

Amy Donahue
Infinity Fellowship Contribution
University of Hawai'i at Manoa
Spring, 2003

Yoga, Free Will and the *Puruṣa-Prakṛti* Problem

It hardly makes sense to talk of a Yoga mind-body problem, since this problem, as we know it in the West, is loaded with peculiarly Western ontological assumptions. Nonetheless, Yoga's dualist ontology, as presented by Patañjali in the *Yoga-Sutra-s*, with its causally conservative *guṇa*-metaphysics, seems to leave no room for free, conscious, non-causally determined will. Insofar as this is the case, there would also seem to be no place for moral action within its soteriology. Therefore, though Yoga does not face a mind-body problem, it faces a comparable difficulty, namely a *puruṣa-prakṛti* problem, and this strikes at the heart of the tradition's moral theory.

Simply put, the difficulty presented by the mind and body in the West is one of interaction. If all material events have only material causes and effects, and if mental events are not material, then mental and material events cannot interact. Yet, we believe that the human mind is non-material and that the human body is material. We also believe that the human mind interacts with the human body. But by the above reasoning, we believe that the human mind cannot interact with the body. Therefore, one of our beliefs is false.

Responses to this problem have been classed in six general categories. The first two include varieties of materialism – eliminativism, which denies the existence of mental states and claims that only material states exist, and reductivism, which acknowledges the existence of mental states but says that these are ultimately reducible to physical states. The third and fourth are varieties of idealism. The third, solipsism or spiritualism, denies the existence of physical states and claims that only mental states exist. The fourth, reductive idealism, acknowledges the existence of physical states but says that these are ultimately reducible to mental states. The fifth, neutral monism, says that there are states but that these are essentially neither physical nor mental. The sixth, dualism, holds on to the mental/physical divide and instead abandons either the claim that they cannot interact

or the claim that they do interact. We can classify these dualist theories as interactionism and parallelism, respectively.

It is impossible to say which stand Yoga would take in this debate. As indicated above, Yoga does not share the presuppositions that have placed the West into this particular ontological muddle. As we will see, the Western terms ‘mind’, ‘body’, ‘material’ and ‘mental’ do not have exact counterparts in Yoga philosophy.

Sāmkhya-Yoga posits the existence of two fundamentally distinct natures – *puruṣa* and *prakṛti*. The distinction between *puruṣa* and *prakṛti* is best understood, I find, in terms of the distinction between subject and object. *Puruṣa* is that which sees, while *prakṛti* is that which is seen. *Prakṛti* is unconscious and in constant flux, while *puruṣa* is unchanging consciousness. Thoughts, desires, perceptions and emotions, therefore, are all aspects of *prakṛti*-c experience, since they are observed and known, while that which observes these thoughts desires, perceptions and emotions is *puruṣa*. Experience never happens to *puruṣa*. *Puruṣa* never thinks, desires, perceives or feels itself. It is always in an important sense removed, untouched, from the objects of its awareness. It is like a kaleidoscope-watcher. *Prakṛti* dances before *puruṣa*, but *puruṣa* does not join in the dance. It only appears to do so in *samsara*.

If we accept Yoga’s dualism, a grammatical change is in order. Rather than saying that “I x”, where x is an experiential action verb, like “think”, “feel”, “perceive”, etc., it will be more appropriate to say “x is occurring”. So if I were to feel joyful, perhaps after receiving good news about a friend, I would express this appropriately as “joy is arising, and this joy that is arising is arising as a result of that immediately prior occurrence of ‘good news’.”

In Yoga, the term ‘*citta*’ is most aptly translated as ‘mind’. As in the West, it is an aggregate concept – it refers not to a single entity but to several collective ‘mental’ functions. Roughly, these are the super-ego (*asmitā-mātra*), intellect (*buddhi*), phenomenal ego (*asmitā*), volitions, dispositions and habits (*samskāra*-s), and thoughts, perceptions and sensations (the products of the functioning of the six sense-receptors,

including mind or *manas*).¹ Each of these is intentional, or object-directed, and composed of the *guṇa*-trifold, which consists of three affective threads – *sattva*, *rajas* and *tamas*. *Sattva* has the affective character of joy and lucidity. *Rajas* has the affective character of discomfort and activity. *Tamas* has the affective character of torpor and rigidity. All of these *citta*-states are more or less dominated by *sattva*, though all *guṇa*-threads are present to some degree in all of *citta*.

