

## EMPTY OR EMERGENT PERSONS? A CRITIQUE OF BUDDHIST PERSONALISM

JAVIER HIDALGO

**ABSTRACT:** *In contrast to Buddhist Reductionists who deny the ultimate existence of the persons, Buddhist Personalists claim that persons are ultimately real in some important sense. Recently, some philosophers have offered philosophical reconstructions of Buddhist Personalism. In this paper, I critically evaluate one philosophical reconstruction of Buddhist Personalism according to which persons are irreducible to the parts that constitute them. Instead, persons are emergent entities and have novel properties that are distinct from the properties of their constituents. While this emergentist interpretation is an interesting and well-motivated reconstruction of the Personalist position, I ultimately reject it on substantive grounds. I distinguish between different kinds of emergentism in the contemporary philosophical literature and show that they fail to support Buddhist Personalism. I thus conclude that Buddhist Personalism is untenable if it's committed to emergentism about persons. This paper also indirectly defends Buddhist Reductionism by showing that it has crucial advantages over Buddhist Personalism.*

**Keywords:** Buddhism, Buddhist philosophy, emergentism, philosophy of mind, reductionism

### 1. INTRODUCTION

Many Buddhist philosophers argue that persons and selves are unreal. On their view, persons are aggregations of psychophysical elements, such as material form, sensation, and mental formations. Yet aggregates are fictions. While we impute persons and other wholes onto phenomena, they don't really exist. Some fictions are useful though. Certain fictions help us to achieve our aims and their acceptance can promote good outcomes. Persons are one such fiction. If we accept that persons are real, then suffering may be reduced. If so, then we can say that persons are conventionally real, although it's false that persons exist from the ultimate perspective that describes what the world is really like. Let's call this set of views: Buddhist Reductionism.

Classical Buddhist philosophers, such as Vasubandhu, seem to have endorsed

---

HIDALGO, JAVIER: Associate Professor of Leadership Studies, University of Richmond, USA. Email: [hidalgoj@gmail.com](mailto:hidalgoj@gmail.com)

Buddhist Reductionism.<sup>1</sup> Contemporary philosophers such as Mark Siderits (2016) also defend this position. Yet Buddhist Reductionists have many critics. Among fellow Buddhists, their ancient opponents were the Pudgalavādins or Buddhist Personalists. The Pudgalavādins were adherents of a group of early Buddhist schools that thrived from approximately the third century BCE until the eleventh century CE.<sup>2</sup> It's hard to say for sure what the ancient Pudgalavādins thought, as most of their literature has been lost. But it appears that the Pudgalavādins held that, although the person (*pudgala*) is neither identical with nor different from the psychophysical elements that make it up, the person is real in some important sense. So, the Pudgalavādins rejected Buddhist Reductionism. Recently, some philosophers have offered philosophical reconstructions of the Pudgalavāda. These philosophers don't merely aim to interpret and clarify the Pudgalavāda position. They also argue in favor of Buddhist Personalism on philosophical grounds. Amber Carpenter, Christian Coseru, and other authors defend Buddhist Personalism. These philosophers contend that alternatives to Buddhist Personalism, such as Buddhist Reductionism, are deficient as well.

In this paper, I'm going to evaluate one philosophical reconstruction of Buddhist Personalism. According to this reconstruction, persons are irreducible to the parts that constitute them. Instead, persons are emergent entities and have novel properties that are distinct from the properties of their constituents. I'll argue that the emergentist reading is an interesting and well-motivated reconstruction of the Personalist position. But I'll ultimately reject it on substantive grounds. I'll distinguish between different kinds of emergentism and argue that none of them can justify Buddhist Personalism. So, I'll conclude that Buddhist Personalism is untenable if it's committed to emergentism about persons. I'll also indirectly defend Buddhist Reductionism by arguing that it has crucial advantages over Buddhist Personalism. I've organized this paper as follows. In section 2, I'll clarify both Buddhist Reductionism and Buddhist Personalism, and explain some of the motivations for Personalism. In section 3, I'll examine different conceptions of emergence in the contemporary philosophical literature and show that they fail to support the Personalist position. In section 4, I'll respond to objections to my argument. Section 5 will conclude the paper.

## 1. BUDDHIST REDUCTIONISM, PERSONALISM, AND EMERGENCE

Like much of the Buddhist philosophical tradition, Buddhist Reductionists draw a distinction between two truths: ultimate and conventional truths. Ultimate truth corresponds to the facts and neither asserts nor presupposes the existence of what is not ultimately real. True propositions about ultimate reality are uncontaminated by our interests and conceptualizations. In contrast, conventional truths are useful fictions.

---

<sup>1</sup> For a discussion of the details of Vasubandhu's position, see Duerlinger (2003). But some authors argue that Vasubandhu's position is not fully reductionist (Hayashi 2016).

<sup>2</sup> The Pudgalavāda included several different schools of ancient Indian Buddhism, the largest of which may have been the Vātsīputrīyas and the Śmītiya (Westerhoff 2018, 55). Pudgalavāda Buddhism never spread far beyond India. As a result, their school died out along with Indian Buddhism as a whole. This is one reason that much of their literature has been lost (Priestly 1999, 2).

According to Mark Siderits (2016, 16), a sentence is conventionally true if and only if it reliably leads to successful practice. Drawing on the ancient Vaibhāṣika school of Abhidharma, Goodman (2005) also suggests that conventional truths are useful metaphors and allow us to express complex ultimate truths. With these points about conventional and ultimate truth in mind, we can now describe Buddhist Reductionism. Buddhist Reductionism is the view that persons are unreal from the perspective of ultimate reality and that, nonetheless, persons are conventionally real. Buddhist Reductionists think that persons lack reality, but also that it's often useful to talk and act as if persons are real.

Buddhist Reductionists believe that persons are unreal in part because they accept mereological nihilism. Mereological nihilism is the view that all composite entities are unreal and that the only things that ultimately exist are simples. Mereological nihilism undergirds the Buddhist Reductionist approach to persons. Persons are composites. Persons are composed of psychophysical elements. Buddhists refer to these psychophysical elements as *skandhas*. The *skandhas* include material form, sensations, consciousness, and other elements. If mereological nihilism is true and persons are composed out of *skandhas*, then persons are unreal. Mereological nihilism also extends beyond persons and applies to all composite objects. Thus, the Buddhist Reductionist believes that macroscopic objects, such as tables and chairs, are unreal too. All macroscopic objections ultimately consist in aggregations of property-particulars or tropes, which Buddhist Reductionists call *dharmas*.

While these metaphysical views sound incredible, it's worth remembering that Buddhist Reductionists believe that it's useful to talk about macroscopic objects as if they existed. Consider the chair that I'm sitting on. The ultimate truth is that this chair is an aggregation of innumerable simples. My chair is composed of an enormously complex set of iron, carbon, manganese, phosphorus, sulfur, and silicon atoms, all arranged in a complex structure. And, of course, atoms are also composites. They're composed of more elementary particles. To spell out the truth about my chair, we'd probably need to talk in terms of fundamental particles, vibrating strings, or whatever. Needless to say, that would make things even more complex. So, the ultimate truth about my chair is going to be hard to understand and communicate to others. It's vastly more convenient to refer to my chair as if it were a whole. In that respect, statements about macroscopic objects can be conventionally true.

