriateness of at least some instances of anger is guaranteed by the fact that the occurrence of this emotion leads, in some specifiable kinds of circumstances, to socially beneficial consequences. We need anger to achieve certain goals; so there must be one important sense in which anger is appropriate. Moreover, since both Śāntideva and I are consequentialists about normative ethics, we ought to regard this form of argument as a reasonable one.

We know, however, that Śāntideva would resist this argument. In his *Training Anthology*, Śāntideva asks and answers a relevant question:

“But suppose you are strongly motivated to bring about benefit to others, or some very weighty benefit to yourself. And suppose that although you are overpowered by anger while scolding somebody else, immediately afterwards, you make a vow not to do the same thing again from now on. If you can prevent some harm through anger, what's the problem?”

Well, there's the problem of giving that reactive pattern an opportunity to work. There's the problem of losing your compassion. And as we will see later on, there's the problem of, while cutting that, cutting your own roots. Even if your anger was helpful to that sentient being, because of the loss of the bodhisattva's compassion, great benefits to sentient beings would be lost as a consequence." (Goodman forthcoming, p. 165, my translation)

Thus, Śāntideva holds that, for a *bodhisattva* (a being dedicated to awakening for the sake of all sentient beings), the deleterious karmic and psychological consequences of angrily reproaching others always outweigh whatever benefits might result from such speech. But some readers may not share his religious and/or psychological premises, and may think that in some situations we need anger to get things done, so that its overall consequences will be positive. For them, I offer the example of a yet-to-be-invented appliance called 'the Moderately Futuristic Toaster'. What follows is a description of this interesting device.

The Moderately Futuristic Toaster will make a wide variety of ready-to-eat preparations out of bread and other ingredients: buttered toast, French toast, grilled cheese sandwiches, and many other things. It has no general intelligence and no consciousness, and its internal workings are deterministic. But it is capable of adjusting how it prepares its products in response to voice commands. However, because of whimsical and mischievous design choices on the part of its maker, it will not respond to commands spoken in a calm tone. The toaster listens only to commands that exhibit sonic properties characteristic of angry speakers. Its settings are adjustable by pushing buttons and turning dials, but this process has deliberately been made

baroque and confusing. The easiest, quickest way to get the toaster to make what you want is to yell at it.

Surely it will be widely agreed that anger towards the toaster has a purpose: you get your French toast made how you want it. But I expect most readers will agree that there's something silly about getting angry with a toaster. It may be a complex machine, but it's just a toaster. If we know how the machine works, we easily see that there is nothing to get angry at: not having a self, the toaster contains no appropriate object of anger. For the owner of a Moderately Futuristic Toaster who is hurrying to make breakfast so that she can get to work on time, the most appropriate and rational response to its peculiar design would be to *feign* anger: deliberately to put on a mock-angry voice that is convincing enough to activate the device, but without feeling any resentment towards it.

Now suppose I can't feign anger well enough to fool the Moderately Futuristic Toaster. The only way to cause the device to make toast properly would be to get angry at it. My inability to feign anger wouldn't make it the case that anger at the toaster would be cognitively appropriate, for it has no self, intelligence or consciousness.

Nevertheless, there might be desirable results if I did get angry at it, and so I would have some reason to allow anger to arise, or to arouse it deliberately. This would be a case of *rational irrationality*, in Parfit's sense; in getting angry at the toaster, even though the toaster is not an appropriate object of anger, I would be somewhat like a parent who foils the threats of a murderous robber by taking a drug that makes him temporarily irrational (Parfit 1987, pp. 6-13).

So if it turns out, for example, that in certain types of negotiations expressions of anger tend to lead to more concessions from the other side—a result found by many psychological studies (Adam and Brett 2015, p. 48)—that would not show that anger is cognitively appropriate. It would show at most that negotiators might sometimes have reason to feign anger, and if they are unable to do so convincingly, to allow anger to arise for its consequences and despite its cognitive inappropriateness. Hence, even if anger is often useful, that would not refute Śāntideva’s argument for hard incompatibilism presented above. And if we take all the relevant considerations together, it should be easy to motivate to everyone’s satisfaction the claim that cases in which anger is genuinely helpful are far rarer than cases in which it is dangerous, seductive, and destructive.

Finally, let’s consider some textual evidence from the *Bodhicaryāvatāra* that has been thought to count against my view. Near the end of the crucial passage in chapter VI in which Śāntideva argues against the cognitive appropriateness of anger, we find this verse:

“If it is argued that to resist anger is inappropriate, for ‘who is it that resists what?’, our view is that it is appropriate: since there is dependent origination there can be cessation of suffering” (*BCA* VI.32; Crosby and Skilton 1995, p. 53).

Mark Siderits cites this verse, also pointing out that “throughout the chapter on forbearance Śāntideva heaps scorn on those who respond to provocation with anger” (Siderits 2008, p. 34).

Now it is hard to deny that, to be worthy of scorn, people must be blameworthy, and to be blameworthy, they must be responsible. So, Siderits concludes,

“while the self-determination argument may prove to be a useful tool in developing forbearance toward others, it does not establish that the reactive attitudes are never justifiable. This in turn suggests that Śāntideva believes people can, at least in certain contexts, be deemed responsible for their actions.” (Siderits 2008, p. 34)

As Siderits interprets Śāntideva, by including verse VI.32, and also by criticizing those who fail to control their anger, Śāntideva indicates that there must be a perspective from which people really are responsible. The rest of Siderits’ article explicates what he takes to be the relationship between these two perspectives.

The evidence from the rest of *BCA* VI shows far less than Siderits seems to assume. Here are, I think, the passages he has in mind, or at least a fair sample of them:

“5. Even friends shrink from him. He gives, but is not honored. In short, there is no sense in which someone prone to anger is well off.” (Crosby and Skilton 1995, p. 50)

“67. Some commit offenses out of delusion. Others, deluded, grow angry...” (p. 56).

“83. How can one who is angry at the good fortune of others possess the Awakening Mind?” (p. 57).

Nowhere in *BCA* VI does Śāntideva express resentment towards those who give in to anger. He does criticize such persons, describing them as deluded, miserable, and lacking the Awakening Mind. But his way of criticizing them does not constitute a form of blaming that would conflict with hard indeterminism. These are (arguably) accurate descriptions intended to be helpful to the reader of *BCA*. Providing them does not in any way commit Śāntideva to the appropriateness of anger or the existence of basic desert.

Moreover, *BCA* VI.32 cannot mean what Siderits thinks it means. Here's the very next verse, VI.33:

“Therefore, even if one sees a friend or an enemy behaving badly, one can reflect that there are specific conditioning factors that determine this, and thereby remain happy” (Crosby and Skilton 1995, p. 53).

This verse is an expression of a hard incompatibilist conclusion. And it starts with a word for ‘therefore’ (Sanskrit: *tasmād*) (ed. Vaidya 1988, p. 94). So, on textual grounds, the preceding verse, VI.32, should not be interpreted as an expression of a perspective that contradicts hard incompatibilism.

Instead, it is a response to an objection against hard incompatibilism: that hard incompatibilism implies we should feel free to act on impulse or to abandon efforts to improve. This objection is misguided. As a descriptive psychological matter, and due to our deep-seated confusion and

irrationality, for us to accept the absence of basic desert and ultimate responsibility might turn out in practice to sap our motivation to restrain anger and cultivate goodness. But not being responsible could not give us a good reason to act destructively or abandon self-restraint and moral discipline.

Few would be fooled if the issue at stake were their own welfare. Suppose I will soon travel to a tropical country where Japanese encephalitis is endemic. Stipulate that if I fail to get vaccinated and contract the disease, I will not be ultimately responsible for this. It would be a mistake to feel guilty about it, and I would not basically deserve the resulting suffering. If we accept these claims, does that in any way weaken the reasons I have to get the vaccine? Fatalism might imply that it will make no difference whether I get vaccinated, since my fate is sealed regardless of what I do; hard incompatibilism could never imply this. Similarly, fatalism might imply there is no point trying to resist my desire to act out of anger; hard incompatibilism could never imply this.

As Śāntideva points out in VI.32, even though hard incompatibilism is true, it is still appropriate to apply remedies to overcome anger and do whatever would benefit sentient beings. If we were free from the illusion of selfhood, we would, in one sense, never ‘act’ at all; meanwhile, we would move flexibly and spontaneously in appropriate and responsive ways. Until then, since most of us aren’t sufficiently moved by reasons to act well, it can be helpful to drive ourselves towards goodness with a variety of psychologically effective techniques and tricks. For those just setting out on the path, a naïve view that uncritically accepts free will and responsibility might be the most helpful approach. For somewhat more advanced students, a doubly asymmetric view

could be valuable medicine. Even though it is never rational to get angry, for a master having to deal with students who are ordinary people, it could sometimes be helpful to feign anger, so as to get the students to shape up.

Many people fear the possibility that responsible agency might turn out to be an illusion. But liberation from suffering into compassion and joy involves many forms of abandonment, and among them is abandoning our sense of being responsible agents. In the meantime, though agency is an illusion, it does have its uses. If we cling to views of free will, they may keep us from being swept away into chaos and misery. But someday, perhaps, the time will come when we can let go of the delusion of free and responsible agency, and relax into selfless and spontaneous freedom.

Note

¹ I had argued that South Asian Buddhists thought the Buddha had ‘omniscience’ (Skt. *sarvajña*, literally ‘complete knowing’ or ‘knowing everything’), and interpreted *sarvajña* to include full knowledge of the future. But if the Buddha knows everything, the future cannot be open; there must already be a fact of the matter about what everyone will do at all future times. Breyer’s reply consists of pointing out that Dharmakīrti rejected a robust understanding of *sarvajña*, holding that all we need to claim is that the Buddha knows everything that is necessary for salvation (Breyer 2013, p. 361). However impressed we may be with Dharmakīrti for advancing this view, we should recognize that it was an unusual position among Indian Buddhists. In particular, Śāntideva, whose views I am trying to interpret, accepts a very ambitious

interpretation of *sarvajña*, one that, plausibly, entails determinism. See, e.g., *BCA* V.31; Crosby and Skilton (1995), p. 36.

References

Adam, H., and Brett, J., 2015. Context matters: the social effects of anger in cooperative, balanced, and competitive negotiation situations. *Journal of Experimental Social Psychology*, 61, pp. 44-58.

Baumeister, R., Masicampo, E.J., and DeWall, N., 2009. Prosocial benefits of feeling free: disbelief in free will increases aggression and reduces helpfulness. *Personality and Social Psychology Bulletin*, 35(2), pp. 260-68.

Breyer, D., 2013. Freedom with a Buddhist face. *Sophia*, 52(2), pp. 359-379.

Crosby, K., and Skilton, A., trans., 1995. *The Bodhicaryāvatāra*. New York: Oxford University Press.

Dennett, D., 1984. *Elbow room: the varieties of free will worth wanting*. Cambridge, MA: MIT Press.

Goodman, C., forthcoming. *The training anthology of Śāntideva: a translation of the Śikṣā-samuccaya*. New York: Oxford University Press.

Goodman, C., 2009. *Consequences of compassion: an interpretation and defense of Buddhist ethics*. New York: Oxford University Press.

Nadelhoffer, T. and Tocchetto, D., 2013. The potential downside of believing in free will (and related concepts): some preliminary findings. In: G. Caruso, ed. *Exploring the illusion of free will and moral responsibility*. Lanham, MD: Lexington Books. pp. 121-140.

Parfit, D., 1987. *Reasons and persons*. New York: Oxford University Press.

Pereboom, D., 2007. Hard incompatibilism. In: J.M. Fischer, R. Kane, D. Pereboom, and M. Vargas, eds. *Four views on free will*. Malden, MA: Blackwell.

Ross, L. (1977). The intuitive psychologist and his shortcomings: distortions in the attribution process. In: L. Berkowitz, ed. *Advances in experimental social psychology*. New York: Academic Press. pp. 173–220.

Siderits, M., 2008. Paleo-compatibilism and Buddhist reductionism. *Sophia*, 47, pp. 29-42.

Thurman, R., 2008 (1976). *The holy teaching of Vimalakīrti*. University Park, PA: Pennsylvania State University Press.

Vaidya, P.L., ed., 1988. *Bodhicaryāvatāra of Śāntideva with the commentary Pañjikā of Prajñākaramati*. Darbhanga: Mithila Institute.

4 Just Another Word for Nothing Left to Lose: Freedom, Agency and Ethics for Mādhyamikas

Jay Garfield

1 Free Will and Theodicy¹

The free will problem appears to modern Western sensibility as obvious and natural, as a perennial philosophical puzzle arising upon reflection. But it is a Western cultural/religious artifact. For it to arise, one needs a *will*, a sense of uncaused *agent causation*, and a thesis of determinism. Only then can one ask the question whether that will is deterministic or capable of causing acts without being caused.

Once the question arises one can perform the philosophical rope trick and ask whether the will is compatible with determinism, if so, how, and if not, whether to opt for libertarianism or determinism. The free will question has engendered a massive literature in the West, remaining active topics of philosophical research. (See ed. Watson 2003; Campbell, O'Rourke, and Shier 2004; ed. Kane 2005; Pereboom 2009; and Meyers 2010 for discussion of this meta-issue.) It is not my purpose to survey that literature or weigh in on these questions.

This problem never arose in Buddhism (*cf.* Davids 1898; Gómez 1975; Harvey 2007), not because Buddhist philosophers were less astute, nor because they solved it, but because the

presuppositions that raise these questions are not satisfied in Buddhism. Such considerations may lead us to a hermeneutic distance that allows us to see the problem as peculiar, and set it aside in favor of more productive inquiry. I hope attention to another philosophical tradition can help here. My plan is to show why the free will problem—as construed in the Christian context, and subsequently in the Western philosophical tradition—cannot arise in the context of Madhyamaka, and how concerns related to those that motivate the problem arise and are addressed within Madhyamaka.

The *will* is ubiquitous in the West, in technical philosophical, religious, and legal discourse, and popular culture. “Did you perform this act of your own free will?” we might ask when deciding whether to blame or to excuse an apparent wrongdoer, or when notarizing a document. We explain our inability to stop smoking or lose weight as weakness of will (another topic with a vast literature: see Hoffman 2008; see also the “Weakness of Will” entry in the *Stanford Encyclopedia of Philosophy*). Even contemporary cognitive science is concerned to locate and understand the will, or our conception thereof (Dennett 1984, 1992, 2003; Libet 1985, 1999; Mele 1995, 2001, 2010).

But what is the ‘will’? We do not come by the idea that we have wills through observation of ourselves or of others. Try introspecting and finding a will. What does it feel like? Nor is it the theoretical posit of any science. Nor has it always been in Western intellectual history that persons took themselves to have wills—faculties of action. Despite the influential translation of *akrasia* (non-control) as *weakness of will*, Aristotle never identified a faculty of will.

The will is the legacy of St. Augustine and his struggle to solve the theodicy problem raised by the Fall of Adam and Eve. If God is the omniscient cause of all, then God caused the Fall, but then God would not be omnibenevolent, since he punished Adam, Eve, and the human race, consequently. To preserve God's omnibenevolence, thought Augustine, Adam's and Eve's disobedience had to be authored by them, not God. Augustine posited a faculty of uncaused (free) action to show how *that* could be the cause, and argued that only action produced by that faculty is morally praiseworthy or blameworthy; all other behavior, being heteronomously caused, is mere natural event. (Stump 2001)

This linkage of morality, personhood, and freedom runs through Aquinas to Kant, grounds Enlightenment political and legal theory, and infuses high and popular culture with a presupposition of the reality of the will and its freedom. It leads us to presuppose we are persons insofar as we are free, and that responsibility requires freedom.

Many take the Fall seriously, and are sanguine about the foundation of this aspect of our culture and the nest of philosophical problems it motivates. But this genealogy of metaphysical freedom suggests need for reassessment. If there are reasons to worry about free will, and reasons to ask what Mādhyamikas think about it, Christian theodicy is not a resource.

2 Why Worry about Free Will?

We do not worry about free will because of cultural concerns with theodicy. The motivations for most modern thought on free will are twofold. The first is metaphysical—to understand agency and personhood, and the distinction between what we *do* and what *happens*. The second—closely connected—is ethical and legal. We distinguish between actions for which we are responsible and events for which, though we may be causally implicated, we are not responsible. This distinction is often taken to be that between free will and caused behavior. Determinism is taken to be a threat on both fronts. All of this is well-worn territory, and none of my reflections are original. It is, however, useful to recall what is at stake in this discussion before turning to Madhyamaka.

Consider the metaphysics. When we take ourselves to be agents, we take ourselves to be capable of directing our actions, choosing between alternatives, acting for *reasons*; we don't view our behavior as *caused* by external events.² This authority over actions makes us who we are. But choice seemingly requires the alternatives we consider are accessible, and that deliberation is an effective consideration of reasons for each—not a sham to which we are spectators. Multiple alternatives are genuine possibilities just in case one can choose any of them, instead of performing causally determined behavior. So, it seems, agency requires exemption from determinism; freely willed actions cannot have sufficient causes, for these entail the explanation is outside the agent. (Plantinga 1967, p. 134)

Similar considerations are advanced in defense of responsibility. The *locus classicus* is Kant's *Foundations of the Metaphysics of Morals*. Kant argues that freedom is a transcendental condition on moral responsibility, and to think of ourselves as responsible, we must regard

ourselves as free. He argues we cannot *know* we are free, we *are* morally responsible, and hence we must *assume* we are free. This freedom is determination of the will not by *causes*, but by *reasons*. (Kant 1785)

Thus, if someone acts in a blameworthy way, and we can find no exculpatory external cause for her action, we take her action to have been free, and hold her responsible. If she is able to defend herself by appealing, say, to mental illness, or a horrible childhood, and if we agree that those are the *causes* of her behavior, we absolve her, arguing she was causally compelled, not free to act. Determination by causes undermines agency; if ethical assessment is possible, we require that agent causation be exempted from determinism. The Kantian transcendental argument joins with the premise that ethical and legal assessment must be possible to yield the conclusion that the will is free.

There is an older analysis of freedom in Locke, with ancestry in Aristotle: for an act to be free (Locke) or chosen (Aristotle) is for its cause to be the intention or desire of the agent, not for the salient cause to be *external* to the agent. This criterion is *absence of constraint*. The advantage of this approach over Kant's is that, instead of uncaused agent causation, it ties responsible agency to the kinds of causes operative: when actions are determined by our intentions they are free; when they are otherwise caused, they are not—the will is uninvolved.³ When I deliberate and decide I'm better off dead and jump from the window, I act freely; when you toss me from the window, my defenestration is unfree: the relevant distinction is captured by the most salient proximal cause of my exit.

This view is problematic, however. Two considerations challenge its ability to articulate a robust free/unfree distinction sufficient to underwrite agency and responsibility. First, suppose you don't toss me from the window, but threaten to torture my children if I don't jump. Plausibly, I jump freely, preferring that to living with the knowledge that my children are being tortured. But, conversely, I did not freely commit suicide; I was driven to it by threat: you caused my death. There is no need to figure out which intuition is better; they are both robust, and a clean distinction may not be forthcoming.

Second is the problem of extended causal chains. I worry about the future of the world, what with global warming, etc., and decide to end it all. It looks free, and if this is an immoral decision I appear responsible because it is free, and neither caused by force nor coerced. The cause is my desire to avoid suffering, but if that desire is uncaused, it is a random occurrence for which I cannot be responsible, and if caused, it must be caused by prior events, many of which (global warming, etc.) lie outside me. I fail to meet the conditions of agent causation. These considerations motivate an uncaused will acting purely on reasons and imply that anything that counts as a genuine action (for which we are responsible) cannot be causally determined.

Freedom seems necessary for these categories of personhood, and inconsistent with determinism. But as Schopenhauer claimed in *Essay on the Freedom of the Will* (1841)—an argument Dennett rediscovered (1984), encapsulating ideas developed by Frankfurt (1969) and Davidson (1980)—free will is not only compatible with, but *demand*s, determinism. Thus, when we say an action is free, we mean it is *caused* by our desires. If not, it would not be *ours* and we would not be responsible for it. When we want freedom, we don't want our bodies and mouths moved

randomly; we want to *cause* them to move as we desire. We want those desires to be caused; randomly occurring conative states do not make us *free*: they make us *insane*—not *responsible*, but *excusable*. We want our desires—the proximal causes of our actions—to be caused by our beliefs, traits, etc., which are desirably caused, and so on. Freedom is not *absence* of determination, but *self*-determination. It is not inconsistent with determinism, but entails it.

Are we chasing our tails? Freedom requires determinism; determinism entails our actions spring from chains of causation originating outside us; authority and responsibility require agent-causation through choice among real alternatives (Plantinga 1967; Pereboom 2009). These three premises suggest freedom in the moral-responsibility-entailing sense is impossible. It seems talk of free will and determinism may be a dead end, a relic of a theodicy of little interest to Buddhists.⁴ I think a Madhyamaka account of action might give us what we care about: making sense of our moral lives and our agency.

This set of problems presupposes another metaphysical doctrine anathema to Buddhists—that of a self or soul as center of agency. For Augustine, the soul's existence requires no argument, and for many of his successors, such as Kant, while the idea of a transcendental subject or agent may require argument, that argument was provided. While these ideas are independent (unfree souls and soulless free agents are conceivable), the soul doctrine is connected to the Western idea of an autonomous entity, capable of initiating its actions, and ultimately responsible. We can then draw the distinction between responsible human agents and non-responsible animals on the basis of soul-possession, and provide the hope for reward or punishment on the ground of the soul's responsibility and post-mortem survival. Since a distinctive feature of Buddhist philosophical

systems is their *rejection* of the view that there is a metaphysical soul (the doctrine of *anātman*), there is no Buddhist basis for formulating the free will problem.

3 *Pratītyasamutpāda* in Action: Why These Problems Cannot Arise for a Mādhyamika

Given that what motivates the free will and determinism puzzle are these Western assumptions, *this* problem cannot arise in Buddhism. A fundamental tenet of Buddhism is *pratītyasamutpāda*, the idea that all phenomena are ‘dependently originated’. In Madhyamaka, following Candrakīrti (1992, 2003), this is glossed three ways.

(1) “All phenomena arise and cease in dependence on causes and conditions.

All things arise in dependence on causes and conditions, and this is the meaning of dependent origination.” (*Prasannapadā* 2b)⁵

(2) “All wholes and parts are interdependent.

Although both from the standpoint of reality and from that of everyday life,

... a chariot cannot be established,

In everyday life, without analysis

It is designated in dependence on its parts....

If the chariot were not to exist,

Without that which possesses parts, there would be no parts either.

Just as when the chariot is burned, there are no longer any parts,

When the fire of understanding consumes the chariot, it consumes its parts as well.”

(*Madhyamakāvātāra* VI: 159-161)

(3) Entities depend for existence on conceptual imputation; mereological interdependence is related to this (*Madhyamakāvātāra* VI: 169), a theme taken up in subsequent verses and comments in this text. (Garfield 1994; Cowherds 2010)

“...Therefore, although dependent origination is generally maintained to be dependence upon conditions..., [t]his is not inconsistent with it also being dependence upon mundane nominal conventions.... In this context, to be recognized in everyday life, the conventional designation is clearly understood without the slightest bit of analysis necessary.” (*Madhyamakāvātāra-bhāṣya*, 259)

The universality of *pratītyasamutpāda* ensures that persons, psychophysical states, and actions arise interdependently. The emptiness or insubstantiality of persons—absence of soul/self—is a central moral and metaphysical insight on the Buddhist path. As Candrakīrti put it:

“In the same way, although in everyday life, the self is maintained to be

The appropriator of the aggregates, it is designated on the basis of

The aggregates, the sensory domains and the six sense faculties.

The appropriated taken as the object and the self as the agent.” (*Madhyamakāvātāra* VI: 162)

“Since it does not exist, it is neither continuous
Nor discontinuous, neither arisen nor ceased;
It has no properties such as permanence,
Existence, nonexistence, identity or difference.” (VI: 163)

“The self is simply whatever it is towards which
Beings constantly develop the attitude of ego-grasping.
The self arises out of the attitude that something is *mine*.
Since it becomes manifest unreflectively, it arises from confusion.” (VI: 164)

So long as one takes oneself to be a substantial center of subjectivity or agency, as opposed to a causally connected stream of momentary psychophysical phenomena, one is mired in primal confusion that makes the cultivation of compassion and the liberation from suffering impossible. Only by recognizing that our identities arise from our imposition of unity and coherence on a complex, multifaceted stream of events and processes can we escape that confusion. (Siderits 2005; Garfield 2010)

This position entails that all actions, thoughts, intentions, and character traits are causally dependent, and any unity we ascribe to ourselves is imputed. Thus, any ethical assessment is of caused events or merely conventionally designated persons. Libertarian agent causation is incoherent in this framework. Buddhist psychology posits no faculty for action—no ‘will’.⁶ (Siderits 1987, 2008) With no theodicy problem to solve, ‘will’ is unnecessary. Actions,

according to Buddhist psychologists, are caused by intentions, but this causation does not require mediation by any special conative faculty. So without an agent, a category of uncaused events, and a will, the ‘free will problem’, and its question of compatibility with determinism, cannot be formulated.

However, Mādhyamikas such as Nāgārjuna, Candrakīrti, and Śāntideva are committed to the view that we are responsible for our situations and destinies: this is the *karma* doctrine. Buddhist texts such as *Ratnavālī*, *Catuhṣatakatikka*, and *Bodhicāryāvatāra* are replete with admonitions to perform or refrain from various actions, and accounts of mental episodes as the primary causes of actions. Reconciling this with *pratītyasamutpāda* is a concern to Mādhyamikas.

4 Some Bad Arguments for Supposed Buddhist Doctrines of Free Will

That it’s impossible to formulate the free will thesis in a Buddhist framework has not stopped recent Buddhist philosophers from doing so. These attempts are motivated by the desire to present Buddhism as a ‘modern’—Western—doctrine that comes to similar conclusions. Here is a brief sampler. Meyers (2010) offers an extensive survey and critique. Davids (1898), Potter (1963), and Bodhi (Ñāṇamoli and Bodhi 1995) argue that *pratītyasamutpāda* is consistent with free will because dependent origination is not deterministic. The idea is that dependent origination specifies only that conditions *occasion* events, not that they *cause* them: they somehow *give rise to* events, but do not *necessitate* them. Given that there is no necessitation, there is room for freedom to choose. This argument looks best in the Madhyamaka context of

Nāgārjuna's critique of causal powers (Garfield 1994, 1995), but is still implausible. All Buddhist philosophers give universal scope to the *pratītyasamutpāda* thesis that 'when this arises, so does that; when this fails to arise, so does that', that any event can be completely explained by reference to prior and simultaneous causes and conditions. In opening his *Mūlamadhyamakakārikā*, Nāgārjuna rejects arising from nothing at all.

A cousin to this argument is Griffiths' (1986) claim that dependent origination entails that causes or conditions are necessary for the arising of events, but not sufficient: initiation of an action may require many conditions, but these are not sufficient. An act of free will is required. This is implausible. First, there is no textual evidence of a Buddhist doctrine of necessary but insufficient conditions, or of the necessity of an uncaused act to potentiate action. But second, even if this constituted a rational reconstruction of Buddhist action theory, it would be incoherent because the posited act of free will would require conditions, and those would be insufficient, resulting in a self-defeating regress.

Jayatileke (1963) argues that *karma* necessitates free will. *Karma* (in the sense of *karmaphala*) is the reward or punishment for action, which would be unjust if action were caused.⁷ Since *karma* is central to Buddhist ethics and a doctrine of reward and punishment, Buddhist ethics requires free will. But many who disagree about Buddhist ethics agree that karmic consequence is *not* reward or punishment, but purely causal consequence (Keown 2001; Goodman 2009; Garfield 2010). Justice is inapplicable to a billiard ball moving after struck by a cue ball. Thus, *karma* cannot motivate free will.

Payutto (1990) argues that since all Buddhist schools agree that action is caused by *cetanā*, which can be translated as *choice*, choice is essential to Buddhist action theory; since choice entails ability to opt for alternatives, Buddhists are committed to free will. *Cetanā* is difficult to translate (Meyers 2010), but broad consensus is that its central meaning is *intent*, *wanting*, *volition*, none of which entails *choice*. Even if it meant *choice*, it would require heavy lifting to argue that Buddhist understanding of choice is causeless. While this is only a sample of arguments to this conclusion, it suggests that it is hard to generate a discourse of freedom in the libertarian, Augustinian, or Kantian sense in Buddhism. For an account of agency and responsibility in Madhyamaka thought, we must look elsewhere.

5 Madhyamaka and Persons: The Two Truths

Central to Madhyamaka philosophy is the doctrine of the two truths⁸ (Newland 1992, 2009; Cowherds 2010). Many Buddhist schools distinguish between two truths, in different ways for different purposes. Sautrantikas and Vaibhāṣikas argue that conventional reality is erroneous because it comprises composite entities, which they regard as illusory, fabricated: forests aren't real; trees are; trees aren't real; leaves and trunks are; and so on. But, for them, these conventional entities reduce to ultimately real, simple, momentary, causally interacting micro-constituents of reality, *dharmas*. Things that might be conventionally true about wholes (say, the persistence of the person) are false, but reduce to true claims about *dharmas* (the momentary micro-constituents of persons).

Perhaps some measure of conventional freedom and responsibility reduces to ultimately impersonal causal processes (Siderits 1987, 2008), but the Madhyamaka account of the two truths neither exempts the conventional from dependent origination nor leaves us with an ultimate truth comprising fundamental constituents. Madhyamaka is not reductionist (Garfield 2006). For Madhyamaka, *nothing* exists ultimately, and to say truly that anything exists is to say that it exists conventionally. As Nāgārjuna puts it,

“That which is dependent origination

Is explained to be emptiness.

That, being a dependent designation

Is itself the middle way.

There does not exist anything

That is not dependently arisen.

Therefore there does not exist anything

That is not empty.” (*Mūlamadhyamakakārikā* XXIV: 18, 19; Garfield 1995)

(See Garfield 1995; Cowherds 2010.) An account of agency and responsibility in Madhyamaka only addresses the realm of dependent origination, of conventional truth.

This applies to persons, and it is the freedom, responsibility and agency of persons we consider. Candrakīrti, in *Madhyamakāvatāra-bhāṣya* (1992; see also Huntington and Wangchen 1989), argues that the person is neither identical to the psychophysical aggregates, nor different from

them, not one of them, nor the collection, nor the owner, controller or possessor of them. None of these can be made intelligible. Instead, the person is a conceptual imputation, a convenient designation, with no reality apart from that designation. Candrakīrti (1992) puts it this way in

Madhyamakāvātāra:

“The self is not the aggregates; and the aggregates
Are not the self. If there were any difference
Between them, such ideas would make sense.
But since there is no such difference, these are just ideas.” (VI: 142)

“The self cannot be maintained to be the possessor of the body;
Because the self does not exist, it cannot be the possessor of anything.
Only where there is difference can there be possession, as when one has a cow.
Or without difference, as in the possession of the body; but the self is neither different
nor non-different from the body.” (143)

“The self is not the body; the self does not possess the body;
The self is not in the body; the body is not in the self;
All four aggregates are to be understood in this fourfold way” (144).

“Therefore, the basis of self-grasping is not an entity.
It is neither different from the aggregates nor the essence of the aggregates.
It is neither the basis of the aggregates nor their possessor.

Instead, it is posited in dependence on the aggregates.” (150)

“The self can thus be said to be no different from a chariot.

It, in the same sense, is neither different from, nor identical with its parts.

Nor does it possess its parts; it does not contain them, and they do not contain it.

Nor is it the mere structure or mereological sum of its parts.” (151)

We are, as Dennett (1992) put it, ‘centers of narrative gravity’. That is not to say that persons or their actions do not exist, but rather to say that our mode of existence is conventional, imputed. (Garfield 2006; Newland 2009)

If Mādhyamikas are to ascribe agency and responsibility, or engage in moral evaluation, they will ascribe agency and responsibility to nominal entities, evaluating actions without ultimately existent agents. One might despair of any discourse of ethics and agency within this framework. But even Mahāyāna Buddhism—perhaps especially Madhyamaka—has the path to liberation at its core, and that involves the cultivation of moral qualities and a commitment to the welfare of all sentient beings, expounded by Śāntideva in *Bodhicāryāvātāra*.

6 Agency and Responsibility in Madhyamaka

For a Mādhyamika, selves are *constructed* through the appropriation of aggregates, through recognizing a body, thoughts, values, dispositions, and intentions as *mine*. In turn, those physical

and cognitive processes are constructed in relation to that appropriating self. That appropriation and narration of a life is not a solo affair. We narrate and construct each other in the hermeneutical ensemble act of social life. (See Hutto (2008), and Bogdan (2011), but as Nehamas (1985) notes, this idea goes back to Nietzsche.) None of us is innocent in our creation; but none of us is *autonomous* in it. Our identities are negotiated, fluid and complex because marked by the three universal characteristics of impermanence, interdependence, and absence of self. It is context-governed interpretive appropriation—not autonomous, substantial selfhood—that sets the metaphysical and moral questions regarding agency and responsibility in Madhyamaka.

To act is for our behavior to be determined by reasons, by motives we regard as ours. For Madhyamaka, it is for the causes of our behavior to be part of the narrative that makes sense of our lives, as opposed to being part of the vast uninterpreted milieu in which our lives are led. This distinction is not *metaphysical* but *literary*, and so a matter of choice, sensitive to explanatory purposes. That means the choice is not *arbitrary*. We can follow Nietzsche here. For what do we take responsibility and for what are we assigned responsibility? Those acts we interpret—or others interpret for us—as ours, as constituting part of the basis of imputation of our identities.

When I consider jumping out the window to avoid living through global warming, etc., the conditions that motivate my act are cognitive and emotional states I take as mine, and which others who know me would regard as mine. The narrative that constructs the conventional self that is the basis of my individuation includes them, in virtue of our psychology and social

practices. This is an action. When you toss me from the window against my will, the causes of my trajectory lie in what we would, on conventional, hermeneutical grounds, interpret as parts of *your* biography—no action of *mine*. The agency lies with *you*, not on metaphysical, but conventional grounds: not on discovery of agent causation in your will, but based upon the plausible narrative we tell of the event and of each other's lives as interpretable characters.⁹

The interesting questions concern intermediate cases of coercion, where you threaten my children with torture if I do not jump. There are two ways to take this case, and many ways to construct a narrative here. In one, I am the passive victim of your blackmail; here we read the causes of my jumping as your actions, not mine: agency is assigned to you, not me. In another, I make a sacrifice in the face of circumstances beyond my control. Here we explain the jumping by reference to my character and desires, locating agency in me, not you. In a more nuanced story we say that while I may not be responsible for the circumstances that forced me to make the sacrifice, when faced with the hard choice I made it. How to choose between narratives in particular legal or moral discourses is interesting and difficult. But the Madhyamaka point is that in asking how best to tell this story and where to assign agency, we are never forced to look to a will, its freedom, or agent causation.

On this Madhyamaka understanding of personal identity established through imputation—a view with affinities to Hume and Nietzsche—we *do* make choices, perform acts that merit moral assessment, assign responsibility to agents for their actions and absolve others, and assess acts morally. But none of this requires talk of free will. These practices can be better understood in the framework of *pratīyasamutpāda*: a choice occurs when we experience competing motives,

consider alternative reasons, some of which could, if dominant, occasion alternative actions, and one set of reasons dominates, *causing* the action, and was *caused to cause the action* by our background dispositions, cognitive and conative states. Some actions are expressive of and conducive to virtue, happiness, liberation and the welfare of others and merit praise, others not. But there need be no more to it.¹⁰ What the post-Augustinian libertarian West buys with the gold coin of free will—at the expense of its metaphysical problems—is bought by the Mādhyamika more parsimoniously with the paper currency of mere imputation.

7 Freedom on the Path; Freedom from *Samsāra*

Mādhyamikas do talk about freedom—from suffering, cyclic existence, the *kleśas* (maladaptive psychological processes). This is a kind of freedom: of actions—mental, verbal, physical—from determination by those aspects of our personality we wish to write out of the narrative. Many of my actions are driven by fear, anger, despair, greed, etc., states I appropriate or others assign me as part of my biography. Consequently, I interpret a great deal of the events I participate in as occasioned by the acts of others—those who threaten, annoy, or compete with me, and react to them accordingly. This is the basis of vice. (Garfield 2010)

The path to liberation, for a self that is mere conceptual imputation, is a path to the authorship of a narrative in which a better self is the protagonist, a self whose actions are conditioned by compassion, sympathetic joy, generosity and confidence, by responsiveness as opposed to reaction. The self I imagine at the higher stages of the path is free in ways the self I construct

now is not. More of its acts are actions it claims to author, and the conditions of those actions are morally salutary rather than counterproductive.

However, the freedom achieved through the cultivation of this path, understood in the Madhyamaka framework of Candrakīrti and Śāntideva, is not a freedom of the *will*, but *authority*—freedom of a conceptually imputed person from the bars of a self-constructed prison, a freedom that demands no indeterminism. And when that freedom is complete, there is nothing left to lose.

Notes

¹ Mādhyamikas are followers of Madhyamaka, later ‘Middle Way’ Buddhism. This paper is a condensation of a longer version for which I express my thanks to the publisher for permission to reproduce here: “Just Another Word for Nothing Left to Lose: Freedom, Agency and Ethics for Mādhyamikas”, in *Free Will, Agency and Selfhood In Indian Philosophy* (Oxford: Oxford University Press, 2014), edited by Matthew R. Dasti and Edwin F. Bryant, pp. 164-85 (www.oup.com). Thanks to Edwin Bryant, John Connolly, Matthew Dasti, Bronwyn Finnigan and Karin Meyers for very helpful comments on earlier drafts.

² That is, external to the *self*, the agent, not the *body*. Being caused to act by a device implanted in the brain or disease would count as externally, *heteronomously*, determined. This parallels the distinction between being determined by *reasons* versus *causes*. So, following Locke and Kant, an action is thought free insofar as we can provide reasons for it, and it is for those reasons that

we undertake it. An action is unfree insofar as it is caused, and its causes are not reasons we give for acting. This distinction is transformed in the Madhyamaka account of responsibility below.

³ Such positions are *compatibilist*. Baker (2003) contrasts libertarianism and compatibilism positions clearly: libertarians require agency not be ultimately exogenous, whereas compatibilists permit it, depending on the details.

⁴ This is not to say the post-Fall theodicy is the only root of puzzles about agency, but that in the West this root gives them their character, and it is plausible that there are traditions in which they do not arise or arise differently.

⁵ All translations are my own, from Tibetan.

⁶ Or anything corresponding to Augustinian *voluntas* or Kantian *Wille*.

⁷ Augustine (1993) makes an argument of exactly this sort.

⁸ The Sanskrit '*satya*' is ambiguous between the English 'truth' and 'reality', an ambiguity not salient in Sanskrit philosophy.

⁹ Not all narratives are equally good. Some make sense of our lives; some are incoherent, facile and self-serving, profound, or revealing. It is possible to disagree about whether a particular event is an action, or about the attribution of responsibility, to wonder about whether we should feel remorse for a particular situation. These are questions about which narratives make the most sense. While these questions may not always be easy (or possible to settle), that they arise saves this view from the facile relativism that would issue from the observation that we can always tell *some* story on which this is an action of mine, and *some* story on which it is not, so that there is no fact of the matter, and perhaps no importance to the question.

¹⁰ This also indicates why, on a Madhyamaka view, persons can be held responsible for actions even when they involve no explicit choice. Choice is not necessary in this account.

References

Augustine, 1993. *On free choice of the will*. Translated by T. Williams. Indianapolis, IN: Hackett.

Baker, L.R., 2003. Why Christians should not be libertarians: an Augustinian challenge. *Faith and Philosophy*, 20(4), pp. 460-478.

Bogdan, R., 2011. *Our own minds: sociocultural grounds for self-consciousness*. Cambridge, MA: MIT Press.

Campbell, J.K., O'Rourke, M. and Shier, D., eds., 2004. *Freedom and determinism: a framework*. Cambridge: Cambridge University Press.

Candrakīrti, 1992. *Madhyamakāvatāra-bhāṣya (dBu ma 'jug pa'i rang 'gral)*. Sarnath: Kagyud Relief and Protection Society.

Candrakīrti, 2003. *Prasannapadā. (dBu ma tshig gsal)*. Sarnath: Gelukpa Student Welfare Committee.

Cowherds, 2010. *Moonshadows: conventional truth in Buddhist philosophy*. New York: Oxford University Press.

Dasti, M.R. and Bryant, E.F., eds., 2014. *Free will, agency, and selfhood in Indian philosophy*. New York: Oxford University Press.

Dauids, R.C., 1898. On the will in Buddhism. *Journal of the Royal Asiatic Society of Great Britain and Ireland*, pp. 47-59.

Davidson, D., 1980. How is weakness of the will possible? In: Davidson, D., *Essays on actions and events*. Oxford: Oxford University Press. pp. 21-42.

Dennett, D., 1984. *Elbow room: the varieties of free will worth wanting*. Cambridge, MA: MIT Press.

Dennett, D., 1992. *Consciousness explained*. Cambridge, MA: MIT Press.

Dennett, D., 2003. *Freedom evolves*. New York: Penguin.

Frankfurt, H., 1969. Alternate possibilities and moral responsibility. *Journal of Philosophy*, 66(23), pp. 829-39.

Garfield, J., 1994. Dependent origination and the emptiness of emptiness: why did Nāgārjuna start with causation? *Philosophy East and West*, 44, pp. 219-250.

Garfield, J., 1995. *Fundamental wisdom of the middle way: Nāgārjuna's Mūlamadhyamakakārikā*. New York: Oxford University Press.

Garfield, J., 2006. Reductionism and fictionalism: comments on Siderits. *APA Newsletter on Asian and Comparative Philosophy*, VI(1), pp. 1-8.

Garfield, J., 2010. What is it like to be a bodhisattva? Moral phenomenology in Śāntideva's *Bodhicāryāvātāra*. *Journal of the International Association of Buddhist Studies*, 33, pp. 33-57.

Gómez, L., 1975. Some aspects of the free will question in the Nikāyas. *Philosophy East and West*, 25, pp. 81-90.

Goodman, C., 2009. *Consequences of compassion*. New York: Oxford University Press.

Griffiths, P., 1986. *On being mindless: Buddhist meditation and the mind-body problem*. LaSalle, IL: Open Court.

Harvey, P., 2007. 'Freedom of the will' in the light of Theravāda Buddhist teachings. *Journal of Buddhist Ethics*, 14, pp. 271-291.

Hoffman, T., 2008. *Weakness of the will from Plato to the present*. Washington, DC: Catholic University of America Press.

Huntington, C. and Wangchen, G.N., 1989. *The emptiness of emptiness: Candrakīrti's Madhyamakāvātāra*. Honolulu, Hawaii: University of Hawai'i Press.

Hutto, D., 2008. *Folk psychological narratives*. Cambridge, MA: MIT Press.

Jayatilke, K., 1963. *Early Buddhist theory of knowledge*. London: Allen and Unwin.

Kane, R., ed., 2005. *The Oxford handbook of free will*. New York: Oxford University Press.

Kant, I., 1785. *Groundwork of the metaphysics of morals*. Translated by M.J. Gregor, 1998. Cambridge, MA: Cambridge University Press.

Keown, D., 2001. *The nature of Buddhist ethics*. New York: Palgrave.

Libet, B., 1985. Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences* 8(4), pp. 529-539.

Libet, B., 1999. Do we have a free will? *Journal of Consciousness Studies*, 6, pp. 47-57.

Mele, A., 1995. *Motivation and agency*. New York: Oxford University Press.

Mele, A., 2001. *Autonomous agents: from self-control to agency*. New York: Oxford University Press.

Mele, A., 2010. *Effective intentions: the power of conscious will*. New York: Oxford University Press.

Meyers, K., 2010. *Freedom and self-control: freedom in South Asian Buddhism*. Ph. D. University of Chicago.

Ñāṇamoli, B. and Bodhi, B., trans., 1995. *The middle length discourses of the Buddha*. Boston, MA: Wisdom.

Nehamas, A., 1985. *Nietzsche: life as literature*. Cambridge, MA: Harvard University Press.

Newland, G., 1992. *The two truths*. Ithaca, NY: Snow Lion.

Newland, G., 2009. *Introduction to emptiness*. Ithaca, NY: Snow Lion.

Payutto, P., 1990. *Freedom, individual and social*. Bangkok: Buddhadhamma Foundation.

Pereboom, D., ed., 2009. *Free will*. Indianapolis, IN: Hackett.

Plantinga, A., 1967. *God and other minds*. Ithaca, NY: Cornell University Press.

Potter, K., 1963. *Presuppositions of India's philosophies*. Englewood Cliffs, NJ: Prentice Hall.

Schopenhauer, 1841. *Essay on the freedom of the will*. Translated by K. Kolenda, 2005. Mineola, NY: Dover.

Siderits, M., 1987. Beyond compatibilism: a Buddhist approach to freedom and determinism. *American Philosophical Quarterly*, 24, pp. 149-159.

Siderits, M., 2005. Freedom, caring and Buddhist philosophy. *Contemporary Buddhism*, 6, pp. 87-116.

Siderits, M., 2008. Paleo-compatibilism and Buddhist reductionism. *Sophia*, 47, pp. 29-42.

Stump, E., 2001. Augustine on free will. In: E. Stump and N. Kretzman, eds. *The Cambridge companion to Augustine*. Cambridge: Cambridge University Press. pp. 124-127.

Watson, G., ed., 2003. *Free will: Oxford readings in philosophy*. New York: Oxford University Press.

5 Negative Dialectics in Comparative Philosophy: The Case of Buddhist Free Will

Quietism

Owen Flanagan

“You sound to me as though you don’t believe in free will,’ said Billy Pilgrim.

‘If I hadn’t spent so much time studying Earthlings,’ said the Tralfamadorian, ‘I wouldn’t have any idea what was meant by free will. I’ve visited thirty-one inhabited planets in the universe, and I have studied reports on one hundred more. Only on earth is there any talk of free will.’” (Vonnegut 1969, p. 86)

Negative Dialectics

‘Negative dialectic’ (Adorno 2003) expresses the idea that not all pressures to debate a topic through the famous, canonical process of thesis and antithesis result in a positive synthetic outcome. I think the request that Buddhism speak about and declare a position on our problem of free will is likely a case where the dialectic will yield an unfortunate result, stating a position of a problem that is better left alone, as the culturally idiosyncratic topic it is. There is a consensus, with a few notable exceptions (Davids 1898; Gómez 1975; Harvey 2007) that classical Buddhism does not theorize what we in Western analytic philosophy and philosophical theology call the ‘problem of free will’. I am wary of asking

Buddhists to talk about our problem of free will because it is a bad and idiosyncratic problem.

The situation is a bit like that of Vonnegut's Tralfamadorians who have been around the cosmos and find discussion on free will only on earth. Even on earth it is a rare problem. Abrahamic lineages talk about free will. Confucians, Daoists, Platonists, Aristotelians, Stoics, Cynics, Epicureans, Hindus, Jains, and Buddhists don't.

Since the greater part of this volume contains arguments for what the Buddhist view of free will is or would be if they had one, take my paper as proposing caution, as an exploration of the possibility that this particular cross-cultural discussion might incur the cost of asking a tradition to entertain and discuss, even to make a place for, a bad idea that it did not previously or traditionally entertain. Here is my view in outline.

1. Discussion of the problem of free will in Western analytic philosophy and the Abrahamic theological traditions—probably in the philosophical case because it inherits the problems of the theological case—requires agent causation to be a contender, a thesis in the dialectic.
2. Any dialectic that starts with powerful pressure to treat the idea of agent causation as a serious option, a thesis, is off to a bad start. It has promise for being a negative dialectic.
3. Buddhism has no need for the agent causation thesis, and never entertains it.

4. Because agent causation is a bad idea that never stops giving, it is best not read into or imported into Buddhism.
5. Therefore, don't import it into Buddhism.

The Hermeneutic Situation

One compelling piece of evidence that something is fishy or, what is different, that something is off in the discussion of Buddhism and free will comes from observing the set of answers considered by the Western philosophers who ask about the Buddhist view on free will, who claim to find or think they are entitled to say what Buddhists say, think, or should say or think about free will. These readings all occur inside a consensus that there is no theorizing inside classical Buddhism about free will (Garfield 2014). Here is a brief tour.

- 1) Charles Goodman's original view (2002) was that Buddhism is a form of hard determinism. Goodman's recent view (this volume) is that it might be useful, a matter of *upāya*, skillful means, given the make-up of certain contemporary Buddhists, raised, for example, in cultures where believing in hard determinism would be depressing, to have them believe or make-believe that they have free will.
- 2) Mark Siderits says Buddhism is a kind of compatibilism, 'paleo-compatibilism': free will exists at the conventional level; it does not exist at the ultimate level (2008).

