

6

Subjectivity, Selfhood and the Use of the Word ‘I’

JONARDON GANERI

I. Asaṅga and Vasubandhu on Self-Consciousness and Conscious Attention to Oneself

A well-known Buddhist philosophy of mind has it that conscious experience is a synthesis of five forms of activity: the processes of registering, appraising, stereotyping, readying, and consciously attending.¹ These philosophers say that the last of these, conscious attention, is relative to a perceptual modality, so that, for example, consciously attending to what is being visually registered is different in kind from consciously attending to what is being registered in touch.² They say as well that one can also consciously attend to what is going on in one’s mind. This is a mode of conscious reflecting (*mano-vijñāna*); it is a way of being self-aware, of consciously attending to one’s own psychological state.

To these six varieties of conscious attention Asaṅga adds a seventh, which he calls simply *manas* ‘mind’/‘consciousness’, or else ‘defiled mind’ (*kliṣṭa-manas*).³ It is, he seems to think, a distinct and more basic mode of being

¹ The five so-called *skandhas* or ‘ingredients’ that combine into individual thoughts or experiences: *rūpa*, *vedanā*, *saṃjñā*, *saṃskāra*, and *vijñāna*. For details, see Ganeri (forthcoming). This flexible doctrine is transformed in various ways by later Buddhist thinkers, beginning with Asaṅga and Vasubandhu in the c. 4th CE.

² *Śaḍvijñānadhātavaś cakṣurādyaśrayā rūpādyaḥ lambanāvijñāptayah* (Vasubandhu, *Pañcaskandhaka* 135).

³ Thereby adding a new sense for this term, which in its everyday use in Abhidharma is simply synonymous with conscious attention (*citta*, *vijñāna*). Kramer (2008) translates *kliṣṭa-manas* as ‘notion of I’, and observes that the incorporation of this new concept represents a modification in the traditional system of the five *skandhas*, a modification that is evident in Sthiramati’s commentary on Vasubandhu’s *Pañcaskandhaka*. She says: ‘In particular the function of *vijñāna-skandha*—the original role of which was

self-aware.⁴ Asaṅga argues that there must exist a non-perceptual modality of self-consciousness, which is distinctively associated with what is ‘mine’, as well as being the support of conscious reflection, and something that contributes to the persistence of one’s sense of self. Itself ethically neutral, it is nevertheless responsible for the four vices to do with the self:

[Question:] How does one know that *manas* in the sense of ‘defiled mind’ (*kliṣṭa-manas*) exists?

[Answer:] Without it, there could be no uncompounded ignorance, i.e. a basic ignorance not yet associated with all the diverse defilements but standing as their base. Besides, conscious reflecting (*manovijñāna*) must also have a simultaneous support, as do the sensory consciousnesses which have such supports in their material organs. Such a simultaneous support can only be the ‘defiled mind’. Also, the very etymology of *manas* has to do with ‘mine’, which can be explained only by the ‘defiled mind’. Also, without it there would be no difference between the non-identifying trance and the cessation trance, for only the latter is free of defiled mind. Also, the sense of an existence of self is always existent in nonsaintly states: there must be some special consciousness to account for the persistence of this sense. The defiled mind is always defiled by the false view of self, pride of self, love of self, and ignorance (about self); but is itself ethically neutral.⁵

Asaṅga’s brother Vasubandhu claims that this mode of being self-aware undergoes a ‘transformation’ into what gets described as a self.⁶ He says