The same points hold for persons. The ultimate truth about persons is difficult to fathom. Like other macroscopic objects, persons consist in an incredibly complex set of relations between simples. For convenience sake, we should refer to persons as if they were unitary entities, even though the ultimate truth is that persons are merely a causal series of simples, arranged person-wise. Treating persons as if they were real can have other good effects. For example, suppose that a person commits a serious violent crime and there's a good chance that he will continue to commit similar crimes. We might hold this person responsible for his crime and punish him. This will prevent him from committing more crimes and it might deter others from doing the same. The ultimate truth is that the criminal is unreal and unreal entities are incapable of bearing

moral responsibility. Yet we should treat the criminal as if he were real and thus culpable of a crime, as doing so will prevent more suffering in the future.<sup>3</sup>

I'm glossing over many nuances here in order to briefly describe the overall contours of Buddhist Reductionism. But let's turn to one key principle that justifies the Reductionist position: mereological nihilism. Mereological nihilism says that there are no composite wholes. There are only simples. Why should we accept such a sweeping metaphysical principle? There are different arguments for mereological nihilism. But one of the most influential arguments for mereological nihilism in the Buddhist tradition is the neither-one-nor-many argument.

The argument goes like this. Consider a chariot. The chariot is made up of parts, like an axle, wheels, a frame, and so on. Here's a question: is the chariot identical with or distinct from its parts? Let's consider the first option, the claim that the chariot is identical with its parts. A moment's reflection suggests that this first option is untenable. After all, the chariot and its parts have distinct properties. The chariot is one and the parts are many. If we accept the indiscernibility of identicals, we must conclude that the chariot and its parts are different from one another.

This leaves us with the second option, the claim that the chariot is distinct from its parts. But one problem with this option is that the chariot seems to borrow all of its properties and causal powers from its parts. Yet it's unclear whether we can rightfully say that a chariot exists if it inherits all of its properties and causal powers from its constituents. Suppose that the chariot crashes into a tree and this causes the tree to fall over. Did the chariot as a whole cause the tree to fall over or did its parts cause the tree to fall? If the chariot and its parts occupy the same region of space, then we might be tempted to say that both the chariots and its parts caused the tree to fall. But now we have a problem of causal overdetermination. Instead of saying that "the chariot as a whole" and the "chariot's parts" caused the tree to fall over, it's simpler to say that only the chariot's parts caused this to happen. Given these difficulties, we should conclude that the only things that exist are the parts that compose the chariot.

So, the whole is unreal and only the parts exist. And the same conclusion can generalize to everything, including tables, chairs, and persons. Here's a general version of the argument:

1. If wholes exist, then either wholes are identical with their parts or distinct from them.
2. Wholes and their parts have incompatible properties.
3. Two entities cannot be identical if they have incompatible properties.
4. So, wholes are not identical with their parts (2, 3).
5. The properties of wholes can be entirely explained in terms of facts about their constituent parts.

---

<sup>3</sup> Yet Buddhists Reductionists deny that reactive attitudes, such as anger and resentment, are intrinsically appropriate. They also reject retributivism, the view that backward-looking considerations like desert can justify punishment. So, reactive attitudes and punishment can only be instrumentally justified. Moreover, there's plenty of room for Buddhist Reductionists to question whether our current practices of blame and punishment are justified. For a further discussion of Buddhist perspectives on blame and punishment, see Goodman (2014). I want to thank an anonymous referee for pushing me to clarify these points.

6. Entities whose properties can be entirely explained in terms of facts about their constituent parts are not distinct from their parts.
7. So, wholes are not distinct from their parts (5, 6).
8. Therefore, wholes don't exist (1, 4, 7).

If the neither-one-nor-many argument is sound, then we have a case for mereological nihilism. If mereological nihilism is true, then a core part of Buddhist Reductionism follows.

Let's now turn to the Pudgalavāda. As I noted in the introduction, the Pudgalavāda were a set of schools of early Indian Buddhism. At one point, the adherents of the Pudgalavāda—the Pudgalavādins—may have made up more than a quarter of the entire Buddhist monastic population in India (Westerhoff 2018, 55). Yet scholars disagree about how to interpret the views of the ancient Pudgalavādins because most of their work has been lost, we often only have descriptions of their views from hostile sources, and the Pudgalavādins likely disagreed among themselves (Priestly 1999). Although I'll refer to the ancient Pudgalavādins, I'll focus my discussion on recent philosophical reconstructions of the Pudgalavāda position. I'll call this position: Buddhist Personalism.

Buddhist Personalists reject the reductionist view that persons are merely conventionally real. Personalists hold that persons are ultimately real in some respect.<sup>4</sup> So, Buddhist Personalists seem to deny that persons are fully reducible to the *skandhas* that make them up. If so, then Buddhist Personalists should reject mereological nihilism. Buddhist Personalists think that persons are composed out of, or reliant on, the *skandhas*. If mereological nihilism were true, then the person would lack ultimate reality, as the person is a composite. And, if Buddhist Personalists should reject mereological nihilism, they should reject the neither-one-nor-many argument.

But where exactly would Personalists object to the neither-one-nor-many argument? My suggestion is that Personalists should reject premise 5, the premise that the properties of wholes can be entirely explained in terms of facts about their constituent parts. If we could explain the properties of wholes in terms of their parts, then we could eliminate the whole. But, if the whole does have novel properties that are irreducible, then premise 5 turns out to be false. We're unable to account for the properties of the whole in terms of the properties of its parts in at least one case, the case of persons. Thus, the neither-one-nor-many argument is unsound.

So, Buddhist Personalists might contend that a whole has properties that are irreducible to its parts. The whole has novel properties that emerge from the elements that make up the person. Contemporary philosophers refer to this view as *emergentism*. Interpreters often suggest that the Personalists are committed to emergentism.<sup>5</sup>

---

<sup>4</sup> But there are already some who disagree with this claim. For example, Dan Lusthaus (2009) argues that the Pudgalavādins believed that persons are merely conventionally existent. If this interpretation is accurate, then Buddhist Personalism and Reductionism appear indistinguishable for reasons I outline in section 3.

<sup>5</sup> One notable exception is Priestly. Priestly (1999, 101) initially agrees that the pudgala is true and ultimate "in the special sense that its existence and functions are not reducible to those of its constituents." However, he later argues that this interpretation is untenable and favors a radically different one. According to Priestly's interpretation, the pudgala is a kind of local instantiation of, or reflection of,

Although Amber Carpenter refrains from attributing a strong emergentist thesis to the Personalists, she claims that person-constituting aggregates have “emergent functionality” and person-constituting aggregates can perform actions and bear responsibility only if they’re taken as wholes (2015, 23-6).<sup>6</sup> In sympathetically reconstructing the Personalist position, Christian Coseru (2020) says that we can interpret Personalism along emergentist or non-reductive physicalist lines.<sup>7</sup> Jonathan Ganeri (2015, 139) argues that Buddhist Personalists “are emergentists: they claim that the *pudgala* is an inexplicable product of diachronic aggregation.” Thích Thiện Châu (1999, 161) contends that “the *pudgala* is something more than the reunion of its constituent elements. It is the essential factor which unifies the process of the life of the person....” Siderits (2016, 111-118) construes the most defensible version of this view as committed to emergentism or non-reductive physicalism.<sup>8</sup>

So, emergentism would allow Personalists to reject the neither-one-nor-many argument and it might be one useful way of interpreting the Personalist position in terms of contemporary philosophical concepts. But why should we accept emergentism about persons? What independent reason do we have to think that this view is plausible? In her insightful reconstruction of the Personalist view, Amber Carpenter argues that persons have characteristics that prevent their complete reduction to *dharmas*. The aggregates that make up the person inherently belong together and allow the person to do things that mere collections of *skandhas* lack the capacity to do. Furthermore, the ultimate existence of persons is necessary to explain karmic responsibility over time.

