

# Early Buddhism: Basic Teachings

In this chapter we will explore the basic teachings of early Buddhism, the teachings of the Buddha and his immediate disciples. This will serve to introduce a set of core principles that all Buddhist philosophers accept. In later chapters we will examine how various Buddhist philosophers developed these core teachings in different ways. But before we get to those basic ideas that are common to all schools of Buddhism, it might be useful to say a few words about the life of the Buddha.

## 2.1

Apart from his career as a teacher, there is little that is known with much confidence about the details of Gautama's life. Until recently, scholars were fairly certain that he lived from 566 to 486 BCE. But recent research suggests that his death may have been as late as 404 BCE. So if we accept the traditional claim that he lived for 80 years, then perhaps his life was lived wholly within the fifth century BCE. He was born in the city-state of Kapilavastu, the home of the Śākya, <sup>1</sup> in what is now the western part of Nepal, near the Indian border. He grew up in relatively comfortable circumstances. But in early adulthood he chose to abandon the settled life of a householder and became a wandering renunciant or *śramaṇa*, someone whose life is dedicated to finding answers to certain spiritual questions.

The *śramaṇas* of sixth and fifth century BCE India represented a new phenomenon in Indian religious life. They rejected key elements of the prevailing Brahmanical orthodoxy as inadequate to their spiritual concerns. The Vedic religion that they challenged was centered on a set of texts, the *Vedas*, that the Brahmin priests considered supernatural in origin and authoritative. These texts enjoin performance of various rituals and sacrifices, both to uphold the cosmic order and to obtain various benefits for the person in whose name the ritual or sacrifice is carried out. But the new set of ideas associated with the notions of karma and rebirth made these older religious practices seem unsatisfying. If after I die I shall just be born into some new life, what point is there in trying to make my present situation more comfortable? Shouldn't I be more concerned with the lives to come after this one? Indeed what exactly is the point of going on to life after life? Is that cycle to go on forever? The Vedic religion seemed satisfactory as long as people held on to conventional views of human life and human happiness. If we each have just this one life on earth (and perhaps an afterlife thereafter), then it might make sense to devote it to things like

---

<sup>1</sup>Hence the epithet he later acquired, 'Śākyamuni' or 'sage of the Śākya'.

sensual pleasure, wealth and power, and the social standing of a virtuous person.<sup>2</sup> But with the advent of new ideas about the nature of human life, the old answers no longer seemed to work. And so the *śramaṇas* sought a new account of human happiness and how to attain it.

Among the many *śramaṇas*, there were some who claimed to have found a solution to the problem of human existence, and offered to teach it to others. Their answers differed, but most shared the idea that true happiness could only be found by overcoming our ignorance about our true nature. And most also agreed that the truly ideal state for us must involve liberation (*mokṣa*) from the cycle of rebirths. The *śramaṇas* also explored a wide variety of techniques for attaining this ideal state they sought. These included various ascetic practices – performing austerities such as fasting, remaining utterly motionless for long periods, abstaining from sleep, and the like. They also included various meditational or yogic practices: learning to calm the mind and focus it in one-pointed concentration, exploring a variety of altered states of consciousness, and the like.<sup>3</sup>

Like other new renunciants, after abandoning his life as a householder Gautama sought to find a suitable *śramaṇa* teacher. According to our oldest accounts, he studied with several, and mastered the theories and techniques they taught, but found these inadequate. He then struck out on his own. Coming across an isolated forest grove, he resolved to devote a full night of concentrated effort to solving the problem of human suffering. Employing a variety of yogic techniques, he entered into four successive stages of meditation, and thereby acquired three sorts of knowledge: recollection of his own past lives, understanding of the general laws of karma, and knowledge of what would come to be called the four noble truths. This knowledge signaled his enlightenment (*bodhi*), his attainment of nirvāna or liberation from rebirth. Having thus attained his goal, he considered whether or not to teach his discovery to others. At first he is said to have been deterred by the difficulty and subtlety of the truths he had discovered. But he eventually concluded that there were some who could grasp these truths and thereby profit from his discovery. So he

---

<sup>2</sup>While the *Vedas* did not teach rebirth, they were not entirely clear on the question of an afterlife. Brahmanical culture of the time also recognized three possible goals in life: sensual pleasure (*kāma*), material wealth and power (*artha*), and virtue and social repute (*dharma*). For each of these goals there was thought to be a special science concerning methods for obtaining it. And a literature developed around each of these sciences. So the *Kāma Sātra*, for instance, is the foundational text for the traditional science of obtaining sensual pleasure.

<sup>3</sup>While the *śramaṇa* movement may have started as a protest against Brahmanical orthodoxy, the Vedic tradition eventually responded to this challenge by developing a number of its own systems for attaining liberation or *mokṣa*. These included such philosophical schools as Sāṃkhya, Nyāya and Advaita Vedānta. These schools are referred to as ‘orthodox’ because they accept the authority of the *Vedas*. In this they differ from Buddhism and the other ‘heterodox’ schools (such as Jainism), which deny that the *Vedas* have any special authoritative status. Through the orthodox schools the Brahmanical tradition was in effect countenancing *mokṣa* as a fourth possible goal in life, in addition to the original three of *kāma*, *artha* and *dharma*.

embarked on the career of a Buddha, one who has solved the problem of human suffering through their own efforts (without reliance on the teachings of others) and imparts that knowledge to others out of compassion.

There is another, far more elaborate account of Gautama's life before his enlightenment. On that account, Gautama is a prince, his father, Śuddhodana, being a powerful and wealthy king. Gautama's conception is immaculate, and he is born not in the normal way but by emerging from his mother's side without breaking her skin or otherwise causing her pain. Immediately after birth he takes seven steps in each of the four cardinal directions; the world roars in response, and blossoms spring up under his feet. A seer tells Śuddhodana that the infant will grow up to be either a Buddha or a world monarch. He will become a Buddha if he sees four things in his youth: an old person, a sick person, a corpse, and a wandering renunciant. If he does not see all four he will become a world monarch. Śuddhodana wishes to ensure that his first-born son becomes a mighty king, so he has Gautama raised in a luxurious palace surrounded by only young, healthy and attractive people. Gautama grows up in these surroundings, marries and has a son. Yet on four successive days while out hunting he sees each of the four sights. He then resolves to become a *śramaṇa*, and makes his escape from the palace at night. He spends several years with a succession of teachers, but only after striking out on his own does he succeed in attaining the goal of liberation. Upon attaining enlightenment, it is Māra, the evil god of death, who tries to persuade him not to convey his discoveries to the world. Other gods then intercede to protect him from Māra's powers and ensure that there is a Buddha in the world.

This more elaborate account of Gautama's early life is the basis of popular depictions of the Buddha in Buddhist art and literature. But this version of the story only emerges several centuries after the Buddha's death. And it clearly reflects the common process whereby the life of a sect's founder comes to be draped in legend. We know, for instance, that Gautama cannot have been a prince nor his father a king, since Kapilavastu was not a monarchy in his day. Likewise the Buddha was quite insistent on the point that he was no more than an ordinary human being. This would seem to explain why the tales of miracles surrounding his birth and enlightenment are absent from the earliest accounts of Gautama's life. Only much later did some of his followers, perhaps out of missionary zeal, transform the story of his early life into a hagiography. Still there are things we can learn from these legendary accretions to his biography. Consider the tale of the four sights, for instance. Why might those who shaped the legend have chosen an old person, a sick person, a corpse, and a *śramaṇa* as the sights that would spur a pampered prince to renounce his life of luxury? Clearly because the first three signify the fact of human mortality, and the existential crisis that results from this fact, while the fourth represents the possibility of averting the crisis. This point will prove useful when we try to understand the Buddha's teachings on suffering.

## 2.2

While there is not much we know with certainty about Gautama's life before his enlightenment, we know a great deal about his career as a teacher after enlightenment. For instance, we know that he first taught his new insights when he encountered five former companion renunciants at Sārnāth, near Vārānasī.<sup>4</sup> We will examine the record of that encounter later, but it might be helpful to begin with an overview. It seems that these renunciants followed a path of extreme asceticism, but when Gautama left them and struck out on his own he abandoned such practices. So they now suspect him of having lapsed into a dissolute life. He thus begins by describing the path he has discovered as a 'middle path' between the two extremes of asceticism and the life of sensual pleasure. He then describes this path as a 'noble eightfold path', listing its eight component practices: right view, right intention, right speech, right action, right livelihood, right exertion, right self-possession and right concentration. This leads naturally to the enumeration of the four noble truths, since the claim that there is such a path is the fourth of the four truths. The four are, in summary form:

- 1 There is suffering.
- 2 There is the origination of suffering: suffering comes into existence in dependence on causes.
- 3 There is the cessation of suffering: all future suffering can be prevented.
- 4 There is a path to the cessation of suffering.

Now the second truth is later elaborated in terms of a twelve-linked chain of causes and effects, the first of which is ignorance. And the ignorance in question will be explained as failure to know three characteristics of reality: impermanence, suffering and non-self. It is thus significant that the Buddha goes on to teach the five renunciants the doctrine of non-self, and moreover that he argues for non-self on the grounds that all the constituents of the person are impermanent. Finally, according to the sūtra that recounts this first teaching, it ended with all five *śramaṇas* attaining enlightenment.

To summarize, in this early episode in the Buddha's teaching career we find reference to the following doctrines and ideas:

- the Dharma as a middle path,
- the eight-fold path,
- the four noble truths,
- the twelve-linked chain of dependent origination,
- the three characteristics of existence.

---

<sup>4</sup>The Buddha's teachings are referred to collectively as the Dharma. (This use of the word is often translated as 'law'; we will encounter other uses of the same Sanskrit term.) The Buddhist tradition refers to the encounter at Sārnāth as 'the first turning of the wheel of the Dharma'.

Let us now look at these in more detail. The doctrine of the four truths plays a central organizing function in the Buddha's teachings, so we should begin there. The first truth, that there is suffering, seems clear enough. And it would be hard to deny that it is true: there is all too much suffering in the world. But this raises the question why the Buddha should have thought it necessary to point it out. In fact, Buddhists claim this truth, properly understood, is among the hardest for most people to acknowledge. This is the first of the four truths because the Buddha thinks it is something about which ordinary people are all in denial. To see why, we need to understand just what is meant here by 'suffering'. And here is where the legend of the four sights becomes relevant. What it tells us is that by this term Buddhists do not mean ordinary pain, such as what we feel when we are injured or sick. Instead they mean existential suffering – the frustration, alienation and despair that result from the realization of our own mortality. Remember that according to the legend, Gautama would not have become a Buddha had he not encountered the facts of old age, disease, decay and death until late in his life. What is it about these facts that makes their recognition significant? Well, we each want our own lives to go well. We want to be happy. And when we want happiness, what we want requires a sense that our lives have meaning, value and purpose. Of course different individuals are made happy by different sorts of things. But when something makes someone happy, that's because they take it to say good things about who they are and where they are going. The difficulty is that once we are forced to acknowledge our own mortality, it becomes difficult to sustain the sense that events can have significance for my life. How can anything contribute to the meaning of my life when in the long run I shall be dead, with the world going merrily on its way without me? Now we all know at some level that some day we will die, yet we still seem to live our lives on the assumption that death can be indefinitely postponed. It is when events show this assumption to be false that existential suffering arises.

Here is one point at which you might think it makes a difference whether or not we accept the doctrine of karma and rebirth. Indeed you might think that the account of existential suffering that has just been given only makes sense if we deny this doctrine. And since the Buddhists accept the doctrine, you might suspect that they must mean something else by 'suffering' than existential suffering, the sense of alienation and despair that comes from recognizing the implications of our mortality. After all, if we live another life after we die, my death can't be the end of me. And if what I do in this life determines what sort of life I get next time around, wouldn't what happens to me now always have meaning for my future existence? So why would existential suffering arise for someone who accepted karma and rebirth? The Buddhist will reply, though, that these suspicions merely illustrate how difficult it can be to grasp the true nature of suffering. The tradition distinguishes among three different layers within the notion of suffering, each more subtle than its predecessor: suffering due to pain, suffering due to impermanence, and suffering due to conditions. It is the last of these that is meant to explain why the fact of rebirth itself constitutes a kind of existential suffering. But to see why Buddhists think this, we

need to say something about the first two ways in which they claim we experience suffering.

The first includes all those experiences that we would ordinarily classify as painful: being cut, burnt or struck, having a toothache or headache, losing a prized possession, not getting the job we'd set our hearts on, and the like. Note that even with such simple cases as a toothache there are actually two levels to the negative nature of the experience. First there is the feeling of pain itself, the immediate sensation of hurting. But there is also the worry that we commonly experience when we have something like a toothache: what does this painful feeling say about who I am and where I am going? Even when we don't put it to ourselves in so many words, this sense of 'dis-ease', of not being at home with ourselves, can permeate our lives when we have some nagging pain, undermining even our enjoyment of ordinary pleasures.<sup>5</sup>

The second form of suffering includes all negative experiences deriving from impermanence. This has much wider scope than one might suspect. As we will later see in more detail, Buddhists claim that everything that originates in dependence on causes must also cease to exist. And since all those things we ordinarily care about are dependent on causes, it follows that they are all impermanent. Now the pain of a toothache could be counted among the experiences that derive from impermanence. We get toothaches because healthy teeth are impermanent. But it is not just getting something we don't want, like a toothache, that is included here. Getting something we do want also comes under the category of suffering as impermanence. Of course it seems counter-intuitive to classify getting what you desire – a car, a job, a child, the esteem of people you care about, happiness for a friend – as a negative experience. But this is why Buddhists call this kind of suffering more subtle than the first. There is suffering in getting what one wants because the desired object is impermanent. So the happiness we feel is always tinged with anxiety about losing it. Indeed the feeling of happiness we derive from getting what we want is itself impermanent. When the novelty wears off, so does the feeling of happiness. Which is why we seem to always be in pursuit of something new. This explains the pattern we follow: always formulating some new goal, some new object of desire, when we get what we previously wanted (or give it up as unattainable). And when we begin to notice this pattern in our behavior, the happiness we feel on obtaining something new begins to drain away.

The last point leads naturally to the third level of suffering, suffering due to conditions. By 'conditions' here is meant the factors that are said to be responsible for rebirth (namely the intentions or volitions that motivate actions and cause karmic fruit). So suffering due to conditions refers to the suffering that results from rebirth. But to revert to the question we asked earlier, why should the mere fact of rebirth

---

<sup>5</sup>'Dis-ease' might be a better translation of the Sanskrit term we are discussing here, *duḥkha*, than is 'suffering'. This term is formed from the prefix *duḥ*, which is related to the English 'dis', plus the noun *kha*, which came to mean 'happiness' or 'ease'.

count as a form of suffering? Some specific rebirths might be quite unpleasant. But if we knew the karmic causal laws, we might be able to avoid those and obtain only rebirths in relatively fortunate circumstances. Why would that still count as suffering? The answer is encapsulated in the fact that re-birth also entails re-death. When we think that rebirth would help us avoid the suffering that is due to our own impermanence (that is, our mortality), we are forgetting that rebirth means re-encountering that very impermanence we wish to escape. Once we take this into account, the prospect no longer seems quite so inviting. Indeed the idea of perpetually going through this cycle – being born, living a life, losing that life and then starting anew – can only inspire a kind of cosmic ennui: what could possibly be the point? What we are now faced with is the requirement that there be an endless succession of future lives in order to sustain the sense that the life I am now living has a point. But if this life gets its point from the next, and that from its successor, and so on, will this really work? Perhaps the doctrine of karma and rebirth, instead of undercutting the claim that sentient beings are subject to suffering, actually reinforces the point.

It might be natural to wonder if the Buddha was not unduly pessimistic. Surely life is not all doom and gloom. And perhaps with a little luck and some good sense, one can live a life that is predominantly characterized by happiness. Of course the Buddhist will respond that this is just what *nirvāna* amounts to. But the opponent will say that seeking *nirvāna* seems a rather drastic step. For this requires abandoning much of what is usually thought to give life value: sensual pleasure, wealth and power, and virtue and repute. Surely at least some people can live lives that are happily devoted to such conventional ends as family, career and recreation. The Buddhist will respond that such pursuits can sometimes give pleasure and happiness. Buddhists do not deny that people sometimes experience pleasure and happiness. They claim, though, that pleasure and happiness are deceptive in nature: being in these states leads us to believe that they can be made to endure, when in fact, for the reasons sketched above, they cannot. And in the long run, they claim, those reasons dictate that the happiness one obtains from such pursuits will be outweighed by the suffering. The pursuit of happiness will become a kind of treadmill, and the sense that we are on this treadmill leads to alienation and despair. For anyone who is at all reflective about their life, it is inevitable that the happiness in their life will be outweighed by the suffering.