Unlike the western ‘mind’, no *citta*-state is independently conscious. They all have a phenomenal character, or feel, but they do not witness this feel. Rather they are witnessed by *puruṣa*, which is not part of the *citta*/mind-complex. In the West, there is no comparable distinction to that which exists between *citta* and *puruṣa*. Contemporary Western philosophers tend to identify phenomenal conscious with *qualia*. In contrast, Yoga has put forth a metaphysics in which pure consciousness is separate from *qualia* – it is that which observes affective experience.

Unlike mind in the Western mind-body debate, *citta*-states are causally determined. The varying balances of their constituent *guṇa*-threads determine their affective, structural and behavioral characteristics at any given moment. The more volatile a state of mind, the greater is its preponderance of *rajas*. Further, *citta*-states are not exclusively private. Highly evolved (or, better in the Yoga context, ‘in-volved’) individuals on the Yoga soteriological path are able to observe the *citta*-states of all beings.

Therefore, while the Western concept of ‘mind’ typically refers to phenomenal, freely determinable and private states, *citta* refers to states that are unconscious, causally determined and non-private. In the West, mental states are considered subjective. In Yoga, *citta*-states are objective.

It is similarly difficult to locate the Western ‘body’ in Yoga. There is no substantial difference between *citta* and the body’s sense and motor organs, which are less subtle evolutes of the same substance, *prakṛti*, that forms *buddhi*, *asmitā* and *manas*. While they

¹ Dasgupta, Surendranath, Yoga as Philosophy and Religion, Motilal Banarsidass (Delhi), 1924, pp. 48-59.

are characterized less by *sattva* and more by *tamas*, they are still *prakṛti*-c substances. Like *citta*, Yoga's human body is a manifestation of the *guṇa*-trifold. Its varying parts and functions are distinguished by the preponderance of their respective *guṇa*-threads. These *guṇa*-threads, as we have already seen, are affective and intentional. This contrasts sharply with the body of the West, which is considered (like meat) to lack all affective characteristics and intentionality.

This *prakṛti*-c nature is not only restricted to *citta* and the human body. The five gross elements – space, air, fire, water and earth – that comprise all that we in the West would call the 'material' world are also *prakṛti*. They are therefore also affective and intentional. Again, their distinctive characteristics are determined by their respective *guṇa*-trifold balances. *Tamas* is the predominant thread of all of the five gross elements, but *sattva* and *rajas* are present in various degrees.

Consequently, whereas the Western concepts of 'body' and 'material' refer to non-intentional, non-affective and objective entities, Yoga's comparable referents are intentional, affective and objective.

On the basis of the above analysis, we can see that it would be exceptionally difficult to pigeonhole Yoga's ontology comfortably into any one of our six categories of responses to the mind-body problem. On the one hand, Yoga seems to be clearly dualistic, since it posits substantial and irreducible differences between *puruṣa* and *prakṛti*. But this seems to be a dualism between classes of mental states, rather than between mind and body. On the other hand, if we acknowledge the 'mental' aspects of *citta*, as indeed we must, and treat the bodily organs and the five gross elements as 'material', then we might be more inclined to characterize Yoga as reductivist. But what type of reductivism should we choose – materialism or idealism – and on what grounds? Gerald Larson once identified Yoga as "a peculiar mix of dualism and reductive materialism." It is dualist insofar as it posits the existence of two fundamentally distinct natures – *puruṣa* and *prakṛti*. It is

reductive insofar as every object in nature is a manifestation of the *guṇa*-trifold.² Roy Perrett also seems to lean in favor of the idea of that Yoga's ontology is a variety of dualism and material reductivism or eliminativism.³ In order to reach these conclusions, both of these philosophers have had to treat '*prakṛti*' as if it were largely equivalent to the Western concept of 'material'. But is this treatment warranted?