Consider a child’s growth and development over time. How should I describe this development? On the one hand, we can describe the growth of a child as mere aggregation and rearrangement of *dharmas*. But this description omits something

---

*Nirvana*. There’s much to recommend in Priestly’s novel interpretation. However, I will set it aside here and focus on the emergentist reading. James Duerlinger’s careful study of Vasubandhu’s criticisms of the Pudgalavādins also deserves mention. Duerlinger (2003, 28) interprets the Pudgalavādins as claiming that persons “ultimately exist without being independently identifiable or being separate substances, since ultimate existence does not require independent identifiability.” There are some hints that this is compatible with an emergentist interpretation. On some forms of emergentism, persons are not separate from the aggregates or identifiable apart from them, but instead they’re emergent properties of aggregates.

<sup>6</sup> It’s hard to precisely categorize Carpenter’s position. On the one hand, she says that Personalists would not go so far as to endorse “classical metaphysical emergentism” (Carpenter 2015, fn. 112). On the other hand, Carpenter says that Personalists would reject Derek Parfit’s view that “a person just is a particular brain and body, and such a series of interrelated events” and “though persons exist, we could give a complete description of reality without claiming that persons exist” (*ibid.*, fn. 69). These commitments seem to make her reconstruction of the Personalist position close to a kind of emergentism.

<sup>7</sup> Coseru (2020, 173) also says: “the personalist takes [mental states] to possess novel causal powers that cannot be ontologically analyzed in terms of the properties of the physical states upon which they supervene.”

<sup>8</sup> Non-reductive physicalism and emergentism may not be distinct positions. According to Jaegwon Kim (1998, 228-9), the main difference between emergentism and non-reductive physicalism is that emergentism offers an account about the relationship between mental and physical properties, rather than merely affirming the negative thesis that mental states are irreducible to brain states. So, in this sense, they’re compatible. For this reason, I’ll focus on emergentism in this paper. For a further argument that non-reductive physicalism collapses into emergentism, see Crane (2010).

important. Carpenter (2015, 30, 25) says: “it is because this order of changes really belongs together in a distinctive way that we are able to conceive of the organism as growing” and that “the very notion of development picks out an internal coherence through change, independently specifiable and indeed making sense as development only by reference to the previous stages.” Carpenter suggests that the causal series that constitutes a child inherently belongs together. These series manifest a kind of coherence that distinguishes a person from a mere object, like a chariot. The *dharmas* that constitute persons have special connections of belongingness that chariots and other entities lack.

Carpenter invokes the notions of emergent functionality and unity to explain the personalist view. Consider the fact that it’s hard to explain how actions are possible without mentioning persons. Here’s a dramatic example: murder. In the *Sāṃmitīyanikāya Śāstra*, one of the few surviving Pudgalavāda texts, we get the following argument: “If non-self were really the non-existence of the self, there could not be the killing beings nor the killer would have anything killed” (Venkataramanan 1953, 177). The point here seems to be that, if we omit talk of persons, then it’s impossible to describe what murder involves. After all, if we avoid reference to persons, then murder only involves the rearrangement of aggregates by other aggregates. Clearly though, something in this description is lost. Aggregates lack the ability to murder. Only persons can do that. Like all actions, killing involves ineliminable reference to a higher-order entity—a person—that commits an action.

Another motivation for the Personalist position has to do with responsibility and karma. Personalists believe that we need to posit the real existence of persons in order to explain how karma and responsibility are possible (Carpenter 2015, 27-30). Suppose that a person, Devadatta, commits a wrongful act. This act accrues negative karma, which will result in a bad rebirth. If Buddhist Reductionism is true, then it’s false that the person who committed the wrongful act is the same person as the one who undergoes a bad rebirth. This is so because it’s false that there are any persons to begin with. There are only ever-changing collections of *dharmas*. And the same point applies to good deeds. If Devadatta’s actions accrue positive karma, then another aggregate of *dharmas* receive the fruits of these good actions.

But there’s now a lack of fit between Devadatta’s actions and the consequence of his actions. It seems inappropriate for another aggregate of *skandhas* to bear the karmic costs or benefits of Devadatta’s deeds. As the *Sāṃmitīyanikāya Śāstra* puts the point: “if the self were totally unreal, then...[t]here would be neither the doer nor the deed, nor any result thereof.... There being no deed and no results, there would not be either birth or death” (Venkataramanan 1953, 177-178). Even if we reject the notion of karma, we can recast this concern in the language of moral responsibility and desert. It’s false that Devadatta can be morally responsible for anything if he’s unreal. Yet the Personalists can accommodate this concern. Because the *pudgala* is ultimately real, it can be the bearer of moral responsibility and desert, unlike an aggregation of impermanent *dharmas*.

So, one philosophical reconstruction of the Personalist position goes like this. First, we should accept the existence of persons in order to explain how action, responsibility,

karma, and development are possible. Mere aggregations of *dharmas* fail to make sense of these phenomena. Second, the only way to explain how persons can ultimately exist is through emergence. If emergentism is true, then mereological nihilism is false and some entities are irreducible to their constituents. Persons may be one such entity. These strands of argument suggest that we have good reason to believe that emergent persons exist.

## 2. THREE EMERGENTIST INTERPRETATIONS OF BUDDHIST PERSONALISM

Let's suppose then that Buddhist Personalism is committed to emergentism. What though does it mean to say that persons are emergent properties of psychophysical elements? There are different understandings of emergence. In this section, I'll examine some of them and consider whether they can justify the Personalist position.

One kind of emergence is weak emergence. Let's say that a property is weakly emergent if this property is unpredictable or unexpected given the properties of their constituents (Bedau 1997). Many complex systems are like this. Consider weather systems. We might understand all of the basic physics about the weather. Perhaps we fully understand in principle how wind, rain, clouds, and other components of the weather work. But we can't fully predict the weather. It has properties that are unexpected despite our understanding of the fundamentals. Other examples of weak emergence might include connectionist networks. Connectionist networks are composed of simple logic units that are connected with one another and these connections are modified by weights. Although the components of connectionist networks are well-understood, the properties that emerge from these networks can be surprising.

Weak emergentism won't support Buddhist Personalism though. This is so because weak emergence is an epistemological notion, not a metaphysical one. Properties are weakly emergent if they're unpredictable in light of the property of low-level properties, where "unpredictable" makes references to the epistemic capacities and expectations of the observer. But weakly emergent properties are still in principle deducible from the properties of its lower-level properties. Consider the weather again. It's hard to predict the weather. But suppose that we understand all of the physics relating to the weather, we have perfect knowledge of all of the facts concerning the weather, and we have unlimited computational power. In other words, imagine that we have godlike omniscience. Well, we could predict the weather then. We could deduce all of the facts about weather systems from lower-level properties.