Here is one last question before we move on: might anti-depressants help? Modern medicine has created a class of drugs designed to help people who have lost all sense of enjoyment in their lives. And the more subtle sense of suffering that we have just been discussing sounds somewhat like this condition. Could a simple pill be an alternative to the arduous task of seeking enlightenment? Here is one possible way the Buddhist might respond to this question. First, they might claim that no pill can alter the facts. Taking a pill might alter how we assess those facts, but that is another matter entirely. For what the pill might actually do is foster an illusion, create the sense that we can continue to ignore those facts. Suppose that by taking an anti-depressant we could avoid the sense that the happiness-seeking project is an endless treadmill. We might then be looking at the same facts that led the Buddha to his

analysis of suffering, but we would be seeing those facts in a different light. The Buddhist would claim, though, that our assessment of the facts would be unrealistic. Taking the pill would simply re-install the illusion that conventional happiness is attainable in the long run. And this, they would hold, is no alternative to facing the facts squarely and taking the appropriate action: seeking nirvāna.<sup>6</sup> It is an interesting question whether the assumption they would then be making is true.<sup>7</sup>

### 2.3

While the first of the four noble truths points out the existence of suffering, the second is meant to explain how it originates. The underlying idea at work here is that by learning the cause of some phenomenon we may become able to exercise control over it. So the Buddha gives a detailed account of the factors he claims are the conditions in dependence on which suffering arises. This account, the twelve-linked chain of dependent origination, is traditionally understood as describing a sequence that takes place over three successive lives. In one life there occurs (1) ignorance (namely ignorance of the fact that all sentient existence is characterized by impermanence, suffering and non-self), and because of its occurrence there occur (2) volitions (*saṃskāra*), understood as the active forces in karma. It is in dependence on these volitions in the one life that there occurs (3) consciousness in the next life. That is, rebirth (in the form of the first moment of consciousness in a new life) occurs because of the desires that led to the performance of actions in the past life. On this consciousness in turn depends the occurrence of (4) a sentient body. That is, it is due to that first moment of rebirth consciousness that the organized matter of the fetal body comes to be a sentient being. On the existence of the sentient body in turn depend (5) the six organs of sense (the organs of the five external senses plus an 'inner sense' that is aware of inner states such as pain). On these depend (6) contact or sensory stimulation. And given sensory stimulation there arises (7) feeling, that is, the hedonic states of pleasure, pain and indifference. Feeling in turn causes (8) desire, and desire leads to (9) appropriation (*upādāna*), the attitude whereby one takes certain

---

<sup>6</sup>This is not to deny that anti-depressants can be genuinely helpful for those suffering from clinical depression. The Buddhist claims that the happiness-seeking project cannot be sustained in the long run. While this might seem like a depressing analysis, remember that they also claim there is a better alternative to that project, namely nirvāna. And they think we should make the effort to seek that better alternative. Someone who is clinically depressed might not be capable of making such an effort. Their sense of the futility of it all might render them unable to do anything to better their situation. A Buddhist might then say that anti-depressants would be useful in their case.

<sup>7</sup>Assume that by taking a pill one could permanently prevent the subtle sense of suffering from arising. Assume as well that the Buddha's analysis is correct, that the happiness-seeking project really is an endless treadmill. Would it actually be better to not take the pill, face up to the facts, and seek nirvāna? The Buddhist claims it would be, but why? What assumption would their answer seem to be based on? And is that assumption correct?



things as being 'me' or 'mine'. In dependence on appropriation there originates (10) becoming. This consists of the volitions that bring about the next rebirth, as well as the psychophysical elements making up the sentient body in that rebirth. In dependence on this there is (11) birth, that is, rebirth into the third life. And in dependence on birth there is (12) old age and death, here standing for all existential suffering.

There are obviously some difficulties in this list. For instance the tenth condition, becoming, seems to involve a repetition of the second, volition, and the fourth, sentient body. It also seems odd that birth into the third life should be listed as a separate condition, while birth into the second life is not. There is another version of the list that omits the six organs of sense, and instead has the sentient body serve as the condition for consciousness. Since consciousness has already been said to be the condition for sentient body, this has the effect of making consciousness both the cause and the effect of sentient body.<sup>8</sup> And there are versions of the list with only ten links, omitting the first two conditions altogether. These and other problems have led some scholars to suggest that our list of twelve results from the fusion of what were originally two or more separate lists.

But let us put such questions to one side, and look instead at the basic logic underlying the list that we have. The idea seems to be this. One is born into this life because in the last life one acted on the basis of volitions that were formed in ignorance of the facts about our existence. Having been born with a body, senses and mind, one comes in contact with sense objects, and this cognitive contact brings about feelings of pleasure, pain and indifference. These feelings trigger desires, and desires that are conditioned by ignorance lead to the stance known as appropriation: taking certain things (including things that no longer exist or do not yet exist) as 'me', and other things as 'mine' or my possessions. It is this stance that fuels rebirth, and this produces the suffering that is associated with all sentient existence.

How, one might wonder, could the first condition, ignorance, occur without there already being a sentient being (something that is not found until the fourth link in the series)? Doesn't ignorance require someone whose ignorance it is? When we wonder this, we are taking this list as an account of the very beginning of the series of lives. But the list should not be taken this way. What is here treated as the first life in a sequence of three is itself the effect of prior conditions that occurred in some yet earlier life.<sup>9</sup> So it is not saying that ignorance occurred before there were mind and

---

<sup>8</sup>It is this version of the list that will later lead some Abhidharma philosophers to hold that two simultaneously existing things can be both cause and effect of one another. This notion of reciprocal causation will become the center of some Abhidharma controversies.

<sup>9</sup>The Buddha says that we cannot discern the very first life in the series of lives we have lived. In the later tradition this is often taken to mean that the series of lives (and so our ignorance as well) is beginningless. But the Buddha's statement might be interpreted another way: while there might have been a very first life in the series, we could never tell which one that is. For it's always possible that although there were earlier lives, we simply can't remember any. Given this difficulty, it is pointless to speculate about whether there is or is not a first life in the series, and what might explain this. Suffering exists in the present life, and such speculation won't help solve that problem.

body. Ignorance comes first on the list because of its key role in producing suffering. In effect what we have in this theory is an account of how ignorance, by bringing about suffering, reinforces and thus perpetuates itself. When the chain of dependent origination is seen in this way, it is even possible to separate it from the doctrine of karma and rebirth. What it then amounts to is basically just the claim that the ignorance occurring at any one point in one's life causes one to act in certain ways that set the stage for both later suffering and continued ignorance.

The third truth, that there is the cessation of suffering, follows directly from the second truth. Ignorance is a remediable condition. Since it is possible to cure our ignorance, it is possible to put an end to the feedback loop that results in suffering. The fourth truth then spells out a set of eight practices that are designed to bring about this cure. They are: right view, right intention, right speech, right action, right livelihood, right exertion, right self-possession and right concentration. These eight are said to fall into three basic kinds: the first two represent wisdom, the next three are the factors of morality, and the final three are the practices that make up meditation. The factors are listed in a way that might suggest a sequential order: start with right view, follow the rules of right conduct, proceed to concentration, then attain nirvāna. But in actual practice the different factors are said to mutually reinforce one another, so that the mastery of each will involve contributions from the others. For instance, one might begin by acquiring a rudimentary grasp of the basic teachings of the Buddha (right view), on that basis form the (right) intention to seek nirvāna, and then set about trying to obey the moral rules set out for lay followers, such as not lying (right speech), not stealing (right action) and not working as a butcher (right livelihood). But when following these moral rules becomes habitual, this has the effect of clearing the mind of certain passions that can interfere with attaining wisdom. So this can lead to a deeper appreciation of the Buddha's teachings (right view), followed by the (right) intention to become a monk or nun. Entrance into the order of Buddhist monks and nuns (the *saṅgha*) brings with it a new set of moral virtues one must acquire. Practice in accordance with these virtues, along with the newly deepened understanding of the Dharma, helps one then begin to engage in meditation. But meditating also makes it easier to attain the required moral virtues. And meditation likewise produces insights into the nature of the mind that further strengthen one's appreciation of the Dharma. And so on.

For our purposes the reciprocal relation between wisdom and meditation is particularly significant. In the context of the Buddhist path, 'wisdom' means the practice of philosophy: analyzing concepts, investigating arguments, considering objections, and the like. So the content of this 'wisdom' is just the Buddhist philosophy that we are examining here. Now we already know that Buddhists claim ignorance is ultimately responsible for our suffering. And wisdom looks like the antidote to ignorance. So it makes sense that Buddhism should claim doing philosophy is necessary for attaining enlightenment. But will doing philosophy be sufficient? Buddhists generally say no. And it's not too difficult to guess why this is. For we also know something about what this ignorance supposedly consists in: the

failure to recognize the three characteristics, the facts of impermanence, suffering and non-self. This failure is exhibited in some fundamental assumptions we make about our lives: that we and the things we want can continue to exist indefinitely, that we can attain happiness by pursuing conventional goals, and that there is a true 'me' for whom this life can have meaning and value. Since almost everything we do is based on these assumptions, we are constantly in the business of reinforcing them. So even if our philosophical practice tells us they are false, it may not be so easy to uproot them. The situation here is like the case of a smoker. They may know perfectly well that smoking shortens their life. But each cigarette smoked reinforces their addiction, making it harder to act on that knowledge. So, the Buddhist says, meditation is needed in order to break the cycle and bring home the knowledge gained through philosophy.

To learn to meditate is to learn to control the mind. That control is then used to examine various mental processes, and to counteract those processes that perpetuate ignorance and suffering. So through meditation one can supposedly confirm that there is no self, by observing how impermanent mental states actually do all the work that we imagine could only be done by an enduring self. We can also see how certain mental states, such as anger and hatred, can reinforce belief in a self and thus perpetuate ignorance. And through meditation we can learn to counteract such states. In the case of anger and hatred, for instance, the adept is taught to cultivate feelings of kindness and sympathetic joy toward ever larger circles of beings, starting with friends and loved ones and eventually extending to those toward whom they feel anger and enmity. So meditation serves as a necessary supplement to philosophy in Buddhist practice. (This is why, even if the Buddhist philosophers are right about things, studying Buddhist philosophy wouldn't bring about liberation by itself.)

At the same time, doing philosophy is said to be necessary if the practice of meditation is to be effective. For one thing, many meditational attainments involve altered states of consciousness. What one is aware of in these states is very different from what goes on in our ordinary experience. This means that we need a conceptual framework to help us sort out our experiences in meditation and figure out their significance. Otherwise we would be confronted with what could only seem like a buzzing, whirring mass of confusion. Doing philosophy is said to help us acquire the conceptual tools we need to make sense of what we encounter in meditation. So, for instance, mastery of the philosophical arguments for the non-existence of a self will make it easier to appreciate the significance of the complex causal connections we find when we closely observe our mental processes. That there are these causal connections will then be seen to confirm that there is no self standing behind the scenes directing our mental lives. And this will bring home the truth of non-self as it applies to our own case. So while meditation is meant to help the practitioner apply the knowledge they acquire through philosophy, philosophy in turn plays an important role in facilitating meditational practice.

Just as there are interesting relationships among the components of the eight-fold path, so it is worthwhile to examine how the three characteristics are related to one

another. Suffering is caused, we are told, by ignorance of impermanence, suffering and non-self. And it is overcome by coming to know fully these three facts about the world. We now have some understanding of what Buddhists mean by the truth of suffering. Suppose they are right in their claims about what suffering is and why it is inevitable. They also claim that everything is impermanent, and that sentient beings are devoid of selves. Suppose these claims are also true. What might they have to do with the claim about suffering? It is tempting to think that impermanence is the chief factor here. On this interpretation, it is the fact that everything is impermanent that makes it true both that suffering is inevitable and also that there is no self. On this account, we wrongly believe that the things we desire are permanent, we become attached to them, and then we suffer when they reveal their impermanence by going out of existence. Likewise we base our lives on the assumption that we have permanent selves, and then suffer when our mortality shows this assumption to be false. The solution is then to learn to live with the fact of impermanence. Suffering will cease when we stop clinging to things and learn to live in the moment.

While this interpretation of the three characteristics is tempting, it is wrong. It is the truth of non-self that is said to be key to understanding suffering's genesis and dissolution. And the interpretation just offered does not take sufficiently seriously the fact of non-self. For what it assumes is that I do have a self, just a very impermanent one. This is the assumption behind the advice that we live our lives in the present moment. This advice would make sense only if there were a true 'me' that could derive value and significance from its experiences, but that existed only for a short while, to be replaced by a new self, someone who is not 'me' but someone else. We are advised to live in the present precisely because it is thought that when we plan for the future instead, we are letting the interests of that future self dictate what this present self does. Now while Buddhism is sometimes understood in this way, this is clearly incompatible with the claim that there is no self. Indeed this turns out to be one of the extreme views the Dharma is supposed to be a middle path between.<sup>10</sup> So this cannot be how to understand the three characteristics.

The doctrine of non-self is widely acknowledged to be the most difficult of all the basic teachings of Buddhism. We will examine it in detail in the next chapter. But we can now say this much about its relation to the other two of the three characteristics. Recall that by 'suffering' what Buddhists mean is existential suffering. And existential suffering arises from the assumption that there is a 'me' for whom events can have significance. Such suffering arises out of the suspicion that the kind of meaning we want is not to be had, that our best efforts at attaining happiness will inevitably be frustrated. And we experience suffering because this seems like such an affront to the dignity of the being we take ourselves to be. Now suppose it could be

---

<sup>10</sup>This is what is called 'annihilationism', the view that while I exist now, when the parts that presently make up me cease to exist, I will go utterly out of existence, typically to be replaced by someone else. The other extreme view is called 'eternalism'. It is the view that the true 'me' is eternal. The theory of dependent origination is said to constitute a middle path between these two extremes.

shown that while there are the experiences that make up a lifetime, those experiences have no owner. There is no 'me' whose experiences they are. In that case the conviction that my life should have uniquely special significance to me would turn out to be based on a mistake. For experiences in my life to have meaning, there must be more than just the experiences, there must be something separate from them for which they have good or bad meanings. Without belief in a separate self, existential suffering would no longer arise. Such suffering requires belief in something whose demand for meaning and significance is violated. It requires belief in a self. Impermanence also plays a role here. It is the fact of impermanence that first awakens us to suffering. And the fact that everything is impermanent will play a major role in the arguments for non-self. But it is non-self that plays the central role. And it is our false belief in a self that Buddhists identify as the core of our ignorance.

## 2.4

What might it be like to be enlightened? The Buddha claims that at the end of his path lies the cessation of suffering. And we've just had a glimpse of how following the path might bring that about. But even if we can make some sense of his path as a cure for suffering, this only tells us what being enlightened is not like. Being enlightened would mean being without existential suffering. Is there anything positive to be said about it? Is it pleasant? Is the enlightened person happy? Or is it just that because it's devoid of suffering, it's the best we can hope for? This would be a reasonable question to ask for someone considering whether or not to follow the Buddha's advice. The 'live for the moment' idea that was just rejected as an interpretation of the three characteristics did at least give an answer to this question. For then the enlightened person would appreciate their present experiences without any concern about what will come in the future. And perhaps this would enhance the enjoyment of any good experiences while diminishing the anxiety that normally accompanies bad experiences. So perhaps on that interpretation being enlightened would be pleasant. But since that is not what Buddhist enlightenment is like, this does not answer our question.

Here is another place where the doctrine of karma and rebirth has a role to play. To become enlightened is to enter into the state of *nirvāna*. The Sanskrit term *nirvāṇa* literally means 'extinction' or 'going out' (as when a fire is said to go out). What gets extinguished is, of course, suffering. But Buddhists sometimes equate this extinction with another sort, namely the end of the series of rebirths. What would that be like? Well, if there is no self, then to say I won't be reborn is to say I will cease to exist. Is this what *nirvāna* is, utter and complete annihilation? If so, then our question is answered in the negative: enlightenment would have no positive result, only the purely negative one of escape from all further suffering. And since this escape looks like a state of pure non-being, an utter blank, it also seems singularly unappealing.