- 3) Rick Repetti (2010) says Buddhism is a kind of sophisticated hierarchical compatibilism *a la* Frankfurt (1971).
- 4) Paul Griffiths (1982) says Buddhists hold (would hold) a libertarian/agent causation view: “Buddhology ... as formally identical with Christian theology” (1994, p. 182; see also Wallace 2011).
- 5) Galen Strawson says Buddhism is either a kind of hard determinism/incompatibilism or otherwise committed to the falsity or, what is different, incoherence of the idea of free will (2010).
- 6) Sam Harris (2012) thinks Buddhism holds that free will is an illusion.

What is going on here? What explains the fact that very smart, conscientious philosophers claim that every position ever entertained in Western philosophy is either there or can be extracted from Buddhism? My argument here situates the source of the lack of consensus to a perfect storm comprised of indifference of the Buddhist tradition to our problem of free will, projection of various preferred solutions to our problem of free will onto some one among the types of Buddhism (hereafter, “Buddhisms”), and lack of clarity and carefulness about method in comparative philosophy.

Topics Better Under- or Un-theorized?

Imagine someone asks: What is the analytic philosophical consensus on karmic rebirth, on the question of whether sentient beings recycle for eons as other sentient beings, animal and human, and that they do so based on the moral quality of their lives? One answer is that there is no consensus inside analytic philosophy about karmic rebirth because karmic rebirth is not a topic we talk about. Another answer, perhaps more intellectually honest, is that it is an implication of things we do talk about, and have opinions about, that there is no such thing as karmic rebirth. In either case, it might be judged best for us not now to talk about karmic rebirth, at least not very much, and not as a sustained and serious topic of discussion in contemporary metaphysics, in theorizing now about what there is and how it is.

This sort of philosophical ethnocentrism would not be the end of the matter. It leaves room for karmic rebirth as a central topic in the history of philosophy, in philosophical anthropology and in comparative philosophy, as part of the important project of getting what other great wisdom traditions have thought about the nature and fate of persons, as well as understanding important folk philosophical views, which are still in the blood and bones of many Hindus, Jains, and Buddhists, and which animate their morals, their folk metaphysics, and their soteriological thinking. And obviously it is a sign of a healthy critical philosophical environment to always be willing to wonder and discuss whether ideas that other communities have explored might be good for us now, in our current situation.

The situation with the topic of karmic rebirth is similar to our topic of free will in the other direction. Free will is a central topic in Western philosophy since roughly the time of St.

Augustine and largely for reasons internal to a certain set of theological and eschatological assumptions. Free will is a topic we were led to inside a certain very particular, parochial language game. Philosophizing inside the Abrahamic tradition(s) compelled asking the question what a person must be like for an all loving, all knowing, all good creator God to justifiably, consistent with his infinite goodness, eternally reward or punish individuals for how they lived.

Here are some canonical statements of the kind of capacity that persons must have in order to solve this problem. Each articulates the libertarian view of free will, which sure looks like the only option that can answer the question in a way that keeps God from being capricious, even cruel. Libertarianism, or agent causation, is not of course the only view of free will in Western analytic philosophy, but it historically defines the shape of our problem.

“But the will is so free, that it can never be constrained... And the whole action of the soul consists in this, that solely because it desires something, it causes a little gland to which it is closely united to move in a way requisite to produce the effect which relates to this desire.” (Descartes 1649, p. 41)

“If we are responsible ... then we have a prerogative which some would attribute only to God: each of us when we act, is a prime mover unmoved. In doing what we do, we cause certain things to happen, and nothing—or no one—causes us to cause those events to happen.” (Chisholm 1964, p. 32)

“Free will ... is *the power of agents to be the ultimate creators or originators and sustainers of their own ends and purposes* [W]hen we trace the causal and explanatory chains of action back to their sources in the purposes of free agents, these causal chains must come to an end or terminate in the willings (choices, decisions, or efforts) of the agents, which cause or bring about their purposes.”

(Kane 1998, p. 4; italics his)

This puffed up view of human agency is invariably on stage as a live option as soon as our topic of free will is. Of course, the dialectical situation typically involves trying to tame this option, making the world safe for compatibilism, for example. The trouble is that no view other than the libertarian view will solve the theological problem. And that problem haunts our discussions of free will. It comes as the inevitable penumbra of our problem of free will. It is always in the shadows as long as our problem of free will is in view.

Here is Nietzsche’s admittedly polemical diagnosis of the *raison d’être* of our conversation, which I think is more or less on the mark:

“Might it not be the case that that extremely foolhardy and fateful philosophical invention, first devised for Europe, of the ‘free will’ of man’s absolute freedom (*Spontaneität*) to do good or evil, was chiefly thought up to justify the idea that the interest of the gods, in man, in man’s virtue, *could never be exhausted*” (Nietzsche 1887, 2nd essay, p. 7).

“The *causa sui* is the best self-contradiction that has been conceived so far: it is a sort of rape and perversion of logic. But the extravagant pride of man has managed to entangle itself profoundly and frightfully with just this nonsense. The desire for ‘freedom of the will’ in the superlative metaphysical sense, which still holds sway, unfortunately, in the minds of the half-educated; the desire to bear the entire and ultimate responsibility for one’s actions oneself, and to absolve God, the world, ancestors, chance, and society involves nothing less than to be precisely this *causa sui* and, with more than Baron Münchhausen’s audacity, to pull oneself up into existence by the hair, out of the swamps of nothingness.” (Nietzsche 1886, §21)

This—“desire for ‘freedom of the will’ in the superlative metaphysical sense”—is not a need the Buddhisms have. There is no creation *ex nihilo*. And there is no omnipotent, all loving, all good God; nor is there eternal reward or punishment. The question of how persons must be, what they must be like, for God not to be capricious and evil is not a question internal to the Buddhisms. There is of course a series of rebirths that track moral quality, but in Buddhism quality of rebirth, as well as eventual release from suffering, is a matter of the operation of impersonal causal laws. Finally, there is no evidence from psychology that anything like the “desire for ‘freedom of the will’ in the superlative metaphysical sense” is a normal or natural human desire absent a certain background theological or philosophical framework.

If this is right, it provides a much-too-quick and admittedly superficial explanation for why the Buddhisms do not theorize free will. The philosophical tradition, core metaphysics, epistemology, and ethical thinking that comprise the Buddhisms, does not require the kind of dialectic that generates our problem of free will. This helps explain why there is close to consensus among scholars that across the varieties of Buddhism there is no direct or systematic reflection on the problem of free will as we conceive it, that the concepts we use in Western philosophy to discuss the topic, forget about full-on theorizing, are not present in traditional East, South, and Southeast Asian Buddhisms. My diagnosis is that there isn't systematic reflection on the problem of free will because Buddhists never had the internal philosophical-theological problems that engendered such theorizing for us. And thus English translations of the Pali Canon or of commentaries in Pali, Sanskrit, Chinese, Korean, Japanese and so on, do not use terms that clearly mark our topic of free will.

Some analytic philosophers think it unfortunate that the Asian Buddhisms do not have views, let alone theories, about free will. My view is that it is good—minimally good dialectically—that Buddhists do not have theories about free will. It shows, among other things, that wondering about free will is not as inevitable as wondering about the furniture of the universe, about chickens, goats, rivers, mountains, fire and rain. Free will is a theoretical object, an explanatory device, not a basic phenomenon or feature of the world in need of explanation. Once free will is theorized as something that humans uniquely must be possessed of in order to accept a certain theological and eschatological picture, specifically an all good, loving, and powerful God who doles out eternal reward or damnation, then there is a lot of further explaining to do about how it, free will, is possible,

and thus the free will industry in Western analytic philosophy is born and then fed. But free will really is something about which Buddhist traditions are unopinionated or, as likely, something which—if we find ways of speaking of it inside the traditional Buddhism—will seem odd, awkward, unsuited, possibly incoherent or meaningless, perhaps, at the limit, just an empty idea. That said, if I were a Buddhist comparative philosopher I would be very curious to understand what we think free will is and why we care so much about it. I have just sketched how that answer will look and where it will come from.

There is an immediate and formidable objection to the tactic I am hinting at, the tactic of leaving well enough alone, where what is well-enough is that Buddhism avoids one of the black holes of Western philosophy, the topic of free will. The objection is straightforward: There are internal and external questions that can be asked of any philosophical tradition. Inside Greek atomism, the question of the nature of subatomic particles doesn't arise, unless, of course, such particles are the atoms. But once atomism meets elementary particle physics, the atomist must speak, offer an account, of phenomena—protons, electrons, etc.—we, the believers in subatomic particles, claim are there, particles that split what was formerly thought to be unsplitable.

One of the amazing, beautiful, and helpful things about the spaces where theories, scientific or philosophical, meet is that they are called to account for what their mates, neighbors, or competitors see or theorize as real, as mattering. So, suppose it is true that the classical Asian varieties of Buddhism do not theorize free will. These varieties now have Western branches, largish footprints in North America and Europe. And thus we are allowed to ask

our questions of the tradition. So we ask what do Buddhists think about free will? Are you/is the tradition (which strand?) committed to libertarianism, compatibilism, weak free will, or strong free will? What about responsibility, heavy or light? We will ask and reconstruct answers that suit us, Buddhism's answers to our questions, answers that are at a minimum not inconsistent with Buddhism. Furthermore, I am certain, it is a guaranteed consequence of smart people reading vast tracts of ancient, hard-to-decipher, often poetic speech that they will find answers to their questions: what the Buddha thought about free will, what he would have said about free will. Seek and ye shall find here; oh, also, you will find every answer, as many articles in this book attest. And thus I realize that I am fighting a rear-guard action insofar as we will ask these questions and demand answers.

The dialectical situation is one in which we can learn a lot. There is a similarity to lessons we might be in the process of learning from the vast amount of recent work asking for Buddhism's ethical theory, which is sometimes based on the lament that for all the richness of the tradition, it does not declare whether it is a virtue theory or a kind of consequentialism or deontology. But note what is assumed here, that a unified ethical theory is desirable, a sensible demand for a philosophy. Analytic philosophers have recently assimilated the Buddhisms to varieties of Aristotelian, consequentialist, and even Kantian ethics. The fits don't work perfectly. There are two directions one can go here. One can lament that Buddhism doesn't have a unified ethical theory. Or, one can wonder whether this latter fact might teach us something important about the ethical domain, for example, that it resists unified theorizing, that our impulse to theorize moral life is the problem, not Buddhism's under-theorizing.

An even better example pertains to the particular question of human rights. The current Dalai Lama is one of the foremost proponents of universal human rights. But as Christopher Kelley (2015) has shown in his important dissertation, there is absolutely no discussion of rights and no conceptual space for Enlightenment-style inalienable rights inside Buddhism. The language of 'intrinsic' and 'inalienable' can find no comfortable home inside Buddhism to warrant agreement that there really are such things as inalienable rights. That said, there could be, indeed there is, a *modus vivendi*, a practical agreement on universal human rights, that Buddhists sometimes lead. But the best way to read the situation is not that Buddhists really believe in universal human rights in the ways and for the reasons that we children of Kant do. Not at all. They don't. And the lesson is not that ultimately we can force them to agree to our conception, to conceive of rights as we do. We cannot, and they cannot (and won't). What we can gain is some agreement from polite people to behave as if they share a consensus, which for all practical purposes they do. The situation might be similar in the case of free will to the extent that there is agreement about how persons ought to feel, think, and be (loving, compassionate), as well as the conditions under which moral and legal institutions will hold people accountable without any shared theory about free will.

Four Frameworks for Comparative Philosophy

One way to think through our question is to think of four kinds or approaches to comparative philosophy or, what is different, four roles for comparative philosophy; call

these *Classical*, *Fusion*, *Cosmopolitan*, and *Liberal* (Flanagan 2011). They are not mutually incompatible.

Classical comparative philosophy tries to enter into a form of life, typically through its canonical texts, and say in an idiom we can comprehend how the world view, form of life, is (sometimes was) for the people who live inside that world, that set of practices and traditions, and who abide the ways of feeling, thinking, and being expressed in the world that those texts describe or aspire.

Fusion comparative philosophy takes two or more classically well-understood traditions and asks what you get when you bring them together. Is it something good, new and interesting, or it is un-blendable like oil and water, or a bad mix like lemon juice and milk? Who would have known that French and Vietnamese cuisine could go so well together? So we work, let us suppose, at trying to mix the Indian virtue of *ahimsa*, non-violence, with American foreign policy discourse to see if something good happens from blending these seemingly foreign ingredients together. Perhaps it does; perhaps it is like oil and water or lemon and milk, a bad mix in different ways for different reasons. Or, we ask, as many East Asians are now asking, whether classical Confucianism and liberal democratic theory can be blended into something new, improved, even better than either alone. Some fusions are conceptually impossible if one makes certain demands on maintaining the integrity of the original ingredients; for example, one cannot be a Christian Trinitarian and a Muslim, and both of these are incompatible with Greek or Hindu polytheism.

Cosmopolitan comparative philosophy is open to, attuned to, opportunities for *fusion*, but looks in the first instance for a solution to this problem: In many great cities, think of NYC and London for starters, people from very different traditions intersect on a daily basis. There is, or needs to be, at a minimum, a *modus vivendi*. For that, it is good to ‘get’ each other, to understand what we are each doing, feeling, and thinking. Sometimes this will involve something very much like *classical* understanding. The Polish Catholics ‘get’ how their Sufi neighbors think and vice versa, although it all seems very alien on both sides, and there is no fusing or merger. Other times, there might be fusions. The African Americans, Irish, and Italians in Queens see among their Chinese neighbors a kind of respect for elders, *xiao*, which they think would make family and civil life go better if they could get some such virtue or set of understandings and practices into the blood and bones of their kids, possibly using Confucian reasons, possibly using reasons closer to home that are now lost or hidden.

Liberal comparative philosophy can make use of *classical*, *fusion*, and *cosmopolitan* methods, but it responds to this sort of situation, and applies especially to the present case of Buddhism and free will. Imagine a young Jewish woman from Brooklyn (I have many such friends who call themselves ‘Bu-Jews’) who is also a Ph. D. candidate in philosophy at the CUNY Graduate School, and who is now a convert to Buddhism. (Imagine, unlike many Americans who say they are Buddhists, she knows something about Buddhism and doesn’t think it just involves meditation.) She asks herself what Buddhism says or would say about the problem of free will. Here I picture the individual responding to the personal question of how her commitments to the letter or spirit of some version of Buddhism sits with her

other commitments, some from her Jewish upbringing, and some from her philosophical training and her exposure to the problem of free will.

Liberal allows her, and especially so in the case of Buddhism, which is a living, non-creedal tradition, to say whatever suits her so long as it does not contradict a core belief in Buddhism. Here the situation is revealing. I think she can say pretty much anything she wants about free will and carry on because the classical varieties of Buddhism are so utterly un-opinionated about the matter. If the shared core across the varieties of Buddhism (Flanagan 2011) consists of the Four Noble Truths, the Eightfold Path, and the four immeasurable virtues (compassion, loving-kindness, sympathetic joy, and equanimity), then every position on free will is compatible. If the shared core also includes, as I think, commitment to the metaphysics of dependent origination, impermanence, and no self, the situation may be the same, although I think agent causation might be conceived as an incoherent option, a nonstarter right out of the gate.

The key feature of *liberal* is that it takes a step into the space where things are not yet fixed for a living tradition. A *liberal* can say traditional Buddhists don't talk about free will, but say that they are a Buddhist, perhaps a new kind of Buddhist: a Westerner born inside an Abrahamic religion, trained in analytic philosophy, now trying to be a Buddhist, and wanting to think through what they are allowed to say about the problem of free will. They get, when they wear their *classical* hat, that the topic is not discussed in traditional South, Southeast, or East Asian Buddhism, and they get when they wear the *fusion* and/or *cosmopolitan* hats, that free will is not a problem that their Thai Buddhist neighbors here in

Brooklyn need help from them on. This person is trying to find out for herself, and other members of this new community, what they can allow themselves to permissibly think or say about a problem they have because of where they are coming from.

Detecting Frameworks

With this matrix of frameworks in place, we can ask what is going on with the discussion of Buddhism and free will in general and each particular view listed in the inventory of positions above. Is a scholar claiming that there are discussions of free will in canonical texts, in the Pali Canon, say? Extracting such discussions would be something *classical* could do if the discussions are actually there, internal to the tradition. Perhaps scholars are claiming that the problem of free will must be there, has to be there, because the problem of free will is visible to everyone everywhere in the way rivers and mountains and rabbits are, in which case *classical* will necessarily unearth such a discussion (just as we better find terms and concepts for rivers and mountains and rabbits), even if we don't see it at first. It is a hermeneutical axiom that it must be there!

We can't be sure whether when the natives say 'gavagai' that they are referring to rabbits, or undetached rabbit parts, or rabbit time slices, but they better say something when those thingamajigs we call 'rabbits' do the thing we call 'hopping' in the things we call 'fields' and 'meadows' (Quine 2013, pp. 23-72). Or, are scholars claiming that *fusion* applies, and that there are some new, more insightful things we can say than the Buddhists say when they

are *not* talking about free will, but are talking about something that reminds us in a free-associative way of our problem of free will, and which if fused with our talk about free will would either improve their original thinking about their non-free will problem and/or ours about our free will problem? Or are inquirers asking, also consistently with *fusion*, that some new and improved set of solutions to our problem of free will emerge from blending Buddhist sources and our sources? As for the last question, if this is the thought, I see no evidence from the literature that any Buddhist insight has usefully advanced our discussions of free will.

Then there are more *cosmopolitan* observations and questions that are not a proper subset of *fusion* observations and questions: It is interesting that whereas, let's suppose, we Catholics think a lot about free will and in terms of free will, my Thai Buddhist neighbors don't think this way. It is kind of amazing to me that they have managed to think about virtue and goodness, reward and punishment, even about afterlives following a similar logic to the way I think without that conceptual apparatus. I wonder if it is necessary for me to think about, wonder and worry about, free will the way I do? Here one could be wondering about what ordinary folk thought about free will or something spiffed up like talk of the problem in Christian philosophical theology or analytic philosophy.

The *liberal* comparativist, unlike the *cosmopolitan*, but maybe like some advocates of *fusion*, is a convert, and she is trying to find the space to move forward with integrity in the light of her new commitments. She brings questions she already knows about, cares about, has conceived her life in terms of, and asks them now in the light of her new core philosophy.

The only constraint is not to say something inconsistent with the new set of commitments, which in the present case is not hard.

I read the dialectical situation as allowing all four sorts of inquiries, observations, and questions. But I am pretty sure that *classical* will not discover in ancient Buddhism any discussions of our problem of free will (Gowans 2015; Garfield 2014). *Fusion* might seem an attractive option if one thinks that the problem of free will is one anyone will come upon and must deal with, like rivers and mountains or rabbits, however these things are metaphysically parsed. But I don't think that. I think free will is more like the topic of original sin, which requires an entire philosophical theology to get, unless one goes free-associative, in which case any tradition (everyone does this) that has things to say about human weakness has a view on original sin.

Finally, I see the need for and room for the *liberal* response among converts who have their feet in the Buddhist tradition as well as in the ones, e.g., analytic philosophy or Christian philosophical theology, that make free will a problem. Westerners impressed by Buddhism *and* analytic philosophy and/or the Abrahamic theology, eschatology, will be of this sort.

My own preference (in the particular case of free will) is the *cosmopolitan* response with a therapeutic goal. Observe that there are smart people who were not sucked into the particular black hole of philosophy we call the 'problem of free will' and learn from them how to avoid it (Flanagan 2002). It is worth avoiding. Get over it and please do not introduce Buddhists to our problem of free will.

I have written elsewhere (2002) that, were I charged with being the benevolent dictator of philosophy, I would forbid using the words ‘free will’ in anything like the libertarian sense, which is required to get the debate going. Once the libertarian conception is in play, then compatibilism is rightly seen as what Kant called a ‘wretched subterfuge’ (1788, p. 332). Nothing good comes from discussing the philosophical problem of free will. It is the black hole of philosophy, a pseudo-problem caused by philosophers taking “language ... on holiday” (Wittgenstein 2001, § 38), a pathological detour from clear thinking engendered by the invention of a heavenly Father in the West whose regimen of reward and punishment only makes sense if we possess this oddity called ‘free will’. Aristotle never discusses free will, although he does discuss the differences between voluntary and involuntary action (1999).

And so it is with Buddhism (Flanagan 2011). There is action, wholesome and unwholesome, and how a life goes depends on the ratio of wholesome to unwholesome action. There are impersonal laws of karma; there is dependent origination; there are agents conventionally, but not ultimately, and so on. But Buddhism has no need for the concept of ‘free will’. It doesn’t need free will, in part, because it does not need agents who conform to what the God(s) of Abraham needs them to be/be like—with immutable souls and magical agentic powers. This is a great advantage.

Some smart philosophers are trying to draw Buddhism into the Western conversation about free will. This is good (I guess) insofar as it engenders cross-traditional *verstehen*.

And it is good if it forces the Western problem to dissolve or evaporate. But it is unfortunate if the agenda is to introduce the concept of free will into Buddhism or to claim that it can be extracted from Buddhism. Buddhism doesn't have the concept of free will and it doesn't need it. It is like the old song lyric, 'Got along without you before I met you, gonna get along without you now'.

This much is evidence that the classical Buddhisms are indeterminate enough to allow multiple interpretations from some one or another of the four frameworks, *classical, fusion, cosmopolitan, or liberal*—indeed every one ever invented in the West on the problem of free will. I propose the following two interpretive principles to help understand this situation. First, if very smart people can find at least six different views on any question, problem, or topic formulated in our terms in some other tradition that is by consensus silent on that question, problem or topic, then infer that it is not the tradition queried that is yielding the answers to the question, but us. So, take *classical* off the table as an explanation of what is happening in this particular inquiry, which leaves *fusion* or *cosmopolitan*. Second, think hard about why a problem that we feel with urgency must have an answer, possibly that we think every great philosophy must take a clear position on, wasn't conceived as urgent, possibly not worth theorizing at all by a great wisdom tradition.

Conclusion: Buddhism and Freedom

Charles Goodman writes,

“Buddhists have always been interested in freedom, but only recently have they begun to think about free will. Concepts closely related to freedom—spontaneity, independence, self-mastery—have been central to Buddhism since its beginnings. Serious Buddhist reflection on the problem of free will and determinism, however, is a product of dialogue between Asian and Western cultures. Unfortunately, this dialogue has barely begun, and very little is known about what a Buddhist position on free will might be like.” (2009, p. 145)

The first point to notice, and it is important to my argument, is that the topic of Buddhism and free will is the result of a dialogue, of a meeting of traditions. It is fair to say that there is not, internal to the Buddhisms, any set of texts or philosophical strands that clearly announce themselves as about our problem of free will. There is lots of discussion of what we call action and freedom but not free will. Second, Goodman rightly uses the phrase ‘what a Buddhist position might be’ to mark this fact that there isn’t one in the classical texts or traditions. Third, and it is remarkable, the Buddhist position is being articulated and offered almost exclusively by Western philosophers, not by Asian Buddhists. Fourth, the fact that Buddhism has always been interested in freedom, self-control, self-perfection, and so on, does not mean that Buddhists were always sitting next to the mine that would reveal that precious vein we call free will. The concept of free will is fraught, freighted and culturally peculiar in ways these other concepts are not.

One source of my admiration for Buddhism as a tradition is that it seems the possibility of proof that one can have a metaphysically and morally serious tradition, which, if naturalized, is science-friendly, and that does not need *our* concept of 'free will'. This much is a certain protection against philosophical projection and ethnocentrism, especially the kind that is already common, where Western analytic philosophy claims to see the right questions and sets the agenda.

One discussion that we can have across traditions, that is about a topic on which Buddhism has always been strongly opinionated, is the topic of freedom, freedom from suffering. That is an important topic, worth talking about. Here I agree with John Dewey when he writes: "What men have esteemed and fought for in the name of liberty is varied and complex—but certainly it has never been metaphysical freedom of the will" (1957, p. 303). But here in this book I find myself part of a conversation I wish we were not having, and that I strongly advise against having. So let's next time talk about cross-cultural views about freedom from suffering, freedom of conscience, freedom from oppression, and freedom from economic injustice. These are serious and real problems. But they have nothing, exactly zero, to do with the problem of free will.

References

Adorno, T., 2003. *Negative dialectics*. New York: Routledge.

Aristotle, 1999. *Nicomachean ethics*. Translated by T. Irwin, 1999. Indianapolis, IN: Hackett.

Chisholm, R., 1964. *Human freedom and the self (Lindley lecture)*. Kansas: Department of Philosophy, University of Kansas.

Davids, R., 1898. On the will in Buddhism. *Journal of the Royal Asiatic Society of Great Britain and Ireland*, pp. 47-59.

Descartes, R., 1649. *The passions of the soul*. Translated by S.H. Voss, 1989. Indianapolis, IN: Hackett.

Dewey, J., 1957. *Human nature and conduct*. New York: Henry Holt and Company.

Flanagan, O., 2002. *The problem of the soul: two visions of mind and how to reconcile them*. New York: Basic Books.

Flanagan, O., 2011. *The bodhisattva's brain: Buddhism naturalized*. Cambridge, MA: MIT Press.

Frankfurt, H., 1971. Freedom of the will and the concept of the person, *Journal of Philosophy*, 68, pp. 5-20.

Garfield, J., 2014. Just another word for nothing left to lose: freedom of the will in Madhyamaka. In: M.R. Dasti and E.F. Bryant, eds. *Free will, agency and selfhood in Indian philosophy*. New York: Oxford University Press. pp. 164-85.

Gómez, L.O., 1975. Some aspects of the free will question in the Nikāyas. *Philosophy East and West*, 25, pp. 81-90.

Goodman, C., 2002. Resentment and reality: Buddhism on moral responsibility. *American Philosophical Quarterly*, 39(4), pp. 359-372.

Goodman, C., 2009. *Consequences of compassion*. New York: Oxford University Press.

Gowans, C.W., 2015. *Buddhist moral philosophy: an introduction*. New York: Routledge.

xx

Griffiths, P.J., 1982. Notes toward a Buddhist critique of karmic theory. *Religious Studies*, 18, pp. 277-291.

Griffiths, P.J., 1994. *On being Buddha: the classical doctrine of buddhahood*. Albany, NY: SUNY Press.

Harris, S., 2012. *Free will*. New York: Free Press.

Harvey, P., 2007. 'Freedom of the will' in the light of Theravāda Buddhist teachings. *Journal of Buddhist Ethics*, 14, pp. 35-98.

Kane, R., 1998. *The significance of free will*. New York: Oxford University Press.

Kant, I., 1788. *The critique of practical reason and of judgment*. In *Great Books of the Western World, Vol. 42, Kant*, 1952. Chicago: Encyclopedia Britannica.

Kelley, C., 2015. *Toward a Buddhist philosophy and practice of human rights*. Ph. D. Columbia University.

Nietzsche, F., 1886. *Beyond good and evil*. Translated by H. Zimmern, 2015. Brooklyn, NY: Sheba Blake.

Nietzsche, F., 1887. *On the genealogy of morals: a polemic. By way of clarification and supplement to my last book beyond good and evil*. Translated by D. Smith, 1997. Oxford: Oxford University Press.

Quine, W.V.O., 2013. *Word and Object*. New ed. Cambridge, MA: MIT Press.

Repetti, R., 2010. Meditation and mental freedom: a Buddhist theory of free will. *Journal of Buddhist Ethics*, 17, pp. 166-212.

Siderits, M., 2008. Paleo-compatibilism and Buddhist reductionism. *Sophia*, 47(1), pp. 29-42.

Strawson, G., 2010. *Freedom and belief*. 2nd ed. Oxford: Clarendon Press.

Vonnegut, K., 1969. *Slaughterhouse-five, or the children's crusade: a duty-dance with death*. New York: Dell.

Wallace, B.A., 2011. A Buddhist view of free will: beyond determinism and indeterminism. *Journal of Consciousness Studies*, 18(3-4), pp. 217-33.

Wittgenstein, L., 2001. *Philosophical investigations*. Malden, MA: Blackwell.

6 Free Will and the Sense of Self

Galen Strawson

1 Feelings and the Causality of Reason: Doings and Happenings¹

Suppose one comes to believe that no one is ever ultimately morally responsible for what they do in such a way as to truly deserve praise or blame, punishment or reward. One realizes that ultimate self-determination (“USD”) is impossible; one endorses skepticism about ‘strong free will’ (“SFW”) of the sort that could ground USD. One accepts that many of what P.F. Strawson (1962) calls one’s ‘personal-reactive’ attitudes to others (gratitude, resentment, anger, admiration, etc.) are in some fundamental respect inappropriate or unjustified, inasmuch as they seem to presuppose SFW.

What then? It depends on what kind of person one is. To think that certain of one’s feelings and attitudes presuppose SFW, and to come to believe that SFW is impossible, may possibly cause one to cease to have these feelings. It seems clear, however, that most people are incapable of giving up the personal-reactive attitudes to others, even when convinced of the impossibility of SFW. Our deep commitment to belief in SFW is not so much irrational as *non-rational*, as P. F. Strawson observes (Strawson 1962). On his view, the true ground of our commitment to belief in SFW lies in our deep susceptibility to the emotional reactions that seem to depend on SFW for their appropriateness. Instead of being supported, they support.

Still, suppose that coming to believe in the impossibility of SFW does cause a change in one's personal-reactive attitudes. Suppose one comes to adopt what P. F. Strawson calls the 'objective attitude' to all human actions (Strawson 1962). If so, this will be something that simply happens to one, and there seems to be a further question about what one should actively do. One can hardly decide to take no notice of what one now believes—that people, including oneself, are never ultimately responsible for their actions. But if forming this belief hasn't caused one's reactive attitudes to change, isn't one then bound to *try* to stop treating people as proper objects of gratitude and resentment, praise and blame?

Say one doesn't want to. Isn't that a sufficient reason not to? It's not clear that it is; or rather, it's not clear that these questions really arise. Suppose one believes there are okapi in San Diego zoo, but hasn't been to check. There's no reason why one should check if one doesn't want to. But the present case is different. One has formed a belief, and nothing remains to be checked. To say that one needn't try to take into account the fact that people aren't really proper objects of reactive attitudes if one doesn't want to seems like saying that one needn't believe something one believes if one doesn't want to. But one doesn't have such a choice; belief isn't subject to the will in this way.

True—and yet people who reach the theoretical conclusion that SFW is impossible rarely seem much perturbed. Their lives continue more or less as normal. Is this acceptable? Shouldn't they do something about it? Can facts about our natural non-rational commitment to believe in SFW somehow justify, as well as explain, their imperturbability, or somehow pre-empt the need for

any such justification? The question is pressing because the argument against SFW seems watertight (see, e.g., Strawson 1994). There's an immovable sense in which we neither are nor can be ultimately responsible for what we do. But we go on thinking of ourselves as if we are.

The argument against SFW doesn't depend on the view that determinism is true; it shows that SFW is equally impossible even if determinism is false. But for simplicity I'll consider the case in which one thinks (rightly) that SFW is incompatible with determinism, and then comes to believe that determinism is true.

What should one do? What might one do? What might happen? It's no good saying to oneself 'I'm determined to go on believing in SFW, and in having these SFW presupposing feelings and attitudes'. To think that this dissolves the problem is to make the mistake of fatalism—the mistake of thinking that nothing one can do can change what will happen. One may be so determined that one does make this mistake, but it's still a mistake, and people who think clearly won't make it. It looks as if such people can't avoid the problem of what now to think, what now to do.

2 Determinism, Action, and the Self: A Thought Experiment

At this point one faces the fact that the heart of one's commitment to belief in the self-determining self doesn't lie in the *interpersonal* reactive attitudes but in certain *self-concerned* reactive attitudes. One's deepest commitment is to the view of oneself as radically self-

determining and responsible. If one tries to overcome this commitment, it seems that one risks thinking oneself out of existence as a mental agent-self, a *mental someone*—where ‘mental someone’ is used to characterize an absolutely central way in which we *experience* ourselves. (I’m not concerned with the question of what if anything a ‘mental someone’ could possibly be, only with the phenomenology of self-experience.)

Why? Because what one naturally takes oneself to be, considered as a mental someone, is a truly self-determining agent of the impossible kind. One takes it that this is an *essential* aspect of what one is: I feel that I—what I most truly am—couldn’t continue to exist and lack this property. So it’s not just as if concentration on the impossibility of SFW may cause me to cease to believe I have a certain property—radical responsibility—whose possession means a lot to me. It’s rather that there may remain nothing that is recognizable as me at all, nothing recognizable as me, the ‘agent-self’, only a bare consciousness-function.

A thought experiment may help. It consists in the rigorous application of the belief in determinism to the present course of one’s life. One does one’s best to think rapidly of every smallest action one performs or movement one makes—everything that happens, so far as one is oneself concerned—as determined as not, ultimately, determined by oneself; contemplate this for a minute or two, say.

* * *

This should have the effect of erasing any sense of the presence of a freely deciding and acting

'I' in one's thoughts; for—it appears—there is simply no role for such an 'I' to play. It may be strangely, faintly depressing, or it may give rise to a curious, floating feeling of detached acquiescence in the passing show of one's own psychophysical being, a feeling, not of impotence, but of radical uninvolvedness. Or the feeling may be: I am not really a person; there isn't really anyone there at all. The thought experiment may make a good exercise for certain schools of Buddhists.

I take this to indicate the respect in which one's natural pre-theoretical sense of self has a strongly libertarian cast, although it also has strong compatibilist elements (Strawson 1986, §§6.3–6.5). One naturally and unreflectively conceives of oneself, *qua* the mental planner of action, as capable of USD. The reason one disappears in the thought experiment is that it reveals that one isn't possible, so conceived.

At the same time, of course, one doesn't—can't—disappear just like that. One's thought naturally and inevitably occurs for one in terms of 'I', and one's conception of this 'I' remains, so far, a conception of a truly responsible self-determining someone. One's attempts to grasp the consequences of determinism fully may succeed in bursts, but they will keep breaking up on the rock of one's natural commitment to a self-conception which is simply incompatible with fully fledged belief in determinism.

When this happens, one may try to continue to think that everything about one is determined, but it won't be striking with full force. And when it isn't striking with full force it probably won't make it seem that one doesn't really exist at all as an agent-self. It's more likely to make it seem

that although one does somehow still exist as a mental someone, and does continue to act in various ways, still one cannot truly be said to do anything oneself, because determinism gobbles up everything, revealing everything one does to be not really *one's own doing*.

Seasoned philosophical compatibilists may have trouble with the thought experiment. They may find it hard to appreciate the force of these points. Perhaps they should imagine facing the following choice: if you agree to submit to twenty years of torture—torture of a kind that leaves no time for moral self-congratulation—you will save ten others from the same fate. Perhaps they should agree to be hypnotized into believing that they really are facing such a choice—hypnotized in such a way that, afterwards, they remember exactly what it felt like. They may then rerun to the thought experiment.

It seems, in any case, that there are two principal poles around which one's thought will oscillate when one is trying to apply the thought that one is totally determined. At one pole, the freely deciding and acting 'mental someone' goes out of existence. At the other pole, the mental someone continues to exist, but can no longer see itself as a freely deciding and acting being. One's thought is likely to oscillate around this second pole when the thought experiment has not been engaged with full force, and isn't having its full effect of strangely dissolving the (sense of) self.

Suppose one focuses on the fact that every moment of one's thinking is completely determined. Whatever thought one then has, one will, pursuing the thought experiment, think of that thought too as determined. And, this being so, no thought will ever be able to emerge as the true product

of the *I*, the putative true originator of thoughts, decisions, and actions. This *I* will perpetually evanesce, however far one pursues the possible regress of thoughts about thoughts about thoughts.

This is just one way in which the thought experiment may develop. Those who reach the theoretical conclusion that SFW is impossible should try it. For them, undertaking it involves nothing more than dwelling with special concentration on something they already believe to be true. Those who learn to maintain the state of mind induced by the thought experiment will be well on the way to a thoroughgoing, truly lived, or as I shall say *genuine* belief in determinism or non-self-determinability (they may be on the way to *nirvāṇa*). It is, however, important to be clear what this involves. One may theoretically accept that one is a product of one's heredity and environment in such a way that one can never attain USD—many of us do—and yet, in everyday life, have *nothing like* the kind of self-conception that is here required of the genuine skeptic about SFW.

3 What Might Happen

Consider Louis and Lucy, incompatibilists who have just come to believe that determinism is true, and hence that there is no SFW. They're struggling to attain a true perspective on their situation. How is Lucy to think of herself as she wonders whether to give money to charity—thinking perhaps of each of her thoughts and movements that it is determined, and thinking that her thinking this is determined in turn, and so on?

We may suppose that she doesn't make the fatalist mistake of ceasing to try to get what she wants because she thinks it's already determined whether she will get it or not—in such a way that she can do nothing about it. She knows her own planning and action are real and effective parts of the continuing deterministic causal process. It's rather that when she does something intentionally which she feels to be reprehensible (say), she may then think: that was determined to happen, and yet if I hadn't done it that too would have been determined. This is a very ordinary thought in philosophy. What is it like to take it seriously in life, trying to apprehend every detail of one's life as determined?

Louis may find that he feels that *he* (i.e. as he automatically conceives of himself in his natural, unreconstructed thought, an agent capable of truly deserving punishment or reward, praise or blame) can do nothing at all. Here he is at the second of the two main poles of serious self-applied determinism. He feels he exists, but that he can't really *act* at all. This is how he puts it, at least. Or rather, this is how he would put it, if it weren't for the fact that, relaxing his application of the thought of determinism to himself, and being a novice, he still feels completely responsible. He feels he simply knows that he knew at the time of action that he could have done otherwise. Both Louis and Lucy are unable to accept that they are at bottom radically exempt from responsibility or praise or blame, because what happened was determined to happen as it did. Yet they also now believe that the way they are, and their decisions, are things for which they are ultimately in no way responsible. And when, see-sawing back, they concentrate again on this thought, they find, again, that they can no longer make sense of the idea of performing actions that are truly *their* actions. For the sense of self they naturally have is simply

incompatible with any deep acceptance of the idea that all they are and do is determined.

Lucy may think as follows: for her to choose to (try to) abandon her personal-reactive attitudes is not really possible, because only a free agent, which she does not now consider herself to be, can really have a reason for action which is really its own reason. There may be a train of practical deliberation going on in her head, but she feels that it's not really *her* thought at all (although it feels just like it, as soon as she stops concentrating on her determinedness), but (because) a determined process.

Louis, equally, thinks he can't reason or deliberate in a way that culminates in a decision which is truly his, and which is such that the ensuing action is something for which he is truly responsible, or indeed something that *he* really did. For he knows that what he thinks of as his choice is determined, however much he may contrasuggestibly change his mind. And so, stuck with his unreconstructed sense of self, he can't think of it as really *his* choice. To talk of freedom here, as compatibilists do, is, both think, to talk of the freedom of the turnspit (Kant 1788, p. 191, Ak 5.97) or the self-sealing tank (Davidson 1973/1974, p. 141). It's the 'wretched subterfuge' of compatibilism, a 'petty word-jugglery' (Kant 1788, pp. 189–90, Ak 5.96), 'so much gobbledygook' (Anscombe 1971, p. 146). It's not really to talk of freedom at all.

So the picture of a thoughtful incompatibilist like Lucy coming to believe in determinism, and then raising the question of what to do about it, may be ill-conceived. The question may be completely unreal for her, so long as she concentratedly applies the thought of determinism to herself. For she may then feel that she can't really choose to do, or do, anything, in the way she

thought. This rejection of the possibility of real choice or action is, certainly, a piece of reasoning on her part. But it too can't be thought to have any practical consequences or rationalize any decision—such as a decision not to choose or decide anything on the grounds that it is strictly speaking impossible to do so. One can't decide not to decide anything on the grounds that one can't decide anything.

This is a strange drama, an enactment of the deep problem of free will. In the end only the exigencies of everyday life will carry them forward. The continual tendency of Louis's unreconstructed thought will be to reinsert *him*—him conceived as a truly praise-and-blame-deserving mental someone—into his thought and deliberation. And continually he will correct this tendency. For nothing, he realizes, can be done by him, so conceived. Nothing can be done by him in the sense that matters to him; things can only happen. So too for Lucy. Whatever she starts to plan and do, it's whipped away from *her*, only to appear as not really her own, by the thought that it's entirely determined.

It seems, then, that a genuine belief in the impossibility of SFW, uneasily coupled with an unreconstructed conception of self, may produce a total paralysis of purposive thought as ordinarily conceived of and experienced. There are more compatibilist ways of thinking and theorizing about deliberation and action that may leave us untroubled by the thought of determinism. It doesn't follow that the present story isn't accurate as a story of what might happen to a newly fledged believer in the impossibility of SFW. What may follow is that we're deeply inconsistent in our characteristically very vague thoughts about freedom, action, deliberation, and ourselves.

4 The True Center of Commitment

There are many respects in which our natural sense of self and of freedom is profoundly libertarian. There are also many respects in which it is entirely compatibilist—many aspects of our general sense of ourselves as free agents that don't seem to be put in question in any way by the impossibility of USD. This natural compatibilism is a very important part of what underlies our general commitment to belief in SFW, and this is so even though compatibilism is incompatible with belief in SFW.

A full description of our natural compatibilism would take a long time.² Here I want to focus again on the question whether our commitment to belief in SFW is grounded primarily in our experience of other people as proper objects of the reactive attitudes, or in our experience of our own agency. I think the second view is correct: the true center of one's commitment to the notion of human freedom lies in one's experience of oneself and one's own agency, one's deep sense of oneself as a self-determining planner and performer of action, someone who can create things, make a sacrifice, do a misdeed.

This is not to deny that one's experience of oneself is deeply determined by one's interaction with others. It's simply (i) to consider two things that develop in us in the course of our social development—our sense of ourselves as truly responsible and our sense of others as truly responsible, (ii) to claim that the nature and causes of these two things can profitably be

distinguished, and (iii) to claim that the former is more important than the latter, so far as our general commitment to belief in SFW is concerned.

A naturalistic explanation of our deep sense of self-determination must connect it tightly with our sense, massively and incessantly confirmed since earliest infancy, of our ability to do what we want to do in order to (try to) get what we want, by performing a vast variety of actions, walking where we want, making ourselves understood, picking up this and putting down that. We pass our days in more or less continual and almost entirely successful self-directing intentional activity, and we know it. Most of these actions are routine or trivial, more or less thoughtlessly performed, but this doesn't diminish the importance of the experience of their performance as a source of the sense of radical self-determinability that we ordinarily have. Even if we don't always achieve our aims, when we act, we almost always perform a movement of the kind we intended to perform, and in that vital sense (vital for the sense of self-determining self-control) we are almost entirely successful in our action.

This experience is central to our natural compatibilism. It gives rise to a sense of freedom to act, of complete self-control, of responsibility in self-directedness, that is compatibilistically unexceptionable and completely untouched by arguments against SFW based on the impossibility of USD. And it is precisely this compatibilistically unexceptionable sense of freedom and efficacy that is one of the fundamental bases of the growth in us of the compatibilistically *impermissible* sense of SFW. To observe a child of two, fully in control of its limbs, doing what it wants to do with them, and to this extent fully free to act in the compatibilist sense of the phrase, and to realize that it is precisely such unremitting experience of self-control

that is the deepest foundation of our naturally *incompatibilist* sense of true-responsibility-entailing self-determination, is to understand one of the most important facts about the genesis and power of our ordinary strong sense of freedom (and of self).

One reason why we advance from the permissible to the impermissible sense of freedom is perhaps negative. Ignorant of the causes of our desires or pro-attitudes, we don't normally experience them as determined in us in any way at all, let alone in any objectionable way, as Spinoza remarked (Spinoza 1677). So too for our characters. We don't think back behind ourselves as we now find ourselves. And even if a desire is experienced in its importunacy as somehow foreign, imposing itself from outside the self, this probably only serves, by providing a contrast, to strengthen our general sense that our desires are *not* determined in us. For if a pro-attitude is experienced as imposing itself, then there must be another one in the light of which the first is experienced as imposing itself, and the second one will presumably not also be experienced as an imposition. It will presumably be a pro-attitude one 'identifies' with and apprehends as part of oneself, and acquiesces in.

A great deal is locked up in this acquiescence. For although it's unlikely to involve any explicit sense that one has been in any way actively self-determining as to character, it does nevertheless seem to involve an implicit sense that one is, generally, somehow in control of and answerable for how one is, even, perhaps, for those aspects of one's character that one doesn't particularly like. As for those pro-attitudes and aspects of one's character that are welcome to one, it's as if the following ghostly subjunctive conditional lurks in one's attitude to them: *if per impossibile* I were to be (had been) able to choose my character, then these are the features I would choose

(would have chosen). This, I suggest, makes a deep contribution to the impermissible sense of true responsibility for themselves that most people have, more or less obscurely, more or less constantly.

It's hardly surprising that the ghostly subjunctive conditional confirms the central, acceptable *status quo*, for there's a fundamental respect in which the 'I' that features in it is constituted, as something with pro-attitudes that imagines choosing its pro-attitudes, by the very pro-attitudes that it imagines choosing.

But this is still not the principal reason why we have the impermissible sense of true responsibility. The principal reason is the nature of our experience of choice. It's simply that we are, in the most ordinary situations of choice, unable not to think that we will be truly or absolutely responsible for our choice. Our natural thought may be expressed as follows: even if my character is indeed just something given (a product of heredity and environment, or whatever), I'm still able to choose (and hence act completely freely and truly responsibly, given how I now am and what I now know); this is so whatever else is the case—determinism or no determinism.

To illustrate this, consider the following story. You arrive at a bakery on the evening of a national holiday. You want to buy a cake with your last ten dollars to round off the preparations you've already made. There's only one thing left in the store—a ten-dollar cake. On the steps of the store, someone is shaking an Oxfam tin. You stop, and it seems quite clear to you that it's entirely up to you what you do next. You are—it seems to you—truly, radically, ultimately free

to choose what to do, in such a way that you will be ultimately morally responsible for whatever you do choose. Look, you can put the money in the tin, or you can go in and buy the cake. You're not only completely, radically free to choose in this situation. You're not free not to choose (that's how it feels). You're fully and explicitly conscious of what the options are and you can't escape that consciousness. You can't somehow slip out of it.

You may have heard of determinism, the theory that absolutely everything that happens is causally determined to happen exactly as it does by what has already gone before—right back to the beginning of the universe. You may also believe that determinism is true. (You may know that current science gives us no more reason to think that determinism is false than that determinism is true.) In that case, standing on the steps of the store, it may cross your mind that in five minutes' time you'll be able to look back on the situation you're in now and say truly, of what you will by then have done, 'Well, it was determined that I should do that.' But even if you do fervently believe this, it doesn't seem to be able to touch your sense that you're absolutely morally responsible for what you do next.

The case of the Oxfam box is relatively dramatic, but choices of this general type are common. They occur frequently in our everyday lives, and seem to prove to us that SFW is both possible and real. The argument that SFW is impossible seems powerless when faced with this conviction, which is reinforced by the point just considered, according to which something in itself negative—the absence of any general sense that our desires, pro-attitudes, character, and so on are *not* ultimately self-determined—is implicitly taken as equivalent to some sort of positive self-determination. We certainly don't ordinarily suppose that we've gone through some sort of

active process of self-determination at some particular past time. Nevertheless, it seems accurate to say that we do unreflectively experience ourselves *rather as we would experience ourselves if we did believe that we had engaged in some such activity of self-determination.*

There are many complexities here, but the main causes of the development of our sense of true (straight-up punishment-and-praise-justifying) responsibility out of our unremitting and compatibilistically speaking unexceptionable sense of complete self-control may be summarized as follows. (1) We tend to think that we have a will (a power of decision) distinct from all our particular motives. (2) In all ordinary situations of choice, we think that we're absolutely free to choose whatever else is the case (even if determinism is true, for example), and are so just because of the fact of our full appreciation of our situation. (3) In some vague and unexamined fashion, we tend to think of ourselves as in some manner responsible for, answerable for, how we are.

All these aspects of the sense of true responsibility directly concern only one's experience of oneself and one's own agency. It's one's commitment to belief in one's own radical efficacy, control, self-determination, and total responsibility (in normally unconstrained circumstances), rather than one's commitment to holding others responsible and treating them as proper objects of reactive attitudes, that is primarily unrenounceable. What on earth is one to think that one is, or is doing, if one thinks one cannot really be responsible at all for what one does? Those who have fully understood what the application of the thought of determinism to themselves involves should be bewildered by this question. And yet it's likely to leave them undisturbed for more than a few minutes.

5 Satkāyadr̥ṣṭi

Is it impossible to abandon commitment to belief in SFW? Perhaps partial if not total erosions are possible. It seems that it isn't equally unrenounceable in all areas. One's commitment to belief in one's absolute punishment-and-praise-justifying responsibility for one's own actions seems more deeply founded than one's commitment to belief in the responsibility of others, even if this difference doesn't show up in a difference of surface strength in everyday life. It seems correspondingly more likely that one might cease to be moved to blame others, on account of loss of belief in SFW, than that one might cease to feel guilty about what one felt to be one's own wrongdoing.

But perhaps one can raise a more general doubt about arguments for unrenounceable commitments to beliefs that appear to be false from some natural point of view. Consider certain Buddhists who argue, on a variety of metaphysical grounds, that our natural notion of the persisting individual self is a delusion. Having reached this conclusion, they set themselves a task: that of overcoming the delusion.