As an alternative, it has been suggested that Yoga might be better classed as a sort of reductive idealism. After all, like mental states, all of *prakṛti*'s manifestations are characterized by affects and intentionality. Further, manifest *prakṛti* evolves from its subtlest states, where it seems most 'mental', to its most gross, where it seems most 'material'. On these grounds, it may seem to make sense, or at least better sense, to classify Yoga as a form of reductive idealism. But this does not eliminate the fact that, unlike Western mental states, *prakṛti* is unconscious, causally determined, non-private and objective. It simply does not make good sense to classify *prakṛti* as either 'mental' or 'material'.

Therefore, whatever problems there may be in Patanjali's Yoga, the mind-body problem is not among them. There is, however, a comparable difficulty in Yoga, a *puruṣa-prakṛti* problem, and once again, it has to do with interaction. If all *prakṛti*-c manifestations have only *prakṛti*-c causes and effects, and if *puruṣa* is non-*prakṛti*-c, then *puruṣa* and *prakṛti* cannot interact. Yet, Yoga believes that all consciousness is *puruṣa* and that all intentional states are *prakṛti*. Yoga also believes that *prakṛti* is oriented towards the liberation of *puruṣa*. But by the above reasoning, *prakṛti* cannot be consciously directed towards *puruṣa*'s liberation. Therefore, Yoga's dualist ontology seems to leave no room for free, conscious, non-causally determined will. If this is the case, and if we recognize the Western truism that an action is moral if and only if the agent of the action might have chosen to act otherwise, there would also seem to be no place for moral action within its soteriology.

² Larson, Gerald James and Ram Shankar Battacharya (eds.), Encyclopedia of Indian Philosophies (Volume IV): Sāṃkhya, a Dualist Tradition in Indian Philosophy, Motilal Banarsidass (Delhi), 1987, pp. 76-77.

³ Perrett, Roy W., "Computationality, Mind and Value: the case of Sāṃkhya-Yoga," Asian Philosophy, Vol. 11, No. 1, 2001, pp. 5-13. Professor Perrett argues only that Yoga could accommodate such a view.

Feuerstein challenges the claim that the ontology of Patanjali's Yoga-Sutras excludes the possibility of free will. In his commentary on sutra II-13, which he translates as "[So long as] the root exists, [there is also] fruition from it: birth, life and enjoyment," Feuerstein writes, "Naturally, this determinacy cannot be all-comprehensive; otherwise there would be no scope for asserting one's will to break out this cycle."⁴ The root that this sutra refers to is that of *avidyā*, which sutra II.5 defines as "the conviction of permanence, purity, happiness and self in what are really impermanent, impure, painful and not self."⁵ As long as this misapprehension of identity persists, it causes birth, life and enjoyment, that is, bondage in *samsara*.

Feuerstein's objection does not seem to me to be well argued. It does not follow that in a strictly determined environment there can be no asserting of will. It could be the case that in a strictly determined environment this asserting of the will is itself causally determined, and hence existent but not free.

Because *puruṣa* is at all times separate from the fluctuations of *prakṛti*, it is difficult to see how *puruṣa* can influence these fluctuations. Sutra II.20 states, "[*Puruṣa*] is sight alone; though pure, he looks on at the thoughts."⁶ In his commentary on this sutra, Sankara claims, "The word 'alone' (*matra*) is to reject all other qualities; it is used to exclude any idea that [*Puruṣa*] is equally the seat of desire and so on, which would contradict the fact that [*Puruṣa*] is sight."⁷ Indeed, desire and aversion are clearly regarded as aspects of *prakṛti*. *Asmitā*, or I-am-ness, arises from *avidyā*. It is the false conjunction of *puruṣa* with the intellect, *buddhi*. By apprehending the observer of experience and the event of experience as a single entity, *asmitā* provides the condition for samsaric experience – the delusive experience of *being* pleased or *being* pained. Pleasurable experiences create memories of pleasure, latent proclivities, which in turn manifest desire. Painful

⁴ Feuerstein, George, The Yoga-Sutra of Patañjali: A new translation and commentary, Inner Traditions International (Rochester, Vermont), 1989, p. 68.

⁵ Translation provided by Leggett, Trevor, The Complete Commentary by Sankara on the Yoga-Sutra-s: A full translation of the newly discovered text, Kegan Paul International (London), 1990, p. 185.