That there is something wrong with this understanding of *nirvāna* is suggested by

the fact that one of the extreme views the Buddha rejects is called ‘annihilationism’. Moreover, when the Buddha is asked about the fate of the enlightened person after death, he says it would not be correct to say they are utterly non-existent. But the explication of these claims will have to wait until Chapter 4. What we can say at this point is that there is more to nirvāna than what happens after the death of an enlightened person. There is also the state of the person between the time of enlightenment and their death. In discussing the goal of their practice, Buddhists draw a distinction between ‘cessation with remainder’ and ‘cessation without remainder’. By ‘cessation’ is meant stopping the accumulation of new karma. And the ‘remainder’ is the residual karma that keeps the present life going. Once that residue is exhausted, this life ends. So they distinguish between nirvāna as the state of a living enlightened person, and nirvāna as the state of the enlightened person after death.<sup>11</sup> If we want to know if there is anything positive to the state of nirvāna, the place to look would seem to be this cessation with remainder.

Unfortunately, there isn’t much in the early Buddhist texts about this state. There is a great deal about how to attain cessation, but not much about what it is like to have attained it and remain alive. Artistic depictions of the Buddha and other enlightened persons often portray them with a serene half-smile on their faces, and this suggests that there is a kind of quiet happiness to the state. But this is not stated explicitly in our sources. Buddhists were not, though, the only Indian philosophers to teach the goal of liberation from rebirth. And among the non-Buddhists there is also a debate as to whether or not liberation is pleasant or joyful. Now this debate concerns post-mortem liberation. It is possible for these schools to have such a debate because they all affirm the existence of a self. So unlike the Buddhists, they all claim that the liberated person continues to exist when their last life is over. Some, though, claim that the self enjoys eternal bliss in this state of post-mortem liberation, while others deny this. Indeed some of the latter go so far as to say that the self feels nothing in this state, that its existence forever after is indistinguishable from that of a rock.

Now all the parties to this debate agree that liberation is the supreme goal for humans. They also agree that ignorance about who we truly are is what keeps us in the unliberated state – by making us pursue inappropriate goals like sensual pleasure, wealth and power, and virtue and repute. Since they all seem to mean more or less the same thing by liberation, this makes us wonder why some would deny that the supremely valuable end has any intrinsically desirable features. Why would they expect anyone to seek a state whose only attraction lies in the absence of pain and suffering? If that were all that was being offered, wouldn’t most people figure they could beat the odds and stick with the strategy of seeking ordinary happiness?

This is not a question that can be definitively answered by examining the texts of these orthodox schools. But a bit of speculation might throw some light on the situation here, and in so doing suggest an answer to our question about Buddhist

---

<sup>11</sup>This is sometimes referred to as *parinirvāṇa*, though strictly speaking that term only applies to the death of a Buddha.

nirvāna. Suppose that, as the *Bhagavad Gītā* says, ‘desire is here the enemy’ (III.37). That is, what keeps us bound to the wheel of *saṃsāra* (the state of perpetual rebirth and consequently suffering) is our desire for things like sensual pleasure, wealth and power, and virtue and repute. Desire for these things is thought problematic because it is based on the false assumption that I am something that could be made better off by having them. Further, suppose that were it not for such desires, and the ignorance about our identity that they both presuppose and reinforce, we would be in a state that is intrinsically valuable. Suppose, that is, that to be liberated from *saṃsāra* is to enjoy true happiness, perhaps even true bliss. There then arises what we could call the paradox of liberation. This paradox involves the following propositions, each of which seems true to the orthodox Indian philosophers:

- 1 Liberation is inherently desirable.
- 2 Selfish desires prevent us from attaining liberation.
- 3 In order to attain liberation one must train oneself to live without selfish desires.
- 4 One does not engage in deliberate action unless one desires the foreseen result of the action.

Taken together, propositions (3) and (4) tell us that no one will set about trying to attain liberation unless they desire it. And proposition (2) tells us that no one will attain liberation unless they seek it. Liberation isn’t something people just fall into through dumb luck: you have to make an effort to overcome ignorance, otherwise it will just perpetuate your bondage in *saṃsāra*. Putting these things together, we get the result that you have to desire liberation to obtain it. And (1) tells us that it’s reasonable to desire liberation. The trouble is, (2) also tells us that if we desire liberation we won’t get it. So although it’s reasonable to want liberation, it’s impossible to get it, so it isn’t reasonable to want it after all. This is a paradox.

There are different strategies we might use to try to resolve this paradox. We might deny (1), though that would then raise the question why anyone should be interested in attaining liberation. Or we might claim that the desire for liberation is not a selfish desire. But this seems implausible if (1) is true. If liberation is such a good thing, then surely my wanting to attain it would count as a selfish desire – a desire to benefit myself. Perhaps, though, not all such desires are selfish in the sense that’s relevant for (2) to apply. Remember that the trouble with desires is that they reinforce the wrong view about who we really are. What if liberation were joyful in a way that didn’t conflict with the facts about who we really are? The difficulty is that even if this were true, those of us who have not experienced this bliss would have trouble thinking of it in anything other than conventional terms. When told that liberation is a state of bliss, we would imagine it to be like sensual pleasure, or the thrill that can come from gaining great wealth and power. We would then end up desiring liberation in just the wrong way – the way that (2) says prevents our attaining it. But this suggests a possible strategy: deny (1) not because it is false but because it is misleading for those with conventional views about what is desirable. For such people what should be

emphasized is not what is positively good about liberation, but the point that to be liberated is to be forever free of pain and suffering. Then they might attain the bliss of liberation without having aimed at it. Their desire would just have been to rid themselves of pain and suffering.

There are situations where this sort of indirect strategy works. Consider the warm feeling we get when we act benevolently, doing something good for someone else. We get this feeling of gratification when our aim is to help others instead of ourselves. But suppose the only reason I ever helped others were because I wanted to have this warm feeling. Then I would never succeed. If my helping someone else were part of a calculated strategy whose ultimate purpose was to benefit myself, I wouldn't get the warm feeling at all. I can't get the feeling by aiming at it. I only get the feeling when I aim at something else – benefitting another person. Does this mean there is a paradox of benevolence? No, we can and do sometimes act benevolently, and thereby get the warm feeling. The best advice to give someone who wants to feel good in this way is that they should become genuinely concerned about the welfare of others. And this is something we can learn to do. We can get the warm feeling indirectly – not by aiming at it but by aiming at something else. There is no paradox of benevolence.

Could something like this be what's going on in the case of those orthodox Indian schools that denied liberation is pleasant or happy? Perhaps they are simply tailoring their advice to the understanding of their audience. Perhaps because their audience would misunderstand the happiness that comes with liberation, and then want it in a way that would prevent their ever getting it, these schools advise their audience to aim at something else, the cessation of suffering. And perhaps we should understand what early Buddhism says about *nirvāna* in a similar way. On this interpretation, the fact that *nirvāna* is depicted primarily negatively, as just the permanent cessation of suffering, and the fact that virtually nothing positive is ever said about cessation with remainder, represent strategic choices. They do not necessarily reflect the nature of *nirvāna*. Perhaps cessation with remainder is a state of true happiness, though this is importantly different from what is ordinarily taken for happiness.

Something like this interpretation may be necessary if the Buddha's path is going to make sense to those who don't accept the doctrine of karma and rebirth. If there is no rebirth, but the Buddha is right that there is no self, then after I die there won't be any suffering regardless of whether or not I attain enlightenment. So telling me that cessation without remainder is devoid of suffering won't motivate me to try to attain enlightenment. I'll only be motivated by facts about cessation with remainder, the state of being enlightened but still alive. And it isn't clear that being told this state is devoid of all existential suffering would be enough. If that were all I thought I'd get out of enlightenment, I might calculate the odds and decide that I'd do better to pursue conventional happiness. It might be that only a positive portrayal of enlightenment as true happiness would motivate me to seek it. And then there is the question whether my desire for enlightenment would get in the way of my ever attaining it. But this is a question to which we will have to return. For we have not yet



considered what it might be like to come to believe that we do not have selves. And coming to believe this is an important component of being enlightened. The Buddhist doctrine of non-self will be the subject of our next chapter. Then in Chapter 4 we will come back to this question of what it might be like to be enlightened.

### *Further Reading*

For more on the details of the Buddha's life and teaching career see Chapters 3 and 4 of A.K. Warder, *Indian Buddhism* (Delhi: Motilal Banarsidass, 1970).

The account of the Buddha's first expounding of his path (S IV.420–4) may be found at *The Connected Discourses of the Buddha*, trans. Bhikkhu Bodhi (Boston, MA: Wisdom Publications, 2000), pp. 1843–47.

For a more detailed account of the reciprocal relationships among the different parts of the eightfold path, see David Burton, *Buddhism, Knowledge and Liberation* (Aldershot, UK: Ashgate, 2004), pp. 62–75.

For a very different account of the nature of cessation with remainder see Paul Griffiths, *On Being Buddha* (Albany, NY: SUNY Press, 1994).

For a discussion of the debate among Hindu schools concerning whether liberation is desirable, see Arindam Chakrabarti, 'Is liberation (*mokṣa*) pleasant?'. *Philosophy East and West*, 33 (1983), pp. 167–82.

The alleged paradox of benevolence, and its resolution, were formulated by the eighteenth-century British philosopher and theologian Joseph Butler. For a discussion of Butler's work see Terence Penelhum, *Butler* (London: Routledge, 1986).

# Non-Self: Empty Persons

The Buddha holds that we experience the suffering of *samsāra* because of our ignorance of the three characteristics: impermanence, suffering and non-self. Of these three, it is the characteristic of non-self that plays the central role in his diagnosis. According to early Buddhism, there is no self, and persons are not ultimately real. This may be put somewhat cryptically as: we are empty persons, persons who are empty of selves. In this chapter we will investigate this claim. We will look at some of the arguments found in early Buddhist texts for the claim that there is no self. And we shall try to determine what it means to say that persons are not ultimately real. But before we can do either of these things we need to determine what it would mean to say that there is a self. The word 'self' gets used in several different ways, only one of which is relevant to the philosophical question the Buddha is trying to answer. We can avoid much confusion about what Buddhists mean by their doctrine of non-self if we begin by getting clear concerning what they mean when they speak of a self.

## 3.1

By 'the self' what Buddhists mean is the essence of a person – the one part whose continued existence is required for that person to continue to exist. This is the definition of 'self' that we will use. But what does it mean? It might be helpful to think of the view that there is a self as one possible answer to the question what it is that the word 'I' refers to. I am a person. And persons are made up of a variety of constituents: parts making up the body, such as limbs and organs, and parts making up the mind, such as feelings and desires. Now persons are things that continue to exist for some time – at least a lifetime, if not longer. But not all the parts of a person must continue to exist in order for that person to continue to exist. I could survive the loss of a finger or toe. And I might lose my desire for coffee without ceasing to exist. So apparently not all the parts of a person are necessary to the continued existence of a person. To say there is a self is to say that there is some one part that is necessary. This one part would then be what the word 'I' really named. The other parts would more properly be called 'mine'; only that one essential part would count as the true 'me'. The alternative to this would be to say that 'I' refers to all the parts collectively. Let us call this alternative the view that 'I' is the name of the person, where by 'person' we mean the whole that consists of all the parts that make up my body and mind over the duration of my existence. So either 'I' is the name of some one essential part of the person or else it refers to the person as a whole. (Of course this applies to the other words we use to refer to persons as well, such as names.)

To say there is a self is to say that there is some one part of the person that accounts for the identity of that person over time. If there were a self, then the person whose self it was would continue to exist as long as that self continued to exist. The self would then be the basis of a person's identity over time. It would be what explained why this present person, me, is the same person as some earlier person. But we need to be careful with the expression 'same person'. For the English word 'same' is ambiguous. When we say 'x and y are the same', there are two things we might mean. We could mean that x and y are qualitatively identical, or we could mean that x and y are numerically identical. To say that x and y are qualitatively identical is to say that they share the same qualities, that they resemble one another or are alike. To say that x and y are numerically identical is to say that they are one and the same thing, that 'x' and 'y' are really just two names for one entity. So there can be cases of qualitative identity but numerical distinctness, as with two t-shirts that come out of the factory looking exactly alike. And there can also be cases of numerical identity but qualitative distinctness, as with a leaf that in summer is green and smooth but in autumn is red and crinkled. We said above that according to the self-theorist, a self is what explains why some person existing now is the same person as someone who existed earlier. The key thing to keep in mind is that here 'same' is meant in the sense of numerical identity.<sup>1</sup>

Like many other things, persons can undergo very significant qualitative changes and yet continue to exist. I can continue to exist as one and the same person, me, even though the properties I now have are quite different from those I used to have. Thanks to the ambiguity of the English word 'same', we can put this as, 'He is the same person but not the same.' When we say this we are not contradicting ourselves. The first 'same' ('the same person') is used in the sense of numerical identity. The second 'same' is used in the sense of qualitative identity; 'not the same' means qualitatively distinct. It is one person, me, who once had the property of liking coffee, but now has the very different property of disliking coffee. A person can undergo qualitative change while retaining numerical identity. Since the self is supposed to be what explains numerical identity over time of persons, perhaps a self could undergo qualitative change. What it could not undergo is numerical change, that is, going out of existence and being replaced by another self.

If there is a self, it is 'what makes me me', 'the true me', that which 'gives me my identity'. These ways of describing what a self is are all open to a common misinterpretation. People often speak of 'discovering their self', of 'finding their true identity'. What they often mean by this is figuring out which characteristics seem

---

<sup>1</sup>The ambiguity of 'same' is often resolved by context. When we say 'x and y are the same P', what is meant is numerical identity. When we say 'x and y are the same', what is often meant is qualitative identity. So I might say that this is *the same leaf* as the one I showed you yesterday, meaning that they are one and the same leaf. Or I might say that this leaf is *the same* as the one that was on this branch last year, meaning that the two leaves are qualitatively identical. Other languages lack this ambiguity. In German, for instance, one says *das selbe* for numerical identity, and *das gleiche* for qualitative identity.

most important or valuable. So someone might say that they have come to realize their identity isn't tied up with physical appearance but with less superficial things like artistic talent or communication skills. Discoveries like this are probably important to personal growth. But they have nothing to do with what the Buddhists mean by a self. We can see this from the fact that even if there is no self, we can still ask which of a person's characteristics are most important to that person's happiness. To speak of a self is to speak of some one part of the person, the part that must always exist as long as the person exists. To speak of an 'identity' that can be 'found' is to speak of characteristics or properties, of what a person is like. There might very well be no single part of the person that must continue to exist in order for that person to continue to exist. (This is exactly what the Buddha is going to argue for.) But it might still be true that some characteristics of a person play a more important role in their life than others. Otherwise it wouldn't make sense to say that a person has 'lost their identity'. Perhaps my life would be less meaningful if I were to lose those traits that now have great importance to me. But it would still be my life. I could survive that qualitative change. I might be a very different kind of person. But I would still be me.

There is another misinterpretation that arises in connection with the idea that the self is what gives me my 'identity'. It is common to think that someone's identity is what sets that person apart from all others. Add to this the idea that one's identity consists in what one is like, one's characteristics or properties. The result is the notion that a self would be what makes one different from everyone else. Now the word 'different' is ambiguous in the same way that 'same' is: there is numerical difference or distinctness, and there is qualitative difference. If it's numerical distinctness that is meant, then it's true that the self would be what makes one different from others. If we have selves, then my self and yours must be two distinct things, not one. But it's not true if what's meant by 'different' is qualitative difference. It is not true that if we had selves, each would have to be unique in the sense of being unlike every other. Two selves could be perfectly alike, like two peas in a pod, and still serve to make one person numerically distinct from another.

The difficulty with the idea that the self must be qualitatively unique is that it once again confuses the notion of the self with the notion of what one is like, one's properties or characteristics. And properties may be shared between two things, whereas numerical identity may not. The leaf on this branch of this tree today might be exactly like the leaf that was here last year – same color, same shape, same pattern of veins, etc. But they are numerically distinct leaves all the same. Perhaps no two persons are ever exactly qualitatively alike. Even twins who share DNA patterns have physical differences, such as different fingerprints. Still there is no contradiction involved in supposing that there might be two persons who are exactly qualitatively alike. Imagine for instance that each of us has lived countless lives in the past. Given the innumerable many beings there may be in the universe, it does not seem unlikely that someone somewhere might once have lived a life just like the one I am now living. Yet that would have been someone else, not me. So if what makes me the

person I am is my self, then my self is not what makes me qualitatively unlike other people.