There are several routes to the doctrine of *satkāyadr̥ṣṭi*, the 'false view of individuality'.³ The Buddhists I have in mind hold that (a), the false sense or conception of self, leads to (b), suffering, because it is essentially bound in with, as a necessary condition of, (c), the having of desires and aversions, which is itself a condition of the possibility of suffering. (b) requires (c)

and (c) requires (a), so to eliminate (a) is to eliminate (b). To realize that there's no such thing as the persisting individual self, to undermine the false view of individuality in oneself, is to cease to be bound by desires, cravings, and aversions, and hence to achieve liberation from suffering. It is, ultimately, to achieve the 'blowing out' of self in *nirvāna*, and thereby to cease to suffer and to fear old age, sickness, and death.

These Buddhists not only have theoretical reasons for believing that their natural sense of self is delusory; they also have powerful practical reasons for trying to improve their grasp of its delusoriness. They recognize, however, that one can't simply abolish one's sense of individuality, by some sort of effortless, rationally motivated, self-directed intellectual *fiat*. Delusions delude, after all, and the ordinary, strong sense of self (and self-determination) is a particularly powerful delusion. They therefore recommend engagement in certain practices whose eventual effect—so they claim—is to cause the delusion to dislimn.

A decision to engage in such meditative practices is presumably motivated by some pro-attitude. One may simply wish to live in truth—without illusion. In the Buddha's case, the originally predominant motive was a desire to overcome fear of old age, sickness, and death. Plainly one's own suffering, decrepitude, and death can be fearful only if one has a sense of oneself as an object, a continuing entity, a person, and has in addition certain desires concerning what happens to that object. If there's no 'I', there's nothing to fear in death and dissolution, because it's precisely the dissolution of the 'I' that is feared. Nor is there anyone there to feel fear, because it's precisely the 'I' that does the fearing.

It seems good, then, to grasp the non-existence of the ‘I’, since it is both the object and the necessary repository of the desires that lead to suffering. If one were to attain a state of desirelessness, one’s desire for truth or correctness of attitude, or one’s wish to escape fear of mortal ills, would lapse with all other desires, so it would no longer be there to be finally fulfilled by the course of action that it set in motion. In this sense, those who attain a state of desirelessness can never give a reason for being the way they are—if giving a reason involves adducing a present desire currently satisfied. Nevertheless, given their love of truth, or their fear of old age and death, their engagement in meditative practices *was* rational—even if they are now (practically speaking) non-rational, and are so as a result of their practice. There’s no paradox here, given that Buddhists standardly use the word ‘desire’ in the narrow sense of craving, not in the entirely general sense of ‘pro-attitude’ (Collins 1997).

This enables one kind of person, at least, to answer the question of what it would be rational to do, given belief in the impossibility of SFW. Someone who had such a belief, and wished to lose any sense of self as a radically self-determining agent, simply in order to achieve a more correct attitude to the world, might do well to adopt the allegedly self-dissolving practice of meditation. A sense of self is not only a necessary condition of fear for one’s future; it is also, obviously, a necessary condition of possession of the allegedly illegitimate sense of oneself as a radically self-determining planner and performer of action.

A decision to adopt the ‘objective’ (SFW-denying) attitude can’t be implemented overnight, but one can perfectly well initiate some practice which may more gradually undermine the supposedly inflexible constraints of the general framework of ideas within which we ordinarily

live. And if we admit the possibility of partial alterations in attitudes or habits of thought to which we are, as things are, deeply committed, this points to the possibility of a progressive abandonment of these attitudes or habits of thought which, gradually achieved, amounts to a total abandonment relative to the original position. It is not implausible to suppose that Buddhist monks and other mystics have succeeded in altering quite profoundly their experience of themselves (and of others) as acting, thinking, and feeling beings. Nor is it implausible to say that they have in so doing achieved what is in certain respects a more correct view of the world, precisely to the extent that they have utterly ceased to regard themselves and others as radically self-determining sources of actions. Philosophers who believe that SFW is impossible, and are committed to the pursuit of truth, should perhaps undertake some practice of meditation. It may enable them to come to appreciate the truth of their theoretical conclusion in a way they cannot achieve by any other means.

Notes

¹ Most of the ideas in this article originally appeared in chapters 5 and 6 of my *Freedom and Belief* (Oxford: Clarendon Press, 1986), for which I thank the publisher for permission to reproduce (www.oup.com); they are recast here to emphasize their connection with the present volume.

² I make a beginning in Strawson (1986), §§6.3–6.5; see also Dennett (1984).

³ *Anātman*, or ‘no-soul’, denotes the corresponding positive doctrine that there is no soul or self. The experiential or phenomenological correlate of the factual or metaphysical error involved in *satkāyadr̥ṣṭi* is called *asmimāna*, or the “‘I am’ idea”. See, e.g., Collins (1982), pp. 94–5, 100–3.

References

Anscombe, G.E.M., 1971. *Causality and determinism*. Cambridge: Cambridge University Press.

Collins, S., 1982. *Selfless persons*. Cambridge: Cambridge University Press.

Collins, S., 1997. A Buddhist debate about the self; and remarks on Buddhism in the work of Derek Parfit and Galen Strawson. *Journal of Indian Philosophy*, 5(25), pp. 467-493.

Davidson, D., 1973/1974. On the very idea of a conceptual scheme. *Proceedings and Addresses of the American Philosophical Association*, 47, pp. 5-20.

Dennett, D., 1984. *Elbow room: the varieties of free will worth wanting*. Cambridge, MA: MIT Press.

Kant, I., 1788. *The critique of practical reason*. Available from: Project Gutenberg.

Spinoza, B., 1677. *Ethics: ethics, demonstrated in geometrical order*. Available from: Project Gutenberg.

Strawson, G., 1986. *Freedom and belief*. Oxford: Clarendon Press.

Strawson, G., 1994. The impossibility of moral responsibility. *Philosophical Studies*, 75(1-2),

pp. 5-24.

Strawson, P.F., 1962. Freedom and Resentment. *Proceedings of the British Academy*, 48, pp. 1-25.

7 What Am I Doing?

Susan Blackmore

I am sitting outside my hut.¹ It's summer, warm enough to put my mat and stool on the flagstones and sit outside, in front of the flower bed.

A blackbird sings on the garage roof; another answers from behind somewhere. There are many birds singing, now I come to notice them, and even a seagull shrieking far above. Bristol isn't far from the sea and the gulls...no, let it go. The buzzing from countless bees and flies, messing about in the flowers, maps out a sonic space around me. The sun feels warm on my arms. I am sitting still: the mind is calming down.

I wonder what I'm doing here. Have I chosen to be in this spot, sitting like this, of my own free will? How much of this am I doing, and how much is just happening? When I'm ready I will look into what it means to act.²

* * *

What am I doing? I am sitting. That is, this body has been sitting here a long time. But does that count as me doing it?

I am breathing. Yes, but the breaths go in and out whether I will them to or not. I can watch, or not, I can decide to breathe faster or hold my breath. But now breaths just come and go. No one is doing them.

I hear birdsong and bees. Yes, but I can't not hear them, and it feels passive. The sounds arise and fall away. I'm not making an effort to listen to them and I don't respond to them now that my mind has settled. So does that count as me doing it?

This is a strange question. Asking what I am doing seems to freeze me in a moment of not doing, of asking and not knowing the answer. I *was* asking the question but now?

I sit, not doing, and wonder. What am I doing?

I could do something else if I chose to, couldn't I? It seems so. This is the essence of free will, and without that there would be no point in doing anything at all, would there?

I am sitting absolutely still, as I am supposed to in meditation. But if I wanted to I could lift one hand, clap loudly, ring the bell, get up and walk away, or run out into the road shouting, "I'm free. I can do anything I like!"

OK, I'll clap.

Clap.

Did I do that freely—for no other reason than that I consciously decided to do it?

Probably not. I thought of clapping because I was asking, “What am I doing?” and casting around for something to do, and there aren't many things you can do sitting in meditation posture, and clapping has a lot of history in Zen, so it's probably a likely candidate to be chosen by this brain at this time, and all of this goes back to why I'm sitting here, and that goes back to...

OK. I give up. I can trace back myriad possible reasons why this happened. But even so, in ordinary language, I would say I did the clapping, not just that this body did it. Did I? And if so, did I do it of my own free will?

* * *

I'm being too intellectual about this, and perhaps that's not surprising. Free will is said to be the most argued-about philosophical problem of all time, and I've read quite a lot of the philosophy.

The basic problem has been apparent for thousands of years both in Western philosophy and in Buddhism. The universe seems to be causally closed: everything that happens is caused by something else. Nothing happens by magical forces intervening from outside the web of causes and effects, for everything is interconnected with everything else.

This means there is no sense to the idea of free will: to the idea that I can jump in and consciously decide to do something without any prior causes, just because I want to. If that happened, it would be magic, implying that conscious actions lie outside the physical web of interconnectedness. Yet I feel as though I can act freely. Indeed this magical view is probably how most people in most cultures have always thought about themselves, imagining a non-physical mental entity that has wishes and desires, can think and plan, and can carry out those plans by acting on the world. But non-physical things cannot act in the physical world without magic, and the more we learn about how the brain works, the less room there is for magical interventions by conscious minds. We are back to dualism and the problem Descartes never solved, and no one else has since.

In Zen, and in the languages in which early Buddhist texts were written, there is no equivalent of the Western concept of free will, but there is plenty about doing and not-doing. On his enlightenment, the Buddha is said to have awakened to the realization that all phenomena co-arise in an interconnected web of cause and effect. This is called ‘dependent origination’ or ‘dependent co-arising’. Everything is part of everything else: nothing has its own self-nature, independent of the rest, including people. If so, then no one can act independently of everything else, hence ‘Actions exist, and also their consequences, but the person that acts does not’. In its

completely different way, Buddhism is saying the same as science – however it feels, there’s no room for magical interventions.

To resolve the problem of free will philosophers have come up with many ways of recasting the problem; for example by treating actions and choices as free if they are not forced, and finding ways in which free action can be compatible with determinism, but I want to stick with the everyday sense of ‘free will’, the sense in which it feels as though ‘I’ have consciously caused things to happen. In that sense the solution seems, purely intellectually, to be obvious. There cannot be free will. It doesn’t make sense.

So what to do? Many people come to a similar conclusion and say, “But I cannot live my life not believing in free will, so I will act ‘as if’ there’s free will”. That seems to satisfy them.

Not me. I am not prepared to live my life pretending the world is otherwise than it is. So I have worked hard systematically challenging the feeling of having free will whenever it arises. Now it rarely catches me out and actions seem to happen on their own. Even so, there is always room for deeper inquiry, and so it was with some enthusiasm that I set aside the time to investigate what it’s like to act, and decide, and do.

* * *

I am sitting outside my hut, and the morning air is fresh and chilly, even though it's mid-summer. The flowers of the feverfew in front of me are back-lit by the early morning sun: small and white with tiny yellow centres. I am calming the mind for half an hour before I start work on today's question. One cat sits beside me, the other on a chair across the orchard. From the corner of my eye I see something move, cat-like though rather large. But both my cats are here. "Let ..."

What is it?

I cannot, or do not, resist. I turn my eyes to look.

A fox slips quietly between them.

I shouldn't have done that. I am supposed to keep looking straight down at the flowers and grass. Was that free will? No. The movement of the fox and my curiosity made me do it. 'My' curiosity. Was this me doing it?

Not yet. I go back to sitting quietly, but thoughts start bubbling up; the foxes died out in Bristol a few years ago and now they've begun to come ... "Let..."

There you are, you see. I did that, didn't I? It was me who jumped in with "Let it go...", and the thoughts went away. But this isn't true. I learnt that trick from John decades ago. That meme, those words, 'Let it come, Let it be, Let it go', is one he infected me with when I began

meditating, and it's been working away in my mind since. But I choose to keep it rather than forget it, reject it, tell myself it's stupid, or any number of other things I might have done. So it is up to me that I keep on using the trick, isn't it?

Not really. The reasons for rejection or acceptance are my personality, genetic make-up, other memes I've picked up along the way, and force of circumstance. All this has led to a person who sits here now, and when thoughts start rattling away whispers "Le...", and they stop.

So where do I come into this? It's all bound up with me. For me to have free will means I do something of my own accord. So who am I?

Ah. That's a familiar one. I'll sit and see who is here for a bit. Perhaps I can see if she's doing anything.

Here it is; the headless body topped with grass and pretty white flowers. It sits still.

* * *

Do I have free will? I hear the words. I ask. I am stumped.

Perhaps I need to work with something easier, with the absolute minimal requirement for free will, which is that I actually do things. I must go back to the simple question.

What am I doing?

I can sit with that. I can stay calm and clear and ask, “What am I doing?”

I sit. The buzzing continues; the cats do not stir.

I am sitting up straight. But it’s not exactly me who’s doing this. It’s such a long-practised habit that this body just gets into that position and stays there. Maybe it makes sense to say that my body is doing it; but am ‘I’ doing it?

All right, I’m making an effort here. I’m paying attention. That really is an effort. Indeed, that is the whole task of meditation: that you have to stay there, minute after minute, hour after hour, and keep on paying attention, not sliding off, not getting side-tracked into worries, fantasies, or imagined conversations. You have to work at paying attention. I am paying attention now. And now.

So that’s the key, isn’t it? If I didn’t make the effort, then it wouldn’t happen. It’s an effort of will; it’s hard work. I am using my will to pay attention, and now, and keeping on paying attention. It’s hard work. So this is what I am doing. The hard work and the effort I feel myself making are proof that I’m doing something.

But oddly enough, I realize, hard work doesn't prove that. As I sit here, I remember being in labor, many years ago. My first child had a very big head, or something else wasn't quite right, and I was in labor for over 24 hours. It was hard work, terribly painful. That's why it's called 'labor', I realized. And who was making the effort? I had this extraordinary sense that I was doing the hard work, but that I had no option. I couldn't say "No. I don't want to have this baby. I won't do it." My body was doing it by itself. I was doing the hardest physical work I had ever done, but was not willing it. The labor was willing itself. Doing, yet not doing.

Come back now, to the garden, sit calmly, pay attention.

Hard work does not prove it's a matter of will, or that I am doing it. So what am I doing?

* * *

Sitting still again. I see the taller flowers, and the branches of an apple tree moving in the chilly breeze. I feel the wind.

Am I doing this? Am I looking at them and seeing them? Is seeing doing?

Yes, but I couldn't do it without them. They are as much doing it as I am. Which is moving: me or them? Is the moving in my mind or in the world?

The trick of turning inwards unfolds again. There are the flowers, legs, arms, half a nose, and then where there should be me inside there are only moving flowers and apple branches. I am not doing this. The looking, seeing, moving, we're doing it together. It's just happening, the universe doing its thing. The body goes on sitting still. The branches keep on waving.

Nine o'clock strikes. End of meditation. I bow. I get up.

Did I do that?

I feel strange. I am used to these mental manoeuvres, yet they still have a deep effect.

I get up and walk attentively, without assuming I'm doing anything. The legs are walking, the grass and flowerbed are slipping by in the space where I should be.

There are raspberries to be picked for breakfast. A hand reaches out, again, and again. It can choose this bush or that. It picks this one, until enough are picked. It's time to go indoors, but which way will she go?

I like making paths, and there are several quite bendy and pointless paths in our garden. There are three ways I can go, one with low branches to duck under, another narrowed by spreading

weeds, and the last clear but longer. Which way shall I go? I try to catch myself in the act of making the decision. Everything slows down, I stand hovering with one foot raised, to see whether I can catch my mind making itself up. If I could catch this moment, or watch this process, I might find out what it's like to act freely, and to know that I am really doing this. A hand reaches out to a just-noticed ripe raspberry, the cats suddenly scamper past, and a foot is already following them. They go to the place on the wall where I always stroke them before going indoors. The hand reaches out, the fur is soft and the cat's head presses against the hand.

So she must have decided to go that way.

This seems to be all that happens; decisions are made because of countless interacting events, and afterwards a little voice inside says, "I did that", "I decided to do that".

Is there any need for that little, after-the-fact, voice?

I must watch some more.

* * *

Do I have free will?

No. I am not separate from the perceptions, thoughts and actions that make up my world. And if I am what seems to be the world, then we are in this together. Me and the world, world/me, are doing all these actions that seem to act of their own accord.

But help!

This means I am not responsible. This is terrible. If I'm not responsible then ...

My mind goes back many years to when I was first practicing mindfulness, and hit upon this fear. We were staying at a campsite in Austria, by a lake, and I was on the little beach with the children. It was a lovely day, and the place was peaceful except for a blaring radio. It was annoying.

I knew what I would usually do in such circumstances: fret, be angry, think about the regulations of the campsite, about telling him to turn it off, feel bad at the thought of doing so, look at his tattoos and worry what he'd do if I approached him, think that if I were a real Buddhist I would feel compassion rather than anger, imagine that if I were a good meditator I wouldn't mind the noise, try not to mind, fail. And so on.

So what happened? Many such thoughts began but each was met with "Let ..." and fizzled out, leaving the grass before my feet, the mud pies the children were making, the feel of the earth on my hands as I joined in, the sounds of birds, the sounds of the radio, the grass again as I walked up the beach and stopped, my voice (in my best German) saying "Excuse me, your radio is very

loud, would you please mind turning it down?”—the man’s face scowling and muttering, his hand reaching out to his radio, the grass and then the mud and stones rough under my feet, and the mud pies and children.

Later I noticed he left. I wondered whether I’d done the ‘right thing’, but “Let ...” and back to the feel of the water on my toes. A moment’s thought was enough to realize that agonising about what to do would not have helped. The world had summed up the options, chosen one, carried it out, and moved on. This action was a result of everything learned and done before. Now was, as ever, fine. And now.

Deep breath. Watch again. The world said, “Deep breath, watch again”. This sensible response was not coming from a little thing inside called ‘me’; it came from somewhere, I don’t know where, from all the past actions of this body, and this brain, and everything it’s gone through. It just happened. There’s nothing wrong with that response.

So is it always like this? Could I just trust the world and this body to work by itself without me doing anything? I realize with some horror that by relinquishing myself to the world, and accepting that actions just happen, I have given up all personal responsibility. I cannot believe in it any more. There’s no one in here making the decisions. They are making themselves. I walked inside, fed the cats, had breakfast, and made loads of little decisions along the way, all with alert attention to what’s happening, and with no sense of myself doing it.

It seems so right. It seems truthful to the way things really are.

* * *

But what about responsibility?

I have played around with this question intellectually since my teens, when I first worked out that free will must be an illusion, but it was only after many years of meditating that I confronted the problem directly.

I was on a Zen retreat at Maenllwyd, practising intensely. Our teacher for the week was Reb Anderson, a Zen master visiting from California, and he was pushing us hard. As the illusion of doing began to loosen its grip, I became frightened. The world was seeping into me and I was disintegrating into the world. I was acting and not acting. This flowing sense of action without an actor felt perfectly natural, but as soon as I started thinking about it I hit the problem: what about responsibility? There could be none in such a world.

I signed up for an interview. The Zen master was an impressive, good-looking man, with shaven head and imposing robes, and this was a formal interview. I walked to the interview room, opened the door and slipped in. I bowed in the prescribed way, sat in the prescribed posture, looked into his shining eyes, and summoned the courage to tell him what I thought: that ultimately no one is responsible for anything.

He chuckled.

“Yes”, he said with a delightfully warm and encouraging smile, “*ultimately*, that’s true”. He seemed to emphasise ‘ultimately’, and I thought of the Zen distinction between the ultimate view and the relative view, wondering whether there’s some other way in which it’s not true.

“Then what do I do about responsibility?” I blurted out.

“You *take* responsibility”, he said (Anderson 2009).³

* * *

Help! Who takes responsibility? Isn’t ‘taking’ responsibility ‘doing’ something? Isn’t taking responsibility an act of will? Doesn’t it require someone who is doing it? Isn’t it freely done? No, in this case he had told me to do it, so it wasn’t free. But I could refuse to do it.

Then who would refuse to do it? I know there is no self, so isn’t taking responsibility just inventing a new false self who is going to have that responsibility? Why would one want to do that if one knew there really was no self?

I am going in circles. Help.

* * *

Over the years, as the sense of free will has slipped away, I have remembered this advice and it has helped.

The illusion of free will does not survive the kind of scrutiny I have given it. It melts away. I no longer feel its pull. People sometimes ask me how I did it; how I gave up free will, but I cannot tell them. I know that I battled intellectually with it for years, but thinking only creates a mismatch between what one believes and how the world seems. I never felt comfortable with this mismatch, and didn't want to go on living as though free will were true when logic and science told me it could not be. I didn't want to live a lie, or a half-truth, or an 'as if'. So this great intellectual doubt drove me to look directly into how decisions are made, and to examine the self that underlies the feeling of acting freely.

I no longer get that feeling. Sometimes a shadow of it arises up—"I've got to decide what to wear for my lecture this evening" or "I don't know whether to accept this work offer". I welcome these as a chance to look again, to investigate what it feels like to make a conscious decision, but all the habits of paying attention and watching what happens dissipate the feeling quickly. It has nothing to cling to.

So it works something like this. An email arrives. It's an invitation to give a lecture in an exciting place, at a prestigious conference. I look in my diary. That day I've agreed to go with my partner to a family event, planned for ages, he'd love me to attend. What to do? I have to decide. I don't like letting anyone down. The lecture is a terrific opportunity. It won't come again. But I've already committed myself.

No, 'I' don't have to decide. There is no inner me who can do so. This whole series of events is part of the play of the world/me as it is, and the decision is too. So the thoughts come, and the feelings of indecision come, and the feelings sway back and forth, and the weighing-up goes on, and it's all just happening, like cars going by and the ticking clock in the background. Then the decision somehow is made, whether it's today or three days later. Eventually the fingers type the replying email and it's done. Then what?

I take responsibility. I don't mean that a little inner me who has free will does so, because that would be to fall back into the endless cycle of the illusion of doing. The little me is fictional. Taking responsibility only means consequences follow and I will accept them. If someone tells me how wonderful the conference was and I missed it I won't be angry that 'I' made the wrong decision. It was made. That's what happened and that's how it is now. If someone is angry with me for being so selfish and mean for not joining the family event I will accept that. That's what happened, these are the consequences. Things just are the way they are. Whether they could have been different I do not know, but I suspect that even asking this question does not make sense. Stuff just happens.

Indeed, the fingers are typing right now. No one is acting. I am not doing anything.

What, then, is the point? What's the point in doing anything?

No point.

Notes

¹ This is a revised, condensed version of a selection from my *Zen and the Art of Consciousness* (Oxford: Oneworld Publications: 2011), pp. 135-149. I thank Oneworld Publications for granting permission to reprint this material here in revised form.

² The ellipses set off immediately below in the text are meant to encourage the reader to pause reflectively; in the original text, there are little drawings.

³ Years later, I met Reb again and we talked about doing and not doing. Now, he says, he would probably say, 'Accept responsibility - without limit'.

Reference

Anderson, R., 2009. Meditation retreat interview. [conversation] (Personal communication, 4 September 2009).

8 Freedom from Responsibility: Agent-Neutral Consequentialism and the *Bodhisattva* Ideal

Christian Coseru

§ 1 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 that 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 nonviolence, 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* nonexistent human suffering (on account of the nonexistence of sentient beings as *conventionally* established), support for a robust notion of phenomenal agency can be found in nearly all major schools of Buddhist thought.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, 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.

§2 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 (*Dīgha Nikāya* 2.19–20, 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 (*Āṅguttara Nikāya* (“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 metaethical 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 well-being (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 (*viz.* 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, Prabhu, and Sharma 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)*, 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 Augustine 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”⁵ 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” (1917-25, 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. Variously 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. Likewise, 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

§3 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’” (*Samyutta Nikāya* 22, 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.” (*Majjhima Nikāya* (“MN”) I, 111-112, in Ñāṇamoli & 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 & 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 versus 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 pp. 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 pp. 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, Gray, and Wegner 2007; Arico, Fiala, Goldberg, and Nichols 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 2016), 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.

§4 Conclusion: 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 versus 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

seem 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 like 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, Morris, Nadelhoffer, and Turner (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’ (*puḍgalavādins*) credit for insisting that important features of personhood are ineliminable (see Carpenter 2015; Priestly 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’ (Davis 1898), ‘volition or conation’ (Aung and Davis 1979), ‘choice’ (Keown 2001), ‘volition’ (Gunther 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 ed. Translated by T. Irwin. Indianapolis, IN: Hackett.

Aung, S.Z. and Davis, C.A.F.R., 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., 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., 2016. Consequentialism, compassion, and Buddhist neuroethics. In: J. Davis, O. Flanagan, and C. Kelley, eds. *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.

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*. London: Macmillan.

Long, A.A., 2001. *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., 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 ed. 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, 1917-25. *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. 217–33.

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.

9 Free Will, Liberation and Buddhist Philosophy

Marie Friquegnon

Three different attitudes toward what we consider free will may be found in Buddhist philosophy: first, the idea that our behavior is not determined by the gods or completely constrained by karma; second, the idea that our behavior is ultimately impersonally determined, the way fire causes smoke; and third, the idea that selfless actions spontaneously arise from enlightened nature, in a way that differs from deterministic behavior. Philosophically, the grounds for these views are as follows:

1. There is no proof that gods or caste cause behavior.
2. In *samsara* (cyclical reincarnational wandering) causal processes govern all things. Therefore, one's actions will be determined by circumstances and our reactions to them. These reactions are themselves governed by our inherited and developed emotional and intellectual characteristics.
3. *Nirvana* is not subject to causality.

In response to the problem of how one can achieve *nirvana* or enlightenment, later (Mahayana and Vajrayana) Buddhists argue that *nirvana* and *samsara* are two perspectives on the same

reality. Unenlightened beings can only perceive *samsara* because they cannot take the perspective of *nirvana*. This is because their vision is obscured. One can, through Buddhist practices, such as meditation, remove these impediments, and one will be able to experience *nirvana*. Then one's actions will be positive, spontaneous, selfless and free. One can only be free if one has freed oneself from the limitations of the self.

Free Will and Buddhist Philosophy

With the exception of free will, almost every philosophical problem in metaphysics and epistemology that has absorbed Western philosophers has been examined in a similar way by Indian philosophers. In the West the problem is seen roughly as follows. 'Hard' determinists and 'libertarian' indeterminists agree on the meaning of free will. Actions to be free must be intentional and avoidable. By 'avoidable' they mean that if one chose to do X at time T_1 , if one then 'rolled back the clock' one could, all things being the same, choose to do Y at T_1 instead, X and Y being two different courses of action. Hard determinists deny that this is possible. Every event must have a cause. When all the causal factors are present the event must necessarily occur. Actions are events; therefore, they necessarily occur.

Even if events were not caused, they would be random. What is random is not free. (This is the 'hard' indeterminist position. The 'hard' incompatibilist thinks free will is incompatible with both determinism and indeterminism.) For example, if while I was teaching a class, and something randomly happened to me, that would not be something I did. Therefore, it would not

be a free action. Following Broad (1934) and van Inwagen (1983), Goodman put this point very well: “After all, if what you do is caused by some random-quantum-mechanical event in your brain, how can you be responsible for it?” (2002, p. 360) However, if an event were intended, a hard determinist would insist that the intention was just another cause in a long chain of causes.

Some indeterminists argue that randomness cannot secure free will because it cannot be authored or controlled. Libertarian indeterminists argue that free actions are not causally necessitated.

There are many forms of indeterminism, but a modern variety would limit the use of the word ‘cause’ to what is statistically predictable. Causal laws are in fact just those generalizations about the past that reflect ‘invariant concomitance’, one type of event always having been found to follow another. But human action is often unpredictable. So we have no evidence that it is always necessitated by causes. So hard determinists and libertarian indeterminists define a free act in the same way. They agree that free will requires indeterminism. But the hard determinist therefore denies the existence of free will, while the libertarian affirms its existence.

Soft determinists, however, provide us with a new definition. Giving up the criterion of strict avoidability, they assert that the definition of a free act is an action that is caused by a rational intention. So, for example, if I chose to stay home and study rather than to go to the movies, that would be a free action. But if a kleptomaniac steals hats although she hates hats and doesn’t want hats, that behavior would not be free. Both behaviors are caused deterministically.

Libertarians such as Campbell are an interesting variety of indeterminist. Campbell believes that although behavior in general is caused, that which is brought about by the self is both free and

avoidable in the strict sense. The evidence for this, he argues, is that the self is able to resist temptation so as to do what is right. In these cases one is acting against one's desires, so one's desires cannot be the causes of those actions. (1957)

Goodman, correctly, I think, argues that soft determinism fails because it implies that no events, strictly speaking, could have turned out differently (Goodman 2002). As Hobbes famously objected, we are free to do what we want, but we are not free to want other than what we in fact want. Libertarianism also fails because, as Santarakshita implies in the *Tattvasamgraha* VII.197 (Santaraskita 1937) and Kamalashila says in his commentary to the root text,

“If the cognitions of BLUE and the rest were the effect of a single such cause as the ‘soul’, which is eternal (continues for all time, past and future), then any order of sequence among such cognitions would be incongruous; as the efficient cause being present, all the effects should appear simultaneously” (Santarakshita 1937, p. 148).

Turning to the Buddhist models of action, we find the field divided in quite a different way, both in relation to levels of understanding and to the different schools. The early teachings of the Buddha concentrate on self-liberation. All one has to do is follow the Eightfold Path, which of course includes right understanding and meditation. From the very beginning in Buddhism, right wisdom was what enables the practitioner to become aware of the human condition. That is, one comes to realize that all is impermanent, and that our cravings and dissatisfactions make us suffer. The key is letting go, since there is nothing permanent to which one can cling. The simplest form is to be found in the *Dhammapada*, where with his dying breath the Buddha

advises his disciples to achieve liberation relying on no one. There is no god, for example, on which one can depend. Nor is one determined by caste or social position. A barber and a king may be equally qualified to follow the path.

But isn't one constrained by karma? Even if one becomes aware of karmic conditions, does one have the power to alter them? Isn't one's will to do so itself constrained by karma? Isn't trying to overcome karma like trying to lift yourself up by your own bootstraps?

The first complication can be illustrated by passages in chapter 6 of the Mahayana text, *The Way of the Bodhisattva*, by Shantideva, an Indian philosopher of the eighth century. Shantideva advises one to dissolve the emotion of anger because one should no more blame or become angry at a person who has harmed one than (i) at bile for causing suffering or (ii) at fire for causing smoke (1997, pp. 78-97). The justification for this is that: (a) all evil actions occur because of ignorance and attachment, and (b) all events occur successively due to a set of causal conditions operating at T_1 that are necessary and sufficient to bring about the state of the world at T_2 .

At first sight this seems to do away with the possibility of self-liberation promised by the Buddha. One will reach liberation only if the causal conditions operating in the world bring us there. Having the right causal conditions or not is just the luck of the draw.

This is an important and difficult issue in Buddhism. Earlier Theravada Buddhist thinkers tended to think of *nirvana* as beyond conception, as completely transcending the phenomenal world. Thus, we have the problem of having to lift ourselves out of *samsara* by our own bootstraps.

Later Mahayana Buddhist thinkers thought of this transcendence as epistemological but not ontological. That is, the thought of the enlightened state of awareness is transcendent, but ontologically, *nirvana* and *samsara* are not distinct states of being, rather different perspectives: *nirvana* is *samsara*. Like a dance and its dancer, *nirvana* and *samsara* are distinguishable, but not separable.

It is easier to think of escaping *samsara* by shifting one's perspective rather than somehow catapulting oneself out of it. Nevertheless, there are still problems. How is one within the samsaric point of view able to assemble the causal conditions that will produce enlightenment?

Buddhist thinkers agree that *nirvana* or enlightenment cannot be produced at will. Mahayanists argue that this is not really a problem because we are already enlightened. But our own enlightenment is concealed from us as clouds block the sun. So, the *Dharma* consists in assembling the correct causal conditions that will dissipate the emotional and intellectual obscurities that prevent us from the realization of ultimate reality.

This view of *samsara* and *nirvana* is both similar and dissimilar to Wittgenstein's duck/rabbit example. In Wittgenstein's example, one can see a figure as either a duck or a rabbit (1980, p. 75). Some people, however, may only be able to recognize it as a duck or as a rabbit. In any case, the duck and the rabbit perspectives are equally correct. But in the case of *samsara* and *nirvana*, *samsara* is the way we see things when we are blinded by obscurations. We see now 'through a glass darkly'. Of course, the duck/rabbit analogy does not hold perfectly, because the perspective of *samsara* is transformed when one is able to take the perspective of *nirvana*. This is because

one's vision is no longer obscured by negativities such as anger and hatred, jealousy and attachment, ignorance and fear. In addition, the duck/rabbit example concerns a perspective on something outside of the subject, whereas the enlightenment perspective is non-dual, transcending subject and object.

But what is there in *samsara* powerful enough to initiate a perspectival shift? Since *nirvana* is *samsara*, there already exists in *samsara* what is called the *tathagatagarbha* or seed of enlightenment. In the Mahayana and Vajrayana traditions, this is what propels all sentient beings towards enlightenment. All the karmic ups and downs are ultimately affected and transformed by this seed.

There are other manifestations of this seed of enlightenment into the phenomenal world. These are the appearances of the buddhas into the phenomenal world, appearances that are illusory in the way we perceive them, but which are rooted in reality at the absolute level, beyond duality, beyond conception.

Illusory as they are on the phenomenal level, they are seen as the way out of the 'bootstraps' problem. In Wittgensteinian terms, they are the ladder that one uses until one achieves enlightenment. After that one throws the ladder away. For the enlightened being, there is no *Dharma* or Buddha.

Now, let's return to the problem of free action. On the most basic level, free action is conceived of as the absence of external constraint. God or membership in a caste do not determine what one can or cannot do.

To repeat, on the psychological level, Shantideva tells us we have no more justification for blaming someone who harms us than for blaming a fire for causing heat or, in another example, for blaming the sky for having clouds (1997, p. 83). This line of thought reflects the view that all wrongdoing is due to ignorance. Should one then draw the conclusion that good as well as bad actions are causally determined? This is sometimes the case. Actions that are morally correct may be caused by selfish desires. Kant's example of the shopkeeper who is honest in order to attract customers is an example of this. Such a shopkeeper, from the Buddhist point of view, would still be chained to his or her desires.

To return to Campbell, what should we think about selfless and heroic actions that demand real moral effort? Are these free? Campbell claims that such actions are self-caused, rather than determined by desires, etc. (1957, p. 5). But Buddhists believe a separate substantial self is an illusion. Who then is the *bodhisattva* (highly evolved Buddhist) that acts unselfishly for the good?

The answer in Mahayana Buddhism lies in the concept of *bodhicitta*, the thought of enlightenment. '*Bodhicitta*', literally the 'awakened mind', is a blend of wisdom, compassion and bliss.

There are both relative and absolute *bodhicitta*. Relative *bodhicitta* is a matter of degree. As Khentrul Tsewang Dongyal Rinpoche, a well-known Tibetan Buddhist lama, once said, “Some days you will wake up, and your *bodhicitta* will be low” (2012). The higher the *bodhisattva*, the greater the *bodhicitta*. Buddhas are claimed to have absolute *bodhicitta*, which is the same as enlightenment.

The *bodhicitta* of a buddha is viewed as beyond duality, beyond subject and object. The *bodhicitta* of a buddha is uncaused and spontaneous. Yet it is not random. It is a kind of holy will. It is similar to the way God is imagined as having compassion even before there are any beings toward which he could show compassion.

So it seems that we have the following anomaly. A *bodhisattva* acts freely insofar as he or she is not motivated by selfish desire, but only by the pure intention of helping sentient beings. The greater the selflessness, the greater the *bodhicitta* and the freer the action. So, a perfectly free action is not caused by the self at all. Of course, on the absolute level, it is not correct to say there is either self or other. The absolute level is beyond duality. But even on the relative level the *bodhisattva*, as he or she gains *bodhicitta*, becomes less and less involved with the illusory self.

What does this mean in practice? It is not simple spontaneity, which can be harmful or crazy. It is not just wisdom, because without compassion wisdom can be only shrewdness. It is not just mindless compassion. One should not, for example, out of sympathy, let a mass murderer out of prison. A *bodhisattva* possessed of great *bodhicitta* is recognizably wise, compassionate and

blissful all at once. Is such a being free or determined? It looks like the truest kind of freedom, yet an advanced *bodhisattva* is said to have irreversible *bodhicitta*, which means he or she is incapable of committing an evil act. It is hard to say if this should be considered a limitation on his or her freedom. Theists say God is always good, yet is free.

To return to the problem of free will, it seems as if one who is not a *bodhisattva* is not free at all. This would be a very good reason to treat wrongdoers with compassionate understanding. As Blackburn states very eloquently:

“As we come to learn about causal regularities lying behind actions and other mental states, we are apt to switch into less moralistic modes. We might blame someone for being depressed all the time, until we learned a chemical story explaining it. We might be angry with someone for being unable to stir himself, until we learn that he has mononucleosis. But according to the determinist, there are always things like this to learn. Quite apart from increasing neurological evidence, we may think of cases where we learn of ‘brainwashing’ or ‘conditioning’. Parents may be inclined to blame their teenage daughter for spending time, energy, and income on valueless cosmetics, but a better reaction would be to understand the social and commercial pressures that paralyze her better judgement and bring this state of affairs about.” (1999, p. 98)

Blame and, likewise, guilt have no place in Buddhism, except as skillful means for the altering of behavior. But regret does, for this signals a realization that one’s actions are not of the best, and that one wants to improve them in the future.

Vajrayana philosophers, mindful of the pervasiveness of primordial wisdom, see even the karmic forces and the three poisons (anger, attachment, ignorance) as part of the process of enlightenment, rather than as external forces that oppress us. This view is similar to Smullyan's, who argues that, since we are not separate entities from the rest of the universe, but rather part of the totality, the so-called 'external' forces acting on us are really our own activity. We have no reason to feel that we are their slaves (1981).

Some, including some Buddhists, are concerned that viewing immoral actions as causally determined may lead to moral laxity, because one can always excuse one's bad behavior. But if one has sincerely entered the Buddhist path, one can be motivated to act from *bodhicitta* by 'tuning in' to the enlightenment qualities that are always with us, because they are part and parcel of our true nature. And when we do fall into bad behavior, the compassionate view we have of others who err morally should also be applied to ourselves.

To summarize, Buddhists resemble soft determinists when they proclaim that we are not coerced by gods or by caste. They resemble hard determinists when it comes to wrongdoing, which is determined by the three poisons. They resemble indeterminists with respect to selfless actions done from *bodhicitta*. So, particularly from the Mahayana point of view, the actions done from *bodhicitta*, strictly speaking, neither are our actions nor do they fail to be our actions. To assert that they must be one or the other, I suspect, is a metaphysical mistake of believing in the reality of the self, part of the more general dualistic division of self and other.

The denial of free will on the relative level should not, I believe, be construed as asserting that there is no freedom. Goodman says the goal should be to “see our true non-existence, abandon the illusion of free will, and abide in real freedom” (Goodman 2009, p. 212).

Further, as the eleventh century teacher Rongzom pointed out, reality is beyond duality. It cannot therefore be claimed that the ultimate is free and appearances are not. So, if there is freedom ultimately, then in reality, all are free (2008, pp. 104-107). This is not, of course, to assert the free will of a non-existent self.

References

Blackburn, S., 1999. *Think: a compelling introduction to philosophy*. Oxford: Oxford University Press.

Broad, C.D., 1934. *Determinism, indeterminism and libertarianism*. London: Routledge & Kegan Paul.

Campbell, C.A., 1957. *On selfhood and godhood*. London: George Allen and Unwin.

Goodman, C., 2002. Resentment and morality: Buddhism on moral responsibility. *American Philosophical Quarterly*, 39(4), pp. 359-372.

Goodman, C., 2009. *Consequences of compassion: an interpretation and defense of Buddhist ethics*. New York: Oxford University Press.

Rinpoche, K.T.D., 2012. *Teaching on bodhicitta*. [Conversation] (Personal communication, 6 June).

Rongzom, 2008. *Establishing appearances as divine*. Translated by H. Kopl. Ithaca, NY: Snow Lion.

Santaraksita, 1937. *Tattvasamgraha with commentary by Kamalashila*. Translated by G. Jha. Delhi: Motilal.

Shantideva, 1997. *The way of the bodhisattva*. Translated by the Padmasambhava Translation Group. Boston, MA: Shambhala.

Smullyan, R., 1981. Is God a Taoist? In: D.R. Hofstadter and D.C. Dennett, eds. *The mind's I*. New York: Basic Books. p. 333.

Wittgenstein, L., 1980. *Remarks on the philosophy of psychology, vol. I*. Chicago, IL: University of Chicago Press.

Van Inwagen, P., 1983. *An essay on free will*. New York: Oxford University Press.

10 Buddhism and Free Will: Beyond the 'Free Will Problem'

B. Alan Wallace

Determinism and Indeterminism, Ancient and Modern¹

The diversity of Indian views concerning causality in the Buddha's lifetime represented the broader, then-prevalent philosophical pluralism, not unlike our world today, and the Buddha's novel responses to those views remain as provocative as ever. Then as now, philosophers fell roughly into two camps, akin to determinism and indeterminism.

Among the former, some asserted that all pleasant, unpleasant, or neutral experiences are due to past karma (*pubbe kata-hetu*) or the will of God (*Issara-nimmāna-hetu*). (For the Western version of the latter type, see Aquinas (1947), pp. 1, 5, 23.) The Ājīvikas maintained the fatalistic doctrine that all actions are predetermined by destiny (*niyati*), over which people lack control (*Dīgha Nikāya* ("DN") I.53). This resembles the deterministic view that there is at any instant exactly one physically possible future (Laplace 1951; van Inwagen 1983, p. 3; Pereboom 2001). This implies that the precise condition of the universe, say, one second after the Big Bang, causally sufficed to produce the assassination of John F. Kennedy in 1963 (Dennett 2004, p. 84). The Buddha rejected such fatalistic views.

Other ancient Indian philosophical schools favored the view that all experiences arise from pure chance, without prior causes or conditions (*ahetu-appaccayā*) (Raju 1985, ch. 3). In some respects, this view parallels that of some libertarians who argue that the indeterminism in quantum mechanics applies to human experience. For us to be the ultimate source of our decisions so that we are truly morally responsible, they insist, there can be no earlier influences sufficient to determine subsequent actions (Kane 1996, 1999).

In response to such views, the Buddha rejected any theory that undermined moral responsibility. He rejected deterministic ideas as supporting ‘inaction’ (*akiriya*)—if one believes one is not responsible for one’s actions, the will to act wholesomely is stifled.

Likewise, he rejected the indeterministic idea that everything arises from chance, without reliance on causes or conditions (*ahetu-appaccayā*) (*Aṅguttara Nikāya* (“AN”) I.173–75 (*cf.* *Majjhima Nikāya* (“MN”) II.214); DN I.28; *Samyutta Nikāya* (“SN”) II.22). And he concluded on empirical and rational grounds that there is no autonomous self that exists *apart from* and controls the body and mind or that exists *among* the psychophysical aggregates (MN I.230–35; SN III.66).

Thus, the Buddha refuted all notions of the self as an unmoved mover, an agent that causes events with nothing causing its decisions (*cf.* Chisholm 1982, p. 32; Foster 1991). Thus, the sense that each of us is an autonomous, nonphysical subject who exercises ultimate control over body and mind without influence from prior psychophysical conditions is an illusion.

Volition and Action in Early Buddhism

At first glance, this position may seem identical to that of certain contemporary cognitive scientists and philosophers of mind. For instance, Daniel Wegner writes:

“It seems to each of us that we have conscious will. It seems we have selves. It seems we have minds. It seems we are agents. It seems we cause what we do.... [I]t is sobering and ultimately accurate to call all this an illusion.” (Wegner 2003a, pp. 341–2; see also Wegner 2003b)

Nowhere in the brain do neuroscientists find any control center that might serve as the neural correlate of an autonomous self, nor do they find any evidence of an independent self that causally influences brain functions. Rather, the brain appears to function according to its own mechanisms, with no independent self presiding over its activities. According to this materialistic view, all causal influences on mental processes occur in the brain, inaccessible to introspection. (Ainslie 2001, p. 40; Dennett 2004, pp. 244, 254)

But these apparent similarities conceal incompatibilities between Buddhism and materialism. Whereas many materialists believe that brain activities causally generate all mental processes, the Buddha declared the opposite: “All phenomena are preceded by the mind, issue forth from the mind, and consist of the mind” (*Dhammapada* I.1). Central to this Buddhist emphasis on the primacy of the mind is the role of the mental factor of volition or

will (*cetanā*), which determines which actions have moral consequences. Indeed, the Buddha equated volition with *karma*: “It is will, O monks, that I call *karma*; having willed, one acts through body, speech, or mind” (AN III.415).

Only voluntary actions produce karmic results, and the magnitude of the moral consequences of one’s actions corresponds directly to the degree of one’s mental balance, intelligence, and understanding. Thus, the moral consequences of the actions of a person who is mentally ill or brain-damaged are relatively light, while those of a person of sound mind and clear understanding are relatively heavy (AN I.249–53). This corresponds to modern principles of jurisprudence. Moreover, it is incorrect to think that previous karma determines *all* experiences.

Although feelings that arise *together with* one’s initial awareness of sensory stimuli are the result of past karma, feelings that arise following such stimuli are not predetermined by past karma but are rather the result of fresh karma associated with the way one responds to those stimuli. So, volitional acts are conditioned by *prior* influences and other factors, such as the quality of one’s awareness, *simultaneous* with it (*Milindapañha* (“*Miln.*”) 134-38; *Visuddhimagga* 532, 535; *Paṭṭhāna* I.1). Here Buddhism asserts some measure of free will in the sense that one can reflect on options and choose the best in terms of its moral suitability (MN I.415-16).

Determinism and Moral Responsibility

Some contemporary scientists and philosophers think determinism—the view that there is at any instant exactly one possible future—is compatible with moral responsibility.

Wegner, for instance, argues that actions are determined by brain activity prior to the conscious experience of making decisions, so consciousness does nothing. If so, conscious will is an illusion, but it is nevertheless the person's guide to his or her moral responsibility for action, and moral action is quite real (Wegner 2003a, pp. 59, 224, 241). But he fails to provide any cogent explanation for how something that is an illusion and doesn't do anything can be responsible for moral action. And his fundamental premise—that conscious will is an epiphenomenal, causally ineffective illusion—has been shown to be inconclusive (Lau, Rogers and Passingham 2007, pp. 81–90; Dennett 2004, pp. 228–42).

Daniel Dennett takes a virtually identical position, and his arguments face the same dilemma. He declares that a person is nothing more than an assemblage of roughly a hundred trillion cells, each of them a mindless mechanism functioning in accordance with the laws of physics and biology. But he writes, "Human freedom is not an illusion; it is an objective phenomenon, distinct from all other biological conditions and found in only one species, us" (Dennett 2004, pp. 2-3; see also p. 305). In an elaborate but specious series of arguments, he tries to assert the existence of 'autonomous human agents' who exercise free will as their ability to control action whenever there are no constraints, coercions, or compulsions that limit their behavior. Yet nowhere does he provide any compelling argument for the existence of a human agent among or apart from the mindless mechanisms that make up a person.

Those who argue for ‘compatibilism’ between determinism and moral responsibility seem to be moved by independent motives. They consider reality explicable in deterministic terms (physics, biology), but feel a psychological imperative to affirm moral responsibility, without which civilization is inconceivable. On the horns of this dilemma, they are forced to introduce morality and purpose into the mindless, deterministic activities of atoms and cells, which is unwarranted by all we know about physics and biology. This makes for bad science and bad philosophy.

As noted, according to determinism, based on classical physics, the precise condition of the universe at the Big Bang sufficed to cause Kennedy’s assassination in 1963, rendering Lee Harvey Oswald a passive cog in the world’s deterministic machinery. Since his actions were predestined billions of years beforehand, it is absurd to speak of his having free will, and irrational to assert he was responsible for them.

Some contemporary Buddhist scholars, while shunning materialism, argue for compatibility between determinism and moral responsibility, citing the Buddhist principle, “When this exists, that comes to be, with the arising of this, that arises” (SN II.28; see Federman 2010; Harvey 2007). Whether the universe is deterministic in accordance with physical causality (materialism) or with mind-matter causality (Buddhism, on a certain reading), it is deterministic, implying that the present is thoroughly determined by the past. If so, Oswald had no more of a choice to not kill Kennedy according to Buddhism than according to materialism. If at any instant there is exactly one physically possible future, as

determinism maintains, then the present is fixed by the past. This offers no wiggle room for freedom or responsibility.

As noted, the Buddha rejected belief in any theory that undermines our responsibility or inspiration for cultivating virtue. The causal relations between actions and consequences are so complex that they cannot be fully comprehended conceptually (AN II.80). So it is vital not to become immobilized by our lack of understanding of the rationale for moral responsibility. The important thing is to recognize the myriad ways in which we are *not* free to make wise choices, to follow courses of action that are beneficial to our and others' well-being, and to devote ourselves to the cultivation of such freedom.