actual perception—was widened through the inclusion of subliminal forms of mind, like the ‘store mind’ (*ālaya-vijñāna*) and the “notion of I” (*kliṣṭa-manas*). The strong emphasis placed by Sthiramati on *vijñāna* is evident, for instance, when he states that ordinary people—those who have not perceived reality—regard the *vijñāna* as the self (*ātman*), whereas they view the other four *skandhas* as “mine” (*ātmīya*)’ (2008: 155). She adds that ‘[i]nterestingly, Sthiramati also mentions alternative concepts of the self, for example that of the Sāṃkhya tradition. According to his understanding, the Sāṃkhyas only regard *rūpa-skandha* [matter] as *ātmīya*, and all the other four *skandhas* as *ātman*. He thus claims that for the Sāṃkhyas the self is not only identical to *vijñāna* but also consists of the other factors accompanying the mind (*caitasika*)’ (2008: 155). Galloway (1980) translates *kliṣṭa-manas* as ‘passional consciousness’, and derives interesting information about the notion from Guṇaprabhā’s commentary on the *Pañcaskandhaka*. See below. Dreyfus and Thomson (2007: 112) translate *kliṣṭa-manas* as ‘afflictive mentation’, and comment that ‘[t]his is the inborn sense of self that arises from the apprehension of the store-consciousness as being a self. From a Buddhist point of view, however, this sense of self is fundamentally mistaken. It is a mental imposition of unity where there is in fact only the arising of a multiplicity of interrelated physical and mental events’.

⁴ See Galloway (1978) for a detailed argument that as a Yogācāra technical term, *manas* should be translated as ‘consciousness’ rather than neutrally through its cognate in English, ‘mind’.

⁵ *Mahāyānasamgraha* 1.7; trans. Anacker in Potter (2003) from the extant Chinese and Tibetan translations.

⁶ Galloway (1980: 18) reports from Guṇaprabhā’s commentary on Vasubandhu’s *Pañcaskandhaka*: [Vasubandhu:] ‘In reality, the consciousness (*manas*) has the storehouse perception for its phenomenon.’

that *manas*—‘consciousness’—is a way of being aware, associating it with the activity of ‘thinking’ (*manana*). It takes the store-consciousness (*ālaya-vijñāna*) as its foundation. It undergoes a transformation into something that we metaphorically call a self, but this transformation is the work of cognitive fabrication, and there is in fact no such thing:

The metaphors of ‘self’ and ‘items’ which develop in so many ways take place in the transformation of consciousness.

Dependent on [the store-consciousness] there develops a consciousness called *manas*, having that as its basis, and having the nature of ‘thinking’.

This transformation of consciousness is a cognitive fabrication, and as it is cognitively fabricated it does not exist.⁷

How are we to make sense of what is going on here? The import of the use of the terms ‘conceptual fabrication’ (*vikalpa*) and ‘metaphorical designation’ (*upacāra*), in connection with the self, is that the end-result of the transformation of pre-attentive self-consciousness is the sort of first-person psychological judgment one would express in the words ‘I am F’. The transformation has made the self into a conceptual thought-content (*vikalpa*), but the expression of that thought-content uses a word, ‘I’ for example, in at most a ‘metaphorical’ sense, or at any rate some usage that is not one of genuine literal reference. (As I will point out below, *upacāra* is not quite metaphor, but nearer to metonymy.)

Let me represent the picture schematically. The claim is that three distinct phenomena are involved in self-consciousness:

1. Conscious attention to one’s own states of mind (*manovijñāna*).

This must have a ‘support’ (*āśraya*). The support is:

2. ‘Self-consciousness’ (*manas*)—a pre-attentive mode of being self-aware.

[Guṇaprabhā:] This means that it phenomenalizes [sees] the storehouse perception as a self.[Vasubandhu:] ‘It is that which is associated with the constant delusion of self (*ātmamoha*), view of self (*ātmadṛṣṭi*), egoism of self (*ātmamāna*), and lust for self (*ātmarāga*), and so on.’[Guṇaprabhā:] It is explained as operating always, and arises as good (*kuśala*), bad (*akuśala*), and indifferent. His saying ‘It is of one class’ means that it has a passionate (*kliṣṭa*) nature. ‘It is continually produced’ means that it is momentary.