Suppose, moreover, that each person is qualitatively unlike every other. This could be true even if there were no selves. Indeed it could be true if there were selves that were all qualitatively identical. This is actually something that many non-Buddhist Indian philosophers hold. On their view, the self is something that is simple or impartite (lacking parts). The self is just the subject of experiences, the part of us that is aware of the different experiences we have. Your self and mine would then be just like those two peas in a pod. It's common to suppose that what makes different people qualitatively different is that they have different experiences. But on this view of the self, the different experiences that people have would not make their selves qualitatively different. Since the self is simple, it cannot be changed by the experiences it is aware of. It is other parts of the person that are changed by those experiences. The experience of eating changes the shape of my body. The experience of smelling coffee changes a desire in my mind. My self is unaffected by these changes, it is simply aware of them. Someone holding this view of the self who also thought that persons are qualitatively unique could say that their uniqueness is explained by facts about those parts of the person that are not the self. Someone who denied the existence of a self could explain the qualitative uniqueness of persons in the same way.

### 3.2

In order to show that the self does not exist, we need to know what we are looking for, and where to look. We now know that a self would be that part of the person that 'I' is consistently used to refer to. So we can tell what to look for by seeing how we actually use words like 'I'. For instance, we say things like 'I was born in New York, now live in the Midwest, and will move to Arizona when I retire.' So if 'I' refers to the self, the self would have to be some one numerically identical thing that continues to exist throughout the past, present and future history of the person. There are more clues to be found in the ways we use this word, but this tells us enough for present purposes. Where should we look? Since the self is supposed to be a part of the person, we obviously need to look among the parts that make up persons. It would be helpful if we had a list of the basic categories of person-parts. This is just what the Buddha provides with his doctrine of the five *skandhas*. (The word *skandha* is here used in its sense of 'bundle'.) These are:

- *Rūpa*: anything corporeal or physical;<sup>2</sup>

---

<sup>2</sup>The literal meaning of *rūpa* is 'form' or 'shape', and you will sometimes see the word rendered as 'form' in translations of Buddhist texts. But as the name of the first *skandha*, *rūpa* actually means 'that which has form or shape', that is, anything material or physical. This is one case where it's best to stick with the Sanskrit original rather than try to come up with an acceptable English translation.

- Feeling: sensations of pleasure, pain and indifference;
- Perception: those mental events whereby one grasps the sensible characteristics of a perceptible object; e.g., the seeing of a patch of blue color, the hearing of the sound of thunder;
- Volition: the mental forces responsible for bodily and mental activity, for example, hunger, attentiveness, and
- Consciousness: the awareness of physical and mental states.

A word of caution is necessary concerning these categories. Their names are here being used as technical terms, with precise definitions. Do not confuse these with the ordinary meanings of these words. For instance, the second *skandha*, feeling, refers only to the three kinds of hedonic sensation: pleasure, pain and indifference (neither pleasure nor pain). It does not include most of the things that are often called ‘feelings’, such as the emotions of anger and jealousy. Those emotions go under the very different *skandha* of volition. Likewise by ‘consciousness’ is here meant just the awareness itself, and not what it is that one is aware of. So when I am conscious of a pain sensation, there are two *skandhas* involved: the pain, which goes under feeling *skandha*, and the consciousness that is aware of it, which goes under consciousness *skandha*. Again, we sometimes use the word ‘perception’ to refer to our beliefs about and attitude toward something. So someone might say, ‘My perception of the new government is that it is weak and will soon fall.’ This is not the sort of thing that would go under perception *skandha*. This is a complex mental state, whereas an instance of perception *skandha* is a simple mental event. A perception in this technical sense is just the occurrence of a sensory content to the mind: the simple thought of a patch of blue or the smell of lemon.

The five *skandhas* are sometimes referred to collectively as *nāma-rūpa* (sometimes translated as ‘name and form’). Here *nāma* refers to the four *skandhas* other than *rūpa*. The literal meaning of *nāma* is ‘name’, but here it means ‘that which can only be named’. The idea is that while *rūpa* can be perceived by the external senses, the members of the four other categories cannot be seen or touched. Because they are not publicly observable, we cannot explain what they are by pointing; we can only communicate about them through the names we have learned to use for these private states. What this tells us is that the doctrine of the five *skandhas* expresses a kind of mind-body dualism. The Buddha is claiming that in addition to those parts of the person that we can see and touch – the parts of the body – there are other constituents that are not themselves physical. Some philosophers today hold the view called ‘physicalism’, according to which all that exists is physical. On this view there is no more to a person than the physical constituents, their body and brain. What we think of as mental events, such as thoughts and emotions, are really just complex brain events. When the Buddha says that in addition to *rūpa skandha* there are the four *nāma skandhas*, he is in effect denying that physicalism is true. On his account, mental events are separate non-physical kinds of things. We will be looking at this claim more carefully later on.

The Buddha uses the doctrine of the five *skandhas* as a tool in his search for a self. He goes through each *skandha* in turn and tries to show that nothing included in that category could count as a self. But this raises a new question: would this really show that there is no self? Isn't it possible that the self exists elsewhere than among the five *skandhas*? In order for the Buddha's strategy to work, he will have to show that the doctrine of the five *skandhas* gives an exhaustive analysis of the parts of the person. We will call this the 'exhaustiveness claim'.

The exhaustiveness claim is this: every constituent of persons is included in one or more of the five *skandhas*.

In the following passage, the later commentator Buddhaghosa argues in support of this claim.

The basis for the figment of a self or of anything related to a self, is afforded only by these, namely *rūpa* and the rest. For it has been said as follows:

When there is *rūpa*, O monks, then through attachment to *rūpa*, through engrossment in *rūpa*, the persuasion arises, 'This is mine; this am I; this is my self.'

When there is feeling ... when there is perception ... when there are volitions ... when there is consciousness, O monks, then through attachment to consciousness, through engrossment in consciousness, the persuasion arises, 'This is mine; this am I; this is my self.'

Accordingly he laid down only five *skandhas*, because it is only these that can afford a basis for the figment of a self or of anything related to a self.

As to other groups which he lays down, such as the five of conduct and the rest, these are included, for they are comprised in volition *skandha*. Accordingly he laid down only five *skandhas*, because these include all other classifications. After this manner, therefore, is the conclusion reached that there are no less and no more. [VM xiv.218]

This at least makes clear that Buddhists recognize the need to support the exhaustiveness claim. But it is not clear how good an argument this is. The idea seems to be that these are the only things we are aware of when we are aware of persons and so come to believe that persons have selves. Is this true? And if it were true, would it show that the exhaustiveness claim is true? We will return to this question.

### 3.3

Let us now look at how the Buddha formulates his arguments for non-self. In the following passage the Buddha is addressing his five former companion *śramanas*, in the episode we discussed in Chapter 2. It contains two distinct arguments. The first is

what we will call the argument from impermanence, since it is based on the claim that all five *skandhas* are impermanent or transitory. But there is also a second argument here.

Then The Blessed One addressed the band of five *śramanas*:

‘*Rūpa*, O monks, is not a self. For if now, O monks, this *rūpa* were a self, then this *rūpa* would not tend towards destruction, and it would be possible to say of *rūpa*, “Let my *rūpa* be this way; let not my *rūpa* be that way!” But inasmuch, O monks, as *rūpa* is not a self, therefore does *rūpa* tend towards destruction, and it is not possible to say of *rūpa*, “Let my *rūpa* be this way; let not my *rūpa* be that way!”

‘Feeling ... perception ... volitions ... consciousness, is not a self. For if now, O monks, this consciousness were a self, then would not this consciousness tend towards destruction, and it would be possible to say of consciousness, “Let my consciousness be this way; let not my consciousness be that way!” But inasmuch, O monks, as consciousness is not a self, therefore does consciousness tend towards destruction, and it is not possible to say of consciousness, “Let my consciousness be this way; let not my consciousness be that way!”

‘What do you think, O monks? Is *rūpa* permanent, or transitory?’

‘It is transitory, Reverend Sir.’

‘And that which is transitory – is it painful, or is it pleasant?’

‘It is painful, Reverend Sir.’

‘And that which is transitory, painful, and liable to change – is it possible to say of it: “This is mine; this am I; this is my self?”’

‘Certainly not, Reverend Sir.’

‘Is feeling ... perception ... volition ... consciousness, permanent, or transitory?’

‘It is transitory, Reverend Sir.’

‘And that which is transitory – is it painful, or is it pleasant?’

‘It is painful, Reverend Sir.’

‘And that which is transitory, painful, and liable to change – is it possible to say of it: “This is mine; this am I; this is my self?”’

‘Certainly not, Reverend Sir.’

‘Accordingly, O monks, as respects all *rūpa* whatsoever, past, future, or present, be it subjective or existing outside, gross or subtle, mean or exalted, far or near, the correct view in the light of the highest knowledge is as follows: “This is not mine; this am I not; this is not my self.”

‘As respects all feeling whatsoever ... as respects all perception whatsoever ... as respects all volitions whatsoever ... as respects all consciousness whatsoever, past, future, or present, be it subjective or existing outside, gross or subtle, mean or exalted, far or near, the correct view in the light of the highest knowledge is as follows: “This is not mine; this am I not; this is not my self.”

‘Perceiving this, O monks, the learned and noble disciple conceives an aversion for *rūpa*, conceives an aversion for feeling, conceives an aversion for perception, conceives an aversion for volitions, conceives an aversion for consciousness. And in conceiving this aversion he becomes divested of passion, and by the absence of passion he becomes free, and when he is free he becomes aware



that he is free; and he knows that rebirth is exhausted, that he has lived the holy life, that he has done what it behooved him to do, and that he is no more for this world.’

Thus spoke The Blessed One, and the delighted band of five *śramanas* applauded the speech of The Blessed One. Now while this exposition was being delivered, the minds of the five *śramanas* became free from attachment and delivered from the depravities. [S III.66–68]

Here the Buddha cites two different sorts of reasons why the *skandhas* are not the self: they are impermanent (‘subject to destruction’, ‘transitory’), and they are not under one’s control (‘painful’, ‘it is not possible to say of *x*, “Let my *x* be this way...”’). To separate out the argument from impermanence from the second argument, let’s ignore the claims about the five *skandhas* not being under one’s control (we’ll discuss this in §4), and focus on the claims about their being subject to destruction and transitory. If we add the exhaustiveness claim as an implicit premise,<sup>3</sup> the argument is then:

- 1 *Rūpa* is impermanent.
  - 2 Sensation is impermanent.
  - 3 Perception is impermanent.
  - 4 Volition is impermanent.
  - 5 Consciousness is impermanent.
  - 6 If there were a self it would be permanent.
- IP [There is no more to the person than the five *skandhas*.]  
 C Therefore there is no self.

This argument is valid or logically good. That is, if the premises are all true, then the conclusion will also be true. So our job now will be to determine if the premises really are all true. But before we can do that, there is one major point that needs clarifying: just what do ‘permanent’ and ‘impermanent’ mean here? Once again the doctrine of karma and rebirth becomes relevant. For those like the Buddha and his audience who accepted the doctrine, ‘permanent’ would mean eternal, and ‘impermanent’ would mean anything less than eternal. This is because if we believe it is the self that undergoes rebirth, and we also believe that liberation from rebirth is possible, then we will hold as well that the self is something that continues to exist over many lives, and can even exist independently of any form of corporeal life. This is probably what the Buddha had in mind with premise (6). And in that case, all that would be needed to show that something is not a self is to establish that it does not last forever – even if it did last a long time. So if, for instance, the *rūpa* that is my body does not last forever,

---

<sup>3</sup>An implicit premise is an unstated premise that must be supplied for an argument to work, and that the author of the argument did not state because they thought it would be redundant – typically because it seemed to the author to be common knowledge that the author and the audience shared. We will follow the practice of putting implicit premises in square brackets.

then it is not my self. And of course my body does go out of existence when I die, so this would be sufficient to show that it is not my self.

What about those of us who do not accept the doctrine of karma and rebirth? To believe in rebirth is to believe that the person exists both before and after this life. If we do not believe in rebirth, then we may believe that the person exists only a single lifetime. In that case, a self would not have to exist any longer than a lifetime in order to serve as the basis of a person's numerical identity over time. So all that 'permanent' in premise (6) could mean is 'existing at least a whole lifetime'. It could not mean 'eternal'. Likewise, to show that a *skandha* is impermanent in the relevant sense, we would have to show that it does not exist for the entire duration of a person's life. Does this mean that the argument won't work without the assumption of karma and rebirth? After all, isn't it true that our bodies last for our entire lives?

Not necessarily. First, we need to remember that the self is supposed to be the essential part of the person, and the body is a whole made of parts. Which of these parts – the organs that make up the body – is the essential one? There doesn't seem to be any single organ that I could not live without. Granted I could not survive without a heart. But as heart replacement surgery shows, I don't need *this* heart in order to continue to exist. If my heart were my self, then when I got a replacement heart I would cease to exist and someone else would then be living in my body. That replacement heart came from someone else, so it would be that person's self. But surely if I chose to have heart replacement surgery I would not be committing suicide! What about the brain? Not only can I not live without a brain; there is no such thing as brain replacement surgery, so I cannot live without this brain. But here the problem seems to be entirely practical, not an 'in-principle' difficulty. If we knew how to reprogram an entire brain, then we might be able to replace a diseased brain with a healthy one while preserving all of a person's psychology. This would be like copying the contents of the failing hard drive of your computer, replacing the hard drive, then reinstalling everything onto the new hard drive.

This brain-replacement scenario might seem too science-fictional to support premiss (1). But there's a second reason someone might give for denying that the body is permanent in the relevant way. This is that all the parts of the body are constantly being replaced – at the level of the molecules that make up our cells. We've all heard it said that none of the atoms that made up our body seven years ago is among those making up our body now. Life processes such as metabolism and meiosis involve the constant, piecemeal replacement of the parts that make up a life-form. After these processes have gone on long enough, all the matter making up a given organ is new: the atoms now making up that organ are numerically distinct from the atoms that made it up earlier. Given this, it could be said that the body and brain I have now are not numerically identical with the body and brain I had seven years ago. *Rūpa* would then be impermanent in the relevant sense.

We have been discussing how to interpret premise (6), the premise that a self would have to be permanent, and how premise (1), which says that *rūpa* is impermanent, might be true in light of our interpretation of (6). Our general practice

in examining arguments will be to first look at what reason there might be to think that the premises are true, and then to evaluate the argument overall. How might someone defend the remaining premises, (2)–(5)? These are not affected by the question of karma and rebirth in the way that premise (1) is. For regardless of whether we interpret ‘permanent’ to mean eternal, or just to mean lasting a single lifetime, the four *nāma skandhas* will all count as impermanent. This is the point the Buddha makes in the following passage:

It would be better, O monks, if the uninstructed worldling regarded the body which is composed of the four elements as a self, rather than the mind. And why do I say so? Because it is evident, O monks, that this body which is composed of the four elements lasts one year, lasts two years, lasts three years, lasts four years, lasts five years, lasts ten years, lasts twenty years, lasts thirty years, lasts forty years, lasts fifty years, lasts a hundred years, and even more. But that, O monks, which is called mind, intellect, consciousness, keeps up an incessant round by day and by night of perishing as one thing and springing up as another.

Here the learned and noble disciple, O monks, attentively considers dependent origination: ‘this exists when that exists, this originates from the origination of that; this does not exist when that does not exist, this ceases from the cessation of that’. O monks, a pleasant feeling originates in dependence on contact with pleasant objects; but when that contact with pleasant objects ceases, the feeling sprung from that contact, the pleasant feeling that originated in dependence on contact with pleasant objects ceases and comes to an end. O monks, an unpleasant feeling ... an indifferent feeling originates in dependence on contact with indifferent objects; but when that contact with indifferent objects ceases, the feeling sprung from that contact, the indifferent feeling that originated in dependence on contact with indifferent objects ceases and comes to an end.