The Buddhist Ideal of Freedom

A modern definition of freedom is the capacity to achieve what is of value in a range of circumstances (Maxwell 1984). The Buddhist tradition emphasizes that ordinary sentient beings are *not entirely* free, but constrained by mental afflictions such as craving, hostility, and delusion; as long as we live under these afflictions, we remain in bondage to their resultant suffering. But the Buddha taught that suffering and its causes are not intrinsic to the mind, for in every being there exists a 'brightly shining' (*pabhāssaram*) dimension of awareness that, though veiled by adventitious defilements, may be revealed through spiritual practice.

Theravāda Buddhist commentaries identify this radiant mind as the naturally pure ‘ground of becoming’ (*bhavaṅga*), the resting state of the mind that is not included among the six modes of consciousness, namely the five physical senses plus ordinary mental consciousness. This dimension of consciousness manifests in dreamless sleep and at death, and when the (waking) mind momentarily reverts to it between periods of engaging with cognitive objects (AN I.61; Harvey 1995, pp. 145–6, 155–79). Ordinarily, one has no recognition of this state of awareness, but it can be vividly apprehended in highly focused, stable, meditative attention (*samādhi*), withdrawn from all objects, sensory and mental. This ground of becoming described in early, Theravāda Buddhism resembles accounts of the substrate consciousness (*ālaya-vijñāna*) in the later, Great Perfection (*rdzogs chen*, Dzogchen) tradition of Tibetan Buddhism (Wallace 2006a, pp. 14–8, 95–6; 2007, pp. 45–8).

This brightly shining mind may be understood as the unconditioned state of awareness present after an *arhat*, one who has achieved *nirvāṇa*, passes away. Such consciousness, transcending the five psychophysical aggregates, is said to be non-manifest (*anidassanaṃ*), timeless, and unconditioned (DN I.223; MN I.162). Unborn—not created by prior causes—and not the consciousness of anything other than oneself, it must be present in each sentient being before achieving *nirvāṇa*. It is beyond the conceptual mind, so its possible influence on the minds of ordinary sentient beings is unimaginable.

Such transcendent, pristine awareness appears similar to the Buddha nature (*buddha-dhātu*) presented in Mahāyāna Buddhism and to the pristine awareness (*vidyā, rig pa*) taught in the Great Perfection tradition. This primordial consciousness is considered the

source of our yearning for liberation, and may be the ultimate ground of freedom for all beings (Paul 1980, XIII). Because its nature transcends the conceptual, it does not lend itself to rational analysis, and its way of impacting the mind and the natural world lies outside philosophy. It may be known directly through non-dual awareness, but cannot be an intellectual object.

Spiritual practice resembles the process of refining gold contaminated by impurities. The first step is to cultivate a wholesome lifestyle, avoiding injury. On this ethical basis, one gradually balances the mind by cultivating focused attention, for, as Śāntideva cautioned, “a person whose mind is distracted lives between the fangs of mental afflictions” (*Bodhicaryāvatāra* VIII.1). When the mind is subject to attentional imbalances such as laxity and excitation, it is as if one’s psychological immune system is impaired: all kinds of mental problems can easily overwhelm it.

Cultivating focused attention bears on morality and free will. William James declared,

“In what does a moral act consist when reduced to its simplest and most elementary form? . . . [I]t consists in the effort of attention by which we hold fast to an idea which but for that effort of attention would be driven out of the mind by the other psychological tendencies that are there.” (James 1958, p. 126, my italics.)

And Renouvier (1912), admired by James, understood free will as the sustaining of a thought because one chooses to when one might have other thoughts.

With the development of sustained, vivid attention, one's awareness may be focused on one's feelings, desires, thoughts, and intentions as they arise. As the *arhat* Nāgasena taught King Milinda, the practice of mindfulness entails directing attention to wholesome and unwholesome tendencies and recognizing them as such so one may cultivate the former and reject the latter (*Miln.* 37-8). Such discerning, metacognitive awareness allows for the possibility of freely choosing whether to allow a desire to lead to an intention or let an intention result in verbal or physical action. Free will depends on the ability to recognize impulses arising involuntarily and to choose which to accept or reject (Wallace 2006b, pp. 77-127).

Without monitoring our mental states, we are bound to succumb to detrimental, habitual conditioning, with attention compulsively focusing on attractive appearances (*subhanimitta*), reinforcing craving, and disagreeable appearances (*paṭigha-nimitta*), reinforcing hostility (AN I.3, I.200-1; SN v. 64-5). Such misguided attention is prone to lead one to view as permanent what is impermanent, as satisfying what is unsatisfying, and as a self what is not-self (*Vibhaṅga* 373). To overcome such delusional ways of viewing reality, one must add to the cultivation of meditative quiescence (*samatha*) the development of insight (*vipassanā*) through the close application of mindfulness (*satipaṭṭhāna*) to the body, feelings, mind, and phenomena (SN v. 156). Only through the unification of meditative quiescence and insight can one gain complete freedom from mental afflictions and resultant suffering, revealing the innate purity of the brightly shining mind.

The Middle Way beyond Determinism and Indeterminism

One may devote oneself to the path to freedom from suffering and its causes without knowing whether one can exercise free will in a manner not determined by prior circumstances. However, it may be helpful to have a working hypothesis for free will to be actualized.

One sticking point for any Buddhist affirmation of free will is the nature of the self, or agent. Buddhism denies any autonomous, controlling self among or apart from the psychophysical constituents. The same analysis applies to all phenomena. For instance, a chariot is not any of its parts, does not exist independent of them, nor does their collection constitute it (SN I.135; *Miln.* 25). The chariot comes into existence only when the label 'chariot' is pragmatically designated on the basis of those parts. Likewise, the term 'I' is imputed on the body and mind, which are not, themselves, a real self. 'I' come into existence only when conceptually designated as such. When most of us use these concepts and conventions, including the words 'I' and 'mine', we grasp onto the referents of those labels as being real, independent of our conceptual projections; this is the delusional basis for all mental afflictions, such as craving and hostility. Those who are free of delusion still use those concepts and words, but are not fooled by them. (SN I.14; *Itivuttaka* 53)

This analysis applies to the body, mind, and all their constituents. Thus, absent delusion, the self is no more and no less real than any phenomenon (*Sutta Nipāta* 937; MN III.31;

Mūlamadhyamakakārikā (“MMK”) V, VIII; Lamrimpa 2002; Jinpa 2002). Therefore, just as we meaningfully speak of a chariot performing certain functions, we refer to the self as an agent who makes decisions and engages in voluntary activity. But the challenge of determinism remains: If all decisions and actions are determined by prior causes and conditions—physical or mental—how can there be free will?

The definition of determinism noted earlier allows no such freedom. Fatalism is its unavoidable implication, as later events are set in stone by prior conditions. Although some look to quantum indeterminacy to escape fatalism, it is difficult to see how chance allows for a coherent picture of a human agent exercising free will. Most interpretations of determinism and indeterminism assume a metaphysical realism whereby mind-independent objects exist and admit of a complete description that is true if it corresponds to them (Putnam 1990, p. 30).

The Middle Way (Madhyamaka) propounded by Nāgārjuna rejects the reification of time and causality that underlies most versions of metaphysical realism (MMK I, V, XVII). All causally conditioned phenomena arise through dependent origination (*pratītya-samutpāda*), in dependence on (1) prior causes, (2) their own parts and attributes, and (3) conceptual designation. A chariot arises in dependence on (1) the materials used to make it and the carpenter’s assembling it, (2) its components, and (3) the conceptual designation ‘chariot’ imputed to this assembly. The first mode of dependence entails prior causes and conditions resulting in a subsequent product. The dependence of the chariot on its parts is

simultaneous: the whole and the parts exist simultaneously. And the chariot as the designated entity comes into existence with its conceptual designation.

For all phenomena the basis of designation is never identical to the object imputed on that basis. Thus, a chariot is imputed on its chassis, wheels, etc., but none of those parts—individually or collectively—constitute it. The chariot comes into existence with the imputation of that label, but other designations are possible, such as ‘fire wood’. Thus, the entities in our world arise in dependence on conceptual designations, and exist relative to the conceptual framework they are embedded in, not intrinsically or independent of conceptual frameworks. There is freedom in the present to view the world through different conceptual frameworks, and here free will may enter our experience. By shifting our way of framing appearances and making sense of them within our cognitive framework, we alter the very nature of the world as it arises from moment to moment relative to our way of viewing it.

The relativity of all phenomena with respect to our cognitive reference frame is put to use in many Buddhist practices to overcome mental afflictions and cultivate wholesome mental states and behavior. The Tibetan Buddhist genre of ‘mind training’ (*blo sbyong*) is explicitly designed to transform all circumstances, felicitous and adverse, so they arise as aids to spiritual maturation. By conceptually designating events in ways that support virtue rather than afflictions, one alters the world one inhabits: this constitutes a fundamental freedom of choice (ed. Jinpa 2006).

According to the Middle Way, time itself has no inherent nature independent of conceptual designation. While past events influence the present, the way we designate the past determines how it arises relative to our present reference frame. There is an asymmetry between past and present, according to Buddhism and physics. We can think whatever we like about a piece of rotten fruit, but it won't reverse the decomposition process. Likewise, we cannot change the past, but we can shift it *relative* to our reference frames, as there is no past, present, or future independent of reference frames. Thus, the past may impact us variably, depending on how we conceptually designate it now. By designating the past differently, the nature of past events shifts relative to those designations.

Drawing an analogy in modern physics, Eugene Wigner commented, "We do not know of any phenomenon in which one subject is influenced by another without exerting an influence thereupon" (1983, p. 178). By reifying time, we assume the past influences the present but is uninfluenced by the present, and that the present influences the future but is uninfluenced by it. Such unidirectional influence runs against the grain of current scientific understanding. Likewise, the Madhyamaka view denies the inherent existence of all three times, supporting the view that they can all influence each other, relative to the reference frame from which they are designated.

John Archibald Wheeler explained this in terms of quantum physics:

"It is wrong to think of that past as 'already existing' in all detail. The 'past' is theory.

The past has no existence except as it is recorded in the present. By deciding what

questions our quantum registering equipment shall put in the present we have an undeniable choice in what we have the right to say about the past.” (1983, p. 194; see Wallace 2007, pp. 76-80)

For example, the systems of measurement used by cosmologists serve a crucial role in bringing about what appears to have happened early in the universe. Wheeler concludes:

“Useful as it is under everyday circumstances to say that the world exists ‘out there’ independent of us, that view can no longer be upheld. There is a strange sense in which this is a ‘participatory universe.’” (1983, p. 194)

More recently, Stephen Hawking and Thomas Hertog proposed that there is no objective history of the universe independent of systems of measurement and conceptual inquiry (2006; Bojowald 2006). Instead, there are many possible histories from which scientists select based on their methods of inquiry. According to Hawking, every possible version of the universe exists in a state of quantum superposition—as a set of possibilities rather than concrete realities. When we make a measurement, we select from this range of possibilities a subset of histories that share the specific features measured. To relate this to the Middle Way, this is freedom to choose the bases of designation on which to designate a history of the universe as we conceive it, based on that subset of possible histories. Thus, we may exercise free will not only to establish our past, but to frame our present and sow the seeds of our future.

The 'empty' or non-inherent nature of time is incorporated in Tibetan Buddhist Vajrayāna practice, when one 'takes the fruition as the path' (*'bras bu lam 'khyer*). This means that, while unenlightened, one cultivates 'divine pride' (*lha'i nga rgyal*), regarding oneself as a buddha, on the basis of the buddha one will become. Likewise, one develops 'pure perception' (*dag snang*), viewing the environment and its inhabitants as manifestations of enlightened awareness (*dharmakāya*)—in emulation of the pure perception of a buddha. Here, one draws the transformative power of one's future enlightenment into the present, understanding that the future is not inherently real and separate from the present. Based on a realization of emptiness and the Buddha nature of all beings, one may enable the future to influence the present.

Another way of interpreting divine pride is to identify one's Buddha nature, pristine awareness, as the basis of designation for one's identity now. The bases of designation of one's sense of personhood are ordinarily one's body and mind. When one refers to oneself as having past and future lives, the basis of designation for one's identity is one's substrate consciousness, which, according to the Great Perfection teaching, provides reincarnational continuity. When one assumes the identity of a buddha, in divine pride, the basis of designation of self is one's timeless Buddha nature. In the practice of the Great Perfection, one nonconceptually rests in this timeless, pristine awareness, allowing actions to arise spontaneously and effortlessly, aroused by the interplay of one's intuitive wisdom and the moment-to-moment needs of sentient beings. In this way, one realizes a trans-temporal kind of freedom.

Again, the Buddha rejected the philosophical extremes of fatalism and chance and discouraged followers from embracing any view that might undermine their inspiration to devote themselves to an ethical life in pursuit of liberation. In pragmatic terms, as ordinary sentient beings we do not have free will to achieve what is of value within our range of circumstances inasmuch as our minds are dominated by mental afflictions. But the Buddha declared that these sources of bondage are not inherent to our existence, but may be dispelled through sustained, skillful practice. The Middle Way shows how free will may operate within the nexus of causal relations through time. Teachings on the Buddha nature reveal the ultimate source of our freedom. And the Vajrayāna tradition, including the Great Perfection teaching, demonstrates how the freedom implicit in the teachings of the Middle Way and the Buddha nature may be put to use in the swift realization of liberation, enlightenment.

Note

¹ This paper has been revised for this volume from my “A Buddhist View of Free Will: Beyond Determinism and Indeterminism”, *Journal of Consciousness Studies*, vol. 18, no. 3-4 (2011), pp. 217–33; I thank *JCS* for permission to reproduce it here.

References

Ainslie, G., 2001. *Breakdown of will*. Cambridge: Cambridge University Press.

Aquinas, T., 1947. *The "Summa Theologica" of St. Thomas Aquinas*. Translated by Fathers of the English Dominican Province. New York: Benziger Brothers.

Bojowald, M., 2006. Unique or not unique? *Nature*, 442, pp. 988–90.

Chisholm, R., 1982 (1964). Human freedom and the self. In: G. Watson, ed. *Free Will*. Oxford: Oxford University Press. pp. 26-37.

Dennett, D., 2004. *Freedom evolves*. New York: Penguin.

Federman, A., 2010. What kind of free will did the Buddha teach? *Philosophy East and West*, 60(1), pp. 1-19.

Foster, J., 1991. *The immaterial self*. London: Routledge.

Harvey, P., 1995. *The selfless mind: personality, consciousness, and nirvana in early Buddhism*. Surrey: Curzon Press.

Harvey, P., 2007. 'Freedom of the will' in the light of Theravāda Buddhist teachings. *Journal of Buddhist Ethics*, 14, pp. 35-98.

Hawking, S.W. and Hertog, T., 2006. Populating the landscape: a top-down approach. *Physical Review D*, 73, pp. 123527-1-123527-9.

James, W., 1958 (1899). *Talks to teachers on psychology, and to students on some of life's ideals*. New York: Norton.

Jinpa, T., 2002. *Self, reality and reason in Tibetan philosophy: Tsongkhapa's quest for the middle way*. London: RoutledgeCurzon.

Jinpa, T., ed., 2006. *Mind training: the great collection*. Compiled by S. Gyalchok and K. Gyaltzen. Translated by T. Jinpa. Boston, MA: Wisdom.

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. 217-40.

Lamrimpa, G., 2002. *Realizing emptiness: Madhyamaka insight meditation*. Translated by B.A. Wallace. Ithaca, NY: Snow Lion.

Laplace, P., 1951. *A philosophical essay on probabilities*. Translated by F.W. Truscott and F.L. Emory. New York: Dover.

Lau, H.C., Rogers, R.D., and Passingham, R.E., 2007. Manipulating the experienced onset of intention after action execution. *Journal of Cognitive Neuroscience*, 19(1), pp. 81-90.

Maxwell, N., 1894. *From knowledge to wisdom: a revolution in the aims and methods of science*. Oxford: Blackwell.

Paul, D.M., 1980. *The Buddhist feminine ideal—Queen Śrīmāla and the Tathāgata-garbha*. Missoula, MT: Scholar's Press.

Pereboom, D., 2001. *Living without free will*. Cambridge: Cambridge University Press.

Putnam, H., 1990. *Realism with a human face*. Edited by J. Conant. Cambridge, MA: Harvard University Press.

Raju, P.T., 1985. *Structural depths of Indian thought*. Albany, NY: SUNY Press.

Renouvier, C., 1912. *Essais de critique générale*. Paris: A. Colin.

Van Inwagen, P., 1983. *An essay on free will*. Oxford: Clarendon Press.

Wallace, B.A., 2006a. *The attention revolution: unlocking the power of the focused mind*. Boston: Wisdom.

Wallace, B.A., 2006b. *Contemplative science: where Buddhism and neuroscience converge*. New York: Columbia University Press.

Wallace, B.A., 2007. *Hidden dimensions: the unification of physics and consciousness*. New York: Columbia University Press.

Wallace, B.A., 2011. A Buddhist view of free will: beyond determinism and indeterminism. *Journal of Consciousness Studies*, 18(3-4), pp. 217–33.

Wegner, D.M., 2003a. *The illusion of conscious will*. Cambridge, MA: MIT Press.

Wegner, D.M., 2003b. The mind's best trick: how we experience conscious will. *Trends in Cognitive Sciences*, 2, pp. 65-69.

Wheeler, J.A., 1983. Law without law. In: J.A. Wheeler and W.H. Zurek, eds. *Quantum theory and measurement*. Princeton, NJ: Princeton University Press. pp. 182-213.

Wigner, E.P., 1983. Remarks on the mind-body question. In: J.A. Wheeler and W.H. Zurek, eds. *Quantum theory and measurement*. Princeton, NJ: Princeton University Press. pp. 168-81.

11 Degrees of Freedom: The Buddha's Implied Views on the (Im)possibility of Free Will

Martin T. Adam

1 Some Basic Distinctions¹

Much of the psychological impetus driving philosophical discussions of free will derives from ordinary, commonly entertained intuitions concerning actions we perform in circumstances of duress or coercion. In most instances our practice is that we don't hold the agent fully responsible for such acts. This corresponds to our personal experience and is intuitive. Do unto others, we say—and who has not felt dismay or anger at being held morally responsible for doing something that was forced upon one? And yet, even in such cases, a subjective sense persists that one's actions remain one's own; there lingers a feeling of personal responsibility. And so the question arises: Just how free need one be for personal responsibility to obtain? And in what sense free? Such questions form the very substance of the ancient and apparently unending thread of Western philosophical discourse referred to as 'the free will problem'. I seek to expand the boundaries of this conversation by opening it up to a perspective on freedom originating beyond Western intellectual horizons—namely, that of early Buddhism, as represented in the Pāli Canon of the Theravāda tradition. To that end, I examine some of the implications of the *Anattalakkhaṇa Sutta* (*Discourse on the Characteristic of Non-self*), comparing these with Harry Frankfurt's account of free will. I argue that the Buddha's position

on human freedom is unique, implying the denial of a metaphysically free will while asserting moral responsibility and the possibility of spiritual freedom.

One basic distinction Western philosophers draw is between *empirical* and *metaphysical* freedom. Empirical freedom refers to the ability to act as one wants. Formulated negatively, it can be understood in terms of the absence of constraints obstructing an individual's ability to do as they like. The notion of 'constraint' can be understood as either *external* or *internal* to the agent. Philosophers have distinguished different sets of constraints in spelling out different understandings of freedom. Political philosophers, for example, have focused on external restrictions such as those imposed by governments, political classes, and material conditions. Psychologically minded thinkers have emphasized internal constraints such as compulsions, obsessive thoughts, depression, confusion and so on.

For philosophers working in the area of metaethics, however, it is the idea of *metaphysical* freedom that has seemed most germane. Metaphysical freedom, like empirical freedom, can be understood negatively as an absence of constraints. In this case, however, the constraint is understood *in abstracto*—as causality itself. Moral responsibility is thought to require some kind of freedom from, or exception to, the necessity and universality that characterize the normal cause and effect operations of nature (van Inwagen 1975). Attaching a clear meaning to such a notion has, however, proven problematic. Two basic approaches have been attempted. The first asserts that a metaphysically free will would entail that some of one's actions or decisions are *uncaused*. This approach has been thoroughly criticized as implying randomness rather than freedom (Dennett 1984), and will not be dealt with here. The present discussion will, however,

involve another kind of account, one that has proven much more resilient. In this view, to assert metaphysical freedom is to assert that at least some of one's actions or decisions are *self-caused*.

One final distinction must be observed here. We can enumerate three principal subjects to which 'freedom' has been predicated, *viz.*, persons, wills, and actions. Conceptually, freedom of *the will* seems to stand between freedom of *the person* and freedom of *action*. Authors often slide between these three things, assuming that the predication of freedom to one *eo ipso* implies a statement of the same truth-value for the others. However, this is not the case within the basic Buddhist soteriological framework.

2 The *Anattalakkhaṇa Sutta*

The *Anattalakkhaṇa Sutta* is the Buddha's second sermon, delivered to his first five disciples. Here the Buddha systematically argues against the possibility of identifying the self with any of the five psychophysical aggregates that constitute a person: form, feeling, perception, volitional formations, and consciousness. While this *sutta* (discourse) is not normally considered as addressing free will, its teachings have implications for free will. For the Buddha suggests that none of the aggregates can be identified with the self because none is *subject to control*.

Beginning with the body or form (*rūpa*) the Buddha states:

“Bhikkhus, form is non-self. For if, bhikkhus, form were self, this form would not lead to affliction, and it would be possible to have it of form: ‘Let my form be thus; let my form

not be thus.’ But because form is non-self, form leads to affliction, and it is not possible to have it of form: ‘Let my form be thus; let my form not be thus.’”² (SN III 66)³

An identical line of reasoning is offered for each of the aggregates. To appreciate the implications for free will we need to see that the Buddha is relying on a conceptual connection between the notions of *self* and of *control*. If there were a self, he asserts, it would be that aspect of the person over which one would have control. We do not have control over any of the aggregates. The five aggregates are all that constitute a person. Therefore, there is no self (Pāli: *anattā*; Sanskrit: *anātman*).

o

Bhikkhu Bodhi makes some observations about the basis of this argument. The selflessness of the aggregates is demonstrated:

“...on the ground that they are insusceptible to the exercise of mastery (*avassavattitā*). If anything is to count as our ‘self’ it must be subject to our volitional control; since, however, we cannot bend the five aggregates to our will, they are all subject to affliction and therefore cannot be our self.” (Bodhi, 2000, pp. 1066-67)⁴

Thus, if there were a self, we would be able to control its states. In the above passage concerning *rūpa*, we would choose not to suffer and to be well in our bodies if we could; this is our natural wish and predisposition. Nevertheless, we remain afflicted and disposed to affliction. Suffering is inherent to *rūpa*. We cannot will it away. If *rūpa* were the self, we would be able to.

Importantly, the sense in which we are said to lack control over *rūpa* is one of *direct* control over

its *states*, in particular its state of being subject to *affliction*. In the above passage there is no denial that we *can do* as we wish with respect to the actions we perform with our bodies; what is denied is that we *can be* as we want with respect to the presence or absence of affliction. The wish that the Buddha describes as impossible to fulfill is ‘Let my form *be* thus, let my form *not be* thus’, not ‘Let my form *do* thus, let my form *not do* thus’. If free will is simply understood as the empirical ability of persons to act voluntarily, to do as they want, the Buddha’s position here does not imply any denial of this. All it suggests is that we cannot directly will away the suffering associated with the first aggregate. In fact, the Buddha’s teachings are premised on the idea that *it is possible* to do something about suffering; indeed we *can* eliminate it. But we cannot simply do away with it *directly*.

Are we then to conclude that Buddhist doctrine implies a qualified free will, one in which we can *do* as we will if not actually *be* as we will *immediately*, according to our wishes? Is this the end of the story? Actually, the Buddha’s implied position turns out to be considerably more complex than this.

To understand how this is so, we need to revisit the concept of ‘the will’. Let us follow others in tentatively identifying the English language concept *will* with the Pāli concept *cetanā* (Harvey, 2007, p. 47). However inexact this match may be, the concept of ‘the will’ must correspond to *some* aspect(s) of the five aggregates—and this is actually all we need to proceed with our argument. *Cetanā* is considered part of the fourth aggregate, *saṅkhāra*. The latter term is commonly translated as ‘volitional formations’, a heading meant to capture those mental events that direct one’s actions—physical, mental and vocal. It would appear, then, that *volitional*

formations constitute the aggregate in virtue of which action is *voluntary*. Keeping this in mind allows us to raise a deeper question regarding the will's freedom. For, as mentioned, an analysis identical to that carried out on *rūpa* is applied to each aggregate in turn—including *sankhāra*.

“Volitional formations are non-self. For if, bhikkhus, volitional formations were self, they would not lead to affliction, and it would be possible to have it of volitional formations: ‘Let my volitional formations be thus; let my volitional formations not be thus.’ But because volitional formations are non-self, volitional formations lead to affliction, and it is not possible to have it of volitional formations: ‘Let my volitional formations be thus; let my volitional formations not be thus.’” (SN III 67)

Thus, it would appear that the aggregate that includes the will is itself not subject to control. Following our analysis with respect to *rūpa*, the lack of freedom here consists in our inability to make *sankhāra* unafflicted directly by wishing it to be so.⁵ This seems a critical consideration; it suggests that the mental factors determining the morality of actions are themselves not subject to control. The mental states that direct our actions—desires, attitudes, and values we identify with—are *themselves* not under control. Thus, it appears that we are unfree with respect to what we *will*, rather than with regard to what we *do*.

If this is indeed the implication, then it seems that the Buddha would likely not have disagreed with the assertion famously attributed to Schopenhauer (1985): “A man can do what he wants, but not want what he wants” (quoted in Einstein 1982, p. 8). The Buddhist analysis suggests that the issue of free will is not simply a first-order problem as to whether we can do what we want.

There is a much deeper concern—one that turns on *second-order* considerations as to whether we can be what we want to be, or, put another way, whether we can have the wills we want to have. The issue of the will's freedom is a question regarding whether we have freedom with respect to our own constitutions. The Buddha's answer appears to be negative. While it may be that we can be judged empirically free to the extent that we can *do* as we want, we are not metaphysically free in the sense of being able to directly determine the constellation of factors that the mind identifies with, and out of which our actions emerge. The reason for this assertion is clear: the will is not subject to control because, quite simply, there is no one over and above the shifting configuration of mental factors to do the controlling. There is no controller. There is no one (*i.e.*, no final independent unity) holding the reigns. There is no self.

3 Harry Frankfurt Meets the Buddha

Second-order considerations are critically important in the well-known analysis of free will provided by Frankfurt:

“Besides wanting and choosing and being moved *to do* this or that, men may also want to have (or not to have) certain desires and motives. They are capable of wanting to be different, in their preferences and purposes, from what they are. Many animals appear to have the capacity for... ‘desires of the first order’, which are simply desires to do or not to do one thing or another. No animal other than man, however, appears to have the capacity

for reflective self-evaluation that is manifested in the formation of second-order desires.”
(1971, p. 7)

Frankfurt’s observations concerning the self-reflective powers of human beings seem directly pertinent to the Buddhist analysis, where they find an obvious resonance in the human ability to reflect upon and have desires concerning the aggregates. Frankfurt aims to provide a coherent account of free will in terms of the capacity to form second-order volitions about one’s first-order desires. He identifies the will with the first-order desire that actually moves, or would move, an individual to act. This Frankfurt terms the agent’s ‘effective desire’.⁶

“[The notion of the will] is the notion of an *effective* desire—one that moves (or will or would move) a person all the way to action. Thus the notion of the will is not coextensive with what an agent intends to do. For even though an agent may have a settled intention to do X, he may none the less do something else instead of doing X because, despite his intention, his desire to do X proves to be weaker or less effective than some conflicting desire.” (1971, p. 8)

Frankfurt’s account turns on the notion that one acts freely only if one wants to be moved by the desire that moves one to act. If one does not want to be moved to act by that desire, but is moved by it, then the will is unfree. Frankfurt employs the example of an unwilling drug addict. In analyzing the addict’s condition one must understand that the agent is the subject of conflicting first-order desires: he both wants and does not want to take the drug. In indulging his habit, however, he is being moved to act in a way that he wishes not to. His desire to take the drug on

these occasions, because it is effective in moving him to act, is to be identified with his will. And in this case it is unfree. It is unfree precisely because the agent has a negative second-order volition against it, *i.e.*, a desire that this desire not move him to act.⁷ In cases where this is not so, which is to say, in cases where one wants to be moved by the desire that is effective in moving one to act, the will is free. (pp. 9-11)

Frankfurt's version of free will makes sense of some common intuitions regarding our everyday actions. Most of us, most of the time, are moved to act by desires we want to move us. Hence, on Frankfurt's analysis, most of our actions are free. This way of thinking about things makes sense of those instances in which we 'feel free' in acting and are therefore willing to take responsibility for what we do. Our actions reflect our choices and the values we identify with. In brief, they reflect 'who we are' (or at least who we take ourselves to be). We do think of such actions as freely willed.

On the other hand, Frankfurt's account is not without its counterintuitive aspects. As we have seen, the identification of the will with one's effective desire entails a denial of free will to Frankfurt's addict. This runs against our intuition that persons are always in possession of a free will—even when their actions are compelled. In such cases we usually say that one is acting against one's own will, which is thought of as remaining free even when one is forced to act against it.

There is, in fact, another well-attested understanding of the will that would support this latter intuition. According to this understanding, in saying that one wills something, there is no

implication of effort. If, *contra* Frankfurt, we conceive of the will as the desire (or set of desires) that one *most identifies with*—as opposed to one’s effective desire—we can maintain that while the unwilling addict’s *action* is not free, his will, which he is unable to act upon, remains so. The notion of will is here connected to one’s deepest wishes and values—even one’s self-concept. The manner in which Bodhi speaks of the will above seems to reflect this usage: the will is identified with a very deep desire indeed, the desire to be free from affliction—ineffective though this is. In this way of speaking, persons can lack free will only in cases where they lack a desire (or set of desires and preferences) that they identify with—a circumstance that would seem applicable only to the unconscious (or, just possibly, the enlightened).

Philosophical discussions of free will appear to be divisible into these two different ways of conceiving the will. Obviously, these two conceptions of the will imply different ways of talking about *free* will. It is, therefore, essential to be clear which concept is being assumed. It would seem that we are faced with a choice of locutions. In one, freedom of the will is conceptually bound to freedom of action: one’s will is free if and only if one’s action is free. In the other, freedom of the will is tied to freedom of the person, and indeed to the very concept of identity and personhood. In the latter manner of speaking, it is possible for one to act unfreely even while retaining one’s free will.

Two further difficulties with Frankfurt’s account seem relevant to our concerns. The first is that an individual’s second-order desires and volitions are not consistent through time. In some cases they are in direct conflict from one time to the next. Desires change depending on a great variety of internal and external conditions. We are inconsistent as to what we want our will to be. Which

of one's various 'selves' is to be identified as one's true self? On what basis? This issue is clearly relevant in the Buddhist context.

A moment's reflection reveals a second problem: an infinite regress threatens to develop when freedom is made to turn on the presence of higher order volitions. If the will's freedom depends on a second-order volition, do we not then require a third-order volition to ensure the freedom of the second? We seem to be faced with the prospect of an infinite regress to higher-order volitions, each needed to guarantee the freedom of the ones below.

One could, of course, respond to this by saying that as a point of empirical fact all we ever really do have are desires of the first and second order or, at most, the third. If we choose to speak of further higher-order desires, it is not clear we would be referring to anything. At some point there is no further 'I want'; we simply find ourselves with certain basic desires, values and preferences that are not chosen or even consciously entertained. Incompatibilist determinists argue that the causes that give rise to these mental states are not subject to control; if one traces them back far enough, they are *impersonal* in nature (e.g. historical, genetic, cultural, etc.). Even if our present awareness can reflect on and evaluate our choices, the thoughts, values and desires entering into such evaluations are ultimately beyond our control. The Buddhist position would appear to accord with this perspective. Whether one identifies the will with one's effective desires or the desires that one most identifies with, in the end there is no final, independent person where the chain of causes and conditions find their origin. In the last analysis it is not possible to have it of the will, 'Let my will be thus, let my will not be thus'.

4 The Foundation of Morality

If this is so, should we conclude that the Buddhist position, like that of the incompatibilist determinist, undermines the foundations of moral responsibility? If there is no self to which responsibility may ultimately be attributed, is there no moral responsibility? Interestingly, from the Buddhist perspective the answer to this question is negative. In fact, the Buddha's teachings imply a very unusual view (from a Western perspective): while the will is not metaphysically free, morally responsibility is just a fact about the way things are.

Although ultimately there is no self, persons' actions do have results that accord with the moral character of those actions. Just as *moral causality* is one kind of causality operating in the universe, so too *moral responsibility* is simply one kind of causal responsibility. Like it or not, results flow from actions; happiness and suffering are the results of moral (*kusala*) and immoral (*akusala*) actions. Such action (Pāli: *kamma*; Sanskrit: *karma*) is distinguishable as mental, physical, and vocal behavior willingly done (i.e. accompanied by *cetanā*); *cetanā* is the key factor in determining moral responsibility. *Freedom* of the will is not. The point is that the action is voluntary, not that the will is metaphysically free in some way. Universal causality is not considered a constraint or obstacle to moral responsibility from the Buddhist perspective; it is, rather, a requirement.⁸

5 Degrees of Freedom

Freedom in Buddhism is not understood as a quality of the *will*. If there is no independent source of volitions over and above our mental, physical and vocal actions, then there cannot be free will in any ultimate sense. It is precisely from the higher perspective that the will can be seen to be unfree. Our lack of free will logically follows from the Buddhist position on the ontology of the self. There is no independent self. Just as the self is known to be a delusion, so too must free will be seen. No self, no free will.

This is a difficult point. It is not, perhaps, irrelevant that the *Anattalakkhaṇa Sutta* was not addressed to an audience of ordinary people (*puṭhujjanas*), but to a small group of ‘learners’ or disciples in higher training (*sekha*)—individuals who had already attained the higher perspective that sees things as they really are. There is an important sense in which such individuals, beginning with the ‘stream-enterer’ (*sotapanna*), are free already. They are free from the delusion of self.

The notion of the *sekha* is defined in terms of having undergone a transformative insight into the truth of no-self. The five disciples are said to have experienced this insight some days earlier, upon hearing the Buddha’s first sermon, the *Dhammacakkappavattana Sutta*.⁹ Upon hearing the second, it is said that they became fully liberated beings or *arahats* (Bodhi, 2000, p. 1066).

These considerations provide a clue as to how freedom in Buddhism might be best understood.

‘Freedom’ is a predicate of persons and consists in an absence of suffering and its causes. It is dependent on the state of knowledge of the agent. The ultimate aim of Buddhism is freedom

from suffering and rebirth—realities that, first and foremost, are to be *understood*. Thus, freedom implies knowing, and then abandoning, the causes and conditions that give rise to suffering and rebirth (*patīccasamuppāda* (Pāli): dependent origination; *pratītyasamutpāda* (Sanskrit)). The delusion that there is a self lies at the basis of this chain of causes. The insight that there is no self allows the mind to become free.

Different degrees of knowledge and mental purity are attributable to the various kinds of agent within the Buddhist soteriological framework; corresponding levels of freedom may be attributed to them accordingly.¹⁰ The *puṭhujāna* cannot be called a ‘free’ person, operating within the deluded perspective of being an independent actor in control of her life in *saṃsāra* (cyclical delusional wandering). Although such an agent may be reflexively aware of her actions, and although such actions may be voluntary, they occur in the context of the basic delusion of ‘self’, whence they are regarded as originating independently. Thus, the ordinary person’s mind is inevitably trapped in delusion, conflicting desires, and suffering.

The *sekha* is free in one important respect, having eliminated the delusion of self, and with it, it should be noticed, any notion of possessing an independent will. Being irreversibly oriented away from suffering and its causes, the *sekha* can be described as consistently having the desires she wants to have. An internal order has been established; such a person cannot do otherwise than act in a way that leads to *nibbāna* (Pāli; Sanskrit: *nirvāṇa*). Freedom here is clearly a function of knowledge, rather than of a capacity to do or choose otherwise. Although the mind of the *sekha* remains obscured to some degree, constrained by residual mental fetters (*saṃyojana*), the complete freedom of *nibbāna* is assured.¹¹

If we apply Frankfurt's analysis of the will, an *empirically* 'free' will can be attributed to the *sekha* insofar as she has the effective desires she wants to have. Thus, one could say her *actions* are free, and this dovetails the Buddhist view that they are informed by the realization of *anattā* and the prospect of *nibbāna*.

The *arahat* has realized *nibbāna*. She is free from all fetters and any trace of self-centered desire; indeed, because she no longer reaps the results of her acts and will not be reborn, there is an important sense in which she is seen as *free from action* itself.

Conclusion

Freedom in Buddhism can best be characterized negatively in terms of freedom from constraints upon a person—either internal or external, depending on one's focus. That is, to the extent different categories of agent are free from the internal constraints of delusion and other fetters, so too are they free externally in relation to *samsāra*. The *puthujjana* is not free from the delusion of self or from *samsāra*. The *sekha* is free from the delusion of self but not yet free from all fetters and from *samsāra*. The *arahat* is free from all internal constraints and thus also from *samsāra*. Such a person is describable as being *free from action*. Indeed, being free from all self-centered desire, the *arahat* can be described as *being free from the will*—as opposed to *possessing freedom of the will*. To see *this* is to recognize that Buddhist perspectives on freedom

emerge from a very different set of paradigms than those that inform most Western philosophical discussions of the free will problem.

Notes

¹ This paper significantly revises some of the arguments of my 2011 paper on the same topic, “No Self, No Free Will, No Problem: Implications of the *Anattalakkhana Sutta* for a Perennial Philosophical Issue”, *Journal of the International Association of Buddhist Studies*, special volume, pp. 239-265; I thank *JIAS* for kindly granting permission to reprint elements of that paper here.

² Translations are those of Bodhi (2000).

³ “SN” abbreviates the *Samyutta Nikāya*.

⁴ He also writes “...the aggregates are suffering because they tend to affliction and cannot be made to conform to our desires” (Bodhi, 2000, p. 842).

⁵ Sayadaw (1996, p. 49) indicates how we would change our volitional formations if we could: we would make them all wholesome (*kusala*) and not unwholesome (*akusala*).

⁶ The concept of *will* as effective desire is traceable as far back as Locke (1959), pp. 313-315.

⁷ Frankfurt defines second-order *volitions* as a type of second-order desire. Second-order desires are simply desires concerning one’s first-order desires. Second-order volitions are second-order desires that have as their object the *efficacy* of one’s first-order desires. This is an important qualification, as one can have a second-order desire for a first-order desire without wanting the latter to be effective. We can imagine, for example, that the drug addict wants to have the desire to give up drugs while simultaneously wanting this desire not to be effective. “If I didn’t want to

give up drugs at least a little bit”, he might reason, “then friends wouldn’t sympathize and lend me money”. (1971, p. 9-11)

⁸ One might well ask how it is that if the aggregates are ultimately beyond our control we could ever begin to strive for the ending of suffering. The Buddha’s response is found in the *Mahali Sutta* (SN III 70). See Adam (2011).

⁹ Hence, the unstated assumption in the Buddha’s second sermon—that the five aggregates are all that a person is.

¹⁰ See Adam (2005, 2008) for discussion of different classes of agent in relation to key moral vocabulary, principally *kusala-akusala*, *puñña-apuñña*, and *sukka-kaṇha*.

¹¹ Ten fetters are progressively eliminated along the supramundane path; corresponding degrees of freedom are attributable to the subcategories of *sekha*, *i.e.*, the stream-enterer, once-returner (*sakadāgāmin*), and non-returner (*anāgāmin*).

References

Adam, M.T., 2005. Groundwork for a metaphysic of Buddhist morals: a new analysis of *puñña* and *kusala*, in light of *sukka*. *Journal of Buddhist Ethics*, 12, pp. 62-85.

Adam, M.T., 2008. Classes of agent and the moral logic of the Pali Canon. *Argumentation* 22, pp. 115-124.

Adam, M.T., 2011. No self, no free will, no problem: implications of the *Anattalakkhana Sutta* for a perennial philosophical issue. *Journal of the International Association of Buddhist Studies*, special volume, pp. 239-265.

Bodhi, B., trans., 2000. *The connected discourses of the Buddha: a new translation of the Saṃyutta Nikāya*. Somerville, MA: Wisdom.

Dennett, D.C., 1984. *Elbow room: the varieties of free will worth wanting*. Cambridge, MA: Bradford Books.

Einstein, A., 1982. *Ideas and opinions*. New York: Three Rivers Press.

Frankfurt, H., 1971, Freedom of the will and the concept of the person. *Journal of Philosophy*, 68(1), pp. 5-20.

Harvey, P., 2007. 'Freedom of the will' in the light of Theravāda Buddhist teachings. *Journal of Buddhist Ethics*, 14, pp. 35-98.

Locke, J., 1959 (1690). *An essay concerning human understanding*. Edited by A.C. Fraser. New York: Dover.

Sayadaw, V.M., 1996. *The great discourse on not self*. Bangkok: Buddhadhamma Foundation.

Schopenhauer, A., 1985. *On the freedom of the will*. Oxford: Basil Blackwell.

Van Inwagen, P., 1975. The incompatibility of free will and determinism. *Philosophical Studies*, 27, pp. 185-199.

12 Buddhist Paleocompatibilism

Mark Siderits

1

My topic is a view that I call 'Buddhist Paleocompatibilism'.¹ This is something that I've talked and written about before (1987, 2008), and I should apologize in advance for the fact that I shall not have much to say about it that I have not said, in one way or another, before. What I hope to do is explain the view more clearly than I have in the past. It is a view that I find interesting and perhaps plausible, and I think it serves as a good illustration of the sort of thing that can happen when we bring two distinct philosophical traditions into conversation with one another.

While I call the view 'Buddhist Paleocompatibilism', I should say right at the beginning that this is not a view that (as far as I know) any Indian Buddhist philosopher actually held. Since we now know of no classical Indian philosopher who claimed that determinism and moral responsibility are incompatible, no Buddhist philosopher would have felt compelled to defend the opposing compatibilist view.² Buddhists were forced to deal with a problem that is in some ways related to the problem of determinism and responsibility—the problem of reconciling karmic justice with the Buddhist doctrine of non-self. And it is their response to this problem that first led me to think that if they were faced with the problem of determinism and responsibility, paleocompatibilism might be their answer.³ So, while the view has a Buddhist heritage, it is not strictly

speaking a classical Indian Buddhist view. But neither is it a Western view. While the problem it addresses is modern and Western, some of the tools it employs are distinctively Buddhist. And therein, I think, lies much of its interest.

Paleocompatibilists make two key claims:

1. There is no sound argument for the incompatibility of determinism and moral responsibility.
2. In the absence of any compelling incompatibilist argument, we should accept the common-sense view that persons are generally morally responsible for their actions.

While (2) might seem uncontroversial, there are difficulties involved in spelling out what moral responsibility would look like given various possible ways of establishing the antecedent of (2). But I shall skip over those now and focus on the defense of (1). The strategy here uses the doctrine of two truths.⁴ That doctrine has it that there are two ways a statement may be said to be true, ultimately and conventionally.

The defense of (1) will be that the form of determinism said by incompatibilists to be incompatible with moral responsibility could only be true ultimately, while the claim that persons are morally responsible is true conventionally, and there can be no entailment relations across the barrier between ultimate and conventional truth, so nothing about responsibility could follow from the thesis of determinism. Any argument

meant to show that the truth of determinism would undermine moral responsibility has to fail.

2

Let me begin to explain all this by saying more about the two truths. Ultimate truth is easy to explain. It is just what most people think of as truth: the property a statement has when it represents reality as being a certain way and reality mind-independently is that way. But the stipulation that correspondence be to how things mind-independently are means that many statements we take to be true are not ultimately true. Take the statement, "I am in Sinyang Hall". Since buildings and other such composite entities are not mind-independently real, no statement about them could be ultimately true. That composite entities are not ultimately real is established by the 'neither-identical-nor-distinct' argument: the building cannot be identical with the atoms, since it is one and they are many; but neither can it be distinct from them, since it cannot be wholly located where each atom is (it being too large), and the hypothesis that it is only partly located where each atom is leads to an infinite regress (since the building can only be partly located where each of its atom-sized spatial regions is, thus necessitating that it have parts, etc.).⁵ It is important to add that while "I am in Sinyang Hall" is not ultimately true, it is not ultimately false either. For there being ultimately no Sinyang Hall, the statement "I am not located in Sinyang Hall" could not be ultimately true.

While the statement lacks ultimate truth-value, the statement is conventionally true.

Explaining what this means is rather more difficult. How, one wants to know, could

there be any other way for a statement to be true than by corresponding to mind-independent reality? Well, “Hamlet is a Danish prince” is true, we say, yet there is no one in the world to whom the name ‘Hamlet’ refers. The statement cannot be really, literally true, yet we still think there is some sense in which it is true. That Hamlet was a Danish prince is, we say, ‘true in the story’. By this we mean that if the statements that make up the story were fact and not fiction, then this statement would be true in the literal sense. Conventional truth is a little like that. There really are no buildings, just atoms arranged in various ways. But given the way that the atoms around me are arranged, if in addition to atoms there were also such things as buildings, then “I am in Sinyang Hall” would be ultimately true. There aren’t such things as buildings. We only think there are because it’s useful for us to think of atoms arranged a certain way not simply as the many things they actually are, but as composing one big thing such as a building. If this pretense of ours reflected mind-independent reality, the statement would be ultimately true. Our pretense does not reflect reality—indeed, given the neither-identical-nor-distinct argument, it cannot. But just as the statement about Hamlet does bear a certain complex relation to marks on a page, so the statement about Sinyang Hall bears a complex relation to the atoms around us. That is why it turns out to be useful, and why we say it is true when, strictly speaking, it is not.

Now take the statement that Hamlet lives in Denmark. We know it is not literally true, but could it be ‘true in the story’ or ‘fictively true’? Opinions about this vary. The difficulty is that we think Denmark is a real place while Hamlet is not a real person, and it is not clear how a real place could be related to an unreal person in the right way to make the story true. If you think that a real place can figure in the truth-maker for a statement that is only fictively true, then you will think the statement can be fictively

true in a perfectly straightforward way. But if not, then you will need to hold that the true-in-the-story “Hamlet lives in Denmark” is not about the real Denmark but some other sort of place entirely. Something similar happens with conventional truth.

We agree that “I am in Sinyang Hall” is conventionally true, and we can also say that a statement like “There are such-and-such atoms arranged in such-and-such a way” is ultimately true. But what about “Sinyang Hall is made of such-and-such atoms”? Since it says something about Sinyang Hall, we know it is not ultimately true (or false either), but could it be conventionally true? Here too, opinions will vary. The difficulty in this case is that if we can talk of Sinyang Hall and those atoms in the same breath, this will lead to contradictions. Perhaps the easiest way to see why is to think of the *sorites* puzzles that will emerge when we begin to look into how many randomly chosen atoms we can remove before Sinyang Hall ceases to exist.⁶ And there are other ways to show why permitting such talk leads to dire logical consequences.⁷ Still, there are those who would say that there can be a conventionally true statement referring to both Sinyang Hall and the atoms.

Although Buddhist accounts of the two truths do not typically discuss this question, if they were asked, then, no doubt, many Buddhists would say just this. But this is because these Buddhists believe conventional truth harbors contradictions within itself, and so must be transcended. These Buddhists would not see a reason to try to insulate conventional truth from the contradictions that arise when you let in the referring expressions of ultimate truth. Other Buddhists, such as Dharmakīrti, would disagree. They would say that we cannot account for the usefulness of conventionally true

statements if they inevitably give rise to contradictions. So the insulation between the two discourses must be two-way.

To summarize, there are two things we might say about the conventional truth-value of “Sinyang Hall is made of such-and-such atoms”. We might say that it is conventionally true. That policy will lead to there being good reason to accept various contradictions. We might welcome this result as showing the inherent instability of our ordinary ways of talking and thinking. On the other hand, we might take the resulting contradictions to show that conventional truth must be reformed. The proposed revision is that we not allow referring expressions from ultimate discourse to be employed in conventional discourse. Buddhist Reductionists might take either of these two stances. What I want to explore is what happens if we take the second, according to which that sentence has no truth-value, since it is simply meaningless.

Now if that is the view we take of the relation between the two truths, then it follows quite straightforwardly that no argument the conclusion of which belongs to conventional discourse can have a premise that belongs to ultimate discourse. No statement about Sinyang Hall follows from any statement about atoms. That will sound counter-intuitive to many, but Buddhists will not be put off by this. All Buddhists agree that common sense is profoundly mistaken about any number of important facts about the world, so the intuitions shaped by common sense cannot always be trusted. What Buddhists disagree about is whether it is possible to revise common sense in such a way as to rid it of its liability to lead to contradictions. The Buddhist revisionists I am discussing now think it is. That is the point of their two-way semantic insulation. If they can then show that the relevant formulation of determinism could only be ultimately

true, while the claim that persons are morally responsible for some of their actions is conventionally true, then they will have a way of establishing (1).