⁷ *Triṃśikākārikā: ātmadharmopacāro hi vividho ya pravartate | vijñānapariṇāme ’sau || Tvḥ 1a-c || tasya vyāvṛtīrarhatve tadāśṛitya pravartate | tadālambaṃ manonāma vijñānaṃ mananātmakam || Tvḥ 5 || vijñānapariṇāmo ’yaṃ vikalpo yadvikalpyate | tena tannāsti || Tvḥ 17a-c ||*. The translation is from Anacker (1984), slightly modified.

This is subject to 'transformation' (*pariṇāma*). What it is transformed into is:

3. First-person psychological judgment—thinking 'I am F', for some psychological predicate F. The use of the word 'I' here, though, is in some sense not a genuine referring use.

It will help if I begin by stating the conclusions for which I want to argue. I will argue that these claims should be understood as follows. My possession of a first-person perspective, a perspective on my own mental life, has to be underwritten. What underwrites it is the fact that my mental life presents itself to me, in a primitive and pre-attentive way, as being mine. This same primitive mode of being self-aware is rendered in such a way that it seems to justify me in making assertions of the form 'I am F'. In fact, it is never the case that assertions of such a form are true of a self. Uses of 'I' never literally refer to a self. I will argue that this final claim is ambiguous, and distinguish the reading Vasubandhu wishes to give it from another, in my view more promising, idea.

2. Pre-Attentive Consciousness: *manas* and Mineness

The proposal we are examining might be expressed as the conjunction of three propositions:

- [1] There is a pre-attentive mode of self-awareness through which my experiences present themselves to me as mine.
- [2] First person psychological judgment draws upon additional conceptual resources, ones not available on the basis of [1] alone.
- [3] First person psychological judgments do not actually involve genuine reference to a self.

Let me examine these propositions in turn.

The ability, not just to have a world in view, but also to reflect upon the fact that one does, seems to be an essential part of what it means to be conscious. Sidney Shoemaker says that

It is essential for a philosophical understanding of the mental that we appreciate that there is a first person perspective on it, a distinctive way mental states present themselves to the subjects whose states they are, and that an essential part of the

philosophical task is to give an account of mind which makes intelligible the perspective mental subjects have on their own mental lives

(Shoemaker 1996: 157).

It is to this task that our Buddhists address themselves when they say that conscious attention to one's own mental life (*mano-vijñāna*) must have a support, which they claim is a pre-attentive mode of being self-aware (*manas*). I think that the point of this argument is easy enough to understand as long as we remember that it is impossible to think about one of one's own mental states, a particular feeling of hope for example, and yet not be sure whose mental state it is. There is no question of having a first person perspective on one's mental life, without that mental life presenting itself to one as one's own. In a much-quoted passage, Peter Strawson says:

It would make no sense to think or say: This inner experience is occurring, but is it occurring to me? (This feeling is anger; but is it I who am feeling it?) Again, it would make no sense to think or say: I distinctively remember that inner experience occurring, but did it occur to me? (I remember that terrible feeling of loss; but was it I who felt it?) There is nothing that one can thus encounter or recall in the field of inner experience such that there can be any question of one's applying criteria of subject-identity to determine whether the encountered or recalled experience belongs to oneself—or to someone else.

(P. F. Strawson 1966: 165)

If I cannot be mistaken about whose inner experience it is that I am experiencing, this is because no identification of a subject, and so no possibility of mis-identification, is involved at all. What I am suggesting, then, is that our Buddhist philosophers explain the 'immunity to error through misidentification' (Shoemaker 1984) of self-ascriptions, by acknowledging that when my experience presents itself to me as my own, no representation of myself as a subject takes place. Asaṅga and Vasubandhu postulate, instead, the existence of a primitive mode of self-awareness, a basic awareness of the contents of my inner life (my 'store-consciousness') as mine. And this, in turn, is what makes it possible for me to have a first-person, rather than merely a third-person, perspective on my mental life.