Just as, O monks, heat comes into existence and flame into being from the friction and concussion of two sticks of wood, but on the separation and parting of these two sticks of wood the heat sprung from those two sticks of wood ceases and comes to an end; in exactly the same way, O monks, a pleasant feeling originates in dependence on contact with pleasant objects; but when that contact with pleasant objects ceases, the feeling sprung from that contact, the pleasant feeling that originated in dependence on contact with pleasant objects, ceases and comes to an end. An unpleasant feeling ... an indifferent feeling originates in dependence on contact with indifferent objects; but when that contact with indifferent objects ceases, the feeling sprung from that contact, the indifferent feeling that originated in dependence on contact with indifferent objects ceases and comes to an end. [S II.96f]

Of course the Buddha knows that reflective people are more likely to consider the mind the self than the body. In the Western philosophical tradition this is just what Descartes did. He concluded that the true ‘I’ is not the body but the mind – a substance that thinks (that is, is conscious), endures at least a lifetime, and is immaterial in nature. Many Indian philosophers reached somewhat similar conclusions. The Buddha’s point is that the conclusion that the mind endures at least

a lifetime rests on an illusion. For what we call the mind is really a continuous series of distinct events, each lasting just a moment, but each immediately followed by others. There is no such thing as the mind that has these different events, there are just the events themselves. But because they succeed one another in unbroken succession, the illusion is created of an enduring thing in which they are all taking place.

The eighteenth-century British philosopher David Hume said something similar in response to Descartes. Descartes claimed to be aware of the mind as something that is aware, that cognizes, perceives, wills, believes, doubts – that is the subject of all one's mental activities. Hume responded that when he looked within, all he ever found were particular mental contents, each of them fleeting, and never an enduring substance that has them. He concluded that it is just the relations among those mental events that make us invent the fiction of the self as an enduring subject of experience. The Buddha claims something similar. And like Hume, he uses the relation of causation to support his claim.

In the last chapter we saw how the doctrine of dependent origination is used to explain the origin of suffering. In the passage we are looking at, that doctrine gets put to a different use. Dependent origination is the relation between an effect and its causes and conditions. Where this relation holds, the effect will arise when the causes and conditions obtain, and the effect will not occur when the causes and conditions do not. The Buddha asserts that all the *nāma skandhas* are dependently originated. He uses the example of feeling, but this example generalizes to the other kinds of mental events as well. Consider the feeling of pleasure I derive from eating my favorite kind of ice cream. This feeling originates in dependence on contact between my sense of taste (located in the taste buds on my tongue) and the ice cream. Before that contact there was no feeling of pleasure, and when the contact ceases so does the feeling. I may have a feeling of pleasure in the next moment, but that occurs in dependence on a new event of sense-object contact – say, when I take my next bite of ice cream. So that feeling is numerically distinct from the first, for it has a different cause. One feeling has gone out of existence and been replaced by another. Now the senses are by nature restless, always making contact with new objects. This means that there will be an unbroken stream of feelings and other mental events. It is easy to mistake this stream for a single enduring thing. But the Buddha claims that if we attend to the individual events making up this stream, then seeing how they are dependently originated will help us overcome the illusion of a persisting subject of experience.

The appeal to dependent origination is meant to show two things: that there is no such thing as the mind over and above the mental events making up the mental stream; and that each of those events is very short-lived. Suppose we agreed with the Buddha on the first point. How successful is this appeal with regard to the second point? It is relatively easy to agree that feelings of pleasure and pain are transitory. We don't really need to use dependent origination to prove this. And since they are transitory, they could not be the self. Likewise for perceptions. But what about volitions? Granted my desire for some new soft drink may last only as long as the effects of the commercial I just saw. But we also seem to have volitions that endure,

such as my desire for coffee. To this it could be replied that this is an acquired volition, one that I did not always have and might very well get rid of. So the opponent must look for volitions that seem to endure a whole lifetime. They might suggest what are sometimes called ‘instinctual desires’, such as the desire to escape life-threatening situations. Might this not be a volition that is permanent in the relevant sense? The Buddha will reply that what we are then describing is not one enduring volition, but rather a pattern of recurring volitions, each lasting only a brief while before ceasing. This is shown by the fact that I am only aware of a desire to escape danger when I perceive a threatening situation. The desire thus originates in dependence on a specific sense-object contact event, and ceases to exist when that event ceases. The opponent will then want to know what explains the pattern of recurring volitions. What the opponent suspects is that this pattern can only be explained by supposing that there is one enduring volition, a permanent desire to escape life-threatening situations, that is always present in me. My perception of a life-threatening situation brings the volition out into the part of my mind that is illuminated by consciousness, but it persists even when I am not aware of it.

Since we have no evidence that the Buddha was ever presented with this line of objection, we don’t know how he would have responded. But later Buddhist philosophers do show us how it might be answered. What we have here is a certain phenomenon – a pattern of recurring desires over the course of a person’s lifetime – and two competing theories as to how to explain the phenomenon. Call the opponent’s theory the ‘in-the-closet’ theory, since it claims that some desires continue to exist hidden away in a dark corner of the mind when not observed. It explains the phenomenon by claiming that it is a single continuously existing volition that manifests itself at different times as the desire to duck a falling safe, the desire to dodge a runaway car, etc. The Buddhist dependent origination theory, by contrast, claims that these are many numerically distinct desires. It explains the pattern by appealing to the ways in which the parts of a person’s body are arranged. Consider the thermostat that controls the heat in a house. It is because of the way in which the parts of the thermostat are put together that whenever the temperature goes below a certain threshold, the thermostat signals the furnace to go on. It is not as if the signal for the furnace to go on waits in the thermostat’s closet until the room gets too cold. By the same token, the Buddhist would say, it is because of the way that the human body is organized that a danger stimulus causes a danger-escaping volition.<sup>4</sup> Now this seems like a plausible explanation. It makes sense to suppose that, for instance, it is because of the way in which certain neurons in the brain are arranged that we have this desire to escape whenever we sense danger. But the in-the-closet theory also seems plausible to many people, so which should we choose?

---

<sup>4</sup>No Buddhist text actually says this. This represents an extrapolation from what members of the Sautrāntika school of Abhidharma say about continuity of karmic seeds during meditational states in which there is no consciousness. Their approach to that problem is dictated by their overall aversion to talk of dispositions or powers as realthings.

There is a principle that governs cases like this. It is known in the West as Ockham's Razor, but Indian philosophers call it the Principle of Lightness, for it dictates that we choose the 'lighter' of two competing theories. The Principle of Lightness may be stated as follows: given two competing theories, each of which is equally good at explaining and predicting the relevant phenomena, choose the lighter theory, that is, the theory that posits the least number of unobservable entities.

To posit an unobservable entity is to say that something exists even though we never directly observe that thing. Now you might think that positing an unobservable entity is always a bad idea. Why believe something exists when no one can see or feel it? But modern physics tells us that there are subatomic particles like electrons and protons, and no one has ever seen or felt such things. Does that make modern physics an irrational theory? No. What the Principle of Lightness tells us is that we should only posit unobservable entities when we have to, when there is no other way to explain what we observe. We accept the theory that says there are subatomic particles because no other theory does as good a job of explaining the phenomena. In the case of the phenomenon of recurring desires, though, things are different. We said that the in-the-closet theory and the Buddhist dependent origination theory give equally good explanations of this phenomenon. But the in-the-closet theory posits an unobservable entity that the dependent origination theory does not. The former theory says that volitions continue to exist in our minds even when we are not aware of them. The latter theory speaks instead of patterns of neurons in the brain – something that can be observed. This makes the latter theory lighter, and so it is the theory that we ought to choose.

The Principle of Lightness would help the Buddhist answer the objection about seemingly permanent volitions. It can also be used in defense of premise (5), the premise that says consciousness is impermanent. In the following passage the Buddha claims that consciousness also originates in dependence on sense-object contact:

Just as, O monks, fire is named from that in dependence on which it burns. The fire which burns in dependence on logs of wood is called a log-fire. The fire which burns in dependence on chips is called a chip-fire. The fire which burns in dependence on grass is called a grass-fire. The fire which burns in dependence on cow-dung is called a cow-dung fire. The fire which burns in dependence on husks is called a husk-fire. The fire which burns in dependence on rubbish is called a rubbish-fire. In exactly the same way, O monks, consciousness is named from that in dependence on which it comes into being. The consciousness which comes into being in respect of color-and-shape in dependence on the eye is called eye-consciousness. The consciousness which comes into being in respect of sounds in dependence on the ear is called ear-consciousness. The consciousness which comes into being in respect of odors in dependence on the nose is called nose-consciousness. The consciousness which comes into being in respect of tastes in dependence on the tongue is called tongue-consciousness. The consciousness which comes into being in respect of things tangible in dependence on the body is called body-consciousness. The consciousness which comes into

being in respect of *dharmas* in dependence on the mind is called mind-consciousness. [M.I.259–60]

To this someone might object that we experience consciousness as some one thing that endures. That when I first see and then take a bite of ice cream, it is one and the same consciousness that is first aware of the color of the ice cream and is then aware of the taste of the ice cream. The Buddhist would respond by pointing out that there are periods in a person's life when there seems to be no consciousness at all occurring. If the opponent were to claim that consciousness continues to exist even then – only in the closet – the Buddhist could reply that their theory of dependent origination gives a lighter explanation of the apparent continuity of consciousness.<sup>5</sup>

But the Principle of Lightness would also help the Buddhist defend their claim that the mind is an invented fiction. As both the Buddha and Hume point out, we are never actually aware of the mind as something standing behind such mental events as feeling, perceiving and willing. We are just aware of the feelings, perceptions and volitions themselves. So the mind is unobservable. And it is the causal relations among these mental events that the Buddha says explain all the facts about our mental lives. So the mind becomes an unnecessary, unobservable posit.<sup>6</sup>

Why, though, should we accept the Principle of Lightness? The idea behind this principle is that what makes some statement true has to be objective: the truth of a statement is not determined by such subjective factors as our interests, or limitations in our cognitive capacities, but rather just by facts that are independent of our interests and limitations. The thought is that when it comes to finding out what the facts are, we should let the world outside our mind dictate what it is that we believe. To think that factors in my mind could determine what the facts are would be to indulge in magical thinking. By the same token, we could say that positing unobservable entities is inherently suspect. Why believe that something exists when no one could possibly observe it? Because saying so makes it easier for us to explain what we do observe? This is letting what seems to us like a good explanation determine what we say the mind-independent facts are. This is letting our cognitive limitations determine what statements we believe are true. Magical thinking. The Principle of Lightness says we should resort to positing unobservable entities only when the world tells us we have no alternative.

---

<sup>5</sup>The Buddha's argument in the passage we just looked at is slightly different. It depends on the claim that there are six distinct kinds of consciousness, corresponding to the six senses and their respective objects. These twelve items (vision and the visible, hearing and the audible, etc.) are collectively referred to as the *āyatanas*.

<sup>6</sup>Remember, though, that early Buddhism is dualist. One can deny the existence of the mind and still be a dualist. The most familiar form of dualism is substance dualism, the view that there are two kinds of substance, physical substance and mental substance. Descartes was a substance dualist. Buddhists deny the existence of the mind. But they affirm the existence of mental events, such as feeling and perception, as things that are distinct from the physical (*rūpa*). While early Buddhism denies substance dualism, it affirms what could be called event dualism.

We are now done with our review of the explicit premises in the argument from impermanence. There still remains the one implicit premise, the exhaustiveness claim. If we accept this, then it seems we must say the argument from impermanence succeeds in establishing that there is no self. There is one important objection to the exhaustiveness claim. Many find this claim unacceptable because it leaves unexplained the sense we have that there is an 'I' that has a body and various mental states. If the exhaustiveness claim were true, then while there would be a body and various mental states such as feelings and desires, these would not be the body and mental states of anyone or anything. They would be ownerless states without a subject. And this strikes many as absurd. Is this a valid objection to the exhaustiveness claim, and so to the argument from impermanence? We will defer this question. We turn instead to the second argument contained in the passage we have been investigating, the argument from control. This argument also relies on the exhaustiveness claim. Examining this argument will help us better frame the important objection to the exhaustiveness claim. We will then be better positioned to determine whether we should accept this claim, and with it the arguments that turn on it.

### 3.4

The argument from impermanence starts from one way in which we use the word 'I'. The argument from control starts from another. We often say things like, 'I felt okay about my hair today, but my nails look pretty ratty; I need to do something about them.' This tells us that we think of the 'I' as something that evaluates the states of the person and seeks to change those it finds unsatisfactory. Let us call this the executive function. Then if there is a self, the self would be that part of the person that performs the executive function. Recall that in the passage we looked at earlier, the Buddha says of each *skandha* that it cannot be the self because it is sometimes other than we want it to be. This makes it sound as if he is assuming that we would have complete control over the self, so it would always be perfect in our own eyes. And why would this be? If the self performs the executive function then it tries to control the other parts of the person. But why must it have complete control over anything? And isn't there something odd about supposing that it exercises control over itself? Isn't the point of the executive function to exert control over other things? So far the argument does not look very promising.

But there is a different way of understanding the argument. Consider the Anti-Reflexivity Principle: an entity cannot operate on itself. This principle is widely accepted among Indian philosophers. As supporting evidence they point to the knife that can cut other things but not itself, the finger that can point at other things but not at itself, etc. Are there counter-examples to this principle, cases that show it not to be universally valid? What about a doctor who treats herself? The difficulty with this case is that when the doctor removes her ingrown toenail, it is not the ingrown toenail



that does the treating, it is other parts of the doctor. Those who support the principle claim that all seeming counter-examples will turn out to involve one part of a complex system operating on another part. So there are no counter-examples, and the principle is valid.

Suppose this is right. Then if the self performed the executive function, it could perform that function on other parts of the person, but not on itself. This means I could never find myself dissatisfied with and wanting to change my self. And this in turn means that any part of me that I can find myself disliking and seeking to change could not be my self.<sup>7</sup> Suppose, for instance, that I thought my nose might be my self. My nose would then be the part of me that performs the executive function. When I evaluate the different parts of my body and mind, it would be my nose that did this. When I decided I didn't like something about my hair, or tried to rid myself of some habit I disliked, this would be the nose's doing. The one thing the nose could never do is dislike and try to change itself. So if I ever found myself wanting to change something about my nose, that would show that my nose is not my self. And of course I do dislike it when my nose itches; I try to make it stop by scratching it. Therefore my nose is not my self. The argument as a whole will then go like this:

- 1 I sometimes dislike and seek to change *rūpa*.
  - 2 I sometimes dislike and seek to change feeling.
  - 3 I sometimes dislike and seek to change perception.
  - 4 I sometimes dislike and seek to change volition.
  - 5 I sometimes dislike and seek to change consciousness.
  - 6 If the self existed it would be the part of the person that performs the executive function.
- IP [There is no more to the person than the five *skandhas*.]  
 C Therefore there is no self.

Does this argument work? The first five premises seem to be true. There doesn't seem to be any observable part of the person that I could not find myself dissatisfied with and wanting to change. (Whether I succeed in changing it is another matter, but that's not relevant here.) We've seen how the anti-reflexivity principle comes in: if the self is the one part of me that's at work when I evaluate my states and try to change those I find unsatisfactory, then it is the one thing I could never evaluate and seek to change. So it looks like the argument does prove its conclusion provided the one implicit premise is true – that there is no more to me than the five *skandhas*.

At this point it may strike you that there is something very peculiar going on here. On the one hand we have an argument designed to show that there is no part of the person that is the controller – no part that performs the executive function. Yet in this

<sup>7</sup>This way of interpreting the argument is suggested by the fact that the Sāṅkhya school of orthodox Indian philosophy gives an argument for the existence of the self that uses the same basic ideas (though put to very different ends). See *Tattvakaumadī* on *Sāṅkhya-kārikā* XVII.

very argument we have premises stating ‘I sometimes dislike and seek to change...’ To say that I dislike and seek to change something is to say that I perform the executive function. Yet according to the conclusion of the argument, there is nothing that performs the executive function. If there really were no one in charge, then wouldn’t the evidence that is being used to show that no one is in charge really be bogus? Doesn’t the evidence presented in the premises actually require that the conclusion be false?

This suspicion can be developed into a very powerful challenge to the exhaustiveness claim. Here is how it goes. Suppose that the five *skandhas* contain all the parts of the person that we ever observe. We agree that we sometimes dislike and seek to change each of the *skandhas*. And we also agree that whatever is performing this executive function cannot perform it on itself. The conclusion then seems inescapable that there must be more to the person than just the observable parts, the five *skandhas*. And this ‘something else’ must be the self, the part that performs the executive function. This would explain how it is possible to exercise control over all the observable parts of the person without violating the anti-reflexivity principle. The controller is itself unobservable. This would also explain why Hume and the Buddha were unable to find a self when they ‘looked within’. The self is the observer, and by the anti-reflexivity principle, it cannot observe itself. It can only observe the other parts of the person, the five *skandhas*. The exhaustiveness claim is false: there is more to the person than the five *skandhas*. Not only do the Buddha’s two arguments not succeed in proving there is no self. The evidence they present actually turns out to support the view that there is a self.