Similarly, let's suppose that persons are weakly emergent. This just means that it's difficult to predict the properties of persons from knowledge of their more fundamental properties. However, it's in principle possible to deduce all of these facts from facts about the *skandhas* along with whatever laws govern them. This is consistent with Buddhist Reductionism because Buddhist Reductionism is an ontological position about what exists, not an epistemological view about what's predictable or unexpected about persons. So, if Buddhist Personalism only implies weak emergence, then it's



compatible with Buddhist Reductionism. We'll need another understanding of emergence in order to distinguish Buddhist Personalism from Reductionism.<sup>9</sup>

Another possibility is epiphenomenal emergence. Let's say that person-properties P are epiphenomenally emergent if P are metaphysically irreducible to lower-level properties and that P lack causal powers. On this view, persons are ultimately real entities. They also have properties that are distinct from those of the *skandhas* that constitute them. Yet person-properties are causally inert. They don't affect the *skandhas*. If Buddhist Personalists accept epiphenomenal emergence about persons, then we can distinguish their view from Buddhist Reductionism. Epiphenomenal emergence is inconsistent with mereological nihilism because it says that some composite properties ultimately exist. They just happen to lack causal powers.<sup>10</sup>

But epiphenomenal emergence doesn't seem like an attractive option for Buddhist Personalists. There are several problems. One long-standing problem with epiphenomenalism is how we could acquire knowledge about epiphenomenal properties. While properties in the world affect epiphenomenal properties, these epiphenomenal properties fail to affect anything. Yet it's plausible that, in order to know about something, we need to causally interact with it. This is how we usually acquire knowledge about the world. To know what's in front of you, photons need to hit your eyes and have causal effects on your brain. Buddhist epistemologists, such as Dignāga and Dharmakīrti, would agree that valid knowledge requires causal interaction.<sup>11</sup> So, if persons are epiphenomenal, then it's uncertain how we could know anything about them.

Another question about the epiphenomenal emergence of persons is: what novel properties could persons possess if they lacked causal powers? Consider an analogy with phenomenal properties or qualia. Some philosophers endorse epiphenomenalism about qualia. Epiphenomenalism about qualia says that phenomenal properties have distinctive properties—they have the property of what-it-is-to-be-likeness. Yet these properties play no causal role in the world. Maybe it's coherent to hold that qualia have properties and lack causal powers. But is this coherent for persons? What properties would persons have if they had no causal powers? Consider Carpenter's reconstruction of Buddhist Personalism. She claims that persons have the property of emergent functionality, such as the property of being able to perform actions, even though the aggregates don't have this power. But the capacity to perform actions seems to be a causal power. If we rule these properties of persons out, it's hard to see what's left.

---

<sup>9</sup> My argument also applies to Hayashi (2016). Hayashi argues that Vasubandhu endorses weak emergence about persons and he contends that this makes Vasubandhu's position distinct from Siderit's reductionism. If my argument here is correct, then there's no real difference between Vasubandhu and Siderit's views.

<sup>10</sup> I'm unaware of any Buddhist philosopher who endorses epiphenomenal emergence as a reconstruction of the Personalist position. But I'm discussing this view for the sake of completeness. I want to consider all possible emergentist views that one could adopt and see if any of these views can help justify Buddhist Personalism.

<sup>11</sup> More generally, the Pramāṇa theorists, both Buddhist and non-Buddhist, agreed that at least one important way of acquiring knowledge is through perceptual awareness, and this required contact between an object and a sense faculty. See: Dunne (2004, 23).

Furthermore, many Buddhist philosophers consider causal efficacy to be a necessary and sufficient condition for existence.<sup>12</sup> On their view, only properties that have causal powers are real. If persons lack causal efficacy, then they would fail this test. So, Buddhist philosophers who endorse the causal efficacy criterion of reality should conclude that epiphenomenal persons are unreal.

Buddhist Personalists seem to only have one option left: strong emergentism. I'll say that properties are strongly emergent if they arise from lower-level properties<sup>13</sup> and they possess novel causal powers that are irreducible to the causal powers of the lower-level properties.<sup>14</sup> Strong emergence is usually thought to require downward causation. By this, I mean that the emergent property can exert a causal influence on lower-level properties. Return to the example of phenomenal properties. Some philosophers think that phenomenal properties are strongly emergent. They claim that, when we organize physical objects in the right way, phenomenal properties emerge. But, unlike epiphenomenal emergentists, strong emergentists about phenomenal properties want to say that these properties have irreducible causal powers. They can affect the physical processes from which they arise. Now, suppose that persons are strongly emergent. If so, then persons emerge from lower-level properties, such as psychophysical elements. Persons also have novel causal powers that none of their psychophysical elements possess, like the power to perform actions, and these powers can impact the *dharma*s that serve as their base.

Strong emergence looks like the right kind of emergence that Buddhist Personalists need to distinguish their view from Buddhist Reductionism. Strong emergence can also explain the properties that Personalists sometimes attribute to persons. Another advantage of the strong emergentist reading is that it can support the Pudgalavādins' doctrine that persons were inexplicable (*avaktavya*). Remember that the Pudgalavādins thought that persons were neither identical with nor different from the *skandhas*. Instead, the relationship of the person to the *skandhas* is inexplicable. If persons are strongly emergent, then the fact that persons emerge from the *skandhas* is a brute fact. There's no way to explain this emergence (Kim 2005, 103). It's unclear whether this is roughly what the ancient Pudgalavādins had in mind when they claimed that persons are inexplicable. Nonetheless, the strong emergentist reading of Buddhist Personalism gives us one argument for this claim.

However, the view that persons are strongly emergent faces problems. These problems have to do with the relationship between persons and the *skandhas*. What's the nature of this relationship? Perhaps the Personalists are right that this relationship is inexplicable. Yet we can still ask whether the person supervenes on the *skandhas* from which it emerges or not. Let's say that properties P1 supervene on properties P2 just in case any difference in P1-properties requires a difference in P2-properties. More simply: no P1-difference without a P2-difference. Supervenience is evidence of

<sup>12</sup> Siderits (2016, 125) refers to the causal efficacy criterion of reality as a "pan-Buddhist position."

<sup>13</sup> What does it mean to say that properties arise from lower-level properties? One interpretation is that this is a relation of nomological necessity. It's a law of nature that, if we organize lower-level properties in a certain way, then a higher-order property emerges from them.

<sup>14</sup> This definition of strong emergence draws on Chalmers (2006) and Baysan (2020).

dependence and determination. If P1 supervenes on P2, then P1 likely depends on P2 or is determined by P2 in some way (Kim 2005, 34).

Again, consider mind-body supervenience. Most people accept that the mind supervenes on the brain. So, changes in the mind require changes in the brain. When your mind goes from a state of calm at one point to a state of rage a minute later, there's some difference in your brain at these times. This indicates that your mental properties depend on the physical properties of your brain. What's the nature of this dependence? There are different possibilities. Perhaps your brain states cause mental properties, constitute them, or mental properties are identical to brain states.

Can we describe similar relationships between persons and their *skandhas*? Let's suppose that person supervenes on the *skandhas*. So, a difference in person-properties requires a difference in *skandha*-properties. This suggests that the person depends on the *skandhas*. And, if the person depends on the *skandhas*, then some of the properties of the person are likely determined by the properties of the *skandhas*.<sup>15</sup>

Yet Buddhists agree that the *skandhas* are in constant flux. Consider one of the Buddha's most influential arguments against the enduring self. In the *Samyutta Nikāya*, the Buddha considers each *skandha* and contends that these *skandhas* are impermanent and tend toward destruction. But, for something to be the self or serve as its basis, this thing must be enduring. The Buddha concludes that the *skandhas* cannot be the self. For my purposes, the important point is that the *skandhas* are transitory. If the *skandhas* are impermanent and the person supervenes and depends on them, then why isn't the person impermanent as well? At first glance, it seems that the person should go out of existence along with the *skandhas* on which it depends. So, person-*skandha* supervenience and dependence give us reason to believe that the person is fleeting.<sup>16</sup> This is consistent with punctualism about persons, the view that persons last for only a short, even momentary, amount of time. But a radically impermanent person is unable to fulfill the functions that Personalists want the person to perform. A punctualist person lacks the capacity to bear responsibility, karmic consequences, and develop over time.