In Sartre's theory of consciousness, I might note in passing, there is a proposal that is in some respects comparable. Sartre speaks of a pre-reflective self-awareness, which 'has no need at all of a reflecting consciousness in order to be conscious of itself. It simply does not posit itself as an object'

(Sartre 1957: 45). Dan Zahavi has redescribed it as 'an immersed non-objectifying self-acquaintance' (Zahavi 2005: 21). Sartre argues that an infinite regress will ensue if such a mode of self-acquaintance is not acknowledged, and it is interesting to observe that we find the infinite regress argument used too by one of Vasubandhu's immediate followers, Diñnāga, in a defense of reflexivism (Ganeri 1999; see also the contributions by Evan Thompson and Mark Siderits to this volume).

3. First Person Psychological Judgment

Vasubandhu speaks of a 'transformation' of basic self-awareness into explicit self-ascription, a transformation based on conceptual fabrication (*vikalpa*) and justifying only a 'metaphorical' use (*upacāra*) of the language of self. What is difficult is to understand how it can be thought wrong to make the transition from being aware of oneself as being in a certain mental state to explicitly asserting that one is. Vasubandhu's thesis is that this transition demands a new conceptual resource (one which is not in fact available). Is that thesis true?

Zahavi, for one, does not see any difficulty with this transition. He says:

Contrary to what some of the self-sceptics are claiming, one does not need to conceive of the self as something standing apart from or above experiences, nor does one need to conceive of the relation between the self and experience as an external relation of ownership. It is also possible to identify this pre-reflective sense of mineness with a minimal, core, sense of self. . . . In other words, the idea is to link an experiential sense of self to the particular first-personal givenness that characterizes our experiential life; it is this first-personal givenness that constitutes the *mineness* or *ipseity* of experience. Thus, the self is not something that stands opposed to the stream of consciousness, but is, rather, immersed in conscious life.

(Zahavi 2005: 125)

Zahavi's 'minimal self' precisely consists in a 'pre-reflective sense of mineness', and it appears to follow that, to refer to oneself in the first person, nothing more is required than that one's experience be given 'immediately, noninferentially and noncriterially' (2005: 124) as mine. He also says, however, that 'this form of egocentricity must be distinguished from any explicit

I-consciousness. I am not (yet) confronted with a thematic or explicit awareness of the experience as being owned by or belonging to myself. The mineness is not something attended to; it simply figures as a subtle background presence' (2005: 124). So there is, after all, a transition, but it is a transition which involves only *paying attention* to the mineness inherent in my experience, not in the exercise of any new conceptual resource.⁸

I think that Vasubandhu's response would simply be that if someone wants to use the words 'minimal self' as a synonym for the *manas*, then nobody will object to him doing so. On the other hand, if the implication is that the 'minimal self' does what a self is meant to do, then it is too minimal to count. It seems to me that the minimal self does not do one of the things that a respectable concept of self must, and that is to individuate thinkers. It is true of you and me alike that our experience is given to us with an immersed mineness. The property 'being a thought of one's own' is a property like 'being a divisor of itself', which is equally true of every number; the reflexive pronoun is just a place-holder. Zahavi says that 'the particular first-person givenness of the experience makes it mine and distinguishes it for me from whatever experiences others might have' (2005: 124). That choice of words suggests that he thinks that first-person givenness is individuating of individual selves.⁹

François Recanati (Recanati 2007) has an interesting account of the transition we are interested in. He draws a distinction between implicit and explicit *de se* thoughts, a *de se* thought being 'a *de re* thought about oneself, that involves a particular mode of presentation, namely a first person mode of presentation' (2007: 169). He continues:

As Frege wrote in 'The Thought', 'everyone is presented to himself in a particular and primitive way, in which he is presented to no one else'. I call the 'special and

⁸ Indeed, in Zahavi (1999), Zahavi makes it a 'minimal demand to any proper theory of self-awareness' that it 'be able to explain the peculiar features characterising the subject-use of "I"; that is, no matter how complex or differentiated the structure of self-awareness is ultimately shown to be, if the account given is unable to preserve the difference between the first-person and third-person perspectives, unable to capture its referential uniqueness, it has failed as an explanation of self-awareness' (1999: 13).