3

This brings me to why I call the view 'paleocompatibilism'. The strategy will be to show that moral responsibility is a property only persons could have, while the relevant form of determinism concerns the parts of persons, such as volitions. Now the first modern Western compatibilists, Locke and Rousseau, both answered the incompatibilists by saying that one should ask not whether the will is free, but whether the person is free to do their will. Since by 'freedom' they meant something that is necessary for moral responsibility, they thus held that moral responsibility pertains to the person as a whole and not to one of its parts. And this much seems right. It is persons we hold accountable for actions, that are the objects of our praise and blame, gratitude and resentment—not hands, feet, hearts, lungs, brains.

Of course, the incompatibilists to whom Locke and Rousseau were responding were dualists who thought of the will as a faculty of the mind. And faculties, it can be said, are not parts of the substances in which they inhere; they are powers or abilities of those substances. So, if it is a mistake to ask if the will is free, it is not the mistake involved in confusing a whole and one of its parts. While a present-day incompatibilist might not want to affirm that the mind is a substance distinct from the body, they might still claim that the will is not a part of the person, but a faculty of the person. But now we must ask what it is about the exercise of this faculty of willing that leads to the thought that the

person, whose willing this is, is not responsible. Answering this question will take us back to wholes and parts.

The thought that if determinism holds universally, then the will is not free, stems, I submit, from the idea that the event of willing to do A is caused not by the person but by some prior event (which is in turn caused by some prior event, etc., back to a time before the person was on the scene). The kind of causation at work here is event causation: one event causing another. We also think of persons as causes. But persons are not events; they are substances. When we say that a person caused some happening—say that Kim caused the death of Lee—we think of the effect as an event (the demise of poor Lee), but the cause not as another event (the act of pulling the trigger, or the act of intending to pull the trigger), but as the thing behind that act, the agent Kim. Agency is a variety of substance causation. And substance causation is not a kind of causation that operates alongside or in competition with event causation. It represents an altogether different model of how the world behaves. Incompatibilist arguments exploit the fact that the two models do not work in tandem. The paleocompatibilist explains this disharmony by claiming that substance causation concerns wholes while event causation involves parts.

The mereological nihilism at the heart of Buddhist Reductionism has no substances in its final ontology. The atoms that are *all there really is* to what we think of as Sinyang Hall are not indivisible material particles. A particle is a substance, an enduring thing that bears various qualities (such as having a certain mass and charge). And the same argument that shows Sinyang Hall not to be ultimately real shows that no substance, no matter how small and simple in structure, could be ultimately real. The ultimately real

atoms are tropes, particular momentary occurrences of qualities such as resistance and heat.⁸ A material particle is a bundle of tropes. And of course it is we who do the bundling, for our own convenience; that is why material particles are not ultimately real.

As for the substance known as mind, Buddhist Reductionists analyze that into mental tropes: occurrences of such mental events as feelings of pleasure and pain, desirings, willings and cognizings. The belief that such events must have an owner arises from the fact that we have conceptually constructed a substance, the mind, as what holds a bundle of such tropes together. Descartes was wrong: from the occurrence of this cognizing event it does not follow that there is any 'I' that is cognizing. That this cognizing event occurs is ultimately true. That I exist is conventionally true. There are no entailment relations between the two discourses.

There is more that could be said about the *cogito*, but our present topic is such mental events as desirings and intendings. Incompatibilist arguments typically invite us to think that if determinism is true, then it must hold for events like these. And then a sort of vertigo sets in. We know how to answer questions about the cause of an action, such as the question of who caused Lee's death. Suppose we agree it was Kim. Presumably this means Kim wanted Lee to die and formed the intention to bring about Lee's death. But when we go on to ask about the cause of the intention, the question becomes somewhat puzzling. Are intendings things that I 'do' in the same sense in which I might do the killing of Lee? I know what it is to want someone dead, and to form an intention to bring this about. But how do I cause it to be the case that I intend Lee to die? By intending that I intend for Lee to die? How do I bring it about that this occurs? It begins

to look as if when we ask about the cause of the volition or the intention, only another event will serve, and not the substance that is the person. Buddhist Reductionism explains why this should be.

Mental events like desires and intentions are among the parts that make up the whole known as a person. This means that such events cannot be said to have persons as their causes. A mere conceptual fiction cannot be the cause of something ultimately real. The causes of such events must instead be impersonal events. So the incompatibilist turns out to be right after all in saying that if the psychological factors involved in performing an action are causally determined by prior events, then the springs of action can always be traced back to thoroughly impersonal forces. What the incompatibilist neglects to mention is that this is because when we speak of these mental events and their causes, we can only be speaking about ultimate truth, which is by nature necessarily impersonal. Nothing whatsoever follows about persons.

4

In defending his agent-causal version of incompatibilism against a Davidson-style objection, O'Connor (2003) confronts the difficulty that arises when we try to combine our talk of agents with what we say about events internal to the agent. It might prove useful to consider how he handles it, for this may throw some light on how paleocompatibilism both accommodates certain strands of the agent-causation approach and rejects its incompatibilism. Agent causation theories typically claim that while actions, as events, have causes, the cause of a basic action is not another event but

rather an agent, an enduring substance. On O'Connor's account, causation is an irreducibly primitive production relation. Causal powers are the properties of particulars, and causation is the manifestation of these powers. There are then two kinds of causation, event causation and agent causation.

Event causation may be thought of as the manifestation of the particular's having its nature in certain specific circumstances (one event) through the production of the effect (another event). Agent causation is the manifestation of the agent's agent-causal power in conferring on the agent the ability to produce at will some effect or other from a range of options. Only particulars that are in certain respects self-determining have agent-causal powers. Thus, a physicalist version of the theory will require that such powers be emergent properties of complex systems with the capacities of representing possible courses of action to themselves and of having beliefs and desires concerning those courses of action. (O'Connor 2003, p. 262)

On O'Connor's account, the basic actions produced by agents are initiations of actions through the formation of intentions. This means that I am the agent of the action of killing Lee by virtue of my producing the requisite intention, which is the event through which my agency flows to the successive events in the series. Intendings are indeed things that I do on this view. The Davidson-style objection, tailored to this formulation of the agent-causal theory, concerns the question what is the relation between agent and action, the action here understood as the primitive action of forming the determinate intention that will in turn initiate the freely-willed action (e.g., the arising of the executive intention to kill Lee, from which the grasping and firing of the gun flow). Specifically, is this relation itself an event or not? If so, then either this event is an

action or it is not. If it is an action, then the supposedly primitive action that is its result is not primitive, and we have the start of an infinite regress. If this event is not an action, then there is a causing by the agent that is not a doing by the agent. If, on the other hand, this relation is not an event, then since a causing is an event, it is no longer clear what it means to say that the agent caused the action.

O'Connor responds by grasping the second horn of Davidson's dilemma: the relation of the agent to the action is not a distinct event, but is instead identical with the action, understood as the agent's production of the intention. The relation—the agent's causing the action—is a complex state of affairs, consisting of the agent's being the agent-cause of the action. And this is not the sort of thing, O'Connor claims, that could itself be caused (2003, pp. 271-2). The threatened infinite regress is avoided by in effect making the producing relation identical with one of its relata.

The Nyāya school of classical Indian philosophy developed a similar strategy as a way to avoid various Bradley-style regresses. Included in their ontology are such things as universals, substances, and a relation of inherence that connects potness to particular pots. Asked how inherence is in turn related to particular pots, they answer that there is no distinct relation between inherence and a particular, that inherence is a self-linking connector, so that the relation between inherence and a pot is just the inherence itself. (This seems to be what Russell had in mind (1927, p. 263) when he sought to stop the original Bradley regress by claiming that relations are not at all like the substances that they relate.) Nyāya likewise holds that there are such things as absences, for instance, the absence of a particular pot from this table. But absences, they claim, only exist when there is an existing counter-positive: there can be the absence of the pot from this table

only because that pot exists elsewhere (or elsewhen). Asked about the relation between the absence and its counter-positive, their reply is that this relation just is the absence itself, which is a self-linking connector. O'Connor's reply to the question about the causal relation between the agent and the action is similar: by in effect making the action a self-linking connector, he hopes to thereby avoid a Bradley regress.

O'Connor's strategy has a distinguished pedigree, but can it succeed? The notion of a self-linking connector has gotten mixed reviews. The reception is somewhat warmer where we are antecedently inclined to say that the connection in question is across different ontological categories. To suppose, for instance, that some extra tie is needed to connect inherence with substances is to think of inherence as substance-like in being just another particular without intrinsic relations to other particulars. If we instead think of inherence as belonging to a category distinct from that of substance (as does Nyāya), then we may be more inclined to accept the claim that it is a self-linking connector. Those who wish for a more austere ontology containing fewer categories will resist such claims. Buddhist Reductionism represents an extreme form of such austerity, maintaining that we can make do with just the category of momentary trope-occurrences, all else being conceptually constructed on the base of these elementary events. If so, then the notion of a self-linking connector will turn out to be both incoherent and superfluous. But assessing the plausibility of so Spartan an ontology as that of Buddhist Reductionism is no easy task.

It might, however, be worth looking into the motivation behind this resort to self-linking connectors. For if it should turn out that the problem it is meant to address can be resolved in a simpler way, that will be one consideration against it. We started with

the difficulty that arises when we 'pop the hood' on my killing Lee in order to ascertain whether I am responsible for this regrettable event. Presumably, what we find when we look under the hood will tell us whether this is genuinely my act or not. The stock answer is that it is my act provided we find that it flowed from an intention wholly my own. How does an intention come to be wholly my own? If we say that it must reflect my character, then we may ask why we should take the relevant elements of my current character as truly my own. If they arose as the result of earlier elements identified as expressive of my nature, we may ask the same question of those earlier elements. Eventually, this quest will lead us back to a time before I was born. It is this prospect that leads the agent causation theorist to claim I can be responsible for the action only if the event of the intention's occurrence is the effect not of a prior event but of an agent, me.

That this line of thought is on the right track might be confirmed by considering the case of the 'Swampman double' who, when the lightning strikes the muck, arises as a fortuitous molecule-for-molecule duplicate of me as I am just before forming the requisite intention.⁹ Should Swampman last long enough to then form the intention and shoot Lee (who happens to be in the swamp that day) before dissolving back into the primordial ooze, there would be no judgment of responsibility. Such judgments require an enduring agent, and Swampman does not endure long enough to count. This is why, says the agent-causal theorist, responsibility drains away when we pop the hood and look for some inner event that might explain the causal chain leading up to the action. As long as our explanations are couched exclusively in terms of event causation, we will never find an enduring agent. This is why we must turn to causation of an altogether different sort.

The paleo-compatibilist will agree with this last point, but draw a different conclusion: that agent causation and event causation belong to the distinct discourses whereby we speak of wholes and of their parts, respectively. Given the semantic restrictions on these discourses, there can be no problem of explaining how an enduring substance can serve as cause of an event's occurrence at one time rather than another. And it is this dating problem that lies at the heart of the present difficulty. The agent-causal theorist's explanation fails to satisfy because it tries to account for the timing of the formation of the intention in terms of the agent's being in certain circumstances just before the intention was produced. In this case, it looks like we should say it is those circumstances that are the cause, with the result that the agent once again drops out of the picture. As long as agent causation and event causation are thought of as competing ways of explaining the occurrence of an event, event causation will always win out, given its ability to solve the dating problem.

The alternative picture offered by the Buddhist Reductionist places event causation at the ultimate level and substance causation at the conventional level. There being no substances at the ultimate level, event causation consists in the relation of universal concomitance and ordered succession between elementary event-types. Substance causation, the sort of causal relation asserted at the conventional level, involves the manifestation of a substance's powers in the production of events (understood as the arising of new properties in this or another substance). When the substance in question is an agent, we say this manifestation can occur 'at will'. This should not be taken to mean that the agent produces a willing that in turn brings about the initiation of the action. It means instead that it is up to the agent to bring about the action, that the agent

produces the action when they choose to, they ‘just do it’ when they see fit. To seek an explanation of the timing here is to pop the hood, to treat the agent as a system analyzable into components whose properties do the real explanatory work.

To those accustomed to the sorts of causal explanations found in science, it may sound odd to claim that the concept of causation at use in conventional discourse is that of substance causation. But this is not really all that untoward a view. If we follow the lead of Locke (1693, I.xxi.4), we should say that the child’s concept of causation begins with its learning to exercise control over its limbs. Seen in this light, it is hardly surprising that the child should attribute agency to inanimate objects. And even when we come to realize that such objects lack the representational and deliberative capacities necessary for agency, we continue to speak of them as the causes of various changes in the world. Just as Kim may be said to be the cause of Lee’s death, so cars, or power plants, or even trees (on Ronald Reagan’s view) may be said to be the cause of the polluted state of today’s air. While current common sense also allows talk of event causation (understood as strictly a relation between events), we may speculate that this has come about under pressure from the ascendancy of scientific explanation. We might even claim that the concept of substance causation is the more primitive concept not just ontogenetically, but phylogenetically as well—that it is what Perry (2010, p. 94) calls a ‘natural’ concept.

If determinism is true, then all mental events are causally determined by other events. We have seen why when this thesis is properly understood it turns out to be ultimately true. We can also see why the claim that persons are sometimes morally responsible for their actions could only be conventionally true. Given the semantics of the revisionist form of Buddhist Reductionism I am discussing, no conventionally true statement is entailed by any ultimately true statement. This completes the defense of (1): nothing about moral responsibility can follow from the thesis that all mental events are causally determined by other events. But it might be useful to illustrate (1) in action by looking at what it tells us about Benjamin Libet's experimental results.

In a series of experiments Libet purportedly showed that when subjects are instructed to spontaneously initiate hand movement on the basis of a sudden urge (i.e., an intention formed at the very last moment), the neural events that begin the process of movement occur before the subject becomes aware of intending to move their hand (1985).¹⁰ This is sometimes taken to show that the event of consciously deciding (the formation of a conscious intention) cannot have caused the initiation of movement. Critics contend that the experimental design is flawed, and that the data do not support this conclusion. But let us waive such questions. Suppose we agree that the subjects reliably report the time when they become aware of intending to move their hand, and that in this case the intention must be a conscious event. Since the initiation of movement precedes the conscious intention, the latter cannot have caused the former. (Of course, if we follow Dennett in thinking of consciousness as a property that mental events only acquire retroactively (2005), pp. 131-57), this conclusion does not follow; but on his view nothing is intrinsically a conscious mental event, so nothing can be a real cause by virtue of being a conscious state.)

The real question, though, is why this should be thought to matter. People often find these results spooky. And the spookiness is sometimes said to consist in this being empirical evidence that there is no free will. What people probably mean by this is that in the experiment the subjects were not responsible for moving their hands—that the movement was not really something the subjects *did*. The question is why anyone would interpret the results of the experiment in this way. The subjects did, after all, do what they were asked to. They produced a series of hand movements in a relatively spontaneous way: in such a way that an observer would not have been able to reliably predict just when the next movement was likely to occur from anything coming before. Of course, the subject is not aware of having decided to initiate movement until after the initiation process has begun. But this does not show that the subject did not initiate the movement. That would follow only if being the agent of the movement requires that one be the agent of the intending. And it might be a mistake to think of intending this way, as an action performed by an agent.

It is this conception of intending that seems to stand behind the thought that intentions must be formed consciously. The idea seems to be that persons come to have intentions by engaging in a process of deliberation about their situation and their standing desires and then choosing from among the range of available intentions the one that best matches the outcome of the deliberation process: springs of action not formed in this way are not truly chosen by the agent, and so are not genuine intentions of the agent. What else could explain the requirement that the formation of the intention be a conscious process, if not the idea that this must be under the control of the agent, and that control requires consciousness? Surely not the evidence of introspection.

We are often hard pressed to say, even with our most weighty actions, just when we decided to do this rather than that. The descriptions we give of the moment of choice are often no more than retrospective reconstructions. That this way of thinking about intending might be mistaken is also suggested by the fact that it introduces yet another action of the agent into the account, the choosing that yields the intention. As we saw earlier, an infinite regress threatens. We avoid this regress by keeping separate our talk of the agent's responsibility and our talk of the occurrence of the intention. The first concerns a big thing; the second concerns a small thing. Talk of big things and talk of small things do not mix.

Gallagher (2006) has a somewhat similar view of the controversy about Libet's results. He also thinks we are looking in the wrong place for 'free will' when we look at the neurological precursors to the initiation of hand movement. The notion, he says, "applies to intentional actions themselves, described at the highest pragmatic level of description" (p. 117). Neurological accounts of motor control are simply at a level of description (the 'subpersonal' level) too low to be relevant to the question of responsibility, just as a detailed account of the workings of the parts making up your car is not relevant to the question whether I should ride in it (p. 115). Instead, the question of 'free will' or responsibility only arises at the level at which it makes sense to speak of persons and the things they think of themselves as achieving through their intentional acts: doing the laundry, killing Lee, complying with the instructions in the psychology experiment.

Gallagher also holds that responsibility requires that intentions be formed through a conscious process. Although this may hold for the sorts of actions that we take as paradigms of 'free will', I am not sure it holds across the board. But what he says in defense of his claim is important. This is that the timeframe required to see the sorts of intendings involved in responsible agency is different from the one used in examining processes of motor control. The time between the neural events that initiate hand movement and the occurrence of the movement is measured in hundreds of milliseconds. With the action of killing Lee the intending can count as such only by being placed within a context of awareness of desired outcome (that Lee be dead) and opportunity (there is a gun before me) through initiation of movement (grasping the gun and pulling the trigger) through monitoring the movements (ascertaining that the gun is grasped by the handle and not by the barrel), to awareness of the goal's having been achieved (seeing the falling of Lee's body as Lee's death). The intending involved here is what it is only by being situated within a temporally extended process; it is not the sort of simple neural event that might count as what initiates the process of muscle contraction.

The Buddhist Reductionist will agree, and go on to claim that as a temporally extended event, the intending involved in the killing of Lee requires an enduring person as its subject.¹¹ The simple, short-lived mental event that initiates the process of muscle contraction requires no subject. By the mereological nihilism at the heart of Buddhist Reductionism, simple momentary events are themselves ultimately real. Temporally extended processes, on the other hand, can only be construed as properties of conventionally real substances. The intention that Libet studied is the wrong sort of thing to be involved in determinations of responsibility.

So much, then, for the defense of (1). I said earlier that there might be difficulties in connection with the second paleocompatibilist thesis:

2. In the absence of any compelling incompatibilist argument, we should accept the common-sense view that persons are generally morally responsible for their actions.

The difficulties I had in mind stem from the way in which the paleocompatibilist argues for (1) and thus tries to establish that there is no compelling argument for incompatibilism. The belief that persons are generally morally responsible for their actions is surely a part of common sense, something most people believe unreflectively. The Buddhist Reductionist view that underlies the defense of (1) is, however, profoundly revisionist. The common sense view has it, for instance, that persons really exist, whereas Buddhist Reductionism denies this. This raises the question whether the paleocompatibilist is actually affirming the common-sense view that persons are sometimes deserving of moral praise and blame.

One way of bringing out the difficulty here is to ask the Buddhist Reductionist why it is that if strictly speaking there are no persons, it should be useful to hold that there are, and attribute moral responsibility to them. The response will be consequentialist in nature: our institutions of moral praise and blame help maximize overall utility, and

these institutions require that we think of ourselves as persons. The common-sense view of desert is not, however, consequentialist. This can be seen, for instance, in the strongly retributivist intuitions most people have about punishment. So, the responsibility and desert delivered by paleocompatibilism are importantly different from those of common sense.

This is the familiar complaint about compatibilisms of every sort—that they deliver something other than what was promised. (Hence Kant’s complaint that compatibilism is ‘a wretched subterfuge’.) Neocompatibilists are often quite open about this. They claim that the notion of desert behind our common-sense view of responsibility is simply incoherent, and must be replaced by something saner. Here we have a situation that is common in philosophy: we find ourselves with conflicting intuitions and must, somehow or other, make the best of a bad lot.

Perhaps the paleocompatibilist can do slightly better than the neocompatibilist. Many people have the intuition that the person who does an evil deed thereby acquires the very real property of deserving blame and (other forms of) punishment. The neocompatibilist must say that this is simply a mistake. The paleocompatibilist has a different response: what common sense says is true. Of course it is, in their eyes, only conventionally true. Still, if we are to speak of persons at all, then this is what we must say. Moreover, this real desert could not be said to exist if we were to say that the psychological events involved in the production of an action are causally determined by prior events. The incompatibilist is right about this as well. The mistake lies not in our intuitions but in what the incompatibilist makes of them. The intuitions behind our views about responsibility uphold judgments that are true conventionally, while the

intuitions behind our views about determinism support claims that are true ultimately, so there can be no conflict here. The responsibility on offer is that of common sense. The conflict lies not in our intuitions but in what philosophers make of them.

So where does this leave us? There is some reason to believe that (1) and (2) are true. If they are, then it can be the case both that persons are sometimes morally responsible for their actions and that determinism is true. We have seen how that could be. I shall end by repeating what I said at the beginning, that I think this view is interesting and might even be true. Much more work is needed to figure out if it is true. And that work requires the active participation of scholars of Asian philosophy.

Notes

¹ This paper is a version of my 2013 paper by the same title, “Buddhist Paleocompatibilism”, *Philosophy East & West*, 63(1), pp. 73–87, which has been partly revised for this volume. I thank *PEW* for granting permission to reprint the paper here.

² The question whether determinism and moral responsibility are compatible has of course been the subject of much recent debate. The question is this: supposing that all events, including those psychological events involved in human deliberation, decision and action, are caused by earlier events in accordance with strict (non-stochastic) causal laws, would it then follow that no one is ever justifiably subject to moral praise or blame for their actions? It is widely called the issue of ‘free will’, but because many neo-

compatibilists find this term problematic, I shall employ the longer and perhaps more accurate label.

³ This was the train of thought behind Siderits (1987), though I did not then call the view 'paleocompatibilism'. There is, however, one passage in the *Nikāyas* that I believe can be read as at least hinting at something like this view. At *Āṅguttara* i.173 ff.

(*Mahāvagga, Tittthāyatana Sutta*), the Buddha discusses three views that he claims lead away from the practice leading to liberation: that everything a person experiences is the result of their past actions, is the result of God, is uncaused. These views are said to all lead to a state of soteriological paralysis. On the first two views, the present 'I' is a mere conduit of causation, channeling decisions made earlier over which I have no present control. On the third view, nothing I presently decide can be counted on to have any effect on the world. In all three cases I am left an impotent observer of the events making up my life, unresponsive to thoughts concerning what should and should not be done. Both determinism and the view that nothing is caused are equally corrosive of prudential responsibility. One might then expect the Buddha to seek wiggle room in an indeterminism lying between determinism and utter causelessness, one holding that certain human choosings are not determined by prior causes. But the 'middle path' he teaches on this occasion does no such thing. Instead, he explicates the formula of 'dependent origination', according to which suffering arises in dependence on a series of psychophysical events beginning with ignorance. In other contexts in which the Buddha invokes dependent origination as a 'middle path' in order to dissolve an apparent dilemma, Abhidharma exegetes took him to be utilizing the distinction between the two truths. If we were to read the *Tittthāyatana Sutta* in that way, the Buddha would be a paleocompatibilist.

⁴ The distinction between two kinds of truth, conventional and ultimate, was first developed by Abhidharma philosophers. It is central to the view of persons I call ‘Buddhist Reductionism’. Not all Buddhists accept Buddhist Reductionism. And not all Buddhist Reductionists accept the formulation of the distinction between the two truths that paleocompatibilism relies on. Many Buddhist philosophers do, however, draw the distinction in the way I am about to describe, e.g., Vasubandhu, Buddhaghosa, Saṃghabhadra, Dignāga, Dharmakīrti, and perhaps even some Pudgalavādins.

⁵ The argument is more fully described at Siderits (2007), pp. 104-111.

⁶ See Unger (1979) for the classic formulation of these difficulties.

⁷ For evidence that Buddhist Reductionists were aware of the sorts of semantic difficulties that arise when the contentious relatives—the referring expressions of the two discourses—are not kept apart, see Mikogami (1979), pp. 83-6. For discussion of some further considerations behind making the semantic insulation between the two truths two-way, see Siderits (2009).

⁸ Support for the claim that the *dharmas*, the ultimate *reals* of Buddhist Reductionism, are tropes can be found in Ganeri (2001), pp. 101-02; also see Goodman (2004).

⁹ For the origins of the Swampman-style thought experiment, see Davidson (1987).

¹⁰ A useful discussion is Pockett (2009), pp. 15-19.

¹¹ See Metzinger (2003), pp. 422-26 for a useful account of the cognitive architecture at work in the self-ownership of actions—what is required for the agent to view the action as their own and thus as something for which they can be held accountable.

References

Davidson, D., 1987. Knowing one's own mind. *Proceedings and Addresses of the American Philosophical Association*, 60, pp. 441-58.

Dennett, D., 2005. *Sweet dreams: philosophical obstacles to a science of consciousness*. Cambridge, MA: MIT Press.

Gallagher, S., 2006. Where's the action? Epiphenomenalism and the problem of free will. In: S. Pockett, W. Banks, and S. Gallagher, eds. *Does consciousness cause behavior?* Cambridge, MA: MIT. pp. 109-124.

Ganeri, J., 2001. *Philosophy in classical India*. New York: Routledge.

Goodman, C., 2004. The *treasury of metaphysics* and the physical world. *Philosophical Quarterly*, 54, pp. 389-401.

Libet, B., 1985. Unconscious cerebral initiative and the role of conscious will in voluntary action. *Behavioral and Brain Sciences*, 8(4), pp. 529-566.

Locke, J., 1693 (1995). *An essay concerning human understanding*. Amherst, NY: Prometheus.

Metzinger, T., 2003. *Being no one*. Cambridge, MA: MIT Press.

Mikogami, E., 1979. Some remarks on the concept of *arthakriyā*. *Journal of Indian Philosophy*, 7(1), pp. 79-94.

O'Connor, T., 2003 (1995). Agent causation. In G. Watson, ed. *Free will*, 2nd ed. Oxford: Oxford University Press. pp. 257-84.

Perry, J., 2010. Wretched subterfuge: a defense of the compatibilism of freedom and natural causation. *Proceedings and Addresses of the American Philosophical Association*, 84-2, pp. 93-113.

Pockett, S., 2009. The neuroscience of movement. In: S. Pockett, W. Banks, and S. Gallagher, eds. *Does consciousness cause behavior?* Cambridge, MA: MIT. pp. 9-24.

Russell, B., 1927. *An outline of philosophy*. London: Allen and Unwin.

Siderits, M., 1987. Beyond compatibilism: a Buddhist approach to freedom and determinism. *American Philosophical Quarterly*, 24, pp. 149-159.

Siderits, M., 2007. *Buddhism as philosophy*. Aldershot, UK: Ashgate.

Siderits, M., 2008. Paleo-compatibilism and Buddhist reductionism. *Sophia*, 47, pp. 29-42.

Siderits, M., 2009. Is reductionism expressible? In M. D'Amato, J.L. Garfield, and T.J.F. Tillemans, eds. *Pointing at the moon: Buddhism, logic, analytic philosophy*. New York: Oxford University Press. pp. 57-69.

Siderits, M., 2013. Buddhist paleocompatibilism. *Philosophy East & West*, 63(1), pp. 73–87.

Unger, P., 1979. There are no ordinary things. *Synthese*, 41, pp. 117-54.

13 'Shifting Coalitions' and the Responsibility of Persons

Ben Abelson

1 Buddhist Reductionism and the Extreme Claim

One of the philosophical doctrines commonly thought most central to Buddhism (in opposition to the Hindu *ātman*, the soul or self), is *anātman*, the 'no-self' view, the idea that there is no persisting essence of any individual, no substantial 'I' standing beneath or behind an individual's constantly arising and perishing parts and properties. The Buddha's teachings, as recorded in texts known as *Nikāyas* (baskets or collections) that comprise the Pāli Canon, analyze human experiences, along with the beings ordinarily thought to have those experiences, into *skandhas* (Sanskrit; Pāli: *khandhas*) or 'aggregates', of which there are five kinds: four that are mental in nature, usually translated from Pāli as 'feeling', 'perception', 'volitional formations', and 'consciousness', and one physical, 'form' (as derived from material elements). (Bodhi 2005, p. 307) Coming to understand that no *skandha*, nor anything else, is one's *self*, is essential to realizing the soteriological aim of Buddhism—the attainment of enlightenment and cessation of the cycle of suffering and rebirth. Given the apparent conceptual connection between having a self that stands behind one's thoughts and actions and the exercise of free agency typically thought to be necessary for moral responsibility, philosophers are hard pressed to reconcile such responsibility with the no-self doctrine.

Mark Siderits is one philosopher who has made a concerted effort at such reconciliation, with some success. In *Personal Identity and Buddhist Philosophy* (2003, ch. 3), Siderits develops and defends a view about the nature of persons that he calls ‘Buddhist Reductionism’. According to this position, persons are not ultimately real, because what is conventionally referred to as a ‘person’ is only a series of distinct, momentary collections of psychophysical constituents or *skandhas*. Nevertheless, persons have conventional existence because grouping some such momentary psychophysical elements into collections and those collections together into temporally extended series has a certain utility. This utility is grounded synchronically in the spatiotemporal contiguity between the (physical) elements and diachronically in the causal connections that obtain between the momentary collections of elements.

Buddhist Reductionists are, according to Siderits, committed to an ‘Impersonal Description’ (“ID”) thesis. This is the claim that “we can give a complete description of reality without either asserting or presupposing that persons exist” (2003, p. 35). Siderits defends Buddhist Reductionism against the charge that the ID thesis implies the intuitively unpalatable ‘extreme claim’, which is:

“...that four central features of our present person-regarding practices cannot be rationally justified: interest in one’s own survival, egoistic concern for one’s future states, holding persons responsible for their past deeds, and compensation for one’s past burdens” (p. 37).

Endorsing the extreme claim would make one an eliminativist, not a reductionist, about persons.

The eliminativist agrees with the reductionist that there is no ‘self’ separate from a body and brain and no similarly independent further fact about persons beyond facts about the person’s psychophysical constituents and their relations. However, unlike the reductionist, the eliminativist insists that the denial of such a self or further fact entails that many of our common-sense attitudes about persons, those targeted by the extreme claim, are irrational. Another way to say that reductionism implies the extreme claim is to say that a reductionism that is not also eliminativist is internally inconsistent—that reductionism necessarily slides into eliminativism, because one cannot consistently deny the existence of a separate self while adhering to the rationality of our normal attitudes concerning persons.

My aim here is to evaluate the objection that the ID thesis implies the extreme claim regarding the third of the four features listed above, concerning responsibility. I consider Siderits’ account of Galen Strawson’s (1986) version of the objection and Siderits’ proposed answer to it on the part of the Buddhist Reductionist, which appeals to the concept of ‘shifting coalitions’ of self-revision in an individual. I argue that the shifting coalitions idea successfully disarms part of Strawson’s version of the objection, and so can account for a modest kind of responsibility (though not quite the robust sort Siderits has in mind) that is compatible with both Buddhist Reductionism and causal determinism (and so does not require a separate self or libertarian—indeterminism-requiring—free will).

Nevertheless, while Siderits succeeds in reconciling a compatibilist notion of responsibility with the metaphysics of Buddhist Reductionism, it (and any account of responsibility whatsoever) will still be in tension with the soteriological aspect of the Buddha’s teachings because it requires

the appropriation and identification of one's mental states as one's own, which is for the Buddha a source of suffering and an impediment to the only sort of freedom he explicitly recognizes: liberation from attachment.

2 Reductionism and Responsibility

The distinction between reductionism, non-reductionism (or inflationism) and eliminativism is illustrated in the Buddhist literature through the famous example of the chariot. As Siderits explains it:

“... ‘chariot’ is a convenient designator for a set of parts assembled in a certain way. Thus while there are ultimately no chariots, there are those wholly ‘im-chariotal’ facts into which all chariot-talk may be reductively analyzed; it is these facts that explain the utility of our talk of the fiction... Given this utility we may say that while the chariot is ultimately unreal, it is conventionally real. This will be the reductionist view of chariots. The non-reductionist will claim that chariots are both conventionally and ultimately real—that in addition to the parts of which chariots are composed, ultimate reality also contains some sort of separately existing chariot-essence. And the eliminativist will claim that chariots are both ultimately and conventionally unreal—that our talk of chariots is misleading and should be replaced by some entirely new way of conceptualizing collections of chariot parts.” (2003, p. 7)

Like the chariot, persons are seen by the Buddhist Reductionist as conceptual constructions that are only conventionally, not ultimately, real. This is the Buddha's self-described 'middle path', developed by the Abhidharma schools, between the eternalism of the non-reductionist Nyāya and Sāṃkhya (orthodox Indian) schools that believed in a transcendent self called, respectively, '*ātman*' and '*puruṣa*', and the annihilationism of the eliminativists, for whom the denial of an eternally and separately existing self entails that the duration of any individual person's life is instantaneous.

Furthermore, reductionism about persons is one interpretation of the seemingly paradoxical claims made by the monk Nagasena in his famous dialogue with King Milinda. The king asks, given that there is no separate self, whether or not infant Milinda is the same person as adult Milinda. Nagasena replies saying that "the adult is neither the same person as the infant, nor is he a distinct person" (Siderits 2003, p. 35). According to the Abhidharma commentators, what he means is that the adult and infant ultimately are distinct entities (so long as it makes sense to think of them ultimately as entities at all), but are conventionally one and the same.

With respect to responsibility, the argument that reductionism implies the extreme claim goes as follows. Being responsible requires that we are capable of 'purposeful action', which in turn necessarily involves what Siderits calls a "thick sense of agency, one that involves seeing ourselves as standing over and above our desires and other psychological states" (2003, p. 62). So if one accepts the reductionist view, then one could not rationally retain that 'thick sense of agency' and therefore could not act purposively in the way required for being morally responsible. Siderits states the argument, which he attributes to Strawson (1986), as follows:

“[I]n order to see ourselves as purposeful agents, we must be prepared to acknowledge responsible ownership of our actions; this in turn requires that we see them as stemming from a character that we have actively shaped. But this we cannot do unless we view ourselves as possessing a self that transcends any of our particular values, projects, and life-plans, a self that is so positioned as to be able to endorse any or all of our pro-attitudes. Hence, the argument concludes, our sense of purposeful agency requires belief in just the sort of conception of ourselves that reductionism denies.” (Siderits 2003, p. 62)

According to Siderits, Strawson’s argument depends on the claim that in order to be free in a way where we are truly responsible for our actions we must think of ourselves as free in that way, which requires that we acknowledge full authorship of the intentional causes of the actions in question. To do so I must be “truly responsible for those of my beliefs, desires, dispositions, policies, decision-making procedures and the like that resulted in my choosing to perform the action” (Siderits 2003, p. 63). However, if determinism is true, then these beliefs, desires, etc. that make up my present nature were “determined by factors (of either heredity or environment) that were set before I was born and over which I therefore have no control” (p. 63), and so the fact that my actions flow from my present nature cannot by itself account for my being responsible for them.

What I need, in order to think of myself as free/responsible, is to think of myself as something that transcends any of the particular mental states that are the antecedent causes of my actions. This self must actually transcend the contents of any individual mental state, particularly “any

value or pro-attitude that happens to be contained within” it (p. 63), for it must be the sort of thing that can potentially be aware of all my mental states, past or present, that causally contribute to my actions.

However, for the reductionist, no individual mental state or series of discrete states can play this role because of their transitoriness. A present momentary mental state cannot have been the subject of previous states, particularly those that contributed to causing it. Furthermore, responsible agency requires normatively assessing the mental states that lead to one’s actions, but this assessment can only come from the self, not from any particular, contingent mental states with their deterministic causal history.

If one wholeheartedly accepts reductionism, Strawson argues, and therefore ceases to think of oneself as possessing the kind of transcendent self explicated above, then one’s sense of oneself “as a truly self-determining planner and performer of action” (Siderits 2003, p. 64) is undermined. This conclusion reflects the worry that accepting reductionism leads to Micawberism, a conception of oneself as a mere passive observer in one’s life-narrative which leads to one waiting to see what happens rather than taking an active role in shaping one’s character and actions.

Before offering his own response to Strawson’s argument, Siderits first considers and rejects one sort of possible move the reductionist could make. As Siderits puts it:

“The Reductionist might be tempted to retreat to a weaker conception of freedom [and

responsibility], according to which it is enough that the action ‘come from within’ the agent, regardless of how the agent came to have the particular beliefs, desires, dispositions etc. from which the action flowed. But this temptation should be resisted, since we do expect agents to take responsibility not just for their actions but also for their own character... Being responsible for my actions means being responsible for being the sort of person who would perform those actions. Any account of freedom that omits this is justly criticized as too weak.” (2003, pp. 64-65)

An acceptable account of responsibility cannot be so broad that it makes one responsible for actions that are not self-consciously chosen, but only for the sort of actions that can be performed by a being that can reflect on its own character and evaluate its own desires. Assuming that only persons can be responsible for their actions, such an account should distinguish, in a principled way, on the one hand, the kinds of actions a mere animal is capable of, and, on the other, the kind only persons can perform (Frankfurt 1971). So the challenge, as Siderits sees it, is to explain how a reductionist might rationally believe herself to be responsible not only for her actions but for the character that results in those actions: How could a person be responsible for their character, understood as the total aggregation of their effective mental life at a time, if they just are that aggregation (however conceived as their character, etc.)? How is the self-determination necessary for responsibility possible without a transcendent self that does the determination? Why isn’t “an agent-self that is transcendent yet contentful” necessary for the “responsibility-entailing freedom that requires self-determination”? (2003, p. 65)

3 The Shifting Coalitions Model of Self-revision

To answer the challenge posed by these questions, Siderits invokes the shifting coalitions conception of self-revision that he develops earlier in the book. There, Siderits is concerned to show how persons could be capable of self-scrutiny, self-control and self-revision without violating what he calls the ‘anti-reflexivity’ principle that ‘an entity cannot operate on itself’ (2003, p. 27). The aforementioned capacities seem to require a separate self as subject and chief executive, with its particular mental states comprising its object, because if the mental states that play the role of the subject that scrutinizes and revises are ever themselves objects of revision, which any state seems potentially capable of being, then some mental states would have to serve as both subject and object, but that would violate the anti-reflexivity principle:

“For if each of them is a potential object of the executive function, and an entity cannot operate on itself, then it seems that none of them could be the one enduring subject that performs this function” (p. 26).

However, Siderits rightly points out that the principle is only violated if the same mental states are subject and object *simultaneously*. A reductionist can offer an account of self-revision that appeals to shifting coalitions of mental states playing the role of chief executive at different times, such that each coalition could be the object of revision at the times when it is not the subject. According to the reductionist, the temptation that leads to positing a transcendent self—a temptation to be resisted—is to take a particular set of mental states that play the subject role relatively frequently and hypostatize them into an enduring subject: “Thus arises the notion that a

person has an essence—that some constituents are more central to the existence of the person than others” (p. 27).

Holding this view seemingly requires one to deny that the set of mental states taken to be the subject is itself subject to revision. But if one has no way of revising the chief executive, then one cannot be responsible for the way that executive scrutinizes, controls, and revises one’s other mental states and behavior. “For instance, when I decide to curb my bedtime snacking I may be employing a particular standard of acceptable body shape, which I may subsequently decide is politically problematic and morally questionable” (p. 26). This isn’t such a problem for the non-reductionist, who can claim that the self, being independent from the psychophysical elements that are subject to deterministic laws of physics, is the sole source of independent valuation in a person (and may be propped up by a conception, usually religious, of an infallible conscience)—a view which is endorsed by libertarians about free will.

However, if one denies, for good reason—namely, that we have no evidence for such a thing and the concept of it may be internally incoherent in various respects—that such a self exists, and says that the executive function is played by some subset of psychophysical states which are, like any others, causally determined, then those states must also sometimes be subject to revision by other elements of the person/brain. Siderits offers the shifting coalitions view as a solution to this problem.

“If I am to be capable of revising my own character, then I require a stock of beliefs and desires on the basis of which I may critically evaluate and seek to reform various of my

dispositions and tendencies I am called upon to monitor. It may now seem as if, were they to constitute a part of the 'I' that performs self-revision, then the anti-reflexivity principle would be violated. But what this picture omits is the possibility that a given stock of beliefs and desires might serve as a basis for a particular bout of self-criticism, yet some among these stand under subsequent scrutiny on the basis of a distinct (though perhaps overlapping) stock of beliefs and desires... On one occasion my anal-compulsive disposition might lead to extirpation of the desire to smoke. Yet, subsequently a wish to be more accommodating to others might lead to an effort to curb my anality. At one time the anal disposition belongs to the coalition making up the 'executive', later it falls out of this shifting coalition." (p. 65)

The shifting coalitions approach posits a kind of feedback loop between psychophysical states that allows one to have a sense of self-determination that depends on nothing that is undermined by reductionism. Each coalition that at one time plays the role of executive can be at another time the object of a different coalition's scrutiny as well as control and revision. Even if the activity of each coalition is causally determined, the fact that there are internal checks and balances, and that one cannot be aware of all the facets of one's psychology at once, yields a rationally tenable sense of self-determination. If a person is a system of such shifting coalitions, and each one of those is shaped by the activity of others that were themselves once objects of scrutiny by some other coalition of elements in the system (though any set may have dropped out of the system at any point), so that each coalition has been available for assessment and revision, then the system as it is at any point may act in a way for which it can reasonably be held responsible in the future.

However, the above account would not satisfy Strawson or others who think that reductionism undermines self-determination. Recall that Strawson thinks that to see oneself as self-determining, one must think of oneself as ultimately responsible for one's character as well as one's actions. Even if one's character is formed internally by the feedback mechanism of shifting coalitions, Strawson would argue that the way in which this system functions is determined by factors before one's birth. For instance, consider Siderits' example of the kind of self-revision an agent must see him or herself as responsible for:

“So my miserable childhood resulted in a predisposition to behavior that causes trouble for myself and others? Others tell me to stop kvetching. I agree, and set about trying to reform and improve my character.” (2003, p. 64)

Strawson would object that in such a case whether I am or am not the ‘sort of person’ who would respond that way to the criticism of others, or who can find the strength within myself to push back against the forces of my upbringing, is not really up to me. In other words, it is still a matter of deterministic luck whether or not I have the right coalitions with the necessary strength to bring about a particular act of self-revision.

Siderits claims that the idea that Micawberism follows from reductionism is due to the assumption that without believing in a transcendent self, one can only see oneself as a ‘mere conduit of causation’, which entails the denial “that free, responsible agents could believe themselves to be constituted by shifting coalitions of volitions” (2003, p. 65). For Siderits, the

type of argument Strawson is pushing underestimates our ability to “carry out the sort of self-monitoring and self-revision required for full responsibility” and our capacity for ‘ironic engagement’ (Siderits 2003, p. 66). ‘Ironic engagement’ is characterized as the ability to “induce and maintain a belief in a useful fiction while knowing it for what it is” (p. 109).

So, even while recognizing that we are merely bundles of causally determined psychophysical elements, we can adopt and genuinely believe in the useful fiction of self-determination. What makes the fiction useful is that it accurately distinguishes two types of phenomena, *i.e.*, cases where one’s actions are compulsive or automatic, and ones where one’s actions are caused by inner states that have been subjected to self-revision under the shifting coalitions model. This genuine phenomenal distinction justifies our belief in responsibility.

Nonetheless, that is not the kind of full, ultimate responsibility that Strawson is interested in. Nothing short of a genuinely transcendent agent acting outside of the deterministic causal matrix could fit that bill. However, it is also not the weak sort of responsibility that Siderits rightly rejects, where one’s actions simply ‘come from within’ the agent. The sort of responsibility made possible by the shifting coalitions strategy and ironic engagement is distinguished by the recognition that actions ‘come from within’ in different ways. The ones that result from a process of dynamic self-revision are the ones for which we are responsible.

Siderits’ account of self-revision therefore helps to ground a kind of determinism-compatible conception of responsibility that is consistent with reductionism, though it may not be quite the robust sort he is aiming for. Still, he has succeeded in demonstrating how resources from the

Buddhist tradition can be employed to further explicate positions taken in the Western discussion of personal identity, free will and responsibility by explaining how a reductionist can give a coherent account of the best, roughly Frankfurtian, compatibilist conception of self-determination. I mean ‘compatibilist’ not in the sense that acceptance of determinism is compatible with all of our pre-reflective attitudes about praise, blame, revenge, and punishment, but only that accepting determinism need not undermine our sense of ourselves as purposive agents who, at least sometimes, act because of reasons that we are able to reflect on and modify, and therefore perform actions for which we can be reasonably held responsible.

4 Identification and Attachment

Siderits succeeds in reconciling the metaphysical aspect of the ‘no-self’ doctrine with a compatibilist conception of responsibility. However, there remains a further conflict between a different aspect of the Buddha’s teachings and any notion of responsibility whatsoever. This conflict concerns not the metaphysics of *anātman*, but its soteriological implications. In putting forth the no-self view, the Buddha is not primarily concerned with the mereological relations between the *skandhas* and the experiences of persons that they constitute, but rather with the attitudes that persons take toward their experiences and psychophysical characteristics. Any conception of responsibility requires, if not the actual identity of the individual with her thoughts and actions, then at least the appropriating or identifying with those actions and thoughts as her own. For the Buddha, such appropriation and identification constitutes the clinging to what is ephemeral that is the essence of suffering and the primary obstacle to enlightenment.

To understand this point it is necessary to take a closer look at the *skandhas* and the role they play in the Buddha's soteriological teaching. In the *Samyutta Nikāya* ("SN"), he states the First Noble Truth, the reality of suffering, by reference to the five aggregates:

"Now this, monks, is the noble truth of suffering: birth is suffering, aging is suffering, illness is suffering, death is suffering; union with what is displeasing is suffering; separation from what is pleasing is suffering; not to get what one wants is suffering; in brief, the five aggregates subject to clinging are suffering" (*Dhammacakkappavattana Sutta* V 420-24, SN 56:11; Bodhi 2005, p. 78).

Clinging to the five aggregates is suffering itself, for the Buddha. According to Bodhi's analysis, clinging occurs in two different modes, *appropriation* and *identification*. One clings to the aggregates first by taking possession of them, 'this is mine', and then by using them as a basis for self-definition, 'this is what I am, this is myself'. One who is not on the path toward enlightenment, according to the Buddha,

"...regards form as self, or self as possessing form, or form as in self, or self as in form. He regards feeling as self... perception as self... volitional formations as self... consciousness as self, or self as possessing consciousness, or consciousness as in self, or self as in consciousness. That is how identity view comes to be." (SN 22:82, 100-103; *Majjhima Nikāya* 109, III 15-19; Bodhi 2005, p. 340)

This passage is a diagnosis of how the mistaken conception of an enduring self comes to be. Individual thoughts, sensations and other psychophysical states are appropriated by the person who experiences them and are then taken to be part of the identity of that person. The supposed person doing the appropriating and identifying, at least according to the shifting coalitions model, is just another set of states. So we have one coalition clinging to others as its possessions and then taking those others to be part of its own essence. For Siderits, so long as one realizes that this is all that's going on, that no state really is essential to a person's identity, that at every moment of self-scrutiny it may be a different state doing the scrutinizing, and that any state may be (at a time that it is not scrutinizing) the object of scrutiny, then one can still rationally assert the conventional existence of persons and hold them, oneself included, responsible for their actions.

However, such a view does not do justice to the soteriological import of the above diagnosis. The problem is not just that we mistakenly take the subject of our self-scrutiny to be independent of our fleeting psychophysical states when it is at each moment just another set of those states, nor that we also take some fleeting psychophysical states to be essential to our identities. The more fundamental problem is that we think of any of those states as *belonging* to ourselves, either essentially or merely contingently, for any duration whatsoever. The shifting coalitions model requires that we think of our *skandhas* as in some way part of us. Each set of elements that make up the *skandhas* may be sometimes in the subject role and sometimes in the object, but they and their activities are taken to have some bearing on who we are and what we do, such that we can be responsible for the effects of those activities. Such appropriation of the *skandhas* is just the sort of clinging that is, for the Buddha, the basis of suffering and the barrier to enlightenment.

Siderits has made a valiant effort at making Buddhism safe for our conception of persons as responsible agents. However, no conception of responsibility can be reconciled with the soteriological aims of Buddhism because all such conceptions require that thoughts and actions be appropriated by and identified with a self, even if that self is regarded as temporary and fleeting. To hold myself responsible for some actions and the intentions behind those actions I must regard them as *my* actions, and to do so, rationally or not, is already to become attached and to suffer in that attachment. To hold someone else responsible for some actions is to believe that they should suffer in the same way and is, therefore, to lack compassion for them.

The upshot is that even if we can consistently think of ourselves—and others—as responsible agents, while maintaining reductionist commitments, the view of the Buddha is that we should not. Both the soteriological aim of non-attachment and the ethical ideal of universal compassion are served by abandoning all conceptions of personal identity, responsibility, and the retributive model of reward and punishment that is often founded upon them.