This is by far the most serious objection we have encountered to the Buddhist arguments for non-self. Can the Buddhists mount a successful response? They will begin by pointing out an error in the opponent’s characterization of the situation. In spelling out their objection to the controller argument, the opponent said that the argument’s conclusion is that there is nothing that performs the executive function. But this is not what the conclusion of the argument says. It says there is no self that performs the executive function. This leaves it open that there might be something else performing that function. Or rather, that there might be several somethings performing that function. What the Buddhist has in mind is that on one occasion one part of the person might perform the executive function, on another occasion another part might do so. This would make it possible for every part to be subject to control without there being any part that always fills the role of controller (and so is the self). On some occasions a given part might fall on the controller side, while on other occasions it might fall on the side of the controlled. This would explain how it’s possible for us to seek to change any of the *skandhas* while there is nothing more to us than just those *skandhas*.

Consider this analogy. In a monarchy, there is the monarch, and there are his or her subjects. A monarch is not their own subject; a ruler rules over others, not themselves. Now in the case of Great Britain, it is true that every living British citizen has been the subject of a British monarch. But it is also true that Queen Elizabeth II is

a British citizen. How is this possible? If she is a British citizen, that means she has been the subject of a British monarch. But she is the British monarch, and by the anti-reflexivity principle she is not her own subject. Does this mean that there is some unobservable meta-monarch presiding over the UK? Of course not. Queen Elizabeth was the subject of her father, King George, when she was still Princess Elizabeth before her father's death.

This shows how it is possible for the following propositions all to be true:

- 1 There is no more to the person than the five *skandhas* (the exhaustiveness claim).
- 2 I can perform the executive function on each of the *skandhas*.
- 3 An entity cannot operate on itself (the anti-reflexivity principle).

They can all be true because it need not be the same part of the person that performs the executive function on every occasion. So on one occasion my nose might form a coalition with other parts of me and perform the executive function on my hair. On another occasion a coalition with different members might perform the same function on my nose. We will call this the 'shifting coalitions' strategy; it will prove useful to the Buddhist in other contexts as well. In effect, the Buddhist is claiming the opponent has forgotten the second possible meaning of 'I'. The opponent saw this word in premises (1)–(5) of the argument from control, and assumed it meant a self, some one thing that exists as long as the person does. They assumed that when we say I can dislike and seek to change all the *skandhas*, it must be one and the same thing that evaluates and seeks to change all of them. But 'I' might also refer to all the parts of the person taken together. It might refer not to a self but to the person.

The Buddhist is not yet out of the woods though. For one thing, we already know that the Buddha says the person is not ultimately real. We don't yet know just what that means, but it certainly doesn't sound like good news for the shifting coalitions strategy as a way around the objection. What's more, if 'I' refers to the person, then the person should be one thing, not many. 'I' is the first-person singular pronoun; 'we' is the first-person plural. Yet the shifting coalitions strategy requires that it be different things that perform the executive function at different times. How is it that these distinct things all get called by a single name for one thing?

The Buddhist has an answer to this question. It is that 'I' is what they call a 'convenient designator', a word that refers to something that is just a useful fiction. The person is that useful fiction. The person is a whole made of parts. And wholes are not themselves real things, only the parts are. I think that 'I' must refer to one and the same thing every time I use it because I have forgotten that the person is a useful fiction. I have forgotten that 'I' is just a useful way to talk about all the parts taken together.

This is the basic strategy the Buddhist will use to answer the challenge to the exhaustiveness claim. But we need to investigate that strategy in much greater detail. Before we begin that task it would be good to summarize the state of play to this point. The Buddha gave two arguments for non-self, the argument from

impermanence and the argument from control. Both arguments relied on the exhaustiveness claim, which says there is no more to the person than just the five *skandhas*. This claim was crucial to both arguments, since they both proceed by showing that there is some property of a self that all the *skandhas* lack. Showing this would not show there is no self if there could be more to the person than just these *skandhas*. The opponent objects that the exhaustiveness claim cannot be true if it is true that we can exercise some degree of control over all five *skandhas*. Indeed the opponent takes this fact to show that there must be more to the person than the five *skandhas*. The first Buddhist response is to point out that if the *skandhas* took turns performing the executive function, then all five could be subject to control without violating the anti-reflexivity principle. To this the opponent objects that in that case there would be not one controller but many. The second Buddhist response will be that there is a single controller, the person, but the person is only conventionally real. We now turn to an examination of just what this might mean.

### 3.5

The text we are about to examine comes from a work called *Milindapañha* or *The Questions of King Milinda*. It is a dialogue between a king, Milinda, and a Buddhist monk named Nāgasena. Milinda is an historical figure. He lived in the second century BCE, was of Greek ancestry (his Greek name was Menandros), and was a ruler in Bactria (in present-day Pakistan) after its conquest by Alexander the Great. Milinda probably did discuss Buddhist teachings with Buddhist monks, but we don't know if there was a Nāgasena among them. The work was composed early in the first century CE, and it is probably not the transcription of an actual conversation. More importantly, it is not an early Buddhist work; it does not record the teachings of the Buddha and his immediate disciples. It is still useful for our purposes, though. For it is recognized as authoritative by a number of different Abhidharma schools. So its views represent a consensus position among a wide variety of commentarial traditions on the teachings of the Buddha.

The passage we are going to look at represents the first meeting of Nāgasena and Milinda. Notice how the conventional practice of exchanging names leads right to a substantive philosophical dispute.

Then King Milinda drew near to where the venerable Nāgasena was; and having drawn near, he greeted the venerable Nāgasena; and having passed the compliments of friendship and civility, he sat down respectfully at one side. And the venerable Nāgasena returned the greeting; by which, verily, he won the heart of King Milinda.

And King Milinda spoke to the venerable Nāgasena as follows: 'How is your reverence called? Sir, what is your name?'

'Your majesty, I am called Nāgasena; my fellow-monks, your majesty, address me as Nāgasena: but whether parents give one the name Nāgasena, or

Sūrasena, or Vīrasena, or Sīhasena, it is, nevertheless, your majesty, just a counter, an expression, a convenient designator, a mere name, this Nāgasena; for there is no person here to be found.'

Notice that his point here is not that his parents could have given him any of those other names instead. While this is true, it's not philosophically significant. His point is rather that whatever name he was given is just a useful way of labeling something that is not actually a person:

Then said King Milinda, 'Listen to me, my lords, you five hundred Yonakas, and you eighty thousand monks! Nāgasena here says thus: 'There is no person here to be found.' Is it possible, pray, for me to assent to what he says?'

And King Milinda spoke to the venerable Nāgasena as follows: 'Nāgasena, if there is no person to be found, who is it then that furnishes you monks with the priestly requisites – robes, food, bedding, and medicine, the needs of the sick? who is it that makes use of the same? who is it that keeps the precepts? who is it that applies himself to meditation? who is it that realizes the Paths, the Fruits, and nirvāna? who is it that destroys life? who is it that takes what is not given him? who is it that commits immorality? who is it that tells lies? who is it that drinks intoxicating liquor? who is it that commits the five crimes that constitute "proximate karma"? In that case, there is no merit; there is no demerit; there is no one who does or has done meritorious or demeritorious deeds; neither good nor evil deeds can have any fruit or result. Nāgasena, neither is he a murderer who kills a monk, nor can you monks, Nāgasena, have any teacher, preceptor, or ordination.'

If there are no persons, there can be no one who gives alms to monks, nor can there be monks who embark on the path to nirvāna. Likewise there can be none who commit evil deeds. These and other absurdities are what Milinda thinks follow from Nāgasena's claim:

When you say, 'My fellow-monks, your majesty, address me as Nāgasena,' what then is this Nāgasena? Pray, sir, is the hair of the head Nāgasena?'

'Indeed not, your majesty.'

'Is the hair of the body Nāgasena?'

'Indeed not, your majesty.'

'Are nails ... teeth ... skin ... flesh ... sinews ... bones ... marrow of the bones ... kidneys ... heart ... liver ... pleura ... spleen ... lungs ... intestines ... mesentery ... stomach ... faeces ... bile ... phlegm ... pus ... blood ... sweat ... fat ... tears ... lymph ... saliva ... snot ... synovial fluid ... urine ... brain of the head Nāgasena?'

'Indeed not, your majesty.'

'Is now, sir, rūpa Nāgasena?'

'Indeed not, your majesty.'

'Is feeling Nāgasena?'

'Indeed not, your majesty.'

‘Is perception Nāgasena?’  
 ‘Indeed not, your majesty.’  
 ‘Is volition Nāgasena?’  
 ‘Indeed not, your majesty.’  
 ‘Is consciousness Nāgasena?’  
 ‘Indeed not, your majesty.’  
 ‘Are, then, sir, *rūpa*, feeling, perception, the volition, and consciousness unitedly Nāgasena?’  
 ‘Indeed not, your majesty.’  
 ‘Is it, then, sir, something besides *rūpa*, feeling, perception, volition, and consciousness, which is Nāgasena?’  
 ‘Indeed not, your majesty.’  
 ‘Sir, although I question you very closely, I fail to discover any Nāgasena. Verily, now, sir, Nāgasena is a mere empty sound. What Nāgasena is there here? Sir you speak a falsehood, a lie: there is no Nāgasena.’

Notice that Milinda goes through each of the different parts of the body first, before coming to *rūpa*, or the body as a whole; in each case he asks if this is what ‘Nāgasena’ is the name of. He next asks about the four *nāma skandhas*. Nāgasena says ‘no’ in each case, though he doesn’t say why. We can imagine that he has the same reasons as those the Buddha gave in his two arguments for non-self. The next possibility Milinda suggests is the five *skandhas* taken collectively. It is noteworthy that Nāgasena denies this as well. The last possibility is that it is something distinct from all five *skandhas*. Nāgasena’s denial is tantamount to the exhaustiveness claim: there isn’t anything else. Finally, note that Milinda takes this all to mean that ‘Nāgasena’ is a ‘mere empty sound’, a meaningless bit of nonsense. This is not what Nāgasena said the name is. He called it a ‘convenient designator’. These two views about what the name is have very different consequences. If Milinda is right that ‘Nāgasena’ is a mere empty sound, then all the absurd consequences Milinda mentioned will follow. As we’ll see in a bit, though, they don’t follow if Nāgasena is right and the name is a convenient designator.

Nāgasena now tries to get Milinda to see the difference between a name’s being a mere empty sound and its being a convenient designator. He does this by turning Milinda’s own reasoning back on him, applying it to the word ‘chariot’. This reasoning leads Milinda into absurdities. Milinda will then realize that the way out of those absurdities involves distinguishing between a word’s being a mere empty sound, and its being a convenient designator. The absurdities don’t follow if we think of the word as a convenient designator:

Then the venerable Nāgasena spoke to King Milinda as follows: ‘Your majesty, you are a delicate prince, an exceedingly delicate prince; and if, your majesty, you walk in the middle of the day on hot sandy ground, and you tread on rough grit, gravel, and sand, your feet become sore, your body tired, the mind is oppressed, and the body-consciousness suffers. Pray, did you come on foot, or riding?’

'Sir, I do not go on foot. I came in a chariot.'

'Your majesty, if you came in a chariot, tell me what the chariot is. Pray, your majesty, is the pole the chariot?'

'Indeed not, sir.'

'Is the axle the chariot?'

'Indeed not, sir.'

'Are the wheels the chariot?'

'Indeed not, sir.'

'Is the chariot-body the chariot?'

'Indeed not, sir.'

'Is the banner-staff the chariot?'

'Indeed not, sir.'

'Is the yoke the chariot?'

'Indeed not, sir.'

'Are the reins the chariot?'

'Indeed not, sir.'

'Is the goading-stick the chariot?'

'Indeed not, sir.'

'Pray, your majesty, are pole, axle, wheels, chariot-body, banner-staff, yoke, reins, and goad unitedly the chariot?'

'Indeed not, sir.'

'Is it, then, your majesty, something else besides pole, axle, wheels, chariot-body, bannerstaff, yoke, reins, and goad which is the chariot?'

'Indeed not, sir.'

'Your majesty, although I question you very closely, I fail to discover any chariot. Verily now, your majesty, the word chariot is a mere empty sound. What chariot is there here? Your majesty, you speak a falsehood, a lie: there is no chariot. Your majesty, you are the chief king in all the continent of India; of whom are you afraid that you speak a lie? Listen to me, my lords, you five hundred Yonakas, and you eighty thousand monks! King Milinda here says thus: "I came in a chariot;" and being requested, "Your majesty, if you came in a chariot, tell me what the chariot is," he fails to produce any chariot. Is it possible, pray, for me to assent to what he says?'

When Nāgasena accuses Milinda of telling a lie, he is just driving home to Milinda the consequences of following Milinda's reasoning about the name 'Nāgasena' when that reasoning is applied to the case of the word 'chariot'. Nāgasena is being a skillful teacher. He wants Milinda himself to come up with the resolution of the difficulty. This is just what happens next:

When he had thus spoken, the five hundred Yonakas applauded the venerable Nāgasena and spoke to King Milinda as follows: 'Now, your majesty, answer, if you can.'

Then King Milinda spoke to the venerable Nāgasena as follows: 'Nāgasena, I speak no lie: the word "chariot" functions as just a counter, an expression, a convenient designator, a mere name for pole, axle, wheels, chariot-body, and banner-staff.'

‘Thoroughly well, your majesty, do you understand a chariot. In exactly the same way, your majesty, in respect of me, “Nāgasena” functions as just a counter, an expression, convenient designation, mere name for the hair of my head, hair of my body ... brain of the head, *rūpa*, feeling, perception, the volition, and consciousness. But ultimately there is no person to be found. And the nun Vajira, your majesty, said this before the Blessed One:

‘Just as there is the word “chariot” for a set of parts,’

So when there are *skandhas* it is the convention to say, ‘There is a living being.’

‘It is wonderful, Nāgasena! It is marvelous, Nāgasena! Brilliant and prompt is the wit of your replies.’ [MP 25–28]

Notice how Milinda agrees that ‘chariot’ is not a mere empty sound, but a convenient designator, a useful way of referring to the parts when they are put together in a certain way. So when Milinda said he came in a chariot, what he said was true, he was referring to something real – just not a chariot. But why is this? Why not simply say that ‘chariot’ is the name of a chariot? The answer is that a chariot is actually not a real thing. The parts are real, but the whole that is made up of those parts is not. The whole can be reduced to the parts, it isn’t anything over and above the parts. This is the view known as ‘mereological reductionism’.<sup>8</sup>

This was the view of early Buddhism. This view was systematically developed and argued for in Abhidharma. We will examine the argument when we investigate Abhidharma (in Chapter 6). In early Buddhism we just have what looks like a kind of ontological bias against wholes:<sup>9</sup> wholes are not really real, only the parts are. There is a way of making sense of that bias though. Consider a set of all the parts needed to make a chariot. Suppose those parts are arranged in what we would call the ‘assembled-chariot’ way: rim attached to spokes, spokes connected to felly, felly connected to axle, axle to body, etc. In this case we have one word that we apply to the set, ‘chariot’. Now suppose those parts are arranged in a different way, the ‘strewn-across-the-battlefield’ way: rim partly submerged in the mud, one spoke wrapped around a tree root, another spoke lying on the ground three meters north-east of the first, etc. In this case we do not have a single name for the set. The best we can do is ‘all the parts that used to make up the chariot’. This difference is reflected in another difference. In the first case we think of the parts as one thing; in the second case we think of the parts as many things. Why this difference in attitude? Is it just because in the first case the parts are all in immediate proximity to one another? But if the parts were all jumbled together in a heap, we still wouldn’t think of them as one

---

<sup>8</sup>Mereology is that part of metaphysics concerned with the relation between the whole and the parts. So mereological reductionism is the view that whole and parts are related by way of the whole being reducible to the parts.

<sup>9</sup>Ontology is that part of metaphysics concerned with determining the basic kinds of existing things. When philosophers speak of ‘an ontology’, they mean a list of the basic categories of existents. So for instance the doctrine of the five *skandhas* represents an ontology if the exhaustiveness claim is true.



thing, we'd think what we had then was just a bunch of parts in a pile. No, the difference in our ontological attitude (thinking of them as one thing in the one case but as many things in the other) stems from the fact that we have a single word for the parts in the first case but not in the second. And why do we have this single word in the one case? Because we have an interest in the parts when they are arranged in that way. When the set of parts is arranged in the assembled-chariot way, they serve our need for a means of transportation 'across the hot sandy ground'.