⁶ *Ibid.*, p. 239.

⁷ *Ibid.*

experiences create memories of pain and tendencies to avoid such future experiences. Desire and aversion therefore clearly depend upon *avidyā* and *asmitā* for their existence, and cannot belong to *puruṣa*. The desire to influence *prakṛti*'s fluctuations would therefore seem to be alien to *puruṣa*, and without desire, in what meaningful sense can there be will?

One could argue that will has nothing to do with desire and thereby posit a distinction between desire, which is always directed towards ends, and will, which is constant and free of intention, precisely like *puruṣa*. However, this distinction would be difficult to sustain. It would ultimately either reintroduce intentionality to will or collapse will into another state. If I were to talk of the will of a king, for instance, and suggest that while this king desires nothing, his will is constant, it would be fair for someone to ask me to clarify what I mean by 'the king's will'. Suppose that I respond that the king wills universal compassion. My questioner could press me further. Do I mean that he demands universal compassion, or that his will is a manifestation of it? If the king demands universal compassion then his will is intentional; it is oriented toward universal compassion as an ends. On the other hand, if his will is a manifestation of universal compassion then it is indistinguishable from that compassion. It would not be will at all but simply an affective way of being.

We have seen that *prakṛti* consists of the *guṇa*-trifold, which is characterized by three affective threads. All *prakṛti*-c experience can therefore be reduced to *guṇa*-trifold experience. More importantly, for my present purposes, all affective experience is reducible to *guṇa*-trifold experience. Therefore, if free will is a property of *puruṣa*, and if free will is not intentional, then, on the basis of the above analysis, *puruṣa* would have to possess affective characteristics. Clearly, this cannot be the case, since all affective experiences are *guṇa*-trifold experiences, which are *prakṛti*-c experiences, and *puruṣa* is entirely unrelated to *prakṛti*. Therefore it must be the case either that free will is not a property of *puruṣa*, or that free will is intentional. No property of *puruṣa* can be intentional, since intentionality is a characteristic only of *prakṛti*. Therefore, free will cannot be a property of *puruṣa*.

It would seem impossible for *puruṣa* to will a change in *prakṛti*'s functioning. Therefore, if we are looking to accommodate free will in Yoga's ontology, we will have to look elsewhere, to *prakṛti*, the other side of Yoga's subject-object dichotomy.

I have described *sattva* above as the *guṇa*-trifold characteristic of joy and lucidity. *Sattva* also enables mental determinations. It is *buddhi*'s *sattva* predominance that enables it to act as a limiting force on *puruṣa*'s awareness. It creates borders around awareness, so to speak, so that awareness seems to become 'of' certain objects.⁸ This functioning also creates the illusion that *buddhi* is the seat of awareness, which in turn leads to the development of *asmitā*, or I-am-ness.

The question will naturally arise as to who exactly suffers this illusion. The answer, in general terms, is *citta*, the mental apparatus and functioning of which *buddhi* is part. More precisely, though, it seems that no entity suffers the illusion. Rather, in the grammar advocated earlier, it is more appropriate to say that the illusion occurs, that is, that delusive thoughts arise.

The arising and fluctuating of thoughts indicate the presence of *rajas*, the *guṇa*-thread of discomfort and activity. Changes in the balance of *rajas* cause thoughts to arise, appear and vanish. In Yoga evolution theory, it is *rajas* that causes the psycho-physical apparatus to become manifest.⁹ Despite *rajas*' active role in formulating experience though, it would be incorrect to suggest that *rajas*, or any other *guṇa*-thread, is conscious. Strictly speaking, nothing conscious guides intelligence. *Puruṣa* is by nature of unlimited scope and is hence necessarily undirected. The *guṇa*-trifold limits *puruṣa*'s intelligence and serves as the objects of its awareness. *Prakṛti* cannot therefore possess intelligence, for it is only after *prakṛti* has appeared to impose limits on *puruṣa* that the illusion of guided intelligence occurs.

⁸ Dasgupta, p. 49.

⁹ Ibid., p. 48.