References

Bodhi, B., ed., 2005. *In the Buddha's words: an anthology of discourses from the Pāli canon*.

Boston: Wisdom Publications.

Frankfurt, H., 1971. Freedom of the will and the concept of a person. *Journal of Philosophy*,

68(1), pp. 5-20.

Siderits, M., 2003. *Personal identity and Buddhist philosophy: empty persons*. Hampshire:

Ashgate.

Strawson, G., 1986. *Freedom and belief*. Oxford: Clarendon Press.

14 Psychological versus Metaphysical Agents: A Theravāda Buddhist View of Free Will and Moral Responsibility

Peter Harvey

A Permanent Self Could Not Be a Free Agent of Actions

A strand of the current discussion of ‘free will’ and Buddhism concerns the fact that most forms of Buddhism do not accept a self in the sense of a permanent essence of a person. This is then generally seen as equivalent to the non-acceptance of a self as the agent of actions, an agent-self, which could be either free or unfree. The connection seems a natural one to make, but it is worth pausing to reflect on its appropriateness.

In teaching that ‘everything (all *dhammas*) [is/are] non-self (*anattā*)’¹ (AN.I.286),² the Buddha meant that they were ‘empty of self and what belongs to self’ (SN.IV.54). Here ‘self’ (Pāli: *atta*; Sanskrit: *ātman*) is used in a specific kind of sense, to refer to a permanent and happy essence of a person. This can be seen in the repeated teaching that it is “not fit to consider that which is impermanent, painful, of a nature to change, as: ‘This is mine (*etam mama*), this I am (*eso ham asmi*), this is my self (*eso me attā*)’” (e.g., SN.III.66). Hence, the kind of ‘self’ in the term ‘non-self’ is best rendered with an initial capital—‘Self’—to signal that it is a specific kind of self. Anything conditioned is seen as impermanent, painful and non-Self, but *nirvāṇa*, the unconditioned, is not seen as impermanent and painful, but still as non-Self, because it is beyond any grounds for the arising of the thought ‘I am’ (Harvey 1995a, pp. 23, 51-3).

Now if there *were* a Self in the above sense, could it be a free agent of action? To answer this, one needs to ask if it could be *any* kind of agent of action. I would argue that the answer is ‘no’. Why? Because the occurrence of a decision or any other action-initiating event in such a putative Self would be a change within it. But if anything has change within it, there is surely no way that further changes can reliably be prevented from happening within it, which would in time bring about its end; so, it could not actually be permanent, in a strong sense. So, a truly permanent Self could not be an agent of action. This seems to be recognized in the Sāṃkhya school of Hinduism, in which the *puruṣa* or inner ‘person’ is not the agent of actions, but is simply the observer of an aspect of conditioned and impermanent (material) ‘nature’ (*prakṛti*) that is the agent of actions (Flood 1996, pp. 234-35).

Admittedly, in the pre-Buddhist *Upaniṣads*, the inner immortal *ātman* or Self is seen as the autonomous ‘inner controller’ (*antaryamin*) of a person’s actions, inner elements and faculties (*Bṛhadāraṇyaka Upaniṣad*, 3.7.3). It would also need to be in full control of itself. But the Buddha argues in the *Anatta-lakkhaṇa Sutta* that no aspect of a person is capable of being perfectly controlled and shaped at will:

“Material form, monks, is non-Self. Now were this material form Self, it would not lead to affliction, and one would be able to effectively say, ‘Let my material form be like this, or not like this.’ But inasmuch as material form is non-Self, therefore it leads to affliction, and one cannot effectively say, ‘Let my material form be like this, or not like this.’” (SN.III.66-8)

(The same formulaic line of reasoning is then applied to feeling, perception, volitional activities, and consciousness.) This must allude to such facts as that the body gets tired, ill and old; we do not feel pleasure all the time, as we might wish; we can act erratically and against our better judgement; and our awareness often wanders, being pulled this way and that by external events or inner emotions. All aspects of body and mind are conditioned in a variety of ways, and cannot be dominated by any controlling factor. Hence, the idea that there is a Self that acts as the ‘inner controller’ of them is negated.

Indeed, in response to various quarrelling brahmins and renunciants who hold that ‘pleasure and pain, Self and the world’ are self-made (*sayam-kato*), made by another (*param-kato*)—presumably meaning made by a Self-essence, or made by a separate God, respectively—both, or neither, the Buddha says that these teachers are bound by problematic and conceited ideas of “‘I-as-maker/doer’ (*aham-kāra-pasutā*) ... ‘another-as-maker/doer’ (*param-kārūpasañhitā*)” (Ud.69-70). That is, the concepts of ultimate makers of the world or ultimate agents of action are unsupportable. The Theravāda view is thus that there is karma (action), but no “doer of karma (*kamma-kārako*)” (*Katthāvatthu*, pp. 53-4). That is, there is *cetanā*/volition, which AN.III.415 defines ‘action’ in terms of, but no specific process that is the *agent* of action, much less a permanent essence that is the agent.

Yet Agency Is Accepted

At MN.III.179-80, it is said that Yama, king of the dead, reprimands an evildoer arriving in hell, saying that a certain deed was done by *him*, and not by any friend or relative, so that *he* must experience its karmic result. This passage need not imply that such a past action was

done by a substantial, still-existent Self, but only that it was done by an earlier portion of the stream-of-states that the person now is, rather than by any other stream-of-states. ‘He’, no one else, is responsible. This shows that agency of, and responsibility for, actions is accepted in Theravāda Buddhism.

Hence, the Buddha criticized the fatalist view of Makkhali Gosāla (AN.I.173-75), who held that there was no cause or condition for the defilement or purification of beings, *i.e.*, no cause or condition *that they could effect themselves*. This is because Gosāla held:

“there is no self-doing or other-doing (*atta-kāre ... para-kāre*), there is no personal doing (*purisa-kāre*), no power, no energy, no personal strength, or personal exertion. All beings ... are without control (*avasā*), without power, without energy, they experience the pleasure and pain of the six kinds of rebirth, changed according to the fixed course (*niyata-*) of circumstances and individual natures (*-saṅgati-bhāva-*) ... [over a fixed series of rebirths] ... Just as a ball of string, when thrown, runs till it is all unravelled, so fools and wise run and wander on till they make an end of suffering.” (DN.I.53-4)

That is, the Buddha rejected a view that saw humans as being passively shaped by an external fixed destiny, with no agency of their own with which to influence their own future.

In a very common sense way, the Buddha also pointed out that, within certain limits, we can control our actions. When a brahmin came to him and said that he had the view “there is no self-doing (*atta-kāro*), there is no other-doing (*para-kāro*)”, the Buddha asks him how one who “himself (*sayam*) comes and himself returns” can say this? He then gets the brahmin to

see that self-doing and other-doing can be seen inasmuch as there is the ‘element of instigation’ (*ārabha-dhātu*) and beings who instigate (actions), likewise with the elements of persistence, exertion, strength, continuation and application (AN.III.337-38).

The Nature of Agency

As a permanent, unconditioned agent of actions is impossible, an agent of actions must be something that is subject to change and have limited influence. It would be enmeshed in the network of conditions that is a person, and like all conditioned things, be non-Self. If it were a permanent Self, it could not *do* anything.

One might like to think that the agent of one’s actions is an essential, permanent Self/I-agent, but this in fact makes no sense if taken literally. Yes, actions are done by the kind-of-person-one-has-been-so far, but this is ‘permanent’ only in an approximate sense, as a cluster of mental and physical process-events with a reasonably consistent, but still changeable, pattern to it. This is the only kind of ‘thing’ that can be an agent of action, and Buddhism does not, and has no reason to, deny its reality. That said, ‘it’ is a misleading term here, for what is being talked about is not a single thing, but an interacting cluster of processes. Yes, a crucial process is *cetanā*—will, intentional volitional impulse—but this is just the most immediate condition for the arising of an action of body, speech or mind. And of course it arises according to conditions, most immediately, sensory contact (SN.III.60, IV.68). Hence, AN.V.113-116 holds that the most immediate condition for misconduct is non-guarding of the sense-faculties (regarding what one focuses on, and how).

So, while the Theravāda *suttas*³ do not accept a permanent Self that is the agent of action, that stands above all conditions, unaffected by them, a will that is unconditioned, totally free, or ‘strong free will’, it does accept that a person, as a cluster of conditioned processes including will, can have agency and responsibility.

The Issue of Freedom and Unfreedom

To what extent and in what way can ‘will’ and its associated mental processes be ‘free’? The crucial thing seems to be a capacity for self-direction. When there are no external threats or inner insanity, can the stream of related processes that is a particular ‘person’ be such as to be able to reflect on possible alternative actions and decide between them, and to make plans and be able to carry them out, at least some of the time? The answer, both in common sense and Buddhism is ‘yes’; hence right ‘resolve’ (*sankappa*) is a factor of the Noble Eight-factored Path.

Buddhism would add, though, that: we are not as in charge of ourselves as we would like to think, as we get captured by, *e.g.*, greed or aversion; we often act in inconsistent ways; we are often deluded and mistaken about what is truly in our best interests; and we delude ourselves in after-the-act self-justifications. That is, our ability for self-direction is often used in a sloppy, ineffective and problematic way. We have a degree of freedom, but are not as free to act in our best interests, or the best interests of others, as we might like to think. But we can overcome these limitations, as seen by the Buddha’s saying that he has the qualities of one with great wisdom, including that:

“He thinks whatever thought (*vittakkam*) he wants (*ākaṅkhati*) to think and does not think a thought he does not want to think; he resolves in whatever way he wants to resolve, and does not have a resolve (*saṅkappam*) he does not want; thus he has attained mental mastery (*cetovasippatto*) over the ways of thought” (AN.II.36-7).

The *Vitakka-saṅhāna Sutta* gives five methods for attaining such mastery of thought (MN.II18-22).

The Buddha clearly taught that we could learn to be freer. But to do that, we need to recognize the ways in which we are currently conditioned in unhelpful ways, and be aware of alternatives and how to access and establish them. That the Buddha criticized all views that saw one’s actions as beyond one’s control—so that bad actions such as killing might be determined by one’s past karma, the creative activity of ‘God’, or no cause or condition other than the fixed course of destiny (AN.I.173-75)—shows that he felt that we could break free of past patterns of conditioning. But doing so was not a matter of going beyond conditions (except at the level of final *nirvāṇa*), but of cultivating a more wholesome, skillful, wise set of conditions. It was a matter of assiduous *training*, involving, among others, four points:

(a) Resolving to keep certain ethical precepts would help cultivate ethical restraint and weaken bad habits of behavior.

(b) ‘Guarding the sense-faculties’ would help reduce the extent to which contact with certain sense-objects and thought-objects automatically trigger habitual reactions that are often unhelpful.

(c) Developing inner calm would help build more inner resilience and mental clarity, bringing about a clearer and freer inner space. Hence it is said that when there is suspension of the ‘five hindrances’ to wisdom—desire for sense-pleasures, aversion, dullness and lethargy, restlessness and worry, and vacillation—one is, respectively, like a person free from debt, illness, prison, slavery, and the dangers of a journey (DN.I.71-3).

(d) Cultivating insight would help one see that the objects of one’s greed and aversion are not truly worthy of these.

A person advanced on the path, strong in these qualities, has an empirical self, *i.e.*, a cluster of mental processes, that is more centered, calm and coherent than most people’s—they are said to have a ‘great self’ or ‘developed self’ (*mahatta, bhāvitatta*, It.28-9, 78-9; Harvey 1995a, pp. 55-8), one that has more effective agency than ordinary people.

In this, mindfulness is a crucial quality, both in enabling a clearer and unbiased view of one’s current and past states, learning from this, and connecting to past wholesome states and thus helping them re-arise in the present. To the extent that we can choose what we bring to mind in memory, we can choose which past experiences and events influence our current actions. Mindfulness helps one be more wise and skillful in how one accesses past memories and hence brings their influence to play on the present. Similar considerations apply in thoughts about the future, in terms of how mindful and skillful one is in bringing to mind future possibilities, and what one resolves to put energies into in terms of planning and promises.

Is a Conditioned Agency Still Agency?

In what one focuses the mind on, for how long, and in what manner—all of which are key conditions for what one does next—there can be an element of choice. This choice, though, will be conditioned by all manner of things: past behavior, how one is feeling at the time, one’s energy level, what one’s ideals are, where one is and with whom, etc. The crucial question here is: do such conditions entail the hard determinist view that determinism negates any idea of genuine choice or agency, and hence responsibility?

None of the factors conditioning a choice will be a single determining cause. As explained by Buddhaghosa in his explanation of the sequence of conditions in the Buddhist causal doctrine, ‘dependent origination’:

“Here there is no single or multiple fruit of any kind from a single cause (*kāraṇato*); nor a single fruit from multiple causes, but only a multiple fruit from multiple causes... But one representative cause and fruit are given in this way, ‘with spiritual ignorance as condition are the volitional activities...’” (Vism. 542)

A condition for something is only ever one among many, such that what we identify as ‘the cause’ is generally only the last condition to fall into place, or a condition of particular interest. A choice will have various factors conditioning it, along with how one currently construes and relates to such factors. And at different times, the items within the same collections of conditions may individually have different strengths.

One can, of course, have done things in the past that would strengthen or weaken certain conditions, with a view to beneficially influencing one's future choices. But that just puts the focus of the question of ability-to-choose further back in time: At age 20, I did this, rather than that, and met a certain range of people, and chose to cultivate particular friends and contacts among these, such that I put energy into X rather than Y, which helped lead me to being the kind of person I generally am now, who is more likely to choose A rather than not-A, and so on.

Hence, when one winds things back, one does not just come to past events and circumstances, but also to more choices. Current choices are always made by the kind of person that one's own past choices as well as circumstances have shaped one into being. And the more that capacity for choice is actively utilized—rather than one passively accepting circumstances, and how one is—the more it grows stronger. Thus, agency and responsibility can *increase*, when the right conditions are nurtured. It is not undermined by the mere fact of conditionality, if self-determination exists. There is hence a form of compatibilist freedom of action, *i.e.*, a freedom of action that is compatible with all actions being conditioned. The relevant freedom may be described as of a variable degree.

Any choice/volition would be the product of the totality of its previous and simultaneous conditions; these would be its determining condition set, which necessitates its arising/existence. In effect, the Abhidhamma⁴ seeks to spell out as many possible conditions as it can for a variety of mental processes, although the Theravāda Abhidhamma at least recognizes that its list of ingredient processes of a particular momentary state of mind may be incomplete (*Dhammasaṅgani*, para. 1).

Hard Determinism and Predictability in Principle?

Given the totality of conditions in the universe at a particular time, are all future events, including a being's choices, stretching over many lives, already determined, as hard determinism holds, so as to be in principle predictable, if all the relevant facts about the present could be known?

In Buddhism, there is no higher kind of being than a Buddha; but could even a Buddha have the knowledge that could enable such a prediction? As Charles Goodman kindly commented on my (Harvey 2007) article:

“great complexity does not conflict with predictability in principle, but only with predictability in practice.... In fact, if we accept that the Buddha knew or even could have known in 500 BCE how you were going to make all your life decisions, we have already and automatically ruled out the view that the future is genuinely open (a view you assert on p. 69).” (Goodman 2009a)

On this topic, though, he elsewhere accepts that the *Tevijja-vacchagotta Sutta* (MN.I.482-83) “does not seem to ascribe to the Buddha any special knowledge about the future” (Goodman 2009b, p. 151). In this *sutta*, the Buddha denies he has continuous omniscience, but only, when he applies his mind to it, unlimited memory of his past lives and ‘divine eye’ knowledge of how other beings are reborn according to their karma, and his own liberating insight.

Goodman continues by saying “But in the *Cūḷasakuludāyi Sutta* [MN.II.31-2], the Buddha claims that he, or anyone else with the divine eye, can answer any question about the future” (Goodman 2009b, p. 151). Yet the ‘divine eye’ is not seen as the ability to know all the future rebirths of a particular being, but simply to know what their next rebirth will be, based on knowledge of the current karma. At DN.III.134, when the issue of whether the Buddha’s great knowledge extends to the future is raised, he claims that it does; but the example of such knowledge that is given is that he knows that he will have no further rebirths.

Admittedly, the post-canonical *Milindapañha* says:

“the Blessed One was omniscient (*sabbaññū*), but knowledge-and-vision was not constantly and continuously (*satataṃ samitaṃ*) present to the Blessed One. His omniscient knowledge was dependent on the adverting (of his mind); when he adverted to it he knew whatever it pleased (him to know).” (Miln.102)

Yet in the early texts, while the Buddha talks about various *past* lives of people, he only ever says how someone will *next* be reborn, or has recently been reborn, or says that someone will have a limited number of future lives due to their spiritual realization: for example, a ‘stream-enterer’ (one having attained a level beyond taking anything as Self or related to a Self) will have seven future lives at most (AN.V.120).

To be able to predict everyone’s future in detail, the Buddha would need to be able to know every conceivable detail of the world at the present moment, so as to be able to work out how this hyper-complex mix of conditions would evolve in the future. But the Buddha’s omniscience is not seen as of the kind that can know everything simultaneously, in one moment (*sakideva*)—he denies the possibility of such omniscience at MN.II.127. It would

thus take *some* time to scan and know the totality of the universe; hence the total state of the universe *at one moment* could not be known even by him. And without this knowledge of a complete past condition set, precise predictability in practice would be impossible even for a Buddha. This must be particularly true of events in the minds of living beings. These are far more complex systems than the weather, for example, but chaos theory explains that complex dynamic systems such as this, though predictable in principle, have limited predictability in practice; for precise starting conditions cannot be exactly known, and small differences can have large effects over time.

What of predictability in principle? Relevant here is the fact that the Buddha did not accept the view that saw the world as finite, nor its opposite, that it was infinite (Harvey 1995a, pp. 83-7). He left this as an undetermined issue, holding both that worrying over it was a time-wasting side-track to the path to liberation (MN.I.429), and also that those with at least the insight of a stream-enterer would not go for either view (SN.IV.395). But only if the world were finite would it contain a finite set of conditions that could be known, on which a precise prediction of the future of the world could be made.

This may be one of the reasons that the Buddha also left as undetermined whether the world is eternal or not. As he saw the world is kept going by the karma of the beings in it (DN.III.84-92), and only unenlightened beings make karma, if one cannot predict if all beings will attain enlightenment, one cannot say if the world will some day end or not. Indeed, when he was asked if all beings will attain enlightenment, he was silent (AN.V.194), leaving it an undetermined question.

So, even the Buddha cannot precisely know everyone's future, and perhaps this is impossible in principle. Nevertheless, all our futures will arise from the immeasurably complex set of conditions that is the universe. There is plenty we do not know, and perhaps things we cannot know about this. So we need to focus on significant patterns of conditions that we can now observe and manipulate. For Buddhism, the most important ones concern the mind, thought and action. These are things that can be directly known, and worked on. We *can* develop greater self-directed freedom within this stream of conditions, so any view that suggests otherwise, and so encourages passivity, is problematic.

Responsibility for Actions, and Blame

In moral discourse, and in a court of law, issues of responsibility and praise and blame are central. The more a person is held to have been free of constraints when doing a reprehensible action, the more they are held personally responsible so as to be subject to stronger blame or punishment.

Hard determinism would seem to undermine all rationales for such blame or punishment other than as a useful fiction that may be used to condition people to act in different ways.

Accordingly, Goodman says:

“An enlightened spiritual teacher might sometimes praise or criticize others, in order to cause them to feel emotions that would be helpful to them at their stage on the path.

But he would never feel any resentment about their mistakes, or regard them as

genuinely responsible for their successes; all praise and blame, for such a teacher, would be merely an expression of skillful means (*upāya*).” (Goodman 2009b, p. 159)

Yet in Buddhist texts, Arahats (enlightened beings) are not beyond giving and being the object of criticism, implicitly blame. It is said that at the council convened after the Buddha’s death, consisting of 500 Arahats, the Arahata Ānanda was taken to task for various failings, such as not having asked the Buddha *which* were the ‘lesser and minor’ monastic rules that might be abolished after his death (DN.II.154). He was declared to have committed ‘an offence of wrongdoing’ for this (Vin.II.288-29).

Certainly, an enlightened teacher would not have resentment when blaming, and would want to be helpful to the blamed person, but then blame is only likely to have beneficial effects if the person it is directed at accepts that they were responsible for the blamed act. And this can be true, not just a useful delusion on their part.

As I see it, in moral discourse, courts of law, and also in Buddhist monastic discipline, holding a person responsible for an action makes perfect sense, so long as they were not insane, knew what they were doing, and knew what the immediate effects of their action would be. What Buddhism does do, though, is to council greater consideration for mitigating circumstances in assessing what *degree* of punishment or blame is appropriate, and also a lack of ill-will or vindictiveness towards a wrong-doer. Blame or punishment should be for the potential benefit of the recipient of these, as much as for the protection of others.

This is not a question of using a ‘skillful means’ falsehood.⁵ It is simply that the only good reason to use blame is due to its beneficial potential. In any case, the concept of blame does

not depend on the idea of a 100% free agent, or agent-essence. It can be sensibly directed at a person seen as a dynamic cluster of events/processes, so long as this cluster includes the capacity to understand the rebuke, and to reflect on behavior in the light of it. That is, blame has a proper place in Buddhist practice so long as it is not given, or taken, as if directed at a permanent Self with an unchangeable fault. Indeed, if this were so, blaming would have no chance of initiating change in the blamed person! So blame and responsibility are not concepts to be transcended as one's insight grows.

In the *Lakkhaṇa Sutta*, it is said that the *Bodhisatta* (the aspiring Buddha-to-be, in the Buddha's immediately previous lives) had great wisdom in that he used to enquire of renunciants and brahmins:

“What is it that is wholesome (*kusalaṃ*), what is it that is unwholesome? What is with fault (*sāvajjaṃ*), what is not? What course is to be followed (*sevitabbaṃ*), what is not?” (DN.III.157)

This sees being ‘with fault’ as closely linked to unwholesomeness, and clearly sees both as not to be pursued. Lance Cousins comments that:

“In the great majority of cases ... whatever other terms are associated with *kusala*, the term which is always present, usually immediately next to *kusala*, is blameless (*anavajja*)” (1996, p. 148).

There is a small uncertainty in how best to translate ‘*sāvajja*’ and ‘*anavajja*’. Cousins sees ‘*anavajja*’, as “(originally) not reprehensible, blameless; (later) faultless” (p. 139). Margaret

Cone, in *A Dictionary of Pāli*, says that ‘*avajja*’ means, as an adjective, ‘*blameable; low, inferior*’, as a noun, ‘*what is blameable; imperfection, fault*’, and ‘*sāvajja*’ as an adjective meaning ‘*blameable; faulty*’ (2001, p. 245). Perhaps a word that captures both aspects of the adjectival meaning is ‘*wrong*’, with ‘*right*’ for ‘*anavajja*’. While ‘*micchā*’ and ‘*sammā*’ are also translated this way, at AN.V.242, the ‘*micchā*’ and ‘*sammā*’ path factors are respectively called, among other things, ‘*sāvajja*’ and ‘*anavajja*’.

It is not surprising, then, that an unwholesome action is also criticized by the discerning or wise (*viññugarahitā*).⁶ Of course, blame from those who are not wise may not be appropriate: the eight worldly phenomena that one needs equanimity in the face of are gain and loss, fame and shame, blame and praise (*nidā, pasamsā*), and pleasure and pain (e.g. DN.III.260). As King Pasenadi says in the *Bāhitika Sutta*:

“...we do not recognize anything of value in the praise and blame (*vaṇṇaṃ vā avañṇaṃ*) of others spoken by foolish and ignorant people who speak without having investigated and evaluated; but we recognize as valuable the praise and blame of others spoken by wise, intelligent and sagacious persons who speak after having investigated and evaluated” (MN.II.114).

So wisdom does not entail going beyond the use of the idea of blame, and hence responsibility and agency.

The Benefit of Associating with the Wise

Indeed, criticism from the wise is much valued in Buddhism:

“It is easy to see the faults (*vajjam*) of others, but hard to see one’s own faults. Like chaff, one winnows another’s faults, but hides one’s own, like a crafty gambler hiding a bad throw.

If you see a wise person who points out faults (*vajja-*) and is an admonisher, you should associate with such an intelligent one as with a revealer of treasure; this will always be for the better, not worse.” (*Dhammapada* vv. 252, 76)

The utterance of another person, which may highlight one’s faults, and one’s own wise, probing attention (*yoniso manasikāra*), are said to be two key conditions for the arising of right view (MN.I.294), and consequently for the seven subsequent factors of the Noble Eight-factored Path. Indeed, if the Buddha had not taught, there would currently be a lot less wise attention and mindfulness in the world; and it is said *he* was inspired to become a buddha by many lives previously meeting and being inspired by a past buddha (*Buddhavaṃsa*, ch. 2A). So, other people are a great influence for good, as well as for ill, on our views and actions. It is in the context of interaction with other people that most questions of responsibility lay. While self-direction shows our agency, direction from others can help us improve this.

Conclusion

Overall, I do not think that either the non-Self teaching, or the teachings that everything other than *nirvāṇa* is conditioned, mean that we do not have a degree of freedom of action, and that

concepts of responsibility and blame become unfounded when a person has deep insight into these teachings. Rather, such insight, and related aspects of the Buddhist path, make us stronger agents of action, with greater responsibility and freedom, but also with greater compassion for the way we all bring suffering on ourselves and others by inept actions rooted in an ignorance of our conditioned nature. Hence, I infer that I am a ‘semi-compatibilist’ who believes in weak free will, and see this as the implication of the teachings of the *suttas* preserved by Theravāda Buddhism.

Notes

¹ The term ‘*dhamma*’ is Pāli for (conditioned) ‘phenomena’; ‘*dharma*’ in Sanskrit. Cf. ‘*Dhamma*’ and ‘*Dharma*’, respectively, for the ‘Buddhist teaching’, among various interpretations (the Law, etc.).

² Translations used are the author’s own, with references to page numbers in the Pali Text Society edition of the Pāli texts, preceded by volume number where appropriate. See list of in-text abbreviations following the Notes section, with references to Pali Text Society editions and translators.

³ A ‘*sutta*’ (Pāli; Sanskrit: ‘*sutra*’) is, in early Buddhism, a ‘saying of the Buddha’, and a ‘discourse’ in later Buddhism as well as broadly within Indian philosophy.

⁴ The ‘Higher *Dhamma*’ (Pāli; Sanskrit: Abhidharma).

⁵ In any case, in the Pāli *suttas*, the Buddha says that he only teaches, from what he knows to be true, what is spiritually helpful to a particular person, whether or not it is pleasant to hear (MN.I.395). He makes no mention of false-but-useful teachings (Harvey 1995b).

⁶ There is also reference to “virtues dear to the noble ones (*ariya-kantehi sīlehi*)” (SN.V.343).

Abbreviations

AN. *Aṅguttara Nikāya* (Bodhi 2012)

DN. *Dīgha Nikāya* (Walshe 1996)

It. *Itivuttaka* (Masefield 2001)

Miln. *Milindapañha* (Horner 1963/1964)

MN. *Majjhima Nikāya* (Ñāṇamoli & Bodhi 1995)

SN. *Saṃyutta Nikāya* (Bodhi 2005)

Ud. *Udāna* (Masefield 1994)

Vin. *Vinaya Piṭaka* (Horner 1938-66)

Vism. *Visuddhimagga* (Ñāṇamoli 1999)

References

Bodhi, B., trans., 2005. *The connected discourses of the Buddha*. Boston, MA: Wisdom.

Bodhi, B., trans., 2012. *The numerical discourses of the Buddha*. Boston, MA: Wisdom.

Cone, M., 2001. *A dictionary of Pāli, vol. I*. Oxford: Pali Text Society.

Cousins, L.S., 1996. Good or skillful? *Kusala* in canon and commentary. *Journal of Buddhist Ethics*, 3, pp. 136-64.

Flood, G., 1996. *An introduction to Hinduism*. Cambridge: Cambridge University Press.

Goodman, C., 2009a. Discussion of the Buddha's omniscience. [email] (personal communication).

Goodman, C., 2009b. *Consequences of compassion: an interpretation and defense of Buddhist ethics*. Oxford: Oxford University Press.

Harvey, P., 1995a. *The selfless mind: personality, consciousness and nirvana in early Buddhism*. London: Curzon Press.

Harvey, P., 1995b. Contemporary characterisations of the 'philosophy' of Nikāyan Buddhism. *Buddhist Studies Review*, 12(2), pp. 109-33.

Harvey, P., 2007. 'Freedom of the will' in the light of Theravāda Buddhist teachings. *Journal of Buddhist Ethics*, 14, pp. 35-98.

Horner, I.B., trans., 1938-66. *The book of the discipline, 6 vols.* London: Pali Text Society.

Horner, I.B., trans., 1963/1964. *Milinda's questions, 2 vols.* London: Pali Text Society.

Ñāṇamoli, B., 1999. *The path of purification: Visuddhimagga.* Onalaska, WA: BPS Pariyatti.

Ñāṇamoli, B. and Bodhi, B., trans., 1995. *The middle length discourses of the Buddha.* Boston, MA: Wisdom.

Masefield, P., trans., 1994. *The Udāna.* London: Pali Text Society.

Masefield, P., trans., 2001. *The Itivuttaka.* London: Pali Text Society.

Walshe, M., trans., 1996. *Long discourses of the Buddha.* 2nd ed. Boston, MA: Wisdom.

15 Emotions and Choice: Lessons from Tsongkhapa

Emily McRae

Introduction¹

B. Alan Wallace has called the Buddhist traditions' view of free will a 'pragmatic position' that builds on an obvious fact of human experience: "there are circumstances under which we are more or less free to make wise decisions that contribute to our and others' genuine happiness" (2011, p. 218). In Buddhist traditions, then, the pressing questions of freedom of will—and the related concepts of autonomy and choice—are more experiential than intellectual: how do we experience freedom in our everyday lives? They are also more normative than descriptive: why is exercising choice a good thing?

I flesh out this 'pragmatic' position on free will in the Buddhist philosophical traditions by focusing on one aspect of human experience that is often assumed to be largely unfree: emotional life. Buddhist moral philosophy, especially Tibetan Buddhist philosophy—which I draw and focus on here—stands out for the attention it pays to the subtlety of emotional experience and the seriousness with which it considers our ability to exercise freedom in our emotional lives. In fact, the intentional intervention in—and cultivation of—our emotional experiences is a foundational part of Tibetan Buddhist ethics. Many of the Tibetan Buddhist mind training (*blo sbyong*) exercises are aimed at reducing the negative emotional experiences of

anger, envy, and hatred and cultivating positive emotions, including love, compassion, and equanimity (*e.g.*, Patrul 1994, pp. 223-237; Tsongkhapa 2000, pp. 50-60; Gyalchok and Gyaltzen 2006, pp. 247-257).

I turn to the 14th century Tibetan Buddhist yogin and philosopher Tsongkhapa, whose account of anger and compassion offers a compelling explanation of the causes and conditions of our emotional experiences and the extent to which they are under our control. Drawing on Tsongkhapa, I argue that our ability to choose emotions is best understood as a capacity for intentional intervention, which depends not only on the strength of the emotion, but our background knowledge of the nature of emotional experiences and our capacity to observe our emotional states as they occur.

First, two qualifications regarding terminology are in order. As has been demonstrated by others (Dreyfus 2001; Heim 2008), there is no concept of ‘emotion’ in Buddhism and hence none of the accompanying concepts, such as the reason/emotion dichotomy, which are so prevalent in Western philosophy. In Tibetan, as in all traditional languages of Buddhism, there is no word for ‘emotion’, although there are words for particular emotions, such as love, anger, compassion, and envy, which are analyzed at length.²

Although there are no theories of emotion in Buddhist philosophy, philosophical reflection about the nature of certain emotions tends to emphasize the cognitive and affective elements of emotional experience, as well as long-term causes and conditions of emotional experience, such as underlying predispositions and habits, one’s environment, and the company one keeps. In

what follows, I draw on these reflections on the nature of particular emotional experiences to investigate the degree of control we have in these experiences and the dispositions that form from them.

Second, one of the aims in this paper is to uncover what ‘choice’ means in the context of emotional life. I use the word ‘choice’ mainly because it is used in the Western philosophical scholarship of the emotions with which I am in dialogue (Solomon 1976a; Rorty 1980; Nussbaum 1990). I take ‘choice’ to refer to a general sense of having control of and facility with our emotional experiences as well as the capacity to directly, intentionally, and through our own power influence our emotional dispositions. In this way, I use a more conventional rather than philosophically technical definition of ‘choice’, for instance, one that already assumes certain metaphysical notions of free will.

Part of what I am attempting to do in this paper is to try to uncover what a Buddhist pragmatic position on free will would look like. In Tibetan there are no words that directly correspond to the Western philosophical concepts of ‘choice’ or ‘free will’.³ But, there is overlap between the more conventional notion of choice, as outlined above, and traditional concepts in Tibetan Buddhism. For instance, the Tibetan word ‘*rang dbang*’—which Tsongkhapa uses in his discussion of managing our negative emotions—connotes self-control, autonomy, and independence. In what follows I hope to show that we can learn a great deal about exercising choice in our emotional lives—in the general rather than philosophically technical sense—by examining Tsongkhapa’s analysis of the possibility of having self-control (*rang dbang*) in the midst of a strong emotional experience.

Causes and Conditions for Afflictive Emotions: Tsongkhapa's Account

In his discussion of the afflictive emotions in the *Great Treatise of the Stages of the Path to Enlightenment (lam rim chen mo)*, Tsongkhapa (2000) draws on three main causes of these emotions: object, subject and basis. The first cause, the 'object', is who or what the emotion is about. Under this heading, Tsongkhapa also includes the judgments we make about the objects of our emotions. If I am envious of you, both you and the judgment I made about you (that you have more than you deserve) are included as part of the object of the envy (pp. 161-163). The object and the judgments made about it are conditions for afflictive emotional responses that are often emphasized in Western scholarship on emotions. Along with the object, emotional experience also requires a subject, that is, the being who is experiencing the emotion. The third cause of an afflictive emotion is what he calls 'the basis', or the basic predispositions towards certain emotional responses and against others. These predispositions are formed by (often complex) previous causes and conditions.⁴

Suppose I am angry with you for (what I perceive to be) a harsh criticism of a paper of which I am (or was, until your criticism) particularly proud.⁵ The object of the emotion, anger, is you and your harsh criticism. I am the subject of the emotion and the basis of the emotion is my own predisposition to be sensitive to criticism and my over-identification with my philosophical ideas. These predispositions, which set me up to be angry in situations where my work is under review, are formed by my experiences in this and previous lives. For example, I may have been

taught or picked up the view that a direct challenge to one's most cherished beliefs is a sign of deep disrespect and I may have seen friends or family members respond with anger in these situations. If I have developed the habit of becoming angry in this kind of situation in the past, then it is likely—if I do not intervene in some way—that I will become angry again in the present circumstance. Other people who do not share my predispositions would not be similarly set up to feel anger in this situation.⁶

Tsongkhapa's inclusion of the basis of emotion gives his account particular explanatory power, since by examining the basis we can better account for emotional experiences that would be difficult to explain by just looking to the object of the emotion and the judgments we make about it. Consider Tsongkhapa's analysis of someone who intentionally (without remorse) harms another. He writes,

“... when the conditions and causes—seeds left by afflictions to which they were previously habituated, a nearby object, and erroneous conceptions—come together, [those who do harm] give rise to the thought to harm, even though the harmdoers do not think, ‘I will feel malice’; whereas if those causes and conditions are not complete, they will never produce the thought to harm, even if the harmdoers think, ‘I will feel malice’” (2000, p. 161).

This analysis recognizes that, for hatred or malice to occur, one needs to have the perception of the right kind of object (‘the nearby object’) as well as a certain judgment (which Tsongkhapa sees as ‘erroneous conception’). But the basic predisposition toward hatred or malice is what

allows certain thoughts and perceptions to ‘stick’ and develop into the intention to harm. As Tsongkhapa points out, the thought ‘I will feel malice’ is neither necessary nor sufficient to give rise to feeling malice and the desire to harm others, since such feelings and desires can arise without the thought ‘I will feel malice’ and one can have this thought without it giving rise to malice. The thought must resonate with the person’s basic predispositions in order for it to give rise to intentions.

Contemporary Western philosophical scholarship (Solomon 1976a; Nussbaum 1990) often focuses nearly exclusively on the judgments or thoughts behind an emotional experience. In the case of malice, these thoughts may include: ‘you are inferior to me’, ‘you do not deserve what you have’, or ‘you have wronged me’. The thought ‘you are inferior to me’ may give rise to feelings of contempt, hatred or malice in some people. But in other persons it may give rise to shame, guilt or pity, which may motivate a desire to help another as a form of compensation. Alternatively, such a thought could give rise to feelings of pride or arrogance without any desire to harm or help another. The same thought and the same perceived object could produce different emotional and conative states in different people. This is difficult to explain if we only focus on the intentionality of an emotion and its accompanying thoughts, beliefs or judgments. Tsongkhapa, however, would explain these differences by pointing to differences in underlying predispositions.⁷

A Puzzle for Tsongkhapa

Tsongkhapa (and most Tibetan Buddhist philosophers) generally sees emotions as fundamentally malleable and subject to intentional cultivation, despite the fact that they issue from a wide range of causes and conditions, some of which are more consciously accessible than others. Yet, despite his commitment to the project of intentionally transforming emotions, he often emphasizes the lack of self-control one has over one's emotional states. In fact, one of Tsongkhapa's tactics for cultivating patience is to see anger as unjustified. The main strategy for doing that is to see the person who has wronged you as lacking self-control. Immediately following the passage on malice quoted above, Tsongkhapa writes,

“These causes and conditions produce the desire to harm; this in turn produces the work of harming; and this produces suffering for someone else, so those harmdoers do not have even the slightest self-control (*rang dbang cung zad kyang med*). Moreover, they have become like servants of their afflictions, because *they are under the control of others, i.e., their afflictions.*” (2000, p. 161, my italics)

This claim seems surprising, especially since it is stated in the middle of a larger discussion about how to manage our emotions, particularly anger and the desire to harm. How, on the one hand, can Tsongkhapa argue that those who feel malice toward us have no self-control and yet our malice towards others must be controlled, managed and eventually eradicated? Do we not also lack even the ‘slightest self-control’ with regard to our anger? Are we not also ‘servants of our afflictions’?

I argue that, with regard to the emotions, choice is best understood as the capacity for intentional intervention in our emotional experiences. Successful intervention in our emotions depends on certain factors, including the knowledge of methods of intervention, some understanding of the nature of the emotion and the depth and breadth of one's awareness of one's emotional state. In the following section, I examine some Western philosophical conceptions of choice or control over the emotions and argue that choice is best understood as intentional intervention. I argue that this understanding of emotional experience can explain the puzzle that I posed for Tsongkhapa.

Emotions and Choice

Historically in Western philosophical ethics, emotions have often been seen as fickle, unreliable and ultimately out of our control. Since the pervasiveness of this view in Western philosophy has been well documented, I will only briefly recount it (Sherman 1997; Solomon 1976a, 2007). Many of the most influential thinkers in the history of Western ethics, despite deep theoretical differences, shared skepticism about the possibility and desirability of intervening in one's emotional responses. Kant, for example, was skeptical of the project of basing morality on emotions at least in part due to the unreliability of emotional responses.⁸ Neither David Hume (1975) nor Adam Smith (1948)—despite their view that emotions, in particular sympathy, are the foundation of morality—present or even imply a program by which we can intentionally cultivate, control or choose sympathy.⁹ Even Aristotle, in his *Rhetoric*, who recognizes the power of orators to trigger certain emotional states, says surprisingly little about the intentional

cultivation of emotions in the *Nicomachean Ethics*, claims that emotions do not issue from choice (2002, 1106a2), and expresses a general skepticism regarding our ability to change our emotional dispositions as adults (Bk. X).¹⁰

However, in contemporary Western philosophical scholarship on the emotions there has been increasing criticism of the traditional view that emotions are out of our control (Solomon 1976a, 1976b, 2003, 2004, 2007; Sherman 1997; Nussbaum 2001). Solomon, for example, even claimed that we can choose our emotions (1976b). But what does ‘choosing’ one’s emotions mean?

Even a skeptic would agree with one basic sense in which we can choose our emotions: we can choose, at least to some degree, the external circumstances that are likely to give rise to emotions that we value and not to emotions that we do not value. This very limited sense of choice is analogous to ‘choosing’ not to get sick; we can only choose to do things that will make it less likely that we will get sick.

This idea of choice, which basically amounts to the avoidance of triggers, is philosophically thin and not always easy to accomplish. Philosophically, this basic sense of choice is not about emotions. Rather, it is about exercising choice over one’s actions and control over one’s circumstances. However, although learning to avoid situations that we have reason to think trigger uncontrollable negative emotions is important, it is not possible (and probably not desirable) to avoid all triggers of negative emotions. Simply by having relationships with others we will be exposed to situations that trigger negative emotions, such as jealousy or clinging infatuation. If, as Tsongkhapa suggests, there is a sense in which emotional dispositions set us up

for similar emotional experiences—consider the case of the angry person looking for someone to knock the chip off his shoulder—avoiding triggers may be difficult. Even hermits meditating in caves cannot escape the triggers of negative emotions.¹¹

Solomon offers a more robust sense in which we may choose emotions. He argues that emotions are essentially judgments. For example, anger is the judgment that one has been wronged and compassion is the judgment that another is unfairly suffering. According to this view, choosing emotions amounts to choosing how we judge situations. He writes,

“By forcing myself to be scrupulous in the search for evidence and knowledge of circumstance, and by training myself in self-understanding regarding my prejudices and influences, and by placing myself in appropriate circumstances, I can determine the kinds of judgments I will tend to make. I can do the same for my emotions.” (1976b, p. 32)

The idea here is not that we choose our emotions simply by avoiding the external situations that trigger them, but rather that we choose them by recognizing the emotion, including the object and cause(s) of the emotion, and challenging the judgments (or, we might add, thoughts, beliefs or images) that are triggering and sustaining the emotional experience. This kind of choice, while not nearly as direct as choosing to lift one’s arm, is still a robust sense of choice that can capture the ways in which we generally think we should take responsibility for our emotional lives.

This approach to choosing our emotions allows for more direct engagement in the formation, maintenance, and longevity of an emotional experience than does the avoidance-of-triggers

approach. On this view, if we are angry with someone, we can choose to reduce or eradicate that anger by analyzing the judgments we have made about the object of our anger and, by critically examining these judgments, we can reduce the duration and intensity of the emotional experience.

One problem with this approach, however, is that it focuses only on engagement with the object of the emotion. As noted, the object of the emotion is not always the cause of the emotion. We may feel anger at someone due to the complex causes and conditions that shaped our personal history (Tsongkhapa's 'basis'), such as our physical and psychological health, childhood experiences, emotional stress or fatigue, or substance abuse. In other words, although an emotion almost always has an object, sometimes the main cause of the emotion is the basis, not the object. When this is the case, analyzing the judgments that we make about the objects of our emotion seems ineffective, except perhaps to show us that the object is not the cause of the emotional experience.

Tsongkhapa's analysis of the three main causes of afflictive emotional experience, however, can give an account of choosing our emotions that is not limited to analyzing the object of the emotion. On his account, our emotional experiences are the result of many causes and conditions—and not only the object (which is most apparent to us). When we believe, as we usually do, that the judgments behind our emotions are true, it seems that the only way to explain and hence transform that emotion is to do something about the object. For instance, if I believe that my anger has that singular cause (the object of my anger), then my attempts to address my anger will focus on the person who I perceive as having wronged me. I may, depending on my

social power and personality, abuse or attack the person or I may turn the anger inward and cultivate resentment and hostility. But, if I open up the causal story of my anger to include, for instance, my sleep-deprivation, my childhood and other formative experiences and the habits I have formed because of them, the company I keep, and the environment in which I spend my days, then I have more avenues for addressing my anger. I could take a nap, seek therapeutic or spiritual practices that address childhood trauma, or befriend more positive people. Because emotions have more causes than simply their objects, methods that address such causes (in addition to the object) will be efficacious.

The intentional intervention in this wide range of causal conditions of emotional experience is, I submit, the best way to think about exercising choice or control over our emotions. As Tsongkhapa's account makes clear, emotions form from a variety of causes and conditions, including our judgments about the objects of our emotions, our environments, health and personal history. The problem with the first two conceptions of choice—as avoidance of triggers and as engagement with judgments—is that they are too narrow in scope. The first restricts focus to one's environment and the second to one's judgments. But if Tsongkhapa is right about the range of causal conditions of emotional experience, then successful intervention in one's emotions will engage all these causal and conditions.

There are certain core features of successful intentional intervention in one's emotional life. These features are highlighted in what I take to resolve Tsongkhapa's puzzle, to which I now turn.

Solving Tsongkhapa's Puzzle

Tsongkhapa's account of afflictive emotional experience allows for many avenues for intervention. This explains his assumption that emotions can be transformed, trained and cultivated. But how do we explain his comments that others who act on their afflictive emotions lack self-control? One possibility is that his claim, that others lack control over afflictive emotions, is a useful fiction designed to facilitate our moral and spiritual development. If the goal is simply to decrease and eventually eliminate our afflictive emotions, then imagining those who harm us as having no control over their actions seems prudent (as long as it is believable enough to be motivating).

But given Tsongkhapa's general commitment to understanding reality and not simply producing desired mental states, it seems unlikely he intended these reflections on a harmdoer's lack of control as useful fictions. I see no reason not to take Tsongkhapa at his word that there is some important sense in which people who harm us have no control over their afflictive emotions, yet we (readers of Tsongkhapa) have some degree of control over ours.

Tsongkhapa's discussion introduces two levels of control over the emotions: the relative self-control of the reader of Tsongkhapa, who has a commitment to practice the meditations Tsongkhapa presents (the 'meditator'), and the relative lack of self-control of the person who has harmed another (the 'harmdoer'). There are two main differences between the meditator and the harmdoer that can explain why it may make sense to say the former has control over afflictive

emotions while the latter does not. The first is that the harmdoer has already harmed someone (in this case, the meditator). The meditator, on the other hand, feels anger or resentment, but has not yet harmed the other person (as the example goes). The second main difference is that the meditator, in virtue of being a meditator and a reader of Tsongkhapa, has exposure to a variety of practical methods that are designed to intervene in our afflictive emotions. This is not to say, of course, that a person who is a meditator could not also be a harmdoer or vice versa. Rather, the differences are between a person who, out of hatred or malice, has already intentionally harmed someone else and a person (maybe the same one at a later time) who feels hatred or anger toward another but has not acted on it, and has access to relevant methods of intervention in emotional experience.

Tsongkhapa seems to take the fact that the harmdoer has already harmed another out of hatred or malice as evidence that the harmdoer lacks control over her afflictive emotions. The idea is that, since people who have strong afflictive emotions and harm others are, in the Buddhist view, perpetuating their own suffering, we can be sure that they lack self-control since, presumably, they do not want to suffer and would not, if they could help it.¹² The fact that they are participating in the misery-producing lifestyle in which people intentionally hurt each other means that they lack control over their afflictive emotions.

That the harmdoer did not intervene in her afflictive emotions (and subsequently harmed someone) means either she did not attempt to intervene in the emotional experiences that preceded her harmful action or she did and was unsuccessful. If she did not consider intervening, either she does not know how (ignorance) or she does but did not on that occasion think to

attempt an intervention (thoughtlessness). In either case, the harmdoer does not have many live choices because either she does not know what her options are or she is not mindful enough to realize she is in a situation in which interventions in her afflictive emotions may be helpful. Alternatively, a harmdoer may have considered an intervention but did not follow through, or attempted an intervention and failed because, for example, she was attracted to the afflictive emotion or identified with it. So, if someone *has* done harm, then that indicates lack of sufficient self-control.¹³

The meditator, on the other hand, is experimenting with interventions at an earlier stage, when the anger or hatred is still forming. Because the meditator presumably has knowledge of possible interventions (because she is reading Tsongkhapa), ignorance of appropriate tactics is not an obstacle to intervention. Similarly, since the meditator is engaging in these meditations, thoughtlessness is not an obstacle to intervention. Furthermore, because the meditator has not yet harmed another, it seems that her afflictive emotions have not reached the pitch of those of the harmdoer. She has three advantages to successful intervention—knowledge, mindfulness, and decreased emotional intensity—that the harmdoer lacked. When we consider these advantages, it seems less asymmetrical to claim that the harmdoer lacks control of her emotions but the meditator does not.

Tsongkhapa's example of the lack of control of the harmdoer and the relative self-control of the meditator reveals some important features of successful intervention in emotional life. First, successful intervention depends not only on the intensity of the emotional experience, but also on our knowledge and ability to pay attention to our experiences. Second, within the same

emotional event, say, anger, there may be points at which we can intervene and points at which we cannot. Anger, like many emotions, is not a monolithic experience. If it is prolonged, there are points when anger increases and decreases. Although we may lack control when anger has surged, we may not when anger begins to diminish (or before it surges). It is helpful, therefore, to understand the nature of emotional experience as changing and amorphous. To say that the degree of control we have in our emotional lives varies does not, therefore, translate into the claim that there are some emotions that we have control over and others we do not. Rather, within any emotional experience there are opportunities for intervention.