At this point you might be thinking, 'Well, of course. We only have a single word for the parts when they are put together in a way that serves our interests. This is no doubt why Nāgasena calls the word "chariot" a convenient designator. Because it's convenient for us to have a way to designate the parts when they're assembled in that way. That configuration is one we're likely to encounter frequently (if we live in a society that uses chariots). And it's one we're likely to want to be able to refer to. It's much easier to tell your servant to fetch a chariot than to ask that they bring a rim attached to some spokes attached to a felly attached to ... By contrast it's much less likely that we'll ever need to refer to the set of parts when it's arranged in the strewn-across-the-battlefield way. And there are only so many words we can learn to use before our brains begin to clog up. If we had to learn a different word for every possible arrangement of those parts our minds would melt. So we only have a single word in the case that serves our convenience. This all makes good sense. But why is it supposed to show that the chariot isn't really real?'

The answer is that our ontological attitude should not be dictated by our interests. Common sense says that the chariot is a real thing. Suppose we simply followed common sense. We would then be thinking of the chariot as one thing, but the same parts arranged in some different way as many things, because it was more convenient for us to think that way. We would be letting our interests dictate what we take reality to be like, and we know where that can lead. Assessing your finances that way can lead to disaster. This is why strictly speaking the chariot is not a real thing. It is just what Abhidharma will call a 'conceptual fiction': something not ultimately real that is nonetheless accepted as real by common sense because of our use of a convenient designator. Here are some other examples of conceptual fictions: a house, a lute, an army, a city, a tree, a forest and a column of ants. The list could be extended indefinitely. Our common-sense ontology is full of things that we think are real, but are also wholes made of parts. The early Buddhist view is that strictly speaking none of these things is really real.

Notice, though, that the word 'chariot' is not a 'mere empty sound'. Nāgasena sees a difference between that status and a word's being a convenient designator. To call a word a mere empty sound is to say it has no meaning. And in this context that would mean that there is nothing that it refers to. So if chariots are not really real, why isn't the word 'chariot' a mere empty sound? We already gave the answer, but it is worth repeating and elaborating on. 'Chariot' does refer to something, but not to what it appears to refer to. Its reference is misleading, for it seems to be the name of a single thing, a chariot, and there really is no such thing. It is, though, a useful way of talking

about a set of parts when they are arranged in a certain way. So when we use the word correctly, there is something in the world that we are talking about. This is different from the case of a word that refers to nothing whatever, such as ‘sky-flower’ or ‘son of a barren woman’. (The Sanskrit equivalents of these expressions are both single words.) Since a barren woman has no children, there is no such thing as the son of a barren woman. So there is nothing that the word is the name of. Using the word ‘chariot’ might help us get what we want, but using ‘son of a barren woman’ never will. The chariot might be a fiction, but it isn’t an utter fiction, like the son of a barren woman. Instead it’s a useful fiction.

In this respect the chariot is like the average college student. Just looking at the form of the expression ‘the average college student’, we might be misled into thinking that it refers to a flesh-and-blood person. It does not. There is no such person as the average college student. So it doesn’t make sense to ask what school they go to, what their major is, or who their parents are. But this does not make the concept useless. For there are real facts that back up what is said about the average college student, facts about all the flesh-and-blood college students. Those facts are very complex, for they involve details about the lives of many people. So for certain purposes it is useful to be able to express them in simplified form. This is just what happens when statisticians come up with the facts about the average college student. The average college student is a fiction, but a useful one. The concept helps fulfill certain interests, like those of college loan officers and credit card companies. And the same goes for the chariot, but not for the son of a barren woman.

### 3.6

There is one last point to make about the passage we have been looking at. Toward the very end Nāgasena says, ‘Ultimately there is no person to be found’. We can now see that he means to call the person a mere conceptual fiction, something we believe to exist only because of our use of a convenient designator. We will have more to say about this in the next section. But we might ask what the force of this ‘ultimately’ is. The answer is that it involves a distinction between two ways in which a statement may be true: ultimately and conventionally. What Nāgasena is saying is that it is not ultimately true that there are persons. He would, however, say that it is conventionally true that there are persons. The distinction may be characterized as follows:

- A statement is **conventionally true** if and only if it is acceptable to common sense and consistently leads to successful practice.
- A statement is **ultimately true** if and only if it corresponds to the facts and neither asserts nor presupposes the existence of any conceptual fictions.

Suppose there is a soft-drink machine in the lobby of the building, and consider the

statement, 'There's a soft-drink machine in the lobby.' You might think that what the statement says corresponds to the facts. But even if there is a sense in which that is correct, still it asserts the existence of a conceptual fiction, the soft-drink machine. Does that mean the statement is ultimately false? No. To call it ultimately false is to be committed to the ultimate truth of the statement that is its negation, 'There is no soft-drink machine in the lobby.' And for that statement to be true it would have to be true that there are or at least could be such things as soft-drink machines. It presupposes the existence of a conceptual fiction. No statement that uses the concept of a soft-drink machine could be ultimately true. Our statement is conventionally true though. Any speaker of English who was informed about the building would agree to it, so it is acceptable to common sense. And its acceptance consistently leads to satisfaction of our desires, such as my craving for a diet soda.

So any statement that uses convenient designators can only be conventionally true. It cannot be ultimately true, or ultimately false either. From the ultimate perspective such a statement is simply without meaning, and so not the sort of thing that could be either true or false. The Sanskrit word (*saṃvṛti*) that we are translating as 'conventional' literally means 'concealing'. And Buddhist commentators explain their use of this term by saying that convenient designators conceal the nature of reality. Words like 'chariot' are misleading because they seem to refer to a single thing when they actually refer to a plurality. If we want a complete description of how things actually objectively are, we should avoid using them. Of course that objectivity would come at a steep price. If we could never use convenient designators in describing the world, then when we wanted to ride over the hot sandy ground we would have to list all the parts that make up the chariot and describe how each is related to the others. That would take a long time. So inevitably we lapse back into using conventional truth.

This is not necessarily a problem though. After all, not just any statement using convenient designators will be conventionally true. The definition said such statements must consistently lead to successful practice.<sup>10</sup> The statement about the soft-drink machine might, but no statement about there being a teletransportation machine in the lobby will. There is no such thing as a teletransportation machine. Isn't it also true that there really aren't any soft-drink machines either? Why should the belief in those non-existent things lead to successful practice? The answer, of course, is that there are all the suitably arranged parts that make up what we call a

---

<sup>10</sup>The definition also mentioned being acceptable to common sense. And some statements that were once acceptable to common sense no longer are. People once believed that the world is flat, but no one does now. But the statement that the world is flat was never conventionally true. Remember that a statement must also consistently lead to successful practice to be conventionally true. The belief that the world is flat leads to the belief that if you sail far enough in the same direction you will reach the edge. But since the world is round, you can never succeed in reaching the edge of the world. Most (though not all) statements that are acceptable to common sense are so because they consistently lead to successful practice.

soft-drink machine. It's because of their interactions that my desire for a cold dose of artificially sweetened carbonated flavored water gets satisfied. And if we wanted to we could probably spell this all out. Usually, though, we don't want to. We just use our shorthand description of the situation: 'There's a soft-drink machine in the lobby.' It's worth remembering, though, that standing behind every conventionally true statement is some (much longer) ultimately true statement that explains why accepting the conventionally true statement leads to successful practice. This connection between conventional truth and ultimate truth plays an important role in what follows.

### 3.7

The distinction between conventional truth and ultimate truth was developed by commentators on the early Buddhist texts in order to solve an exegetical problem. The problem is that the Buddha's teachings seem inconsistent. On some occasions he teaches that there is no self and that what we think of as a person is really just a causal series of impermanent, impersonal states. On other occasions he says nothing of this and instead teaches a morality based on the doctrine of karma and rebirth. The inconsistency stems from the fact that the latter teaching appears to involve the idea that it is one and the same person who performs a deed in this life and reaps the karmic fruit in the next life. So the Buddha seems to affirm in those teachings what he elsewhere denies when he teaches the unreality of the person. Of course we could simply agree that the Buddha contradicted himself and leave it at that. But the commentators saw a way around attributing such a major error to the founder of their tradition: the first sort of teaching represents the full and final truth, whereas the second represents what ordinary people need to know in order to progress toward being able to grasp the full and final truth.<sup>11</sup> Using this distinction, commentators came to say that some sūtras have meanings that are 'fully drawn out' (*nītārtha*), while others have meanings requiring explication (*neyārtha*). The former came to be considered statements of the ultimate truth, the latter were said to be couched in terms of conventional truth.

The original point of the distinction between the two truths was, then, to clarify the early Buddhist view on the person. It was not to help us see that chariots are not ultimately real. It isn't too hard to see that chariots don't belong in our final ontology,

---

<sup>11</sup>This is said to be a manifestation of the Buddha's pedagogical skill (*upāyakaśāla*), his ability to fashion his teaching to the capacities of his audience. Presumably the second sort of teaching is given to an audience that has not yet fully grasped the consequences of rebirth. They thus engage in immoral conduct, which only binds them more firmly to the cycle of rebirth. By teaching them a karmically based morality the Buddha hopes to make them less prone to conduct that reinforces their ignorance. Then they will be better able to appreciate the full and final truth about persons. It is an interesting question whether this practice represents deception on the Buddha's part.

and that we think they are fully real only because of the way in which we talk.<sup>12</sup> It is much more difficult to believe these things about persons. As the following passage from *Milindapañha* makes clear, much work is needed before we can see how this might be true. Nāgasena and Milinda have now been discussing the Buddha's teachings for a while:

'Nāgasena,' said the king, 'is the one who is born that very person, or is it someone else?'

'He is neither that person,' said the elder, 'nor is he someone else.'

'Give an illustration.'

'What do you say to this, your majesty? When you were a young, tender, weakly infant lying on your back, was that you, the person who is now king?'

'Indeed not, sir. The young, tender, weakly infant lying on its back was one person, and the grownup me is another person.'

Milinda's question is whether it is one and the same person who is born and then goes on to become an adult. Two things are worth noting. First, Nāgasena's answer is decidedly odd. How can the adult me and the infant me be neither the same person nor distinct persons?<sup>13</sup> Doesn't one or the other of these two possibilities have to be the case? Second, Milinda's answer is not what we would expect from someone whose views are supposed to represent common sense. Common sense says that adult and infant are the same person. Milinda says they are distinct persons. Here it's useful to bear in mind that Milinda has now been talking to Nāgasena for some time. One thing Milinda has learned is that all the *skandhas* are impermanent and that there is no self. He has concluded that a Buddhist should thus say adult and infant are distinct persons. Nāgasena will now show him why this common misinterpretation of non-self is wrong:

'If that is the case, your majesty, there can be no such thing as a mother, or a father, or a teacher, or an educated man, or a righteous man, or a wise man. Pray, your majesty, is the mother of the zygote one person, the mother of the embryo another person, the mother of the fetus another person, the mother of the newborn another person, the mother of the little child another person, and the mother of the grownup man another person? Is it one person who is a student, and another

---

<sup>12</sup>A 'final ontology' is an ontology that makes no concessions to our interests and limitations, and accurately reflects the objective nature of reality. In early Buddhist terms it would be an ontology that contains no mere conceptual fictions.

<sup>13</sup>It would not be odd if what Nāgasena said was that while adult and infant are not the same qualitatively, neither are they numerically different persons. In fact, most people would say that's true. That baby and I are one and the same (numerically identical) person; but the baby had qualities I now lack, such as cuteness, so we are qualitatively different. This interpretation of 'neither the same nor different' is only possible, though, if we translate what Nāgasena says using the ambiguous English 'same' and 'different'. That ambiguity is not present in the original. It is numerical identity and numerical distinctness that he is denying.

person who has finished his education? Is it one person who commits a crime, and another person whose hands and feet are cut off [in punishment]?’

‘Indeed not, sir. But what, sir, would you reply to these questions?’

Said the elder, ‘It was I, your majesty, who was a young, tender, weakly infant lying on my back, and it is I who am now grown up. In dependence on this very body all these different elements are collected together.’

‘Give an illustration.’

‘It is as if, your majesty, someone were to light a lamp; would it shine all night?’

‘Certainly, sir, it would shine all night.’

‘But now, your majesty, is the flame of the first watch the same flame as the flame of the middle watch?’

‘Indeed not, sir.’

‘Is the flame of the middle watch the same flame as the flame of the last watch?’

‘Indeed not, sir.’

‘But then, your majesty, was there one light in the first watch, another light in the middle watch, and a third light in the last watch?’

‘Indeed not, sir. In dependence on that first flame there was one light that shone all night.’

‘In exactly the same way, your majesty, is the series of psychophysical elements (*dharmas*) connected together: one element perishes, another arises, seamlessly united as though without before and after. Therefore neither as the same nor as a distinct person does this latest aggregation of consciousness connect up with earlier consciousness.’

‘Give another illustration.’

‘It is as if, your majesty, new milk were to change in process of time into sour cream, and from sour cream into fresh butter, and from fresh butter into clarified butter. And if any one, your majesty, were to say that the sour cream, the fresh butter, and the clarified butter were each of them the very milk itself – now would he say well, if he were to say so?’

‘Indeed not, sir. They came into being in dependence on that milk.’

‘In exactly the same way, your majesty, is the series of psychophysical elements (*dharmas*) connected together: one element perishes, another arises, seamlessly united as though without before and after. Therefore neither as the same nor as a distinct person does this latest aggregation of consciousness connect up with earlier consciousness.’ [MP 41 f]

The overall point of the passage is clear enough: the ultimate truth about what are conventionally called persons is just that there is a causal series of impermanent *skandhas*. But there are a number of puzzling features that require close attention. First there is Nāgasena’s examples of the mother, the student and the criminal. What point is he trying to make with these? Remember that Milinda thought the infant and the adult must be distinct persons. He thought this because he realized that the *skandhas* making up the infant are numerically distinct from those making up the adult. So he reasoned that in the absence of a self existing over and above the

*skandhas*, adult and infant have to be two different persons. He is thus implicitly accepting a principle we might name: **Milinda's Principle** – that is, numerically distinct *skandhas* make for numerically distinct persons.

What Nāgasena is doing is showing that we must reject this principle by showing that absurd consequences would follow if we accepted it.<sup>14</sup> It would for instance follow that there is no such thing as a mother. A mother is a woman who conceives and then bears a child and typically raises it to adulthood. So for there to be mothers there must be persons who continue to exist from the time they conceive until the time their offspring is grown. But the *skandhas* making up a person are constantly going out of existence and getting replaced. For instance, the *skandhas* that make up the woman with an embryo in her uterus (second week of pregnancy) are numerically distinct from the *skandhas* that make up the woman carrying a fetus of six months. So by Milinda's Principle, these are distinct persons, and neither one is a mother. Likewise the *skandhas* that make up the person taking exams and the *skandhas* making up the person who receives a diploma are numerically distinct. So by Milinda's Principle the person who gets the degree is not the same person as the one who took the exams for that degree. The one who receives the diploma didn't do the work for it. Similarly the *skandhas* that make up the convicted robber now sitting in prison are numerically distinct from the *skandhas* that held up the flower shop last year. So the prisoner is not the person who committed the crime; they don't deserve to be punished.

Milinda is quick to agree that these are all absurd consequences. But it is important to stop and consider why. When we think of ourselves and others as persons, we are thinking of a person as something that endures at least a whole lifetime. We are, in other words, gathering together all the *skandhas* from birth until death under one convenient designator, 'person'. Why would this practice be useful? The examples of mother, student and criminal show why. If the pregnant woman didn't follow our practice, but followed Milinda's Principle instead, she would not identify with the woman who will later give birth. So she would see no reason to follow her doctor's prenatal healthcare advice. If the student didn't identify with the graduate, she would see no reason to study for an exam that will only benefit the degree-holder. If the criminal didn't identify with the person who robbed the flower shop, he would see no reason to refrain from robbing again after getting out of prison.