An objector might argue in response: "something can change without going out of existence. If the *skandhas* are in flux, that just means that the person changes over time. But this change is consistent with the continued existence of the person." There are two problems with this response. First, this response commits the Buddhist Personalist to a substantialist ontology where an entity can have an essence that survives changes to its characteristics. This would represent a break from the Buddhist tradition of rejecting substance-attribute ontologies (Carpenter 2014, 43-6). Most Buddhist philosophers, at least in the Abhidharma tradition, argue that what we ordinarily call substances are

---

<sup>15</sup> The relationships between the concepts of supervenience, dependence, and determination are controversial. It seems possible for properties P1 and P2 to covary without P1 being dependent on or determined by P2. Nonetheless, it's plausible that supervenience is defeasible evidence of dependence relations between P1 and P2. For discussion, see Kim (1984).

<sup>16</sup> Some of the ancient Pudgalavādins seem to have been aware of this fact. The Treatise of the Saṃmitīya School notes: "if the person were identical with the aggregates, when the aggregates disappear or appear, the person should disappear or appear" (Thiền Châu 1999, 106).

bundles of *dharmas* and that *dharmas* cannot survive changes to their characteristics. Suppose though that we accept a substance-attribute metaphysics and that a person is a substance and *skandhas* are its attributes. It's still plausible that *enough* changes to the attributes of a substance will cause that substance to go out of existence and be replaced with something new. So, even if a substance can survive some changes to its attributes, it's unclear how long persons can survive given that the *skandhas* are continually in flux.

To avoid these problems, a Buddhist Personalist might want to deny person-*skandha* supervenience and dependence. She might argue that it's false that the properties of persons covary with the *skandhas*. And, although the *skandhas* come and go, the person endures. This interpretation has advantages. Perhaps it can make sense of personal identity over time and rebirth. Assume that all of the *skandhas* that make up a person change or perish. If person-*skandha* supervenience and dependence are false, then the person can continue to exist. She can also be reborn, as the person doesn't depend on the *skandhas* from a past life. A person who endures and is reborn can bear karmic consequences of her actions and make progress along the spiritual path across lives.

Suppose that we reject the view that the properties of the person covary with or depend on the properties of the *skandhas*. But now the relationship between the person and the *skandhas* is mysterious. Imagine that the composition of the *skandhas* changes radically or that they're destroyed. If we reject persons-*skandha* supervenience and dependence, then the person may remain unchanged despite these changes to the *skandhas*. So in what sense are persons related to the *skandhas* at all?<sup>17</sup>

Let's consider some different ways in which the person could be related to the *skandhas*. Here's one: changes in the properties of the person only *sometimes* covary with the properties of the *skandhas*. Call this view: option 1. If option 1 is true, we need to explain why person-properties only sometimes covary with *skandha*-properties. Suppose the properties of persons often covary with the properties of the *skandhas* on which the person depends, but this correlation is imperfect. Sometimes changes to the *skandhas* fail to coincide with changes to the person. What explains this correlation between the *skandhas* and the person though? Either the connection between the *skandhas* and the person is inexplicable or there are laws that govern the relationship between the *skandhas* and the person that explain how they're related. But then these laws would just reinstate person-*skandha* supervenience, as these laws would specify the conditions under which the properties of the person covary with those of the *skandhas*. So, option 1 seems inconsistent with the rejection of person-*skandha* supervenience.

Consider another option: the properties of the person never covary with the *skandhas*. Call this: option 2. If the properties of the person never covary with the

---

<sup>17</sup> My argument has some resemblances with Vasubandhu's argument in the ninth chapter of the *Abhidharmakośabhāṣya*. Vasubandhu argues that either the person is identical with the aggregates or separate from them. If identical with the aggregates, then Personalists agree with Vasubandhu's (arguably) reductionist view. If separate from the aggregates, then the Personalists are committed to a soul or substantially real person (Duerlinger 2003, 73).

properties of the *skandhas*, then no change to the psychophysical elements coincides with any changes to the person. And this suggests that it's false that person depends on the *skandhas* in any way. The *skandhas* float entirely free of the person. But now it's unclear how the person could cause anything. For some entity to be strongly emergent, that entity must have causal powers. If we go with option 2, then we must also deny that the *skandhas* could causally influence the person. After all, if changes in property P1 never coincides with changes in property P2, then it's hard to see how P1 could cause anything with respect to P2. This is so because causation requires concomitance. For instance, if brain events cause mental events, then brain events must be concomitant with these mental events. So, option 2 implies that the person is causally isolated from the *skandhas*.

A Personalist might bite the bullet and concede that the *skandhas* fail to cause anything about the person. This raises a new question. What explains when or why a person exercises its causal powers? Nothing about the *skandhas* can answer this question. The person appears to be an unmoved mover. But, if the person is isolated from the *skandhas*, then the person must lack causal powers. To see why, let's consider an influential Buddhist criticism of eternal entities, such as creator Gods and universals. Suppose that you think that God created the universe. But what created God? Nothing. God has always existed. He's eternal. Yet, to be eternal, an entity must exist outside of the causal nexus. An entity that's subject to cause and effect could decay and be destroyed given enough time. And something that exists outside of the causal nexus is unchanging. But, if God is unchanging, then how can God produce effects at any given time? If an entity produces an effect, then there must be some feature of this entity that explains why it produces this effect at this particular time. The problem is that, if an entity is unchanging, then there can be no feature of this entity that explains why it generates the effect at this particular time. It should always be producing this effect.

The same argument applies to persons that fail to supervene or depend on the *skandhas*. Suppose that persons are causally isolated from the *skandhas* and we nonetheless want to say that person A causally influenced the *skandhas* at time T1. What explains why A causally influenced the *skandhas* at T1? If the person is causally isolated, then it's unchanging. And unchanging entities lack the power to cause things at particular times. If persons lack causal powers, then it's false that they're strongly emergent. So, it's incoherent for Buddhist Personalists to both hold that persons are strongly emergent and to deny person-*skandha* supervenience and dependence.

My discussion has ranged widely in this section, so let me recap my central arguments. If Buddhist Personalism is true, then persons are either weakly emergent, epiphenomenally emergent, or strongly emergent. If persons are weakly emergent, then Buddhist Personalism is indistinguishable from Buddhist Reductionism at an ontological level. If persons are epiphenomenally emergent, then it's a mystery how we could acquire knowledge of them and the requirement that real entities be causally efficacious rules out the existence of epiphenomena persons. Finally, let's suppose that persons are strongly emergent. If so, then Buddhist Personalism must either endorse punctualism about persons or reject person-*skandha* supervenience and dependence. Punctualism fails to justify the properties of persons that Personalists want to ascribe

to them. But a view that affirms the strong emergence of the person and rejects person-*skandha* supervenience and dependence is incoherent. Thus, all of the emergentist readings of Buddhist Personalism seem untenable. This is a good reason to reject the existence of emergent persons.

#### 4. OBJECTIONS

##### *A. Causal Continuity and Belongingness*

Here's an objection to my argument against persons as strongly emergent. I argued that, if persons are strongly emergent and persons-*skandha* supervenience and dependence hold, then persons must be transitory entities, as the *skandhas* are in a state of flux. But a Personalist might reject this argument. She could argue that I've omitted a key feature of the Personalist view. On this view, persons can exist over time despite change in virtue of the fact that the *skandhas* stand in the right kind of relationship, such as a relationship of causal continuity. If the relations of continuity obtains between the parts of a causal stream, then the person that emerges from this stream also continues to exist. So, Buddhist Personalism is compatible with both the impermanence of the *skandhas* and the continuity of the person.