Concluding Remarks: Emotions, Choice, and Free Will

I have argued, with Tsongkhapa, that successful intervention in a negative emotional experience depends not only on the intensity of the emotional experience, but also on one's ability to pay attention to the workings of one's mind and body, knowledge of intervention practices, and insight into the nature of emotional experiences. I maintain that this explains Tsongkhapa's seemingly contradictory claims that the meditator can and should control (and eventually abandon) her anger and desire to harm others while the harmdoer is a 'servant to her afflictions'.

But what are the implications of Tsongkhapa's account for understanding free will in the Western philosophical context? Unlike typical deterministic arguments in Western philosophy that make explicit the vast web of causes that preceded any particular action to underscore how *unfree* the action was, Tsongkhapa uses the understanding of the complex preconditions of our

actions in order to explain how we can be *more free* in our actions, thoughts and feelings. Understanding the ways our emotional lives are causally conditioned increases our freedom. Moreover, Tsongkhapa's account suggests that freedom of will is not something that we either have or do not, but can come in degrees: we can gain and lose our freedom. We can gain more freedom, at least in part, by coming to understand the ways in which we are not free (see also Harvey 2007, p. 84). As Peter Harvey has argued with respect to Theravāda accounts of psychological freedom, "Mindfulness of how-I-am-conditioned-so-far seems a crucial ingredient in increasing freedom by reducing one's conditioning, replacing limiting unskillful conditioning by more open-ended skillful conditioning" (p. 85).

Second, Tsongkhapa's approach to exercising choice in one's emotional life highlights the forward-looking moral implications of the free will debate. Backward-looking models of responsibility—placing moral blame or praise for past actions—is not the focus of Tsongkhapa's analysis, although it typically is the focus of Western philosophical approaches here. Rather than asking if we can be held morally responsible for the actions we have performed, the more pressing question for Tsongkhapa (and Buddhist ethicists generally) is whether we will be motivated to strive for virtue and eliminate vice.¹⁴ Buddhist ethics refocuses the question from "Could he have done otherwise and, if not, should we blame (or praise) him?" to "Can I be freer in the ways I think, feel, and act, to help myself and others?" This does not imply backward-looking questions of responsibility are unimportant, but that discussions of free will are relevant to forward-looking considerations of self-cultivation that have been relatively unexplored in Western moral philosophy.

Notes

¹ Parts of this paper first appeared in “Emotions, Ethics, and Choice: Lessons from Tsongkhapa”, *Journal of Buddhist Ethics*, 19 (2012), pp. 344-369; I thank *JBE* for permission to reproduce them here.

² According to Dreyfus, in some circles the neologism ‘*tshor myong*’ is used in order to facilitate communication between Tibetan teachers and Western students, for whom ‘emotion’ is too important a concept to do without (2001, p. 31).

³ I would like to thank an anonymous reviewer at the *Journal of Buddhist Ethics* for helpful comments and questions on the meaning of ‘choice’ in Tsongkhapa’s texts.

⁴ For a more detailed account of the causes and conditions of emotional experience as explained by Tsongkhapa, see Tenzin (1999).

⁵ This example is based on an example given by Solomon (1976a).

⁶ I focus on the emotion of anger because, for Tsongkhapa and other Mahāyāna Buddhist philosophers, anger is a particularly destructive emotion and great attention is given to methods of reducing, eradicating or transforming it. See Cozort (1995).

⁷ Rorty has argued for the explanatory power of the basis of an emotion, which she called the ‘magnetizing disposition’ that orients an emotional experience. She defines it as a “disposition to gravitate toward and to create conditions that spring other dispositions” (1980, pp. 106-107).

⁸ In the *Groundwork*, Kant notoriously overstated his skepticism when he claimed that ‘inclinations’, such as love or sympathy, are “so far from having an absolute worth” that it must be the “universal wish of every rational being to be altogether free of them” (1786, p. 79).

⁹ According to Smith, although the ability to feel love and sympathy is inherent in human nature (even the ‘greatest ruffian’ is not altogether without them), there are natural limits to the degree to which we feel these emotions (1948, p. 73). Yet, no program of cultivating sympathy or intervening in emotional obstacles to sympathy is offered or implied.

¹⁰ There are some exceptions to this general skepticism. In his discussion of friendship, Aristotle considers the possibility that one’s friends, once good, can become bad (2002, 1165b1-b23). He also implies that it is possible that one’s friend may, as an adult, be made good again (1165b19-b20). Since becoming good means habituating one’s action and feelings, becoming good would presumably involve the changing of one’s emotional habits as an adult. However, Aristotle does not state whether these changes are achieved through intentional invention or because of a change in one’s circumstances.

¹¹ In a recent article on meditation retreats for lay people, one participant revealed that in a one-month silent retreat, he suffered from (a completely unfounded) worry that his dog had died in the interim (Stout 2011).

¹² Tsongkhapa writes, “If these beings had self-control, they would not have any suffering, because they could control it” (2000, p. 162).

¹³ This is an interesting point of similarity between Tsongkhapa’s view and Socrates’s position, in the *Apology* 25e-26b and the *Meno* 77b-78b (Plato 2002), that no one knowingly does wrong.

¹⁴ In his discussion of free will in the Theravāda, Harvey claims that “Buddhism accepts ‘freedom of will’ in the sense that before one acts one can and should stop and reflect on things to assess [their] moral suitability” (2007, p. 84); see also Wallace 2011, p. 219).

References

Aristotle, 2002. *Nicomachean ethics*. Translated by S. Broadie and C. Rowe. New York: Oxford University Press.

Cozort, D., 1995. Cutting the roots of virtue: Tsongkhapa on the results of anger. *Journal of Buddhist Ethics*, 2, pp. 83-104.

Dreyfus, G., 2001. Is compassion an emotion? A cross-cultural exploration of mental typologies. In: R. Davidson and A. Harrington, eds. *Visions of compassion*. New York: Oxford University Press. pp. 31-45.

Gyalchok, S. and Gyaltzen, K., 2006. *Mind training: the great collection*. Translated by T. Jinpa. Boston, MA: Wisdom Publications.

Harvey, P., 2007. 'Freedom of the will' in the light of Theravāda Buddhist teachings. *Journal of Buddhist Ethics*, 14, pp. 35-98.

Heim, M., 2008. Buddhism. In: J. Corrigan, ed. *Oxford handbook of religion and emotion*. New York: Oxford University Press. pp. 17-35.

Hume, D., 1975. *Treatise on human nature*. New York: Oxford University Press.

Kant, I., 1786. *Practical philosophy: the Cambridge edition of the works of Immanuel Kant*.

Translated and edited by M.J. Gregor, 1996. Cambridge: Cambridge University Press.

McRae, E., 2012. Emotions, ethics, and choice: lessons from Tsongkhapa. *Journal of Buddhist Ethics*, 19, pp. 344-369.

Nussbaum, M., 1990. *Love's knowledge: essays on philosophy and literature*. New York: Oxford University Press.

Nussbaum, M., 2001. *Upheavals of thought*. New York: Cambridge University Press.

Patrul, R., 1994. *The words of my perfect teacher*. Translated by the Padmakara Translation Group. San Francisco: Harper Collins.

Plato, 2002. *Five Dialogues*. Translated by G.M. Grube. Edited by J. Cooper. Indianapolis, IN: Hackett.

Rorty, A., 1980. Explaining emotions. In: A. Rorty, ed. *Explaining emotions*. Berkeley, CA: University of California Press. pp. 103-126.

Sherman, N., 1997. *Making a necessity of virtue: Aristotle and Kant on virtue*. Cambridge: Cambridge University Press.

Smith, A., 1948. *A theory of moral sentiments*. Oxford: Clarendon Press.

Solomon, R., 1976a. *The passions*. Garden City, NJ: Anchor Press/Doubleday.

Solomon, R., 1976b. Emotions and choice. *Review of Metaphysics*, 27(1), pp. 20-41.

Solomon, R., 2003. Emotions, thoughts, and feelings: what is a 'cognitive theory' of the emotions and does it neglect affectivity? In: A. Harzimoysis, ed. *Philosophy and the emotions: Royal Institute of Philosophy supplement, vol. 52*. New York: Cambridge. pp. 1-18.

Solomon, R., 2004. *In defense of sentimentality*. New York: Oxford University Press.

Solomon, R., 2007. *True to our feelings*. New York: Oxford University Press.

Stout, H., 2011. Getting far, far away from it all. *New York Times*, 4 Dec. p. 1.

Tenzin, G.L., 1999. *Emotions and their impact on health: an analysis of Tibetan and Western approaches*. Ph. D. Emory University.

Tsongkhapa, 2000. *The great treatise of the stages of the path to enlightenment*. Translated by the Lamrim Chenmo Translation Committee. Ithaca, NY: Snow Lion.

Wallace, B.A., 2011. A Buddhist view of free will: beyond determinism and indeterminism.
Journal of Consciousness Studies, 18(3-4), pp. 217-33.

16 Grasping Snakes: Reflections on Free Will, *Samādhī*, and *Dharmas*

Karin Meyers

Introduction

Before one can entertain the question of whether or how there is free will in Buddhism, one must define ‘free will’ and decide which Buddhism(s) one has in mind. There are conceptions of free will that are antithetical to common Buddhist ideas about the mind or person—such as the conception of free will as a specific faculty of the soul or as a power enjoyed by an autonomous agent. The fact that Buddhists cannot accept these forms of free will is of historical interest, but does not tell us much about Buddhist perspectives on freedom and action.

The problem is that these are specific *theoretical postulates* invoked to explain a much more basic fact of the human condition—the fact that human beings *appear* to have *some degree* of choice and control with respect to their external actions and even their internal mental states. We might call these ‘empirical freedom of action’ and ‘empirical free will’, respectively. These are conceptually distinct, but for economy’s sake I use ‘empirical free will’ to encompass both. The crucial point is that the term signals a set of *pre-theoretical* assumptions about human freedom.

Regardless of the metaphysical truth underlying empirical free will (or the conceptual models invoked to explain it), it appears to be at the heart of the free will problem in the West, and is

presupposed by virtually all conceptions of the Buddhist path. However, until recently Buddhists didn't regard its existence a *problem*. This should give us pause to consider from whence our free will problem arose, and caution us against seeking in Buddhist texts analogues to the categories and assumptions that have made free will a problem for *us*.¹

Elsewhere, I have argued if there is something like a problem of free will for Buddhists, it lays in the tension between personal and impersonal levels of discourse and experience, and should be examined in the context of Buddhist soteriology, rather than through the lens of concerns about determinism or moral responsibility (Meyers 2014). I have also argued at length against the thesis that all South Asian Buddhists (or perhaps *any*) were causal determinists and that the Buddhist commitment to empirical free will or to intention (*cetanā*) as a necessary ingredient of action (*karman*) implies some form of libertarianism; I have also illustrated the distortions that arise from imposing our (largely Judeo-Christian) concerns about moral responsibility onto the doctrine of karma (Meyers 2010). However, pressed to arrive at a *conclusion* about 'free will' in Buddhism, and with Abhidharma ('Higher *Dharma*' teaching) in mind, I have also suggested that while the freedoms attributed to persons can be explained in terms of impersonal psychophysical events (*dharmas*)—a view generally consistent with either compatibilism or event causal libertarianism—belief in these freedoms involves a kind of *delusion* about oneself as an agent critical for progress on the path.²

Here, I modify and expand upon this line of argument by tracing a greater plurality of Buddhist views on the relationship between personal and impersonal discourse and experience. I drop any attempt to attach to these familiar labels such as 'compatibilism' and instead focus on the ways

in which they challenge our assumptions. In the first part of the essay, I draw upon the Pāli Nikāyas³ to demonstrate that while empirical free will is assumed, it exists alongside varieties of freedom that we might consider impossible. Next, I draw upon the path outlined in the Nikāyas and Abhidharma⁴ to explain the role *samādhi* (meditative concentration)⁵ plays in the cultivation of these freedoms. I then take a more constructive turn, comparing how empirical free will can be explained in light of various Buddhist views on persons and *dharmas* (phenomena).

Although I cannot examine these views in detail, I attempt to demonstrate how they reveal a greater range of conceptual options for a Buddhist response to the free will problem than has sometimes been appreciated. The comparison also underscores a methodological point: in order to have meaningful philosophical dialogue with Buddhist traditions, we must recognize their diversity and historicity (as well as our own).

Empirical Free Will in the Nikāyas

While later Buddhist schools developed sophisticated explanations for the relationship between appearances and reality, the Nikāyas tend to be more pragmatically orientated. They agree with the broader Indian and Buddhist view that ordinary beings do not see things exactly as they are, but do not offer an ontological theory regarding the relation between what ordinary beings see and what a Buddha or *arhat* (liberated being) sees. In this context, it is assumed that persons are able to choose between one course of action and another and even gain optimal control over their mental states. In other words, the Nikāyas treat empirical free will as a real phenomenon and

basic axiom of the path. There is no palpable worry that this might be inconsistent with the Buddha's other prescriptions to end suffering, such as to regard the psychophysical aggregates as dependently arisen and not-self (*Anattalakkhaṇa Sutta* (SN iii.66-68); Bodhi 2000, pp. 901-903).⁶

Thus, the Buddha exhorts his disciples to reflect on their actions so that they might choose a better course (*Ambalaṭṭhikārahulovāda* (M i.415); Bodhi and Ñāṇamoli 1995, p. 524), or commends the ideal monk for his ability to control his mind and not be controlled by it (p. 310), for having mastery over his mind (AN iv.34; Bodhi 2012, p. 1024)—even to the degree that he does not think any thought that he does not wish to think (AN ii.36; Bodhi 2012, p. 423). As if that were not already considerably beyond what we typically have in mind as empirical free will, the Buddha explains that an adept at *samādhi* can wield manifold supernatural powers thereby creating and controlling matter in the human and divine realms and seeing into others' minds (*Sāmaññaphala Sutta* (DN i.76-80); Walshe 1994, pp. 104-106).⁷ Such powers are not strictly required for individual liberation,⁸ but this does not mean they were unreal or unimportant, even for otherwise 'rational' Buddhist philosophers.

No matter how much we try to domesticate it, Buddhism does not and *could not* take place in a materialist universe. The significance of this with regard to free will is not always appreciated. The fact that consciousness is at least *as* basic and in many cases *more* basic than matter, and *always* more important, results in a consistent prioritization of subjective data (data discerned subjectively and data regarding subjective experience) that is not particularly conducive to the

notion of reality as mechanical or quantifiable—ideas that have contributed to modern conceptions of causal determinism and the free will debate since the 17th century.

As the *Sāmaññaphala Sutta* makes clear, such supernormal powers are worthy of praise, but less valuable than the ultimate fruit of the ascetic life: freedom from the defilements⁹—the emotional, conative, and cognitive distortions that prevent one from seeing the way things really are and bind one to suffering and rebirth. The fact that one can *choose* to take up this path, gain *control* over one’s own mind, and thereby bring about an end to suffering is, in fact, so central to the Buddha’s message that he reserves his deepest censure for those teachers who deny the power of individual effort or initiative (AN i.33, i.287-288; Bodhi 2012, pp. 119, 364-365),¹⁰ and ridicules persons who imagine his teaching to imply that they are unable to act of their own initiative (AN iii.337-338; Bodhi 2012, pp. 900-902).¹¹

None of this implies any particular metaphysical thesis regarding free will, such as compatibilism, libertarianism, or any other ‘-ism’ (except perhaps anti-fatalism), but it does begin to explain why Buddhists did not regard freedom to choose or control our actions and mental states as a metaphysical problem worth worrying about. In a context where the possibility of attaining profound mastery over mind and matter and ultimately transcending the human condition are foundational, a far more pressing concern would have been the fact that people typically do not capitalize on their basic freedom by cultivating the path.

Empirical Free Will, *Samādhi*, and the Path

Although modern Buddhists tend to de-emphasize it, the cultivation of *samādhi* and specifically the *jhānas* (four graduated states of single-pointed absorption) play a central role in the conception of the path in the Nikāyas and the Abhidharma. There is, admittedly, some ambivalence as to whether *jhāna* is necessary for insight (*paññā*). For example, the Nikāyas define right *samādhi* as *jhāna* (e.g., *Saccavibhaṅga Sutta* (MN iii.252), Bodhi and Ñāṇamoli 1995, p. 1101), but the Abhidharma traditions suggest a less intense form of *samādhi* close to the first *jhāna* may suffice.¹² Nevertheless, the cultivation of the *jhānas* defines normative practice of the path in both the Nikāyas and the Abhidharma. Thus, if we want to know how empirical free will relates to the Buddhist path, we ought to consider how it relates to *jhāna*.¹³

Functionally, *jhāna* is a mental state in which a set of defilements called ‘the five hindrances’ (sense desire, ill will, sloth and torpor, restlessness and remorse, and doubt) is in complete abeyance and certain wholesome mental factors are operational. Chief among the latter are the ‘*jhāna* factors’ (applied attention, sustained attention, joy, ease, and single-pointedness). While all five *jhāna* factors are present in the first *jhāna*, grosser factors subside in successive *jhānas* leaving only single-pointedness as the defining factor of the fourth *jhāna*. To enter into and abide in each successive *jhāna* requires a further letting go or dis-identification with the aggregates, beginning with letting go of the gross material senses to enter into the first *jhāna*.¹⁴

The cultivation of the *jhānas* has several interrelated benefits. As a progressive letting go it paves the way for the complete letting go that is *nibbāna*.¹⁵ This helps weaken or eliminate the defilements and thereby enables clear seeing (*vipassanā*) of the way things really are. It is also

the basis for the cultivation of more attenuated states of absorption, mastery of which results in the supernatural powers mentioned above.¹⁶ Although the logic of this path is relatively straightforward, its relation to empirical free will is not.¹⁷

In the typical formula found in the Nikāyas, one exercises empirical free will by taking oneself to the foot of a tree or some such out-of-the-way place to begin cultivating mindfulness (*sati*) leading to *jhāna*.¹⁸ But this freedom is not optimal because it is constrained by the five hindrances. That this is the case becomes painfully obvious the first time one sits down to attend to a meditation object. One might even develop the distinct impression that one is not and perhaps has *never been* in control of one's own mind. But as time passes and one makes repeated efforts to attend to the object, the hindrances begin to subside and the *jhāna* factors start to kick in. As this happens there may be various bouts of distraction, loss of control, and even loss of awareness, but on the whole it becomes easier, more pleasant, and more interesting to focus on the object. It may even feel like one has gained greater control over one's mind, but every now and then one might also observe that even when one places it elsewhere, the mind tends to wander back to the object—*seemingly of its own accord*. Eventually, the mind may become so absorbed in the object that it is drawn into the first *jhāna*. Here there is no sense of separation from the object, and no sense of doing or willing. This absorption may occur spontaneously, but one can train to enter *jhāna* at will,¹⁹ abide in it for a predetermined length of time, exit, and observe with the 'wisdom-eye' and a *supernormal* degree of precision, clarity, and stability the mental factors from the previous moment of mind or any other phenomenon.²⁰

All of these oscillations between identification and dis-identification, and between effort and surrender, are critical and inescapable aspects of the cultivation of *jhāna* and the gradual path as a whole. In other words, the path is a form of intentional *habituation* that has a complex relationship to empirical free will. Although one capitalizes on empirical free will in choosing to take up the practice, at various points in the practice precisely who or what is in control can be difficult to discern (as is who or what is doing the discerning), and upon absorption into *jhāna* any sense of a doer or willer fades completely.²¹ Paradoxically, this temporary loss of a sense of self results in a greater degree of choice and control in everyday action—what we might call an *enhanced* empirical free will. Indeed, all the references to self-mastery mentioned in the previous section are associated with the cultivation of the *jhānas*.

Dharma Theory and the Two Truths

All this shifting back and forth between the personal and impersonal (as well as active and passive) registers might seem imprecise or unsatisfying for the philosopher who seeks a definitive ontology, but it is unavoidable and the Nikāyas make little attempt to defend this way of speaking.²² However, later generations of Buddhists did. They felt compelled to explain how these registers relate to each other in terms consistent with their own evolving epistemological and ontological theories. For the Ābhidharmikas (followers of Abhidharma), these theories concerned *dharmas*, mental and material micro-phenomena conceived as the salient elements of experience, basic building blocks of existence, or both.

Contemporary scholars typically understand *dharma* theory as a thoroughgoing ontological and mereological reductionism based on relatively late works (such as Vasubandhu's *Abhidharmakośa*). Sometimes they also speak as if such reductionism is the implicit intention of the Nikāya teachings on the aggregates and not-self. This reading of the earlier in terms of the later can produce valuable insights, but must be balanced against an attempt to trace the evolution of these ideas and the diversity of their trajectories. With respect to our present inquiry concerning empirical free will, this may reveal a greater range of conceptual options for Buddhist responses. I will sketch some of these here, beginning with a relatively late iteration of *dharma* theory.

In the mature Theravāda and Sarvāstivāda theories, *dharmas* came to be understood in an explicitly ontological sense as the *irreducible* elements of existence. There are important differences between these theories, but on the whole both can be described as a form of *critical realism* that:

“recognizes the distinctness of the world from the experiencing subject yet also distinguishes between those types of entities that truly exist independently of the cognitive act and those that owe their being to the act of cognition itself” (Karunadasa 1996, p. 20).

On this view, *dharmas* are objectively existent and ultimately real while persons are subjectively constructed and only conventionally real. This view can also be described as *conceptually realist*, in the sense that concepts of *dharmas* correspond to real states of affairs while concepts of *persons* do not (Karunadasa 1996, pp. 20ff; Gethin 2005, p. 180). From this perspective, it can

be explained that persons enjoy empirical free will owing to the presence or absence of freedom promoting or inhibiting *dharmas* (primarily mental factors), respectively.²³ Because persons are not part of the final ontology, belief in oneself as an agent is a *delusion*, but a useful and even *necessary* delusion insofar as it generates freedom-promoting mental factors and actions.

(‘Freedom’ here includes empirical free will as well as freedom from suffering.) This belief need not take the form of a philosophical view regarding the existence of a self (and cannot after ‘stream-entry’, the first stage of deep attainment), but remains somewhat operational up to the point of arhathood (attainment of enlightenment).²⁴

On this view, it is conventionally true that persons enjoy empirical free will. Moreover, this free will is *real* in the sense that it can be *explained* in terms of the activities of *dharmas*. However, *dharmas* themselves do not enjoy free will,²⁵ and ultimately there are no persons to enjoy it either. The concepts of *empirical free will* and the *person* are thus useful and even necessary, but potentially misleading with respect to the true nature of things. I take this to be more or less consistent with Vasubandhu’s *dharma* theory and the Sarvāstivādins’, as he presents them.²⁶ (I will discuss the Theravāda view below.)

Not all Buddhists were happy with this reductive view of persons. It is a bit odd, after all, to treat the concept of a *person* as useful or necessary but not give persons a place in one’s ontology—especially in light of the fact that the Buddha constantly spoke as if persons were real. Here I should note that Vasubandhu never says that the concept of a *person* is necessary. In fact, he does a lot of work to show that persons serve no real explanatory function and so are, *ontologically speaking, un-necessary*, mainly in the ninth chapter of his *Abhidharmakośa*

(Duerlinger 2003). But I mentioned it above for good reason: belief in one's agency seems *psychologically* necessary and beneficial for practice; additionally, elements of the path Vasubandhu elaborates—that partake of his final ontology—do not make sense without the *concept* of a *person*, such as the all-important mental factors of shame (*hrī*) and apprehension (*apatrāpya*).²⁷ While psychological and conceptual necessity may not entail ontological necessity, this does help explain why certain Buddhists, namely the Pudgalavādins (Personalists), resisted the reductionist project (long before Vasubandhu arrived).

The Pudgalavādins refused the idea that persons were merely conceptual constructions and felt compelled to defend the idea that they were in some sense real,²⁸ and as such, neither the same as nor different from the aggregates.²⁹ This does not make for a very neat or even coherent ontology (as Vasubandhu ably demonstrates), but may do greater justice to the soteriological necessity of the idea of a person as well as to the mysterious, ever-fluctuating experience of the individual on the path. In other words, the Pudgalavādin alternative exposes the limitations of a reductive ontology when it comes to giving an account of human experience (much as the humanities are supposed to expose the limitations of the natural sciences in giving such an account). On this view, persons enjoy empirical free will but precisely what persons are is inexpressible and how they relate to the conditioned flow of aggregates cannot be determined.

The Theravādins were not prepared to go quite that far, but they also made an effort to defend the value of the person-talk found in the Nikāyas against the encroaching reductionism of their own *dharma* theory. They stressed that personal and impersonal discourses are merely two modes of expressing the *same truth* and *equally effective* in liberating beings (Karunadasa 1996,

p. 24f).³⁰ They also held that persons are potentially misleading conceptual constructions. It has been suggested that the tension between these perspectives can be resolved by adding the further provision that statements regarding persons are true “so long as one does not imagine by ‘person’ a substance enduring in time” (Jayatilleke 1963, p. 365; see also Karunadasa 1996, p. 27). This might reflect an implicit logic at work in the Nikāyas, but the tension is also evidence of a shift from the more practical and empirical orientation of the Nikāyas to the more theoretical and ontological concerns of the Abhidharma (Karundasa 1996; Cox 2004; Gethin 2004a; Ronkin 2005).

While the Sarvāstivādins and Vasubandhu appear to have been centrally preoccupied with populating and de-populating a reductive ontology, respectively, Theravāda *dharma* theory tended to preserve more of the empirical and epistemological orientation of the Nikāyas. They placed more emphasis on how *dharmas* are experienced and known (*i.e.*, through direct perception based on *samādhi*) than on their explanatory necessity or physical or analytical irreducibility (Karunadasa 1996; Ronkin 2005).³¹ This orientation is illustrated in the way that the modern meditation master Pa Auk Sayadaw explains why only material *dharmas* and not aggregations of material *dharmas* (*rūpa-kālapas*), which are merely conceptually constructed, can serve as the basis for insight: with strong *samādhi* the illusion of the latter’s contiguity dissipates too quickly; they simply fade away under observation (2010, pp. 10-11, 124; Catherine 2011, pp. 7219-7224).

This empirical and pragmatic orientation is evident when *dharmas* are introduced in the Nikāyas. There they are conceived as dynamic qualities of experience, particularly those that play a

critical role in the path, such as the five hindrances (Gethin 2004a; see also Cox 2004). This use of the term ‘*dharma*’ amounts to a rejection of a substance metaphysics wherein qualities (*dharmas*) are possessed by an underlying substance (*dharmin*) (Gethin 2004a), but as with the aggregates and the other impersonal analytical rubrics found in the Nikāyas, the emphasis is more on how the world is known or can come to be known through introspective experience than on defining an objective ontology (Gethin 1986). These impersonal rubrics provide a conceptual map and vocabulary for navigating and making sense of experience in which an overt sense of self has receded. At the same time, attending to the individual *dharmas* or aggregates featured on these maps conduces to this dis-identification. If this is right, and we take seriously the outline of the path sketched above as well as the idea that the Buddha conveyed its truth through both the personal and impersonal registers, we arrive at a slightly different perspective on free will: empirical free will is part of ordinary subjective experience and exercising it is critical at various points on the path, but it is not observable or relevant when analyzing the *dharmas*.³²

On this view, worrying about whether we have free will, how we might have it, how we might talk about it, whether we are morally responsible for our actions, and so on, are just so many wrongly grasped snakes (*Alagaddūpama Sutta* (MN i.134); Bodhi and Ñāṇamoli 1995, p. 227).

Conclusion

Buddhist texts admit a variety of powers and freedoms that far exceed what we tend to think of as ‘free will’, or even regard as possible. Serious consideration of these as basic axioms of the

path, together with the fact that the Buddhist orientation towards subjective data makes determinism an unlikely conclusion, helps explain why historically Buddhists were not compelled by something like the free will problem. This should caution us against searching for analogues to our free will problem in Buddhist texts. Nevertheless, Buddhists did think deeply about a variety of human freedoms, agency, action, and reality, and canvassing their views on these topics might help fuel our reflections on free will.

As illustrated above, the cultivation of the path involves a complex play of identification and dis-identification and ever-shifting oscillation between personal and impersonal experience. This is reflected throughout Buddhist discourse in the interplay between the personal and impersonal registers. But Buddhists did not agree on what to make of this. Amongst their attempts to reconcile these registers, we find some theories that resonate with our modern tendency towards ontological reductionism and suspicion that free will is at best a useful delusion belying an impersonal (albeit material) reality. But if dialogue is an opportunity to expand our perspective rather than confirm current convictions, we might also pay attention to those Buddhist voices that resist or precede such reductionism, those that suggest that there is an irresolvable ambiguity between our reality as persons and agents and the impersonal reality of the aggregates, or those that take personal and impersonal experience to constitute distinct but equally relevant perspectives on the truth of conditioned reality.

Notes

¹ In this essay I use the first person plural to refer to persons who, like me, have received their training in the Western academy. I hope other readers will forgive my parochialism.

² This thesis is similar to ‘paleo-compatibilism’ (Siderits 1987, 2005, 2008, 2013), but not identical. Aside from reservations about framing the Buddhist view in terms of causal determinism and moral responsibility, I object to paleo-compatibilism’s ‘semantic dualism’ and ‘semantic insulation’ on exegetical grounds.

³ The Nikāyas are the original Pāli collections (‘baskets’) of the Buddha’s teachings. However, it is safe to assume broad agreement with the Chinese Āgamas on the doctrines discussed here.

⁴ I will use the Sanskrit spellings ‘*dharma*’ (Pāli: *dhamma*) and ‘Abhidharma’ (Pāli: Abhidhamma) rather than switching back and forth between Pāli and Sanskrit.

⁵ ‘Concentration’ is the common translation of ‘*samādhi*’, but connotes a forceful effort absent in *samādhi*. The focus in *samādhi* is the result of the gathering together, strengthening, and balancing of the mental factors, not of gritting one’s teeth.

⁶ References to the Pāli are to the Pali Text Society editions, using standard abbreviations followed by English translations; see “Abbreviations” below. I agree with Adam (2010, p. 248) that this *sutta* (a saying of the Buddha or, more generally, a discourse) implies we cannot directly control the suffering associated with the aggregates, but take its central message to concern the foolhardiness of identifying with the aggregates, rather than establishing the non-existence of a self. As for his conclusion that we cannot will our will, see below.

⁷ ‘Matter’ and ‘form’ are the common translations for ‘*rūpa*’, but it should be noted that *rūpa* is first and foremost the human body as experienced by a subject, which is extended to the subjective experience of other human bodies, and then to other external bodies and phenomena (Gethin 1986). The fact that objective reality is understood to correspond to this subjective

experience—and not the other way round—means we must be careful not to impose our idea of matter as inert stuff confirmed through objective measurement onto the Buddhist idea.

⁸ However, all *bodhisattvas* (advanced, avowed aspirants) must cultivate such powers at some point.

⁹ The term here is ‘taint’ (*āśava*), but has a similar scope to ‘defilement’ (*kilesa*).

¹⁰ Namely, the Ājīvika teacher Makkhali Gosāla.

¹¹ The term here is ‘*attakāra*’ (‘doing by oneself’).

¹² In Pāli the term is ‘*upācāra*’ (‘access’) *samādhi*. In the *Abhidharmakośa* (*Treasury of Abhidharma*), it is ‘*anāgāmya*’ (‘not yet arrived’).

¹³ This description is based on the Nikāyas and Buddhaghosa’s *Visuddhimagga* (*Path of Purification*) (Ñāṇamoli 1991), but at this level of generality does not differ significantly from Vasubandhu’s presentation in the *Abhidharmakośa*.

¹⁴ There is, however, some debate on this point (Meyers 2012).

¹⁵ The term ‘*nibbāna*’ is Pāli (Sanskrit: *nirvāṇa*).

¹⁶ These are the formless absorptions of infinite space, infinite consciousness, no-thing-ness, and neither perception nor non-perception.

¹⁷ The narrative description that follows is based on my limited experience with the practice and substantial instruction according to the Pa Auk method, which defines ‘*jhāna*’ as a more absorptive state than some Theravāda traditions.

¹⁸ A number of scholars have seen two distinct paths of meditation in the Nikāyas, but I follow Gethin (2004b) in taking the *suttas* to describe a single, relatively coherent path that is further elaborated in the Abhidharma.

¹⁹ For the beginner, the main lever of control (aside from faith) is applied attention (*vitakka*)—which involves effort, but, as practice deepens, subtle intention or resolve (*adhimokkha*) for a factor to strengthen or *jhāna* to arise becomes the main lever of control.

²⁰ The attainment of *jhāna* and discerning the nature of things with the wisdom-eye are supernormal psychic powers, but not *iddi* (supernatural) per se.

²¹ The effort associated with applied and sustained thought only becomes clear after some familiarity with the state and/or entry into the second *jhāna*.

²² Aside from the distinction between interpretable (*neyyattha*) and definitive (*nītattha*) statements (AN i.60; Bodhi 2012, p. 151).

²³ See Meyers (2010, 2014) for a detailed explanation and defense of this approach.

²⁴ According to the Nikāyas, there is no more personality view (*sakkāyadiṭṭhi*) after stream-entry, but an instinctual sense of self (conceit or *māna*) persists until arhathood.

²⁵ To say that *dharmas* enjoy free will would constitute what the Theravādin commentators call a breach in convention (*voḥārabheda*) resulting in nonsense (Karunadasa 1996, p. 27).

²⁶ Where they differ is with respect to what kinds of entities count as *dharmas*.

²⁷ These are considered foundational for ethical discipline and thus for the entire path.

²⁸ They are credited with saying that persons are concepts (*prajñapti*) and that they are real (*saccikattha*) and ultimate (*paramattha*). For most Buddhists these would be contradictory positions.

²⁹ For more on this, see the “Pudgalavada Buddhist Philosophy” entry in the *Internet Encyclopedia of Philosophy*; see Lusthaus (2009) for more on the idea of the person as a necessary concept (*prajñapti*).

³⁰ This view may be facilitated by the fact that the Pāli word for convention (*sammuti*) merely means ‘consensus’ and lacks the sense of concealing found in the Sanskrit word (*saṃvṛti*), but it does not seem much different from a view expressed in the *Mahāvibhāṣa-śāstra* (Dhammajoti 2007, pp. 78-79). By contrast, Vasubandhu emphasizes truth as relative to the kind of entity in question (ABK vi.4; ed. Pradhan 1970, p. 334; Pruden 1991, pp. 910-911).

³¹ The ancient yoga masters cited by Vasubandhu and Yaśomitra appear to have a similar view (Dhammajoti 2007, p. 80).

³² It should be noted, however, that mature Theravāda *dharma* theory, while preserving more of an empirical orientation, does promote an ontological reductionism and conceptual realism not dissimilar from that outlined in Vasubandhu’s *Abhidharmakośa*.

Abbreviations

ABK *Abhidharmakośabhāṣyam (Autocommentary on Abhidharmakośa)*

AN *Aṅguttara Nikāya (Numerical Discourses)*

DN *Dīgha Nikāya (Long Discourses)*

M *Milindapañha (Questions to King Milinda)*

MN *Majjhima Nikāya (Middle Length Discourses)*

SN *Saṃyutta Nikāya (Connected Discourses)*

References

Adam, M.T., 2010. No self, no free will, no problem. *Journal of the International Association of Buddhist Studies*, 33(1-2), pp. 239-265.

Bodhi, B. and Ñāṇamoli, B., trans., 1995. *The middle length discourses of the Buddha: a new translation of the Majjhima Nikāya*. Boston, MA: Wisdom.

Bodhi, B., trans., 2000. *The connected discourses of the Buddha: a translation of the Saṃyutta Nikāya*. Boston, MA: Wisdom.

Bodhi, B., trans., 2012. *The numerical discourses of the Buddha: a translation of the Aṅguttara Nikāya*. Boston, MA: Wisdom.

Catherine, S., 2011. *Wisdom wide and deep: a practical handbook for mastering jhāna and vipassanā*. [Kindle eBook] Boston, MA: Wisdom.

Cox, C., 2004. From category to ontology: the changing role of Dharma in Sarvāstivāda Abhidharma. *Journal of Indian Philosophy*, 32, pp. 543-597.

Dhammajoti, B.K.L., 2007. *Sarvāstivāda Abhidharma*. Hong Kong: Centre of Buddhist Studies.

Duerlinger, J., 2003. *Indian Buddhist theories of persons: Vasubandhu's refutation of the theory of a self*. New York: RoutledgeCurzon.

Gethin, R., 1986. The five khandhas: their treatment in the Nikāyas and early Abhidharma. *Journal of Indian Philosophy*, 14, pp. 35-53.

Gethin, R., 2004a. He who sees Dhamma sees dhammas: Dhamma in early Buddhism. *Journal of Indian Philosophy*, 32, pp. 513-542.

Gethin, R., 2004b. On the practice of Buddhist meditation according to the Pali Nikayas and exegetical sources. *Buddhismus in Geschichte und Gegenwart*, 9, pp. 201-221.

Gethin, R., 2005. On the nature of dharmas: a review article. *Buddhist Studies Review*, 22, pp. 175-194.

Jayatilke, K.N., 1963. *Early Buddhist theory of knowledge*. London: George Allen.

Karunadasa, Y., 1996. The dhamma theory: philosophical cornerstone of the Abhidhamma. *The Wheel Publication*, 412/413, pp. 1-17.

Lusthaus, D., 2009. Pudgalavāda doctrines of the person. In: W. Edelglass and J.L. Garfield, eds. 2009. *Buddhist philosophy: essential readings*. New York: Oxford. pp. 275-285.

Meyers, K., 2010. *Freedom and self-control: free will in South Asian Buddhism*. Ph. D. University of Chicago.

Meyers, K., 2012. The pleasant way: the dhyāna-s, insight and the path according to the *Abhidharmakośa*. *Buddhist Philosophy and Meditation Practice, 2nd IABU Conference Proceedings*, pp. 259-277. Available at:

<http://btmar.org/files/pdf/buddhist_philosophy_and_meditation_practice.pdf> [Accessed January 27, 2016].

Meyers, K., 2014. Free persons, empty selves: freedom and agency in light of the two truths. In M.R. Dasti and E.F. Bryant, eds. *Free will, agency and selfhood in Indian philosophy*. New York: Oxford. pp. 41-67.

Ñāṇamoli, B. trans., 1991. *Path of purification: Visuddhimagga by Bhadantācariya Buddhaghosa*. Kandy: Buddhist Publication Society.

Pradhan, P., ed., 1975. *Abhidharmakośabhāṣyam of Vasubandhu*. Patna: K.P. Jayaswal Research Institute.

Pruden, L., trans., 1991. *Abhidharmakośabhāṣyam of Vasubandhu translated into French by Louis de La Vallée Poussin, vol. III*. Berkeley: Asian Humanities Press.

Ronkin, N., 2005. *Early Buddhist metaphysics: the making of a philosophical tradition*. London: RoutledgeCurzon.

Sayadaw, P.A., 2010. *Knowing and seeing*. 4th ed. Singapore: Pa Auk Meditation Centre.

Siderits, M., 1987. Beyond compatibilism: a Buddhist approach to freedom and determinism. *American Philosophical Quarterly*, 24, pp. 149-159.

Siderits, M., 2005. Freedom, caring and Buddhist philosophy. *Contemporary Buddhism*, 6, pp. 87-116.

Siderits, M., 2008. Paleo-compatibilism and Buddhist reductionism. *Sophia*, 47, pp. 29-42.

Siderits, M., 2013. Buddhist paleocompatibilism. *Philosophy East & West*, 63(1), pp. 73-87.

Walshe, M., trans., 1994. *The long discourses of the Buddha: a translation of the Dīgha Nikāya*. Boston, MA: Wisdom.

17 Agentless Agency: The Soft Compatibilist Argument from Buddhist Meditation, Mind-Mastery, Evitabilism, and Mental Freedom

Rick Repetti

As Maria Heim has argued somewhat convincingly (2014), the basic orientation of Buddhist philosophy is significantly orthogonal to the sort of thinking associated with Western philosophical theorizing about free will. In this paper, however, I argue for the view that, despite a lack of explicit theorizing about free will, the Buddha himself was implicitly committed to elements of a free will theory, and his teachings contain a rich set of resources that may be employed in the construction of such a theory. This is not to say that Buddhism *has* a theory of free will, but that it *can* have one—one powerful enough to at least hold its own against the most powerful form of free will skepticism in analytic Western philosophy.

According to (most) Buddhist philosophy, nothing that constitutes us forms a whole, maintains its identity throughout the changes it undergoes, or lasts long enough to constitute a self (MN I.230–235; SN I.135, III.66; M 25).¹ Buddhism denies the *ultimate* reality of the self, but acknowledges the validity of *conventional* person-involving discourse (SN I.14; *Itivuttaka* 53; *Sutta Nipata* 937; MN III.31; Aronson 2004; Adam 2010; Siderits 2003 and this volume; see also Federman 2010; Harvey 2007 and this volume; Meyers 2010, 2014, and this volume). If there is no self, as Goodman insists, there cannot be an autonomous self (2002; *cf.* Repetti 2012b; Meyers, this volume; and Harvey, this volume). Buddhism has remained mostly silent on free

will, with the exception of the Buddha. For these and related reasons (Heim 2014; Gowans, this volume), some think Buddhist discussion of free will is misguided (Garfield and Flanagan, both this volume).

I have reviewed several extant Buddhist theories of free will (Repetti 2010c, 2012a, 2012b, 2014, and the Introduction, this volume), and argued that there ought to be a Buddhist theory of free will (Repetti 2010b, 2015, and in Part I of this volume). In this article, I focus on articulating arguments for the theory of free will I find most consistent with Buddhism and contemporary Western discussions: soft compatibilism, the view that evitabilism, self-regulative ability, and mental freedom (sufficient for moral responsibility) are possible independently of whether the world is deterministic or indeterministic, and, in the case of Buddhism, independently of whether the self is ultimately illusory (the latter of which constitutes a Buddhist variant of ‘semi-compatibilism’). The question whether Buddhist causation is deterministic or not, both, or neither technically does not matter to my view. I leave it to others to dispute such ultimately empirical (Balaguer 2009) matters (*e.g.*, Story 1974; Rāhula 1974; Gómez 1975; Kalupahana 1992, 1995; Garfield 2001; Goodman 2002, 2009; Siderits 2003; Gier and Kjellberg 2004; Harvey 2007 and this volume; Federman 2010; and Wallace, this volume).

The Argument from the Buddha’s Rejection of Inevitabilism

Although the Buddha rejected the notion of an autonomous self and did not have the term ‘free will’, he rejected ideas that negate free will—the notions of inevitable causation by matter

(Wallace 2011), karma (AN I 173-5; M 134–138; *Visuddhimagga* 532, 535; *Patthana* I.1), gods (AN I 173-5; DN I.18;² Federman 2010, p. 14; Wallace, this volume), chance (DN I.28, I.52; AN I.173-175; SN II.22; Gómez 1975; Federman 2010), and fate (DN I.53-54, AN III.338, Raju 1985). These ideas share inevitabilism: the idea that matter or karma causes events mechanistically suggests they are beyond agents' control; likewise, causation by fate, gods, and chance are considered beyond agents' control. We may infer from the Buddha's rejection of these inevitabilist doctrines that the Buddha implicitly accepted evitabilism.

If causation is deterministic (exceptionlessly lawful), even hard determinists—those who think determinism implies there is no free will—concede that it need not imply one kind of inevitabilism, for determinism does not rule out the possibility that certain attempts to avoid certain outcomes are causally effective in avoiding them. To my lights, however, this concession significantly collapses the distinction between soft and hard determinism, since the former think this sort of evitabilism, and other abilities of agents akin to it, suffice for a determinism-friendly free will in the moral-responsibility-entailing sense.

What hard determinists insist on, instead, is a different kind of inevitabilism that focuses on the fact that agents cannot alter *whatever it is that is determined to be the outcome*: if determinism is true, there is a single outcome/event uniquely determined by all prior/simultaneous causal conditions, in every moment, and no outcome in this singular causal chain is evitable (van Inwagen 1983). Call this 'outcome inevitabilism'. I have critiqued deterministic outcome inevitabilism as self-contradictory elsewhere (Repetti 2010a, Repetti 2012b). Nevertheless, the Buddha would likely have rejected *any* form of inevitabilism, including outcome inevitabilism.

For, in rejecting the Ājīvika inevitabilism of Makkhali Gosāla, on the grounds that it implied self-refuting volitional impotence (AN I.173–175, III.338; MN II.214; DN I.28, I.53, II.19-20; SN II.22; Federman 2010), the Buddha implicitly accepts evitabilism. Similarly, in rejecting the (then-prevalent Sāṃkhya³) idea of inevitable causation by *prakṛti* (matter) (Wallace 2011), the Buddha implicitly rejected a strict determinism, *contra* Goodman. Goodman (2012) has objected that in rejecting Ājīvika fatalism the Buddha did not reject determinism, because the two doctrines diverge over causality. But, fatalism and determinism share outcome inevitabilism, and in rejecting Sāṃkhya material causation, the Buddha rejected *causal* inevitabilism, thus determinism.

Similarly, in describing Buddhism as a middle path between the annihilationism of Cārvāka materialism and the eternalism of belief in *ātman* (soul/self) (Gómez 1975; Kalupahana 1976, 1992, 1995; Harvey 2007 and this volume), the Buddha rejected purely materialist and purely spiritual accounts. In rejecting *ātman*, the Buddha rejected *one* libertarian conception that sees the agent as immaterial, outside the causal nexus (Harvey 2007 and this volume; Federman 2010). In rejecting causation by gods (AN I.173ff; Wallace 2011; Harvey, this volume), the Buddha rejected divine inevitabilism, along with the more specific Vedic idea that enlightenment is a function of divine will. Rather, the Buddha taught one can attain liberation/enlightenment, as he did, the most important evitabilist outcome in Buddhism. (For extensive analysis of whether evitabilism is compatible with determinism, see Dennett 1984; Repetti 2010a, 2012b.)

Some scholars have argued that a principal reason Buddhist talk of free will would involve category error is that Western free will talk emerges only within the Abrahamic requirement for theodicy, but no such need arises in Buddhism (Garfield and Flanagan, both this volume), an argument that courts the genetic fallacy, but also seems factually incorrect. In asserting that humanity can attain enlightenment without—against—divine will (his hagiography depicts gods trying to thwart his enlightenment), the Buddha's account is Promethean regarding the human will contra divine will, a kind of reverse theodicy: humans can attain the greatest liberation, enlightenment, and happiness through their own self-regulating efforts in the absence of a fairly libertarian *ātman* or the mercy/will of gods (Federman 2010). The point is that enough reasons for thinking about human agency arise in connection with the Buddha's rejection of the *ātman*/Brahman (soul/God) metaphysics of some, and his rejection of the inevitabilism of other, of his contemporaries. In rejecting the doctrine of the fruitlessness of action (Gómez 1975, p. 81), the Buddha implicitly endorsed evitabilism.

Unlike the other forms of inevitabilism, karmic inevitabilism may be an exception to the claim that Buddhism has not discussed free will throughout its history subsequent to the Buddha's rejection of inevitabilism. The Buddha's rejection of these forms of inevitabilism arguably eliminated need for further clarification in all of the above cases except that of karma, since karma is an opaque, complex concept that plays a central role on the Buddhist path. There are many clarifications about the nature of karma throughout the history of Buddhism that attempt to eliminate the idea that it functions in an inevitabilist manner (Kalupahana 1976, 1992, 1995; Harvey 2007; Federman 2010; Meyers 2010, 2014).

In rejecting causation by pure chance (DN I.28; SN II.22), the Buddha again rejected inevitabilism, for one cannot control or alter what is purely random. Arguably, in rejecting uncontrollable/inevitable chance, the Buddha presaged an argument against a form of libertarianism characterized by indeterminism. Some libertarians argue that because determinism seems to undermine free will by eliminating the possibility of doing otherwise, indeterminism avoids that problem, since the ability to do otherwise is possible in an indeterministic world (Kane 2002). True, but indeterminism brings its own inevitabilism: if the causes of choices/actions are purely random, their occurrence is beyond the agent's control, rendering them inevitable (Pereboom 2001). The Buddha did not have the terms 'determinism', 'indeterminism', 'evitabilism', or 'inevitabilism'. However, he rejected all forms of inevitabilism, implicitly accepting evitabilism.

The Argument from Buddhist-path Abilities

An evitabilist argument may be predicated on the self-regulative abilities required by the Buddhist path. These arguments are spelled out in elaborate detail elsewhere (Repetti 2010b, 2015; see also Meyers 2010, 2014, and this volume; and Hyland 2014). Given these abilities, Buddhism may formulate a theory of agency that stands up against the most powerful arguments for free will skepticism in contemporary philosophy (Repetti 2015), which I'll only summarize, then offer my reply. These arguments include:

- van Inwagen's consequence argument: if choices are necessary consequences of prior events, they are unfree (1975)
- Pereboom's manipulation argument: manipulated choices are unfree, determinism is functionally equivalent to manipulation, and thus determined choices are unfree (2001)
- Pereboom's randomness argument: we cannot claim authorship over random events (2001)
- Strawson's impossibility argument: choices are always conditioned by mental states, so free will is impossible (1994).