⁹ Joel Krueger (this volume) seems to share my reservation about the selfhood of the 'minimal self'. To put the point in an Indian vocabulary, the 'minimal self' is somewhat akin to the impersonal Advaitic *ātman*, present equally in all. Zahavi sometimes, however, appeals to an embodiment criterion, rather than to first-person givenness per se, as what individuates distinct minimal selves, and that would certainly adequately distinguish the notion from the Advaita conception.

primitive' mode of presentation which occurs in first person thoughts 'EGO' or rather 'EGO_x' where 'x' stands for the name of the person thinking the thought.

(2007: 170).

Explicit *de se* thoughts are *de se* thoughts 'the content of which involves an "identification component" through which the object thought about is identified as oneself'. When a subject looks at themselves in a mirror and thinks 'My legs are crossed', they identify themselves under the concept EGO, and ascribe to themselves a property: I am that person whose legs are crossed. An implicit *de se* thought involves no such identification. Recanati says that 'implicit *de se* thoughts are identification-free, and they are *de se* only externally: no concept EGO occurs as part of the lekton [roughly, the content]. The lekton is a personal proposition, without any constituent corresponding to the person to whom a property is ascribed' (2007: 176). What this means is that in an implicit *de se* thought, one simply thinks of one's legs as crossed, a thought that is true because it is indeed one's own legs which are crossed. (The distinction is reflected in language: contrast the anaphoric construction 'He expects that he will be late' with the gerundival construction 'He expects to be late'. See also Perry 1998.) Recanati's point is that it is precisely because no identification of a subject is involved that implicit *de se* thoughts are immune to error through misidentification.

Recanati argues that the concept EGO involved in the notion of an explicit *de se* thought is itself explained by this notion of an implicit *de se* thought. He says:

The notion of an implicit *de se* thought in which the self is not represented is important... , to understand the concept of self that occurs in explicit *de se* thoughts. Indeed, the ability to entertain implicit *de se* thoughts is arguably a necessary condition for anyone to evolve the concept EGO. That is so because, as suggested by Evans, Perry, and myself following them, the concept EGO is best construed as a repository for information gained in a first person way... Now a piece of information is gained in the first person way if and only if it is the content of an implicit *de se* thought. It follows that the first step in an elucidation of the concept of self is a correct analysis of the functioning of implicit *de se* thoughts.

(2007: 177).

The notion of *manas*, as introduced by Asaṅga and developed by Vasubandhu, seems to me to be rather close in spirit to Recanati's notion of implicit *de se* thought. It certainly contains no representation of the self, and

it has as its ‘foundation’ (*ālambana*) all the information contained in the store-consciousness. It is a first personal way of thinking about (*manana*) that information. The description ‘repository for information gained in a first person way’ might seem like a very good description of the joint contribution of *manas* and store-consciousness.¹⁰

Moreover, as we have seen, Vasubandhu speaks of a ‘transformation’ of *manas* into a concept of self, and that echoes with the claim Recanati is making about the evolution of the concept EGO. Finally, Recanati claims that the ability to entertain implicit *de se* thoughts is a necessary—though not necessarily a sufficient—condition for the evolution of the concept EGO, and this too is something Asaṅga and Vasubandhu are keen to stress; indeed, it is the primary motivation for introducing the notion of *manas* in the first place. For, as Asaṅga says in the *Yogācārabhūmi*, ‘That [*manas*] has the mode of taking the store-consciousness as its object and conceiving it as “I am [this]” (*asmitī*) and “[this is] I” (*aham itī*)’ (quoted in Waldron 2002: 42). Note that the concept EGO, as a repository of information gained in a first person way, does individuate thinkers, for no two such repositories will be alike. This concept of self is much less ‘minimal’ than Zahavi’s.