Yet this response falls short. The issue is that we stand in causal relationships with other people too. But it seems false that we bear relationships of personal identity with them. For example, I have caused my children to remember certain events like birthday parties and experience certain emotions like sadness and happiness. But it's false that I'm identical with my children or that they bear responsibility for anything that I've done. Or consider the fact that I'll be causally continuous with my corpse. My body and mind while alive will be causally connected in various ways with the state of my corpse after I die. It would be strange though to conclude that my corpse preserves my identity.<sup>18</sup> The upshot is that mere causal connections are insufficient. The Personalist needs something more to ground relationships of personal identity over time and karmic responsibility.

Here's another possibility. Perhaps the *skandhas* that make up the person belong together in a special way. Amber Carpenter and Jan Westerhoff make this suggestion. Carpenter (2015, 25) writes: "On my hypothesis that the *pudgala* amounts to the fact of certain elements really belonging only to each other, the claim is that this 'belonging to' is basic—ultimately true, and not explicable in any other terms, or by reference to anything else." Drawing on Carpenter's argument, Westerhoff says:

W]hat the Pudgalavāda wants to say when arguing that the person is ultimately real is that the specific kinds of causal connections that unify a person cannot be reduced to something else; in particular, they are not just normal causal relations arranged in a specific pattern. They exist *sui generis*, as causally related elements that belong together ultimately and thereby ensure that the person has a privileged place in the Pudgalavāda ontology (Westerhoff 2018, 57).

---

<sup>18</sup> I owe this point to Harris (2018, 389).

On this interpretation, special connections of “belongingness” unify persons over time and explain why facts about the continuity of persons are not merely conventional truths but ultimate ones. I’ll refer to the special connection between the *dharmas* that make up persons as “the belongingness-relation.”

It’s hard though to make sense of the belongingness-relation. Perhaps for this reason Carpenter says that this relationship is inexplicable. But let’s suppose we can make sense of this idea. We can then reconstruction Carpenter and Westerhoff’s argument as follows:

1. The set of *dharmas* that make up persons share the belongingness-relation.
2. If the *dharmas* that make up persons share the belongingness-relation, then this set of *dharmas* has the right kind of properties to justify ascriptions of personal identity, karmic consequences, and so on, over time.
3. So, ascriptions of personal identity, karmic responsibility, and so on, are justified for persons.

I’ll now raise an objection to premise 1.

To introduce my objection, I’d like to consider a curious story from the *Mahāprajñāpāramitāsāstra*, a Mahāyāna treatise that was probably composed around the 2nd or 3rd century AD and attributed to Nāgārjuna. This story will take a moment to describe. But it will ultimately show that Carpenter and Westerhoff’s proposal is untenable.

The story goes like this (Lamotte 1944, 582-586).<sup>19</sup> A man is traveling and decides to sleep in a deserted house. During the night, a demon appears and is carrying a corpse, which he sets down in front of the man. A second demon then arrives and says to the first one: “That dead man belongs to me; why are you bringing him here?” and the first demon replies: “He is my property; it is I who took him and brought him here myself.” The demons then argue over who has a rightful claim to the corpse. On demon remarks: “There is a man here and we can ask him.” The demons begin to question the man while the man thinks: “These two demons are very strong; whether I tell the truth or I lie, my death is certain; in either case, I can’t escape. What is the use of lying?” The man replies that the first demon brought the corpse.

The second demon is understandably upset at the man’s response. So, the second demon tears off the man’s hand and arm. But the first demon, perhaps out of gratitude to the man for his assistance, takes an arm from the corpse and fits it onto the man, replacing the body parts that the man just lost. The second demon proceeds to tear off each part of the man’s original body while the first demon replaces each with materials from the corpse. Once the demons have replaced the man’s whole body with the parts from the corpse, they eat the man’s original body parts. The demons then leave.

Later the man resumes his journey. After he arrives at his destination, he sees an assembly of Buddhist monks. They ask him who he is and the man responds: “I don’t even know if I am a man or not” and explains what happened. The monks say: “This man knows for himself the non-existence of a self; he will easily be liberated.” The monks declare to the man:

---

<sup>19</sup> This was translated from the French by Gelongma Karma Migme Chodron.

From the very beginning until today, your body was always without self and it is not just coming to the present moment [that that is so]; it is simply because the four great elements were combined that you thought: ‘This is my body.’ There is no difference between your previous body and that of today (*ibid.*, 586).

The narrator recounts how the man goes on to become an arhat, a Buddhist saint.

The author of the *Mahāprajñāpāramitāsāstra* uses this tale to argue against the Pudgalavāda. In particular, the author deploys this story to argue that “there are circumstances where one conceives the idea of self in reference to another. But under the pretence that there are distinctions between ‘that’ and ‘this’, one cannot say that there is a ‘me’” (*ibid.*, 586). I’m going to use this story for a similar purpose. I’ll use it to develop an argument against Buddhist Personalism. To make things easier to follow, I’ll give names to the people in the story. Let’s call the man who is traveling: Devadatta. Let’s refer to the corpse and future man as: Yajñadatta. I’ll also change some of the details and assumptions of the story as I go.

With this story in hand, I can articulate an argument against premise 1 of the above argument for Buddhist Personalism, the premise that the set of *dharmas* that make up persons share the belongingness-relation. Let’s begin with the assumption that the *dharmas* that constitute Devadatta share this relation at the beginning of this story. But what happens after the demons begin to replace Devadatta’s body?

Suppose that, at time T1, the demons have replaced 5 percent of Devadatta’s body with Yajñadatta’s corpse. This 5 percent may not share the belongingness-relation with the rest of Devadatta’s psychophysical continuum. After all, these parts came from Yajñadatta. Presumably though, Devadatta still exists after the demons replace some of his parts. We can survive the loss of some of our constituents. Consider the fact that we’d think that someone who has lost a limb and has it replaced with a prosthesis continues to exist. So, even if the belongingness-relation justifies the unity and continuity of persons, we should agree that some of a person’s *dharmas* can be replaced and that person can be continuous with her past self in the right way. This implies that premise 1 is false. It’s false that all of the set of *dharmas* that make up persons must share the belongingness-relation. But we can modify this premise in the following way. We can say that the important thing for ascriptions of personal identity, karmic responsibility, and the rest is that *enough* of the *dharmas* that constitute a person share the belongingness-relation.

But we now run into a problem. There just isn’t a clear point where Devadatta disappears and is replaced by Yajñadatta. You might judge that Devadatta disappears and Yajñadatta springs into existence once over 50 percent of Devadatta’s body has been replaced. Yet this view is strange on reflection. Let’s suppose that at the beginning of the story Devadatta uncontroversially exists. Now, imagine that the demons replace 49.999999 percent of Devadatta’s body at time T2. But suppose that the demons replace 50.000001 percent of Devadatta at T3. If you think that Devadatta still exists at T2 and he’s replaced by Yajñadatta at T3, then you should believe that a miniscule difference



in the composition of Devadatta's body makes all of the difference about whether he continues to exist.

But consider the fact that Devadatta underwent more major alterations between T1 and T2. By hypothesis, enough of Devadatta's *dharma*s share the belongingness-relation that he remains the same person at T2 as he was at T1. However, if Yajñadatta replaces Devadatta at T3, then an infinitesimally small change in his composition causes him to cease to exist. And this is hard to believe. Why would enormous changes prior to threshold fail to affect Devadatta's continuity while minute differences around the threshold result in his replacement?