Nonetheless, sutra II.18 states, “The seen (i.e., *prakṛti*) ...serves the [dual] purpose of (*puruṣa*’s) enjoyment and emancipation,” and sutra II.21 states, “The essence of the seen is only for the sake of [*puruṣa*].” Clearly, Patanjali is asserting here that *prakṛti* serves two purposes – the enjoyment and liberation of *puruṣa*. Dasgupta has claimed that this apparent teleological orientation of *prakṛti* proves the existence of free will.¹⁰ But an object’s orientation toward a particular end cannot show that the object is conscious, let alone that it has freely selected its orientation, anymore than the theory of gravity can prove that all falling objects are conscious bearers of free will.

It is worth noting that the two purposes of *prakṛti* are *prakṛti*’s only two manifest possibilities. *Prakṛti* either acts for the enjoyment of *puruṣa*, which does not actually experience enjoyment but only appears to do so, or it works for *puruṣa*’s liberation. After liberation, *prakṛti* is no longer manifest. Therefore, it makes sense to speak of *prakṛti*’s two purposes as its only two potential functions, neither of which demands conscious intent.

Unlike the mind-body problem, the *puruṣa-prakṛti* problem does not present Yoga with an inconsistent set of beliefs. It does, however, seem to entail an unsatisfactory conclusion. If *prakṛti* cannot be consciously oriented towards *puruṣa*’s liberation, and if *puruṣa* cannot interact with *prakṛti*, then how can one freely choose to adopt (or abandon) Yoga’s soteriological path? The *puruṣa-prakṛti* problem seems to deny Yoga the possibility of free, conscious, non-causally determined will. Further, if we accept the Western truism that an action can be considered moral if and only if the agent of the action could have chosen to act otherwise, there would also seem to be no scope for morality in Yoga’s soteriology. Even if it were possible for one to will to break out of the cycle of *samsāra*, or to will to cease one’s practice of Yoga, this will itself, as a function of *prakṛti*, would be unconscious and causally determined.

However, if Dasgupta was right, and if Yoga did admit conscious will into its ontological system without changing anything else, it would then be left with a set of inconsistent

¹⁰ Ibid., p. 177.

beliefs – namely that *prakṛti* can and cannot be consciously oriented to *puruṣa*'s liberation. As it stands, Yoga seems to endorse a parallelist variety of dualism in *avidyā*, which is not incoherent, but which does lead to negative consequences – namely, the *prima facie* loss of its potential standing as a valid moral philosophy.

There are, of course, alternatives. Yoga could abandon the claim that fundamentally different substances cannot interact, and arrive thereby at a type of interactionist dualism. However, this would entail the abandonment of the *satkāryavāda* theory of causation, the theory that substances must have similar natures in order to causally impact each other. Abandoning *satkāryavāda* would have its costs since it plays a significant role elsewhere in Yoga and since its critiques of other Indian philosophical systems, especially Advaita Vedānta, depend heavily upon the principle. Instead, Yoga might seek to introduce consciousness to *prakṛti*, most likely to its *sattva-guṇa* thread, in order to create scope for intentional consciousness. This would have other provocative, and maybe even appealing, consequences – it would, for instance, suggest that everything in the phenomenal world, from dolphins and human beings to algae and stones, has a degree of intentional consciousness. More negatively though, it would seem to rob *puruṣa* of its defining characteristic, and could even be said by Yoga to be incoherent, namely because it leads to the objectification of the subject, the impossibility of which seems to be one of the key metaphysical insights underpinning Yoga philosophy and its initial *puruṣa-prakṛti* distinction. In the same vein, we could suggest that Yoga abandon its dualism and frame its ontology instead in either a *puruṣa*-ultimate ontology (somewhat comparable to 'idealism') or a *prakṛti*-ultimate ontology (somewhat comparable to 'materialism'). But would this still be Yoga? We would have to revise most of the Yoga-sūtra-s just to accommodate the change.

As a final solution, Yoga might reject these proposed revisions to its ontology and instead challenge the moral criterion that insists on the presence of free will in any moral theory. In other words, Yoga could acknowledge the fact that there is no scope for free will in its ontology but insist instead that it does not need the concept. But in that case, there would

be no cause for praise or blame in Yoga philosophy. Liberation would simply be something that happens.