A possible Buddhist reply to these arguments is as follows. Buddhist practice cultivates such a heightened form of self-regulative ability that it *doesn't matter* to the *ārya* (advanced meditation practitioner) whether the causal factors influencing her mental states at the choice-moment were produced deterministically, randomly, manipulatively, or otherwise. The reason is that a central skill cultivated along the meditative path is the ability to detect, discern, and disengage from the volitional impetus of any mental state component—the ability to sense and inspect its phenomenological features carefully and mindfully, discern its characteristics and whether or not they are dharmic (productive of mental freedom) (MN I.415), and choose whether or not to act on it (MN I.415–416). (Repetti (2010b) makes the case for these abilities and their relevance to free will skepticism; Repetti (2015) elaborates the argument that these abilities ground a rebuttal to the four leading arguments for free will skepticism.) This sort of control the *ārya* has over her own mind far exceeds the minimal choice-making/alternatives-accessing abilities presupposed in free will literature (Harvey 2007 and this volume; Wallace 2011; Meyers 2014 and this volume).

A distinction from the Western literature may elucidate the claim that Buddhist agency is more powerful than ordinary agency. Frankfurt (1971) distinguished between *freedom of action*, the ability to satisfy one's 'first-order' desires (for objects or experiences), possessed by animals, young children, healthy adults, and even adults with certain disabilities, and *freedom of the will*, the ability to bring one's effective first-order desires into harmony with one's 'second-order' desires (metavolitions, of approval and/or disapproval of one's first-order desires), an ability not possessed by animals, young children, or some of the mentally ill. Rather than *endorse* Frankfurt's theory of free will, or claim that Buddhism does, *contra* Flanagan (this volume) and Tuske (2013), I only *employ* Frankfurt's distinctions to illustrate the superior agential (free-will-theoretic) abilities the *ārya* develops.

Thus, in meditation the *ārya* observes the comings, goings, and patterns among her thoughts, emotions, bodily sensations, desires, and the like (MN I.56-60)—without acting on them—cultivating (among other things) the ability to go *off-line*, relatively speaking, from the typical stimulus/response pairings that dominate the behavior of most sentient beings and engage them immediately with the environmental causal nexus. (See Repetti (2010a) for an elaborate argument in support of the off-line analysis of this aspect of effective agency.) Of course, off-line status is *relative*, particularly in accordance with the Buddhist doctrine of interdependent origination (global phenomenal/causal interdependence), but there is a significant causal/functional difference between being physically engaged in relatively unmediated sensorimotor, affective, volitional interactions with one's environment, and sitting in meditative absorption, withdrawn from sensorimotor activity, observing/detaching from the contents and activities of one's mind, not being pulled or pushed to action by them (SN v. 156). Recent brain

research provides support for such causal/functional differences (*e.g.*, Christoff, Cosmelli, Legrand, and Thompson 2011). As Garfield notes (2001), causal powers play a key justificatory/explanatory role in even the most minimalistic forms of Buddhist metaphysics.

In the process, the *ārya* cultivates a hierarchical will characterized by metavolitions that govern which of her lower-order volitions issue in action and which do not. Individuals whose lower-order volitions issue in action despite higher-order disapproval exhibit volitional incontinence (weakness of will or, if chronic, breakdowns of will, *e.g.*, addiction). The *ārya* exhibits the opposite, strength of will. For in cultivating dharmically-approved volitions, such as compassion, generosity, and altruism, she strengthens lower-order volitions that increase self-regulative ability and mental freedom thereby, and weeds out lower-order volitions that decrease self-regulative ability and mental freedom, such as greed, envy, and hedonistic self-gratification (AN I.3, I.200–201; SN v.64–65; DN I.75; MN I.275; Wallace 2011, p. 227). This is not to suggest that the *ārya* is ‘strong-willed’ the way an egoistic Marine might be, but rather that the features of will that might normally overwhelm the average person increasingly diminish relatively simultaneously with the increase in ability to restrain them.

The *ārya* does employ volition- and mind-training at a level of discipline akin to that of a marital artist or other highly disciplined mind-body expert, and thus is a virtuoso of mental/volitional control,⁴ but as this control increases it is increasingly spontaneous/effortless, on the one hand, and the sense of self as an independently-existing autonomous entity decreases, on the other, and vice versa. When self-regulative control increases, so does the extent to which evitabilism holds

for that individual. Thus, Buddhist-path practices increase the extent to which practitioners exhibit free will, though the need to exercise it diminishes inversely.

The impersonality of agency does not imply that there are no self-regulative/agential abilities. It is *how they are understood* that matters, not *whether they exist*. An enlightened being might never need to exercise self-regulative ability, so it's not clear that enlightened beings actually 'have' agential abilities (*cf.* Coseru, this volume), although some Buddhists claim that practice must continue after enlightenment (Chinul 1983), which implies post-enlightenment agency. I'll offer additional reasons below to think enlightened beings retain agential abilities.

The Argument from Non-Self

We began with the idea that if there is no self, there cannot be an autonomous self. But analysis of the no-self idea might be used to defend free will, ironically. The manipulation argument rests on the premise that manipulation undermines free will. If there is no self, even gods lack a self and thus the ability to determine/manipulate anyone's behavior. Without any self, there can be neither manipulators nor manipulatees, and thus no manipulation argument, technically. The remaining question is whether individuals—however impersonally conceived as closely clustered causal processes—can control themselves sufficiently for attributions of moral responsibility.

To use a bodily-control analogy, some individuals are continent, others not. It doesn't matter whether continent or incontinent individuals possess a self, nor whether continence is caused deterministically or indeterministically. Only those with functional sensorimotor feedback loops from the bladder through volitional centers can control bladder release; absent feedback loops, bladder release is controlled by hydraulics. Only those who can control their bladders are responsible for their release. By extension, only the 'volitionally continent', who can control their actions, respond to moral reasons, guide their actions in accord with higher-order preferences, and so on, are responsible for their actions. *Āryas* cultivate these abilities in abundance.

While it is speculative to form any position on the 'agentless' agency of enlightened beings who lack a false self-sense, I doubt that the same sort of purely impersonal causal conditions that control the weather controlled the Buddha's bladder; nor was he incontinent. Likewise, environmental circumstances alone cannot act on the Buddha's vocal chords, producing speech appropriately formed for its audience the way inputs on a keyboard can produce un-minded/non-intentional responses from a computer. The Buddha exhibited masterful agency in every word, gesture, and action, however spontaneously. Lacking false self-sense and thus ego-volitions, compassion spontaneously guided his actions. But, arguably, they had to be *actions*, not *mere events*.

The relevant question is not whether the individual has a correct understanding of self, but whether she has the appropriate auto-regulative abilities. By analogy, a thermostat presumably lacks a self-sense, but the question about it is whether it regulates the climate system it is designed to control. Likewise, is the individual self-regulative? Manipulation cases imply it is

not the individual self who is causing her own behavior, but another self—but if nobody has a self, true manipulation is impossible, and other ‘selves’ are irrelevant. The question is: where is the control? Conversely, if manipulation—one individual controlling another—*is* conceptually possible in a self-less world, then so is self-control.

To borrow a classic Buddhist teaching, in the same way the wheels, axle, etc. (configured the right way, enabling the chariot to roll) ground the conventional use of the term ‘chariot’ without imputing ultimate existence to a chariot-essence independent of these things (*Milindapañha*), so too the metavolitional/volitional, sensorimotor responsiveness, etc. (configured right, enabling the individual to self-regulate) ground a metaphysically deflated conventional use of the term ‘agent’ without thereby imputing ultimate existence to an agent-essence, *ātman* or trans-empirical self independent of these things. Agency-lite, so to speak, is causally efficacious, but metaphysically insubstantial (Aronson 2004).

Similarly, the consequence argument threatens to undermine the idea that the agent is the originator of her actions, since it locates their origination through distal causal chains predating the agent—as if the causal sequence leading to the agent’s control sphere removes control from the agent, rendering the agent a puppet or robot of pre-agential forces. However, if there is no real agent, then there can be no worry that such a being lacks originative control over her actions because *pre-agential* forces control her. The enlightened being experiences no self, thus *loss of originative control to pre-agential control* is a non-issue. As Kalupahana put it, “The term *nibbāna* (Skt. *nirvāṇa*) conveys the same negative sense associated with the conception of freedom whenever the latter is defined as ‘absence of constraint’” (Kalupahana 1992, p. 91). One

man's drink is another's poison: the 'loss' of self is liberating, but doesn't necessarily remove 'self-control'.

The question—akin to that regarding continence—is whether the individual exerts proximal control over herself sufficient for moral attributions. If determinism is true, deterministic forces predating the agent determine whether she is continent, but if she is, she exerts a form of proximal control over herself that the incontinent doesn't, both of which hold in the absence of a metaphysically substantive self, and only the continent is responsible for her bladder release. Likewise, thermostats may be determined, but differ from avalanches, which lack self-regulative ability the thermostat possesses, though both—lacking cognition and volition—lack anything remotely appearing to function like a self.

Only a being that has free will can suffer its loss: a rock cannot suffer weakness, or loss, of will (Frankfurt 1971). Likewise, it is no deficit in a number that it cannot do as it pleases: it cannot appropriately be said to lack free will. If the absence of a self implies the absence of an entity that could fail to have free will, then there is no implication from non-self to unfree will.

However, questions about proximal causality remain: whether the individual self-regulates is independent of the question of self. The relevant difference between a thermostat and a person here is that the person has greater self-regulative abilities. This causal/functional difference obtains, relatively proportionately, between the *ārya* and the ordinary person. The *ārya* self-sculpts her hierarchical will in accord with the *Dharma*, increasing her self-regulative control thereby, despite her increasing recognition of the ultimate unreality of her self. Thus, self-control (autonomy) is not inconsistent with non-self, *contra* Goodman and Strawson (both this volume).

The Argument from Desert

The argument from desert is similar to the argument from non-self, and also *flips* the reasoning about the agent/self. (Repetti 2012b offers a more elaborate version of this argument.) Goodman, for example, has argued that because Buddhist causation is deterministic and impersonal, nobody is ultimately responsible for what they do, and thus nobody *deserves* praise or blame, though we may have desert-independent reasons for incapacitating those who commit immoral acts; for example, we may quarantine them to protect them from harming the innocent, or to rehabilitate them (Goodman 2009; *cf.* Repetti 2012b).

It seems sensible and humane to quarantine Jones if Jones did X and X is illegal/immoral, without blaming Jones or punishing her, on the assumption that nobody is a true agent or author of her actions. Assuming Jones can control her doing X, it seems reasonable to try to rehabilitate her. Rehabilitative justification thus rests on agential ability, and presupposes a kind of compatibilism: Jones need not do what she does, though her behavior is determined—she can do otherwise in the future, and she/we can intentionally bring about circumstances that make her doing what she's done evitable.

However, if Jones cannot control her doing X, and is not therefore capable of being rehabilitated, it seems reasonable to quarantine Jones to contain the danger of her doing X, and this danger-restricting non-rehabilitative quarantine seems justifiable because Jones did X, and presumably

Jones is likely to do X again (Goodman 2009). Let's examine this more carefully. *Did* Jones actually *do* X? Attributing X to Jones is analogous to attributing causal authorship to someone, say, who suddenly passes out behind the wheel of a car due to the first symptom of an unknown medical condition, crashing into others. If Jones is not capable of controlling her 'doing' X, then Jones is not the *author* of X. Absent that control, arguably, there is no more reason to quarantine Jones than to quarantine anyone. The only justification for singling-out Jones is the assumption that *Jones* controls X. But if Jones controls X, then Jones possesses agential abilities. This is a compatibilist intuition: If Jones is capable of controlling X, that is what justifies differentiating Jones from among everything else regarding preventing X.

If the wind against Jones tosses Jones through the air like a dangerous projectile, then it would be a mistake to attribute to Jones the power of levitation or the fault of recklessly flying, since it is the wind that is responsible for Jones's aerial recklessness, not any failure of self-control in Jones. Jones did not commit an act of aerial recklessness. In fact, Jones didn't 'do' anything. Jones's behavior doesn't count as an action, but as an event. Likewise, if Jones lacks agential abilities, Jones is not the author of Jones's other behaviors, even if they involve elements typically constituting intentional behavior, say, bodily movements, speech, and so forth (as some imagine the Buddha's behavior might be). Jones might be sleepwalking and compulsively cursing, for example.

However, the no-desert idea Goodman and other free will skeptics (ed. Caruso 2013) propose suggests everyone is in the same categories as recklessly-air-bound-somnambulist-cursing Jones, since nobody authors their actions. But that cannot be right. When not air-bound or

sleepwalking, suppose Jones often engages in intentional behavior like the rest of us, who exert proximal self-regulative or agential control over ourselves. Thus, it does not follow that because we lack selves we lack self-regulative abilities sufficient for desert. Agents are unlike air-bound-sleepwalking Jones because they exhibit proximal control over their behavior, rendering their behaviors things they *do*, rather than things that *happen* to them. This distinction stands regardless of the ultimately impersonal nature of individuals. Aristotle didn't discuss free will or the self, but he identified the grounds of praiseworthiness and blameworthiness (on analysis, positive and negative cases of moral desert, respectively) with voluntariness or continence (*enkratia*), which he contrasted with lack of control or volitional incontinence (*akrasia*) (1999).

Relying on the Buddha's equation of karma with volition/intention, and echoing Fischer's (2006) reason-responsiveness criterion for moral responsibility, Wallace makes a closely related point:

“Only voluntary actions produce karmic results, and the magnitude of the moral consequences of one's actions corresponds directly to the degree of one's mental balance, intelligence, and understanding” (2011, p. 221).

Resembling Frankfurt's distinction between those with and without free will, Wallace adds,

“the moral consequences of the actions of a person who is mentally ill or brain-damaged are relatively light, while those of a person of sound mind and clear understanding are relatively heavy” (2011, p. 221, citing AN I.249–253).

Thus, impersonality not only doesn't lead to universal exculpation or non-desert, but analysis of the differences between intentional actions and unintentional happenings reveals that the causal/functional characteristics that make a difference regarding desert are our auto-regulative abilities, irrespective of the impersonal nature of agency (or the character of causation).

***Dharma*-responsiveness, Soft Compatibilism, and Buddhist Semi-compatibilism**

I'm in a position now to use the *āryas'* *Dharma*-responsive abilities to construct a possible Buddhist theory of free will—again, not one necessarily held by any Buddhist, or by me, but simply an available position with potential explanatory purchase. *Dharma*-responsiveness is a species of the reason-responsiveness touted by semi-compatibilists as sufficient for determinism-friendly agency and moral responsibility (Fischer 2006). Reason-responsiveness is the ability of an agent to recognize, understand, deliberate on, evaluate, and voluntarily/effectively react rationally to reasons for action, particularly moral reasons. This skill may obtain in a deterministic or indeterministic universe (Mele 1995), and is thought by compatibilists to be sufficient for a weak-to-moderate form of agency and moral responsibility.

The compatible part of 'semi-compatibilism' is between determinism and reason-responsive moral agency. The incompatible part is between determinism and a more robust or strong form of agency whereby the agent doing *X* *could have done otherwise*, $\sim X$, under *identical* conditions ('contracausally')—impossible in a deterministic world, but conceivable in an indeterministic

one. Semi-compatibilists differentiate between two forms of control corresponding to these weak/strong conceptions of free will: *guidance* control versus *regulative* control (Fischer 2006).

Consider the control a driver exerts over the movements of a vehicle in a deterministic world: by turning the wheel left, the car moves left, and so forth. So, if Jones makes a left turn, she exerts guidance control and 'weak free will' (though her own intentions, beliefs, and so on leading to her making a left are themselves determined), for her voluntarily moving the steering wheel left nonetheless causes the car to go left. However, if we lived in an indeterministic world, it is possible Jones could have turned the wheel right under otherwise identical circumstances. If Jones can do otherwise under identical conditions, she has regulative control, 'strong free will'. Regulative control is impossible in a deterministic world because whatever caused Jones to turn left would *necessarily* cause her to do so under identically determinative conditions.

However, consider driving instructor Smith, accompanying Jones, possessing a second steering wheel, gas pedal, and brakes, which, if applied, override Jones's control, as in some driver's education vehicles. Smith has a kind of *relative* regulative control with respect to Jones, for if Smith sees Jones making a dangerous move, Smith can do otherwise, and steer the car to safety. In a deterministic world Smith's actions are equally determined, so Smith's *relative* regulative control is only *pseudo*-regulative. However, Smith has more actual, causal/functional, *practical* control than Jones, and we can imagine greater forms of pseudo-regulative control consistent with determinism that an agent could have over herself. The *ārya*, in fact, possesses significantly greater relative self-regulative control (Meyers 2014 and this volume), and it is an empirical question whether the world is purely deterministic (Balaguer 2009), and thus whether her

regulative control is ‘merely’ *pseudo*-regulative. Although some Buddhists think the Buddhist understanding of causation is deterministic (Goodman 2002, 2009), most think it is neither purely deterministic nor purely indeterministic (Rāhula 1974; Story 1976; Gómez 1975; Kalupahana 1976, 1992, 1995; *cf.* Griffiths 1982; Garfield 2001; Gier and Kjellberg 2004; Wallace 2011),⁵ what I have termed ‘wiggly determinism’ (Repetti 2010c), but the issue seems irrelevant to the causal/functional mechanics of her control.

In fact, adding an element of indeterminacy will neither decrease nor increase the *ārya*’s practical self-regulative ability because it doesn’t matter whether the conditions influencing her mental state are deterministic or indeterministic, nor whether they are subject to manipulation. She possesses a virtuoso-level self-regulative ability—a powerful practical skill—that appears to be causally/functionally superior to that of the relatively basic ability to simply do otherwise that might be exhibited, say, in Jones being able to have turned right though she turned left. The *ārya* exhibits a robust form of reason-responsiveness that is informed by the *Dharma* and embodied as a regularly-exercised *skill* in the relationship between her cognitive, conative, affective, somatic system and her metacognitive, metavolitional meditative muscles, so to speak. She is able to approve of and act on dharmic volitions, and disapprove of and refrain from acting on adharmic volitions (M 37-38), regardless of their deterministic, indeterministic, or manipulative causal origins. Her highly effective metavolitional control over her own volitional processes is analogous to the driving instructor’s control over the student driver, as she can override her own volitional processes at the metavolitional level. As Wallace notes, a key skill in this regard is her control over her own attention (this volume), the central control lever, so to speak, as well as over how she frames/interprets experience.

'Hard incompatibilism', the most powerful free-will-skepticism, claims free will is not only incompatible with determinism, but with indeterminism. The *ārya* instantiates the opposite: 'soft compatibilism', an evitabilist self-regulative agency that is compatible with determinism and indeterminism (Repetti 2015; see Mele 1995, for a defense of the view that autonomy is compatible with both determinism and indeterminism). Given these powerful agential abilities, non-Buddhists might wish to use these arguments to defend a view of autonomous agency that reifies the agent/self, but Buddhists would avoid this. That's why I emphasize that the Buddhist theory of agency be considered only 'semi-compatible' in a way that differs from Western semi-compatibilism. Western semi-compatibilism asserts compatibility between determinism and weak-to-moderate moral agency (involving guidance control), and incompatibility between determinism and robust or strong agency (involving indeterminism-requiring regulative control). 'Buddhist semi-compatibilism' asserts compatibility between causality (determinism or indeterminism) and a strong form of self-regulative agency that suffices for moral responsibility, but incompatibility with a substantive agent/self. Another way of identifying Buddhist semi-compatibilism, which Siderits does without using the term (1987, 2008, and this volume), is in terms of ultimate versus conventional truths: there is robust conventional agency, but ultimately no self. Meyers and Harvey (both this volume) seem to argue for a similar view.

Such a model coheres well with the Buddha's explicit rejection of inevitabilist doctrines, and his enlightenment insight to the effect that, although most of us are typically very powerfully deluded by primal confusion about the very nature of our own minds, bodies, beliefs, desires, and perceptions, to the point where we are relatively equivalent to somnambulists being thrown

about by powerful winds, there are things we can do to completely reverse the process, wake up, and free ourselves from erroneous beliefs, misguided desires, and self-defeating actions: we can attain total mental freedom *through our own efforts*. Although freedom of the will—per Frankfurt (1971), the ability to have the sort of will one wants to have—is not the goal of the Buddha’s prescription, it is a wholesome side effect (‘collateral non-damage’) that the *ārya* cultivates a highly effective, dharmic, metavolitional/volitional will on the road to agentless agency.

Notes

¹ “MN” abbreviates *Majjhima Nikāya*; “SN” abbreviates *Saṃyutta Nikāya*; and “M” abbreviates *Milindapañha*. Classical Buddhist texts are identified here first by title, subsequently (if reference to them recurs) by initial letter(s) abbreviations, akin to *Nichomachean Ethics* (“NE”), and follow the numbering of the Pāli Text Society editions. Online versions of most are available at www.accesstoinsight.org [accessed January 25, 2016].

² “DN” abbreviates *Dīgha Nikāya*; “AN” abbreviates *Aṅguttara Nikāya*.

³ In the *Buddha-Carita*, Aśvaghoṣa describes one of Siddhartha Gautama’s first meditation masters, Arāḍa Kālāma, as a follower of Sāṃkhya (see “Sāṅkhya” entry in *Internet Encyclopedia of Philosophy*).

⁴ See Meyers (2014 and this volume) on the extent of this control, which is claimed to rise to supranormal if not supernatural levels.

⁵ Meyers (2014) suggests Buddhist causation is consistent with indeterminism.

References

Aristotle, 1999. *Nicomachean ethics*. Translated by T. Irwin, 1999. Indianapolis, IN: Hackett.

Aronson, H.B., 2004. *Buddhist practice on Western ground: reconciling Eastern ideals and Western psychology*. London: Shambhala.

Adam, M., 2010. No self, no free will, no problem: implications of the *Anattalakkhaṇa Sutta* for a perennial philosophical issue. *Journal of the International Association of Buddhist Studies*, 33, pp. 239–265.

Balaguer, M., 2009. *Free will as an open scientific question*. Cambridge, MA: MIT Press.

Caruso, G.D., ed., 2013. *Exploring the illusion of free will and moral responsibility*. Lanham, MD: Lexington Books.

Chinul, 1983. *Tracing back the radiance: Chinul's Korean way of Zen*. Translated by R.E. Buswel. Honolulu, Hawaii: University of Hawaii Press.

Christoff, K., Cosmelli, D., Legrand, D., and Thompson, E., 2011. Specifying the self for cognitive neuroscience. *Trends in Cognitive Sciences*, 15(3), pp. 104-112.

Dennett, D.C., 1984. *Elbow room: the varieties of free will worth wanting*. Cambridge, MA: MIT Press.

Federman, A., 2010. What kind of free will did the Buddha teach? *Philosophy East and West*, 60, pp. 1-19.

Fischer, J.M., 2006. *My way: essays on moral responsibility*. New York: Oxford University Press.

Frankfurt, H., 1971. Freedom of the will and the concept of the person. *Journal of Philosophy*, 68(1), pp. 5-20.

Garfield, J.L., 2001. Nagarjuna's theory of causality: implications sacred and profane. *Philosophy East & West*, 51(4), pp. 507-524.

Gier, N.F. and Kjellberg, P., 2004. Buddhism and the freedom of the will: Pali and Mahayanist responses. In: J.K. Campbell, M. O'Rourke, and D. Shier, eds. *Freedom and determinism*. Cambridge, MA: MIT Press. pp. 277-304.

Gómez, L.O., 1975. Some aspects of the free will question in the Nikāyas. *Philosophy East & West*, 25, pp. 81-90.

Goodman, C., 2002. Resentment and reality: Buddhism on moral responsibility. *American Philosophical Quarterly*, 39(4), pp. 359-372.

Goodman, C., 2009. *Consequences of compassion: an interpretation and defense of Buddhist ethics*. New York: Oxford University Press.

Goodman, C., 2012. Discussion of whether the Buddha's rejection of fatalism implies rejection of (hard) determinism. [email] (Personal communication).

Griffiths, P.J., 1982. Notes toward a Buddhist critique of karmic theory. *Religious Studies*, 18, pp. 277-291.

Harvey, P., 2007. 'Freedom of the will' in the light of Theravāda Buddhist teachings. *Journal of Buddhist Ethics*, 14, pp. 35-98.

Heim, M., 2014. *The forerunner of all things: Buddhaghosa on mind, intention, and agency*. Oxford: Oxford University Press.

Hyland, T., 2014. Mindfulness, free will and Buddhist practice: can meditation enhance human agency? *Buddhist Studies Review*, 31(1), pp. 125-140.

Kalupahana, D.J., 1976. *Buddhist philosophy: a historical analysis*. Honolulu, Hawaii: University Press of Hawaii.

Kalupahana, D.J., 1992. *A history of Buddhist philosophy: continuities and discontinuities*.

Honolulu, Hawaii: University of Hawaii Press.

Kalupahana, D.J., 1995. *Ethics in early Buddhism*. Honolulu, Hawaii: University Press of Hawaii.

Kane, R. 2002. Some neglected pathways in the free will labyrinth. In: R. Kane, ed. *The Oxford handbook of free will*. New York: Oxford University Press. pp. 406-437.

Mele, A., 1995. *Autonomous agents: from self-control to autonomy*. New York: Oxford University Press.

Meyers, K., 2010. *Freedom and self-control: freedom in South Asian Buddhism*. Ph.D. thesis. University of Chicago.

Meyers, K., 2014. Free persons, empty selves: freedom and agency in light of the two truths. In: M.R. Dasti and E.F. Bryant, eds. *Free will, agency, and selfhood in Indian philosophy*. New York: Oxford University Press. pp. 41-67.

Pereboom, D., 2001. *Living without free will*. Cambridge: Cambridge University Press.

Rāhula, W., 1974. *What the Buddha taught*. New York: Grove Press.

Raju, P.T., 1985. *Structural depths of Indian thought*. Albany, NY: SUNY Press.

Repetti, R., 2010a. *The counterfactual theory of free will: a genuinely deterministic form of soft determinism*. Saarbrücken, Germany: LAP Lambert Academic.

Repetti, R., 2010b. Meditation and mental freedom: a Buddhist theory of free will. *Journal of Buddhist Ethics*, 17, pp. 166-212.

Repetti, R., 2010c. Earlier Buddhist theories of free will: compatibilism. *Journal of Buddhist Ethics*, 17, pp. 279-310.

Repetti, R., 2012a. Buddhist reductionism and free will: paleo-compatibilism. *Journal of Buddhist Ethics*, 19, pp. 33-95.

Repetti, R., 2012b. Buddhist hard determinism: no self, no free will, no responsibility. *Journal of Buddhist Ethics*, 19, pp. 130-97.

Repetti, R., 2014. Recent Buddhist theories of free will: compatibilism, incompatibilism, and beyond. *Journal of Buddhist Ethics*, 21, pp. 279-352.

Repetti, R., 2015. Buddhist meditation and the possibility of free will. *Science, Religion & Culture*, 2(2), pp. 81-98.

Siderits, M., 1987. Beyond compatibilism: a Buddhist approach to freedom and determinism. *American Philosophical Quarterly*, 24(2), pp. 149-159.

Siderits, M., 2003. *Personal identity and Buddhist philosophy: empty persons*. Burlington, VT: Ashgate.

Siderits, M., 2008. Paleo-compatibilism and Buddhist reductionism. *Sophia*, 47(1), pp. 29-42.

Story, F., 1976. *Dimensions of Buddhist thought: essays and dialogues*. Kandy, Sri Lanka: Buddhist Publication Society.

Strawson, G., 1994. The impossibility of moral responsibility. *Philosophical Studies*, 75(1-2), pp. 5-24.

Tuske, J., 2013. The non-self theory and problems in the philosophy of mind. In: Steven Emmanuel, ed. *Companion to Buddhist philosophy*. Hoboken, NJ: Wiley-Blackwell. pp. 419-428.

Van Inwagen, P., 1975. The incompatibility of free will and determinism. *Philosophical Studies*, 27, pp. 185-99.

Van Inwagen, P., 1983. *An essay on free will*. Oxford: Oxford University Press.

INDEX

- Abhidharma, 13, 16, 194, 195, 197, 214, 306, 315, 349, 378, 379, 382, 385, 388, 392, 393, 397, 398
- Adam, 13, 82, 89, 94, 262, 279, 280, 360, 392, 397, 401, 425
- adharmic, 58, 59, 422
- agent, viii, ix, xxv, 2, 4, 9, 12, 38, 40, 45, 46, 54, 70, 92, 95, 96, 97, 98, 99, 101, 102, 110, 111, 113, 123, 124, 125, 128, 137, 154, 156, 158, 159, 169, 193, 194, 196, 199, 204, 206, 208, 210, 212, 214, 241, 244, 250, 251, 262, 263, 269, 270, 275, 276, 277, 279, 280, 289, 292, 293, 294, 295, 296, 297, 299, 300, 307, 318, 319, 322, 324, 330, 331, 332, 334, 335, 345, 377, 378, 386, 405, 407, 413, 414, 415, 416, 419, 421, 422
- agent causation*, 92, 96, 97, 102, 110, 111, 123, 124, 125, 128, 137, 193, 194, 292, 295, 296
- agentless agency*, 10, 18, 45, 54, 423
- Ājīvika, 404
- anātman*, 99, 265, 311, 325
- anger, 10, 16, 43, 70, 71, 72, 73, 74, 75, 79, 80, 81, 82, 83, 84, 85, 86, 89, 112, 151, 185, 196, 199, 212, 228, 230, 234, 262, 354, 356, 359, 362, 363, 364, 366, 367, 368, 369, 371, 373
- Āṅguttara Nikāya*, 424
- anti-reflexivity, 319
- arahat*, 277
- arhat*, 47, 247, 248, 380
- Aristotle, xxx, 94, 96, 141, 146, 201, 216, 220, 360, 371, 373, 374, 417, 425
- Aronson, 59, 63, 65, 401, 413, 425
- ārya*, 408, 409, 410, 411, 415, 421, 422, 423
- ātman*, 9, 45, 60, 61, 311, 315, 330, 331, 405, 406, 413
- Augustine, 94, 99, 114, 115, 121, 127, 202, 214, 221, 222
- autonomy, 3, 46, 48, 57, 58, 198, 353, 355, 415, 422, 428
- Balaguer, 402, 421, 425
- Bodhi, 21, 27, 28, 42, 103, 119, 198, 204, 205, 206, 216, 221, 265, 271, 275, 278, 280, 311, 326, 329, 350, 351, 352, 380, 381, 382, 390, 394, 397
- Bodhicaryāvatāra*, 71, 83, 89, 91, 207, 218, 221, 248
- bodhicitta*, 232, 233, 235, 237
- bodhisattva*, 65, 80, 117, 147, 194, 196, 213, 218, 232, 233, 238
- Buddhaghosa, 24, 42, 219, 306, 338, 393, 399, 427
- Caruso, 7, 19, 50, 60, 65, 214, 217, 417, 425
- Cārvākas, 197
- cetanā*, 105, 199, 204, 215, 242, 266, 274, 332, 334, 378
- chance, xxv, xxxi, 11, 61, 129, 190, 219, 240, 251, 256, 258, 345, 403, 406
- Chinul, 411, 425
- compassion, ix, 40, 66, 76, 77, 78, 80, 87, 89, 102, 112, 117, 137, 147, 185, 193, 195, 207, 208, 217, 218, 232, 233, 237, 328, 348, 351, 354, 362, 373, 410, 413, 427
- consciousness, xiv, 11, 19, 28, 38, 65, 66, 67, 81, 82, 115, 154, 165, 205, 206, 207, 210, 216, 217, 243, 246, 247, 257, 260, 264, 299, 300, 308, 309, 311, 326, 332, 351, 381, 393
- consequentialism, 56, 58, 62, 133, 212
- continence, 412, 414, 417
- conventional, xxvi, 1, 3, 9, 13, 14, 38, 40, 52, 53, 54, 57, 58, 101, 105, 106, 107, 108, 110, 115, 125, 207, 209, 284, 286, 287, 296, 297, 306, 312, 327, 355, 401, 413, 423
- Coseru, 8, 9, 10, 11, 14, 15, 19, 55, 58, 61, 65, 193, 210, 217, 411
- Dennett, 73, 89, 93, 98, 108, 116, 171, 172, 212, 217, 238, 239, 242, 243, 244, 257, 263, 280, 299, 308, 405, 426
- dependent arising, 203, 205, 206, 207, 209, 215

dependent origination, xxvi, 27, 28, 29,
 30, 31, 32, 33, 35, 36, 37, 39, 41, 46, 83,
 100, 103, 104, 106, 107, 137, 141, 177,
 251, 275, 306, 338
 desert, 9, 70, 73, 79, 85, 86, 303, 304, 415,
 417, 418
 determinism, ix, xxv, xxvi, xxxi, 4, 7, 11, 12,
 14, 17, 21, 25, 27, 30, 31, 32, 33, 34, 35,
 37, 46, 67, 69, 74, 77, 79, 88, 92, 95, 96,
 98, 99, 102, 115, 120, 125, 143, 149,
 153, 154, 155, 156, 157, 158, 160, 161,
 164, 165, 166, 167, 172, 178, 195, 196,
 198, 203, 204, 206, 209, 215, 222, 225,
 227, 239, 243, 244, 245, 251, 260, 281,
 282, 283, 288, 289, 290, 298, 304, 305,
 306, 309, 313, 317, 324, 338, 340, 344,
 376, 378, 381, 390, 391, 400, 403, 404,
 405, 407, 408, 414, 419, 421, 422, 426,
 427, 429, 430, 431
 Dewey, 144, 146
dhamma, 348, 392, 398
Dhammapada, 228, 242, 347
 dhammas, 330, 398
Dharma, 18, 24, 46, 48, 56, 58, 229, 231,
 348, 378, 384, 397, 415, 419, 421
Dharma-responsiveness, 18, 419
dharmas, 105, 307, 378, 379, 385, 386,
 388, 389, 394, 398
 dharmic, 18, 56, 57, 59, 408, 421, 423
dukkha, 24, 44, 45, 46, 47
 eliminativism, 313, 314
 emptiness, 53, 101, 106, 116, 118, 119,
 207, 208, 255, 259
 enlightenment, 2, 3, 23, 25, 26, 29, 31, 32,
 33, 35, 37, 40, 44, 45, 46, 53, 54, 59, 63,
 177, 193, 224, 229, 230, 231, 232, 234,
 235, 255, 256, 311, 325, 326, 327, 343,
 375, 386, 405, 411, 423
 epiphenomenalism, 196
 evitabilism, 4, 402, 403, 404, 405, 406,
 407, 411
 fatalism, xxxi, 11, 73, 86, 153, 198, 251,
 256, 381, 404, 427
 fate, viii, xxv, 61, 86, 127, 156, 403
 Federman, 60, 65, 245, 257, 401, 403, 404,
 405, 406, 426
 first-order, 56, 57, 268, 269, 270, 278, 409
 Fischer, ix, 12, 19, 90, 418, 419, 420, 426
 Flanagan, ix, xxx, 4, 5, 48, 51, 56, 57, 62,
 65, 122, 134, 137, 141, 146, 147, 194,
 217, 218, 402, 405, 409
 Frankfurt, ix, 12, 13, 19, 57, 98, 116, 125,
 147, 262, 268, 269, 270, 271, 272, 276,
 278, 280, 319, 329, 409, 415, 418, 423,
 426
 free will, viii, ix, x, xi, xxii, xxiii, xxiv, xxv,
 xxvi, xxvii, xxviii, xxix, xxxi, 1, 2, 6, 19,
 21, 25, 27, 34, 35, 36, 41, 44, 45, 46, 47,
 48, 49, 51, 52, 53, 55, 56, 57, 58, 59, 60,
 61, 62, 63, 65, 67, 69, 71, 73, 74, 75, 77,
 78, 79, 87, 89, 90, 92, 93, 94, 95, 98, 99,
 102, 103, 104, 105, 111, 116, 117, 118,
 119, 121, 122, 123, 124, 125, 126, 127,
 128, 129, 130, 131, 132, 134, 136, 137,
 138, 139, 140, 141, 142, 143, 144, 145,
 147, 148, 149, 151, 160, 172, 174, 175,
 176, 177, 178, 179, 180, 181, 185, 187,
 189, 190, 193, 195, 201, 202, 203, 210,
 212, 214, 215, 217, 219, 220, 221, 222,
 224, 225, 226, 233, 235, 236, 238, 243,
 244, 245, 248, 250, 251, 252, 254, 256,
 257, 258, 259, 260, 262, 263, 264, 265,
 266, 268, 269, 270, 271, 274, 277, 280,
 281, 299, 300, 301, 305, 308, 313, 321,
 324, 330, 335, 348, 353, 355, 369, 372,
 376, 377, 378, 379, 380, 381, 382, 383,
 384, 385, 386, 388, 389, 390, 391, 394,
 397, 399, 401, 402, 403, 405, 406, 407,
 408, 409, 411, 415, 417, 418, 419, 420,
 422, 425, 426, 427, 428, 429, 430, 431
 free will skepticism, 47, 401, 407, 408
freedom of action, 264, 271, 339, 348, 377,
 409
freedom of the will, 112, 120, 129, 130,
 144, 147, 196, 198, 264, 271, 277, 281,
 409, 423, 426
 Galen Strawson, ix, 6, 8, 64, 125, 151, 172,
 313
Garfield, 4, 5, 48, 56, 61, 62, 64, 65, 92,
 100, 102, 103, 104, 106, 107, 108, 112,
 116, 117, 124, 140, 147, 200, 214, 215,

217, 218, 310, 399, 402, 405, 410, 421, 426
 God, 4, 25, 34, 35, 60, 94, 120, 127, 128, 129, 130, 131, 141, 202, 228, 231, 232, 233, 238, 239, 305, 332, 336, 406
 gods, xxiii, xxv, 58, 61, 129, 224, 235, 403, 405, 412
 Gómez, 93, 117, 123, 147, 402, 403, 405, 406, 421, 427
 Goodman, xvii, 2, 3, 4, 10, 47, 56, 58, 62, 66, 69, 71, 80, 89, 104, 117, 125, 143, 147, 196, 200, 218, 226, 227, 235, 237, 307, 308, 340, 341, 344, 351, 402, 404, 415, 416, 417, 421, 427
 Gowans, 2, 4, 21, 24, 43, 47, 140, 147, 402
 Griffiths, 104, 117, 125, 148, 421, 427
 hard incompatibilism, 3, 11, 46, 69, 79, 83, 85, 86
 Harvey, 9, 12, 15, 17, 47, 57, 62, 66, 93, 117, 123, 148, 244, 245, 246, 257, 258, 266, 281, 330, 331, 337, 340, 342, 349, 351, 352, 369, 372, 373, 401, 403, 405, 406, 408, 423, 427
 Heim, 215, 219, 354, 373, 401, 402, 427
 higher-order, 272, 410, 412
 Hume, 31, 111, 360, 373
 Hyland, 407, 428
 impermanence, 109, 137
 incompatibilism, 14, 17, 71, 72, 85, 86, 90, 125, 292, 303, 422, 430
 incontinence, 410, 418
 indeterminism, 4, 11, 12, 17, 70, 85, 112, 149, 198, 219, 222, 225, 226, 237, 239, 240, 251, 258, 260, 306, 313, 376, 407, 422, 424
Inevitabilism, 403
 insubstantiality, 53, 101, 207
Itivuttaka, 250, 350, 352, 401
jhāna, 382, 383, 384, 393, 394, 397
 Kalupahana, 402, 405, 406, 414, 421, 428
 Kane, 90, 92, 118, 128, 148, 197, 219, 240, 258, 407, 428
 Kant, 31, 37, 94, 96, 99, 113, 118, 134, 141, 148, 159, 172, 200, 211, 219, 220, 231, 303, 360, 371, 374
 karma, xxv, xxvi, xxx, 16, 26, 29, 31, 34, 35, 39, 51, 54, 55, 61, 103, 104, 141, 197, 217, 224, 228, 239, 242, 274, 332, 336, 341, 343, 378, 403, 406, 418
 libertarian, 36, 37, 60, 105, 111, 125, 127, 129, 141, 155, 161, 211, 225, 226, 313, 405, 406
 Madhyamaka, xxix, 4, 5, 52, 93, 95, 98, 99, 103, 105, 106, 107, 108, 109, 111, 112, 113, 114, 147, 207, 212, 251, 253, 259
 Mādhyamika, 99, 109, 111
Majjhima Nikāya, 240
 McRae, vi, xii, xix, 16, 57, 353, 374
 meditation, xi, xxii, xxvi, xxviii, 3, 7, 8, 17, 19, 26, 45, 47, 49, 56, 57, 63, 68, 117, 136, 169, 170, 175, 176, 181, 183, 225, 228, 259, 371, 383, 389, 393, 398, 399, 408, 409, 424, 428, 430
 Mele, 93, 119, 214, 220, 419, 422, 428
Mental Freedom, 401
 mereological, 100, 108, 290, 302, 325, 385
 metavalitional, 3, 44, 413, 421, 423
 methadone, 4, 6, 54, 55
 Meyers, 9, 16, 17, 38, 43, 57, 92, 103, 105, 113, 119, 194, 198, 199, 220, 377, 378, 393, 394, 399, 402, 406, 407, 409, 421, 423, 424, 428, 429
 Middle Way', 4, 113, 207
Milindapañha, 243, 341, 350, 396, 413, 423
 Moderately Futuristic Toaster, 81, 82
 moral responsibility, ix, xxvi, 9, 12, 13, 15, 17, 19, 20, 25, 61, 65, 66, 67, 70, 90, 96, 116, 147, 173, 194, 195, 196, 207, 209, 213, 237, 240, 243, 244, 245, 263, 273, 274, 282, 283, 288, 298, 303, 305, 311, 378, 391, 402, 412, 418, 419, 422, 425, 426, 427, 431
Mūlamadhyamakakārikā, 104, 106, 117, 250
Nāgārjuna, 102, 103, 106, 116, 117, 207, 251
 Ñāṇamoli, 21, 42, 103, 119, 206, 350, 352, 380, 382, 390, 393, 397, 399
nibbana, xxx

Nietzsche, 109, 110, 111, 119, 129, 130, 148, 149, 194, 221
 nihilism, 290, 302
 Nikāya, 21, 23, 27, 28, 42, 198, 205, 223, 239, 240, 278, 280, 325, 326, 350, 385, 396, 397, 400, 423, 424
nirvāṇa, xxv, xxx, 3, 17, 23, 47, 48, 52, 56, 157, 168, 247, 276, 330, 336, 348, 393, 414
 NOMA, 50, 51, 52
 non-dual, 230, 247
 non-self, 68, 157, 265, 267, 282, 330, 415, 431
 no-self, xxxi, 13, 15, 37, 38, 40, 193, 208, 210, 213, 275, 311, 325, 411
 omniscience, 87, 341, 342, 351
 paleocompatibilism, 14, 282, 288, 292, 303, 305, 306, 310, 400
paṭicca samuppāda, 27
 Pereboom, 70, 90, 92, 98, 120, 239, 259, 407, 408, 429
 Peter Strawson, ix, 7, 8, 64
 phenomenological, 5, 8, 49, 55, 171, 211, 408
 Plato, xxx, 118, 372, 374
pratīyasamutpāda, 27, 99, 101, 103, 111, 275
Pudgalavāda, xxxi, 217, 399
 Rāhula, 402, 421, 429
 Raju, 240, 259, 403, 429
 random, xxiv, 3, 46, 47, 97, 197, 225, 232, 406, 408
 reactive attitudes, 7, 70, 71, 77, 84, 151, 152, 154, 159, 161, 167, 195, 199, 204
 reason-responsiveness, 3, 18, 44, 418, 419, 421
 Reductionism, 117, 195, 205, 290, 291, 294, 298, 302, 303, 306, 307, 311, 312, 313, 314
 reincarnation, 51, 54, 55
 Repetti, x, 3, 4, 6, 9, 10, 11, 12, 17, 18, 19, 44, 56, 60, 67, 125, 149, 215, 221, 401, 402, 404, 405, 407, 408, 409, 415, 421, 422, 429, 430
 resentment, 7, 20, 67, 70, 71, 77, 78, 79, 82, 85, 151, 152, 196, 222, 288, 344, 364, 366
Samādhi, 377, 382
saṃsāra, 52, 56, 276, 277
Samyutta Nikāya, 240, 423
 Śāntideva, 10, 16, 71, 75, 76, 80, 83, 84, 85, 86, 88, 89, 91, 102, 109, 112, 117, 195, 207, 208, 212, 213, 218, 221, 248
 second-order, 57, 268, 269, 270, 272, 278, 409
 semi-compatibilism, 12, 17, 18, 402, 419, 422
 Shantideva, 228, 231, 238
 Siderits, 5, 10, 13, 14, 15, 51, 52, 55, 83, 84, 85, 90, 102, 106, 117, 120, 125, 149, 194, 200, 209, 215, 222, 282, 305, 306, 307, 309, 310, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 327, 328, 329, 391, 400, 401, 402, 423, 430
skillful means, 2, 4, 6, 79, 125, 234, 344, 345
 Socrates, 203, 372
 soft compatibilism, 11, 17, 402, 422
 soul, 9, 22, 40, 45, 50, 60, 99, 101, 128, 146, 171, 218, 227, 311, 377, 405, 406
 Story, 402, 421, 431
 Strawson, v, xi, xii, xx, 6, 7, 20, 46, 47, 67, 149, 151, 152, 153, 155, 171, 172, 173, 196, 204, 222, 313, 316, 317, 318, 322, 323, 324, 329, 408, 415, 431
 stream-entry, 386, 394
 substrate consciousness, 11, 247, 255
 suffering, xvi, xxiv, xxvi, 17, 21, 22, 23, 24, 26, 27, 28, 29, 31, 32, 33, 35, 37, 38, 39, 40, 44, 56, 57, 58, 59, 60, 83, 86, 87, 97, 102, 112, 130, 144, 168, 169, 194, 200, 213, 228, 246, 249, 250, 266, 274, 275, 276, 278, 279, 306, 311, 314, 325, 326, 327, 333, 348, 359, 362, 366, 372, 380, 381, 386, 392
tathagatagarbha, 230
 the Buddha, vii, viii, ix, xxv, xxvi, 9, 11, 13, 17, 21, 22, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42,

48, 51, 61, 63, 65, 87, 119, 133, 168,
 177, 197, 198, 205, 209, 213, 216, 221,
 223, 228, 229, 239, 240, 241, 242, 245,
 246, 247, 255, 256, 257, 262, 264, 265,
 266, 268, 273, 275, 279, 280, 305, 314,
 315, 325, 326, 327, 328, 329, 330, 331,
 332, 333, 334, 336, 340, 341, 342, 343,
 344, 345, 347, 349, 351, 352, 380, 381,
 387, 389, 391, 392, 397, 400, 401, 402,
 403, 404, 405, 406, 412, 417, 418, 423,
 424, 426, 427, 429
 the self, 2, 3, 6, 7, 39, 44, 48, 55, 58, 62, 63,
 84, 101, 107, 108, 112, 113, 146, 154,
 159, 163, 172, 189, 214, 225, 227, 233,
 235, 241, 250, 257, 264, 266, 269, 274,
 307, 317, 319, 320, 401, 402, 407, 417,
 426
 Theravāda, xxix, 15, 52, 53, 66, 117, 217,
 218, 246, 258, 262, 281, 330, 332, 333,
 335, 340, 348, 352, 369, 372, 385, 387,
 388, 393, 395, 427
 Thompson, 47, 66, 67, 410, 426
 tropes, 5, 14, 52, 53, 290, 307
 Tsongkhapa, 16, 258, 353, 354, 355, 356,
 357, 358, 359, 360, 361, 363, 364, 365,
 366, 367, 368, 369, 370, 371, 372, 373,
 374, 375
 Tuske, xi, 57, 68, 409, 431
Two Truths, 52, 105, 384
 ultimate, xxvi, 1, 4, 7, 13, 38, 40, 48, 52, 53,
 54, 56, 58, 71, 78, 79, 86, 106, 125, 128,
 129, 151, 188, 199, 207, 210, 230, 236,
 240, 241, 247, 256, 274, 275, 284, 285,
 287, 291, 296, 306, 307, 315, 324, 332,
 381, 394, 401, 413, 415, 423
 Vajrayana, 224, 230, 234
 Van Inwagen, 238, 260, 281, 431
Visuddhimagga, 24, 42, 243, 350, 352, 393,
 399, 403
 Vonnegut, 122, 123, 149
 Wallace, 10, 11, 17, 57, 125, 149, 197, 222,
 239, 247, 249, 254, 259, 260, 353, 372,
 376, 403, 404, 405, 409, 410, 418, 421,
 422
 Walshe, 23, 42, 198, 223, 350, 352, 380,
 400
 weakness of will, 93, 94, 410
 William James, 248
 Wittgenstein, 141, 150, 230, 238
 Yogācāra, xvii, xix, xxix