In response, you might argue that the proposed threshold of over 50 percent was incorrect. Perhaps another threshold is better. But notice that we can run the same argument with any given threshold. For any threshold between the existence of Devadatta and his replacement with Yajñadatta, we can ask why astronomically tiny changes around the threshold should make the difference between existence and non-existence, continuity or non-continuity. It looks like no good answer to this question is forthcoming.<sup>20</sup>

What's the point of this argument? This argument suggests that, whatever the belongingness-relation amounts to, it fails to justify the unity and continuity of persons over time. The belongingness-relation by itself is insufficient. We need to supplement it with an account of how much of the belongingness-relation must obtain or perhaps the distinctive way in which it must hold over time. But any such account faces the problem of arbitrariness. That is, any specific proposal or threshold that specifies how much of the belongingness-relation is necessary looks arbitrary. So, we should reject the view that the belongingness-relation can justify ascriptions of personal identity and responsibility.

Notice though that, unlike Personalists, Buddhist Reductionists can handle the case of Devadatta and Yajñadatta with relative ease. Buddhist Reductionists would argue that the problems of indeterminacy and arbitrariness only arise because Personalists are assuming that persons are ultimately real and distinct. Once we reject these assumptions, the problems of arbitrariness and indeterminacy disappear. There's no sharp boundary between Devadatta and Yajñadatta because they're unreal to begin with. They're only conventional constructions. Furthermore, we'd expect conventional constructions to be indeterminate. This is so because our concepts and conventions map imperfectly onto the world and this mismatch generates indeterminacy. Thus, our concept "heap" maps indeterminately onto a number of grains of sand. The Buddhist Reductionist says that persons are conventions in the same way that heaps are. So, it stands to reason that persons are indeterminate as well.

How should we resolve indeterminacy in the personhood convention? Buddhist Reductionists would say that we should do whatever works. We should adopt any rule to resolve indeterminacy that would best promote welfare, reduce suffering, or achieve

---

<sup>20</sup> Readers may recognize this argument as a version of Derek Parfit's combined spectrum. For further development and defense of this argument and responses to objections, see Parfit (1986, 236-243), Shoemaker (2002), and Torin and Stuart (2004).

other useful aims. Imagine that the demons have replaced 65 percent of Devadatta's body with Yajñadatta's body. And let's suppose that, at this point, Yajñadatta firmly believes that he's a distinct person and Yajñadatta goes into a furious rage at any suggestion that he's Devadatta. It's useful to now adopt the rule of treating Yajñadatta as a distinct person from Devadatta. Doing so will reduce suffering overall. So, we've resolved the indeterminacy between Devadatta and Yajñadatta pragmatically. In this respect, Buddhist Reductionism has a key advantage over Buddhist Personalism. Buddhist Reductionism can handle indeterminacy about persons, whereas Buddhist Personalism appears to face intractable problems in doing so.<sup>21</sup>

### *B. The Advantages of Buddhist Personalism over Buddhist Reductionism*

But perhaps Buddhist Personalism has other crucial advantages over Buddhist Reductionism. And maybe these benefits justify accepting the Personalist position, despite the problems that it confronts. For example, a Personalist might argue that Buddhist Reductionism is unable to make sense of prudential concern and moral responsibility. According to Buddhist Reductionists, persons are unreal. They're simply convenient ways of referring to aggregations of psychophysical elements. But unreal persons lack the capacity to bear moral responsibility over time. Furthermore, if persons are fleeting streams of *dharma*s, then it's unjustified for us to have any special concern for our own interests and lives. The Personalists will argue further that only ultimately existent persons can bear moral responsibility or have justified prudential concern.

Buddhist Reductionists have a response to this line of argument. They offer a consequentialist justification of self-regard and responsibility. Reductionists argue that it can bring about good consequences if people blame and punish others for their crimes and reward and praise them for their good deeds. It's true that a murderer never ultimately deserves his punishment, as murderers lack ultimate existence. Nonetheless, it's a good thing to punish a murderer because this will likely prevent future suffering by protecting others from this person and by deterring other potential murderers. And Buddhist Reductionists tell a similar story about self-concern. On the whole, it can be beneficial if people identify with their lives and pursue their own interests, within limits. This will lead them to exercise prudence, which will also reduce aggregate suffering in the long-term. Or so Buddhist Reductionists claim.<sup>22</sup>

---

<sup>21</sup> Some Buddhist philosophers endorse enactivism about persons. Enactivists, such as Evan Thompson (2014), argue that persons and selves are processes. More precisely, persons are dynamic and autonomous systems. Moreover, some enactivists are strong emergentists. For instance, Matthew Mackenzie (2011) argues that persons are emergent processes and that they have novel causal powers that are irreducible to their base. My arguments against strong emergentism about persons also seem to apply to MacKenzie's position. At first glance, it's unclear how the person can persist if it's a constantly changing dynamic system. Enactivists might respond that the person can persist because past and future stages of the person share relations of causal continuity or the belongingness-relation. If so, then my arguments in this section tell against the enactivist view as well.

<sup>22</sup> Buddhist Personalists can also argue that we experience ourselves as agents who make choices. This experience of agency is so fundamental and pervasive that we must assume that we're unified agents

A Personalist might dispute various aspects of the Reductionist story. I suppose that a Personalist could argue that it's false that the practices of responsibility and self-concern have these beneficial outcomes. But this line of argument is unlikely to be conclusive without detailed empirical investigation. However, Buddhist Personalists might have a better avenue of attack available to them. They can contend that the Buddhist Reductionist account is circular in some way. Carpenter suggests this objection. She argues that "it is convenient to grasp a group of aggregates together as a person only given some ends; and these ends are conceivable in the first place only by thinking in terms of persons" (Carpenter 2015, 15). Her argument goes like this. Buddhist Reductionists say that we group collections of aggregates as persons because doing so is useful and serves our purposes. Yet we're unable to assume that there are purposes without referring to the existence of persons. Thus, Buddhist Reductionists are caught in a circularity.

But I think that Reductionists can escape this circularity. Here's how. Buddhist Reductionists can maintain that some purposes avoid presupposing the existence of persons. Buddhists in the Abhidharma tradition posit a distinction between *dharmas* with wholesome (*kusala*) moral valences and *dharmas* with unwholesome (*akusala*) valences. For the purposes of illustration, let's suppose that pain states are intrinsically bad *dharmas*. Buddhist Reductionists say that it would be better if there were fewer intrinsically bad *dharmas* like pain states around. Notice that this claim refrains from presupposing the existence of persons in any obvious way. It just says that there should be less pain in the world. Now, assume that it would bring about less pain if we treated certain aggregates of *dharmas* as if they were distinct persons. If true, we've arrived at a non-circular justification for individuating persons. Perhaps this justification is implausible in the end. But it doesn't appear circular.