Our concept of a person has it that persons endure at least a lifetime. If we followed Milinda's Principle we would have to replace that concept with the concept of something that lasted nowhere near as long – perhaps for a day, maybe for just a minute. (It depends on how long individual *skandhas* last, and how many must be

---

<sup>14</sup>This strategy is called *reductio ad absurdum* or reducing to absurdity. The idea is to show that some statement is false by first assuming that it is true and then deducing absurd consequences from that assumption. Since these absurd consequences are presumably unacceptable to everyone, this is supposed to show that we should deny the statement in question. Indian philosophers call this strategy *tarka* or *prasaṅga*.

replaced before we say we have a new whole.) To think of ourselves in that way would not be to think of ourselves as persons as we understand that concept. Let's call the resulting view 'Punctualism', and the new concept of what we are 'P-persons'. What the examples show is that it would be a disaster if we thought of ourselves as P-persons rather than as persons. Our convenient designator 'person' is convenient because it helps us avert this disaster. Why is this, though, if there really are no such things as persons? To think of yourself as a person is to think of yourself as a whole that is made up of all the *skandhas* that occur over a lifetime. And wholes like chariots and persons are mere conceptual fictions, not ultimately real things. So why should it work better to think of ourselves in this way?

The answer to this question lies in the point made in the preceding section. Statements that are conventionally true are ones that work. And for every statement that is conventionally true, there is some (much longer) ultimately true statement that explains why it works. Nāgasena is making this point when he tells Milinda that adult and infant are the same person, and then goes on to say that past and present *skandhas* are united through their bodily causal connections. He is speaking first of what is conventionally true and then of what ultimate truth stands behind that conventional truth. But there was something else Nāgasena said about adult and infant, so let's look at all three of his statements.

- 1 Adult and infant are neither the same person nor distinct persons.
- 2 Adult and infant are the same person.
- 3 There is a causal series running from the 'infant' *skandhas* to the 'adult' *skandhas*.

We noted earlier that (1) seems odd. We can now add that (1) and (2) seem to contradict each other. (1) says that adult and infant are not the same person, while (2) says that they are. But perhaps we can now see a way out of both difficulties. Suppose we were to say that (2) represents the conventional truth, while (1) (and (3) as well) are supposed to be ultimately true. What (1) is meant to remind us of is that at the level of ultimate truth no statement about persons could be true; all such statements are simply meaningless. To ask whether these are the same person or distinct persons is to assume that there are such things as persons. Since this presupposition is false, the question has no answer. Questions of personal identity simply can't arise at the ultimate level.

At the conventional level, though, we can say that I was that infant, that we are the same person. The examples of mother, student and criminal are meant to show why (2) is conventionally true: because it works. And why does it work? As (3) tells us, the ultimate truth is that when the infant *skandhas* went out of existence, they caused child *skandhas* to come into existence, and so on in an unbroken chain until we arrive at the present adult *skandhas*. There are thus many causal connections between the *skandhas* existing at one time in the series and those existing later in the series. This in turn means that what happens to the earlier *skandhas* can influence how things are



for the later *skandhas* in that series. Good eating habits early on make for well-functioning *rūpa skandhas* later. Excessive beer consumption tonight makes for pain sensations tomorrow. The desire to study now can bring about diploma-receiving for later *skandhas* in the series. And so on. So when present *skandhas* identify with past and future *skandhas* in the series – when they think of those other *skandhas* as ‘me’ – they are more likely to behave in ways that make it better for the later *skandhas*. To think of oneself as a person is to have the habit of identifying with the past and future *skandhas* in the series. This is why it is useful that we think of ourselves as persons.

Finally, Nāgasena gives two examples of causal series. The point of the first is clear enough. This is a case where an unbroken chain of closely resembling particulars leads to a conceptual fiction, the one light that shone all night. When we look more closely at what we ordinarily think of as one light that endures an entire night, we see that it is really a series of short-lived flames. Each flame only lasts a moment, for it is composed of incandescent gas molecules produced by the burning of the oil. But when those molecules dissipate, they cause new ones to take their place. For the heat of the first flame causes more oil to burn, producing a new replacement flame. So while each flame only lasts a moment, it causes another to take its place immediately upon its ceasing to exist. The result is what looks like a single thing that endures from dusk till dawn. And so it is conventionally true that there was one light that shone all night. The reality, though, is that there are just the many numerically distinct flames, not the one light that has them. The ultimate truth is that there is just the unbroken succession of flames, each causing the next.

The point of the second example is less apparent. What it illustrates, though, is a case that is in one respect like that of the light, but in other respects is different. Like the series of flames, the series of dairy products is unbroken: there is no gap between the time when there is milk and the time when there is ghee. Unlike the flames, though, the members of this series do not all resemble one another. Milk is white, butter yellow; milk is liquid, ghee is semi-solid. And we use each in different ways. We drink milk, put butter on our toast, and use ghee for frying. By contrast, each flame serves the same purpose for us, to light the room. For this reason we are not tempted to think of the dairy series as just one enduring thing. Instead we think of it as a succession of distinct products. Our ordinary way of thinking about this series is closer to the ultimate truth about causal series than is the common-sense view of the series of flames. But it too is dictated by our interests – the fact that we have different uses for different parts of the series. The point is to learn to look behind our wants and needs and see what is really there, the ultimate truth.

When it comes to the causal series of psychophysical elements,<sup>15</sup> Nāgasena gives an interesting description of the ultimate truth. The conventional truth is that I am a person who has existed for some time. I experience this existence as involving there

---

<sup>15</sup>The *dharmas*: these are the particular entities that get classified under the headings of the five *skandhas*. We will have much more to say about what these are when we come to our investigation of Abhidharma in Chapter 6.

being a 'me' who is aware of the different experiences that this 'I' has. Right now I am aware of reading these words and thinking about these ideas. Earlier this same 'I' was aware of other experiences – eating dinner, listening to music, conversing with friends. The objects that this conscious thing is aware of vary over time, but it is always the same 'I' that is aware of them. There is one thing, the 'I', holding together a plurality, the experiences. This is how things seem to us when we use the convenient designator 'person'. The ultimate truth, though, is that there is a causal series of psychophysical elements. Each exists for a while, then goes out of existence, but causes a replacement element to come into existence. In some cases the replacement resembles what was there a moment ago, as with the flames. Consciousness elements are like this. At each moment there is a new consciousness, but each is qualitatively identical with its predecessor. In other cases what follows an element does not resemble it. A feeling of pleasure gives rise to a desire, and that desire may in turn lead to other kinds of experiences. This is the reality behind the appearance of a person living a life. There is no enduring 'I' who has the different experiences. But neither does this mean that each experience is had by a distinct person, in the way that each stage in the dairy series is a distinct thing. There are just the psychophysical elements and their causal connections. This is the reality that makes it useful to think of the series as a person living a life.

We are now in a position to return to the dispute over the exhaustiveness claim and the Buddha's two arguments for non-self. Both arguments relied on there being no more to the person than the five *skandhas*. The opponent objected to the argument from control on the grounds that our ability to exercise some degree of control over all the *skandhas* shows that there must be more to us than the five *skandhas*. The response was that there could be control over all the *skandhas* if it were a shifting coalition of *skandhas* that performed the executive function. But the opponent challenged this response on the grounds that there would then be many distinct I's, not the one we have in mind when we say that I can dislike and seek to change all the *skandhas*. We can now see how the Buddhist will respond. They will say that ultimately there is neither one controller nor many, but conventionally it is one and the same person who exercises control over first one *skandha* and then another. This is so because the controller is a conceptual fiction. It is useful for a causal series of *skandhas* to think of itself as a person, as something that exercises some control over its constituents. Because it is useful, it is conventionally true. This is how we have learned to think of ourselves. But because this person, this controller, is a conceptual fiction, it is not ultimately true that there is one thing exercising control over different *skandhas* at different times. Nor is it ultimately true that it is different controllers exercising control over them. The ultimate truth is just that there are psychophysical elements in causal interaction. This is the reality that makes it useful for us to think of ourselves as persons who exercise control. Our sense of being something that exists over and above the *skandhas* is an illusion. But it is a useful one.

## 3.8

Does this strategy succeed in defending the exhaustiveness claim against the opponent's attack? I shall leave this question unanswered. Let us move on to a different objection to the arguments for non-self. Perhaps you have long been wondering how the Buddha could have argued for the non-existence of a self given his belief in rebirth. How is rebirth possible if there is no self that gets reborn, that goes from one life to the next?

Notice that this is a very different sort of objection than the one against the exhaustiveness claim. That objection tried to show that a key premise in the two arguments for non-self is false. This one doesn't do that. Instead it tries to show that the conclusion of the arguments (that there is no self) is incompatible with something else that the Buddha believes (that there is rebirth). If these two things really are incompatible, then the Buddhist could respond in either of two ways: by accepting a self, or by abandoning belief in rebirth. Given the centrality of non-self to the Buddha's teachings, the latter might seem the better choice. But the Buddhist will say that we don't need to choose. For there is no incompatibility between non-self and rebirth. This is the point Nāgasena makes in the following:

Said the king: 'Nāgasena, does rebirth take place without anything transmigrating [passing over]?'

'Yes, your majesty. Rebirth takes place without anything transmigrating.'

'How, Nāgasena, does rebirth take place without anything transmigrating? Give an illustration.'

'Suppose, your majesty, a man were to light a light from another light; pray, would the one light have passed over [transmigrated] to the other light?'

'Indeed not, sir.'

'In exactly the same way, your majesty, does rebirth take place without anything transmigrating.'

'Give another illustration.'

'Do you remember, your majesty, having learnt, when you were a boy, some verse or other from your poetry teacher?'

'Yes, sir.'

'Pray, your majesty, did the verse pass over [transmigrate] to you from your teacher?'

'Indeed not, sir.'

'In exactly the same way, your majesty, does rebirth take place without anything transmigrating.'

'You are an able man, Nāgasena.' [MP 71]

In both examples we have a causal process whereby one thing brings about the arising of some distinct but similar thing: a lit candle serves as cause of there being a lit oil lamp, and the teacher's knowledge of the poem serves as cause of the student's knowing the poem. The idea, then, is that rebirth occurs when one set of *skandas*, those making up the person in this life, causes a new set of *skandhas* to come into

existence in a new life. This is not different in kind from the sort of thing that regularly occurs during a single lifetime. The cells in our bodies constantly wear out and die, but give rise to similar replacement cells. Desires, in getting satisfied and so being exhausted, set the stage for similar future desires. The continued existence of a person over the course of a lifetime is just the occurrence of a causal series of impermanent *skandhas*.

There are, of course, important differences between the case of a single lifetime and the case of rebirth. While qualitative changes occur during a life, they are gradual. I might wake up with a few more grey hairs than I had yesterday, but I never wake up to find I've become a cow; it is, though, thought possible to die as a human and be reborn as a cow. Unless I'm riding in a train or flying, I don't go to sleep in one place and wake up in another; typically, though, one is said to be reborn somewhere other than where one died. I can usually remember what I did yesterday, but one doesn't typically remember the events from one's past lives. Still the process of rebirth is governed by causal laws, namely the laws of karma. It is because I did these things out of these desires that I am reborn into this kind of life. In the case of a single lifetime, it is because the distinct psychophysical elements are causally connected that it is useful to collect them all together under the convenient designator 'person'. The same goes for the *skandhas* in distinct lives.

There may be another worry here. Rebirth is supposed to be governed by karmic causal laws. And karma is supposed to represent a kind of natural justice: people get what they deserve, good rebirth for virtuous actions, bad rebirth for vicious actions. And how can it be just if it isn't one and the same thing that performs the action and then gets the reward or punishment? This is something that bothers Milinda:

'Nāgasena,' said the king, 'what is it that is born into the next existence?'

'Your majesty,' said the elder, 'it is *nāma* and *rūpa* that is born into the next existence.'

'Is it this same *nāma* and *rūpa* that is born into the next existence?'

'Your majesty, it is not this same *nāma* and *rūpa* that is born into the next existence; but with this *nāma* and *rūpa*, your majesty, one does a deed – it may be good, or it may be evil – and by reason of this deed another *nāma* and *rūpa* is born into the next existence.'

'Sir, if it is not this same *nāma* and *rūpa* that is born into the next existence, is one not freed from one's deeds?'

'If one were not born into another existence,' said the elder, 'one would be freed from one's evil deeds; but, your majesty, inasmuch as one is born into another existence, therefore is one not freed from one's evil deeds.'

'Give an illustration.'

'Your majesty, it is as if a man were to light a fire in the winter-time and warm himself, and were to go off without putting it out. And then the fire were to burn another man's field, and the owner of the field were to seize him, and show him to the king, and say, 'Sir, this man has burnt up my field;' and the other were to say, 'Sir I did not set this man's field on fire. The fire which I failed to put out was a different one from the one which burnt up this man's field. I am not liable to

punishment.' Pray, your majesty, would the man be liable to punishment?'

'Assuredly, sir, he would be liable to punishment.'

'For what reason?'

'Because, in spite of what he might say, the man would be liable to punishment for the reason that the last fire derived from the first fire.'

'In exactly the same way, your majesty, with this *nāma* and *rūpa* one does a deed – it may be good, or it may be wicked – and by reason of this deed another *nāma* and *rūpa* is born into the next existence. Therefore is one not freed from one's evil deed.' [MP 46]

In the case of the fire, strictly speaking the wood-fire that the man lit to warm himself is not the grass-fire that consumed the other man's field. A fire that depends on one kind of fuel cannot be numerically identical with a fire that depends on another kind of fuel. But since the one fire caused the other, it is conventionally true that the first man burnt the second man's field. Likewise the *skandhas* involved in doing an evil deed are ultimately distinct from the *skandhas* born into the painful circumstances of a *preta*. Suppose I'm the one who did the evil deed. If I die without ever being punished, does the fact that nothing transmigrates mean that I escape getting what I deserve? No. Since these human *skandhas* caused those *preta skandhas*, it is conventionally true that that *preta* will be me, the one who did the deed. I will get what I deserve.<sup>16</sup>

This is how the Buddhist defends the doctrine of karma and rebirth against the charge that it is incompatible with non-self. Of course you might think that karma and rebirth are implausible beliefs that a reasonable Buddhism would abandon. The point here is just that the theory of two truths and the claim that persons are conventionally real may be used to show that rebirth and non-self are not incompatible. If Buddhists ought to stop believing in rebirth, it is not because that belief is inconsistent with their central tenet that there is no self.

There are still some questions that the Buddhist needs to answer. The most important of these is the following. The early Buddhist defense of non-self makes crucial use of the claim that wholes are unreal. This was the basis for their claim that persons are mere conceptual fictions that are only conventionally real. When we discussed the case of the chariot, perhaps it occurred to you that a spoke is also a whole made of parts. A spoke consists of many particles of metal or wood. So if wholes are only conceptual fictions, the spoke can't be ultimately real either. The only things that could be ultimately real would have to be impartite things. And just what are they like? Behind this question may lurk the suspicion that nothing that is genuinely impartite. That would represent a major difficulty for the Buddhist

---

<sup>16</sup>Notice that this case is not different in kind from the case of the convicted criminal that Milinda asked about earlier. That was a case of human justice, while this is a case of natural justice. And in that case justice gets carried out in a single lifetime, while this requires two lives. But the principle is the same: where there are the right kinds of causal connections, it is conventionally true that punishment is deserved even when ultimately distinct *skandhas* are involved.

approach. The Abhidharma movement in Buddhist philosophy represents an attempt to solve this difficulty. In Chapter 6 we will look at some Abhidharma attempts to work out what the ultimately real partite entities are like.

Before we do that, though, we will look at the ethical consequences of the doctrine of non-self. In the last chapter we wondered what it might be like to achieve the Buddhist goal of enlightenment. We now know more about what it would be like. To be enlightened is to know that strictly speaking there is no 'me' but only impersonal impermanent psychophysical elements in a causal series. It is to know that the 'I' is just a conceptual fiction. What might it be like to live with that knowledge? Would it be liberating, or would it be depressing? And how might it affect my behavior toward others? Would it make me more concerned about their welfare? Or would I figure that since there are no persons, I needn't worry about infringing on their rights? Would I conclude that anything goes? These are some of the questions we will address in the next chapter.

### *Further Reading*

The complete debate between Nāgasena and King Milinda may be found in *The Questions of King Milinda*, trans. T. W. Rhys Davids (originally published by Oxford University Press, 1890; reprinted at Delhi: Motilal Banarsidass, 1965).

A recent formulation of reductionism about persons that is like that of early Buddhism is that of the British philosopher Derek Parfit. For exposition of the position and arguments in support, see his *Reasons and Persons* (Oxford: Oxford University Press, 1984), Chapters 11–13.