A Personalist might respond that all experiences must have subjects and so it's incoherent to say that there are pain states without a subject that experiences them. The question of whether experiences require subjects is an old one and I won't be able to resolve it here. But I'd like to point out that, even if experiences require subjects, it doesn't follow from this that the subject is a person. Maybe the subject is some kind of witness-consciousness that merely registers experiences as they come and go. Or, if the subject is a person, perhaps it's only a punctualist entity that goes in and out of existence in a single moment, which is not the kind of person that Buddhist Personalists want to endorse. In any event, even if we grant that experiences require subjects, this

---

rather than a series of psychophysical elements (Korsgaard 1989). Perhaps Buddhist Personalism can explain this experience of agency: we have the experience of agency because we are agents. However, I'm skeptical that this argument succeeds. First, let's assume that we must experience ourselves as agents. It doesn't follow that we are in fact agents. For example, we often can't help but see certain optical illusions, such as the peripheral drift illusion. Yet we know that these illusions are distortions created by our visual and cognitive processes. Similarly, even if we must experience ourselves as agents, this experience could be a cognitive distortion. Second, it's unclear whether we must always experience ourselves as agents. Some Buddhists contend that it's possible to experience a selfless and agentless form of consciousness in deep meditative states. For additional responses to "Kantian" arguments for the self and persons, see: Siderits (2016, 81-92). I want to thank an anonymous reviewer for raising these issues.

fails to show that purposes require persons in any robust sense. So, this rejoinder falls short of establishing that Buddhist Reductionism is circular, at least not in the way that Carpenter alleges.<sup>23</sup>

## 6. CONCLUSION

In this paper, I've argued against Buddhist Personalism. More precisely, I've argued against one philosophical reconstruction of Buddhist Personalism that construes persons as emergent from the *skandhas*. Indirectly, my argument also supports Buddhist Reductionism. If you're deciding whether to accept either Buddhist Reductionism or Personalism, then the fact that Buddhist Personalism faces key objections and Reductionism can avoid similar problems is a reason to favor Buddhist Reductionism. But my conclusion is tentative. I've only evaluated one interpretation of Buddhist Personalism. So, even if the emergentist version of Personalism is untenable, critics of Buddhist Reductionism might still be able to find an alternative version of Buddhist Personalism that's defensible after all.

## ACKNOWLEDGEMENTS

I thank the editors of this journal and anonymous reviewer for their helpful comments. I also am grateful to Ryan Davis and Jessica Flanigan for their help and comments on the ideas in this paper.

## REFERENCES

- Baysan, Umut (2020), "Causal Emergence and Epiphenomenal Emergence", *Erkenntnis* 85.4: 891–904.
- Bedau, Mark A. (1997), "Weak Emergence", *Noûs* 31: 375–99.
- Benovsky, Jiri (2018), *Eliminativism, Objects, and Persons: The Virtues of Non-Existence* (New York: Routledge).
- Carpenter, Amber (2014), *Indian Buddhist Philosophy* (New York: Routledge).
- Carpenter (2015), "Persons Keeping Their Karma Together: The Reasons for the Pudgalavada in Early Buddhism", in *The Moon Points Back*, edited by Koji Tanaka, Yasuo Deguchi, Jay L. Garfield, and Graham Priest (New York: Oxford University Press), 1-44.

---

<sup>23</sup> Coseru (2020, 172) suggests that another argument for Buddhist Personalism is the "too many thinkers argument." Roughly speaking, this argument says that we need to posit the existence of emergent persons in order to avoid the multiplication of conscious subjects. In other words, without emergent persons, there would be too many thinkers in one body. I lack the space to adequately address this argument here. But for a detailed reply, see Benovsky (2018, 65-77).

- Chalmers, David J. (2006), “Strong and Weak Emergence,” in *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion*, edited by P. Davies and P. Clayton (New York: Oxford University Press), 244-52.
- Coseru, Christian (2020), “Reasons and Conscious Persons”, in *Derek Parfit’s Reasons and Persons: An Introduction and Critical Inquiry*, edited by Andrea Sauchelli (New York, NY: Routledge), 160-86.
- Crane, Tim (2010), “Cosmic Hermeneutics vs. Emergence: The Challenge of the Explanatory Gap”, in Cynthia Macdonald & Graham Macdonald (eds), *Emergence in Mind* (Oxford: Oxford University Press), 22-33. <<https://doi.org/10.1093/acprof:oso/9780199583621.003.0002>>
- Duerlinger, James (2003), *Indian Buddhist Theories of Persons: Vasubandhu’s Refutation of the Theory of a Self* (New York, NY: Routledge).
- Dunne, John D. (2004), *Foundations of Dharmakīrti’s Philosophy* (Boston: Wisdom Publications).
- Ganeri, Jonardon (2015), *The Self: Naturalism, Consciousness, and the First-Person Stance* (New York, NY: Oxford University Press).
- Goodman, Charles (2005), “Vaibhāṣika Metaphoricalism”, *Philosophy East and West* 55.3: 377–93.
- Goodman, Charles (2014), *Consequences of Compassion: An Interpretation and Defense of Buddhist Ethics* (New York, NY: Oxford University Press).
- Harris, Stephen E. 2018. “Promising Across Lives to Save Non-Existent Beings: Identity, Rebirth, and the Bodhisattva’s Vow.” *Philosophy East and West* 68 (2): 386–407.
- Hayashi, Itsuki (2016), “Persons as Weakly Emergent: An Alternative Reading of Vasubandhu’s Ontology of Persons”, *Philosophy East and West* 66.4: 1218–30.
- Kim, Jaegwon (1984), “Concepts of Supervenience”, *Philosophy and Phenomenological Research* 45.2: 153–76.
- Kim, Jaegwon (1998), *Philosophy of Mind* (1st edition) (Boulder, CO: Westview Press).
- Kim, Jaegwon (2005), *Physicalism, or Something Near Enough* (Princeton: Princeton University Press). <<https://doi.org/10.2307/j.ctt7snrs>>
- Korsgaard, Christine M. (1989), “Personal Identity and the Unity of Agency: A Kantian Response to Parfit”, *Philosophy & Public Affairs* 18.2: 101–32.
- Lamotte, Étienne (trans. 1944), *The Treatise on the Great Virtue of Wisdom of Nāgārjun* (Louvain: Bibliothèque du Muséon).
- Lusthaus, Dan (2009), “Pudgalavāda Doctrines of the Person”, in *Buddhist Philosophy: Essential Readings*, edited by William Edelglass and Jay Garfield (New York: Oxford University Press), 275-96.
- MacKenzie, Matthew (2011), “Enacting the Self: Buddhist and Enactivist Approaches to the Emergence of the Self”, in *Self, No Self? Perspectives from Analytical, Phenomenological, and Indian Traditions*, edited by Mark Siderits, Evan Thompson, and Dan Zahavi (New York: Oxford University Press), 239-273.
- Parfit, Derek (1986), *Reasons and Persons* (Oxford: Oxford University Press).
- Priestly, Leonard (1999), *Pudgalavada Buddhism: The Reality of the Indeterminate Self* (Toronto, Ontario: Centre for South Asian Studies, University of Toronto).

- Shoemaker, David W. (2002), “The Irrelevance/Incoherence of Non-Reductivism About Personal Identity”, *Philo* 5.2: 143–160. <<https://doi.org/phil02002529>>
- Siderits, Mark (2016), *Personal Identity and Buddhist Philosophy* (2 edition) (Burlington, VT: Routledge).
- Thiền Châu and Bhikshu Thích (1999), *Literature of the Personalists of Early Buddhism*, trans. by Sarah Boin-Webb (Delhi: Motilal Banarsidass Publications).
- Thompson, Evan (2014) *Waking, Dreaming, Being: Self and Consciousness in Neuroscience, Meditation, and Philosophy* (New York: Cambridge University Press).
- Torin, Alter, and Rachels Stuart (2004), “Epistemicism and the Combined Spectrum”, *Ratio* 17.3: 241–255. <<https://doi.org/10.1111/j.0034-0006.2004.00252.x>>
- Venkataramanan, Krishniah (1953), “Sāṃmitīyanikāya Śāstra”, *Visva-Bharati Annals* 5: 155–243.
- Westerhoff, Jan (2018), *The Golden Age of Indian Buddhist Philosophy in the First Millennium CE* (New York, NY: Oxford University Press).