So where is the difference? Both agree that there are explicit *de se* thoughts involving the concept EGO, but while Recanati thinks that the truth conditions of such thoughts are such that they are often true, Vasubandhu thinks that they are always false. He thinks this because he thinks that the concept EGO, which has certainly evolved in the manner described, is an empty concept, like the concept PHLOGISTON or the concept PEGASUS.

¹⁰ The history of the Yogācāra concept of the store-consciousness is rather complicated. Originally conceived merely as a vehicle for the perpetuation of mental forces when the normal six types of awareness are absent, it was a technical solution to what would otherwise be a difficulty in the Yogācāra theory of individual persistence. Dreyfus and Thompson (2007: 112) say of it that ‘[t]his continuously present subliminal consciousness is posited by some of the Yogācāra thinkers to provide a sense of continuity in the person over time. It is the repository of all the basic habits, tendencies and propensities (including those that persist from one life to the next) accumulated by the individual’. Reaffirming this description in his article in this volume, Dreyfus offers the translation ‘basic consciousness’, and discusses an illuminating range of associations and resonances. The detailed studies by Lambert Schmithausen (1987) and Hartmut Buescher (2008) have revealed much greater complexity in the use of the notion in early Yogācāra than scholars had previously acknowledged. Schmithausen comments that ‘it may well be that *ālayavijñāna* was, initially, conceived as a kind of ‘gap-bridger’, but hardly in such a way that its occurrence in ordinary states had been denied’ (1987: §2.13.6). Items in the store do not themselves carry a feeling of mineness, but ground the feeling of mineness which attaches itself to the stream’s conscious self-attention. They comprise a sort of database for the mind, information which can be drawn upon in the activity of bringing the states of the stream into conscious attention.

Our Buddhists think that the evolution of the concept EGO brings with it all manner of moral defilements, and one form of justification for that claim is that the concept rests in this way on an error. Sthiramati’s comment on the first of the 30 *Verses* bears the point out: he says that the concept of self presents only an apparent (*nirbhāsa*) referent, just as the perception of someone with an eye-disease presents only apparent hairs and circles. It is ‘metaphorically designated’ (*upacaryate*) because it is said to be there when it is not, as if one were to use the word ‘cow’ when there is an ox. Sthiramati’s example, incidentally, shows that the notion of *upacāra* is much closer to metonymy than to metaphor, as traditionally understood.¹¹

With P. F. Strawson’s assertion that ‘no use whatever of any criteria of personal identity is required to justify [a person’s] use of the pronoun “I” to refer to the subject’ (P. F. Strawson 1966: 165), Vasubandhu would appear to dissent. For his view seems to be that the use of the pronoun ‘I’ never refers to a subject of experience. Strawson’s point is that we don’t need any extra conceptual resource in order to make explicit self-references, and in particular we don’t need a criterion of identity. The pronoun ‘I’ is not a term which we can correctly use only if we have successfully identified its referent, because if it were then there would be the possibility of error through misidentification relative to it. Strawson infers that ‘I’ refers to its subject without there being a criterion of identity. Perhaps what Vasubandhu would do would be to agree with this argument but contrapose it. His point would then be that all genuine reference involves the identification of a referent, and given that there is no question of such an identification in the case of the first person, the pronoun ‘I’ cannot be a genuine referring term. In saying that it is instead a ‘metaphor’ (*upacāra*), there seems to be a gesture at the possibility of a different, non-referential, account of its use. To say ‘I feel hopeful’ is, it might be thought, to speak non-referentially of the existence of a hope which presents itself pre-attentively as mine. I will argue that, although this is not actually Vasubandhu’s strategy, it is nevertheless a viable one.

¹¹ *tam ātmādirirbhāsam rūpādirirbhāsam ca tasmād vikalpād bahirbhūtam ivopādāyātmādyupacāro rūpādirharopacāras cānādikālikāḥ pravartate vināpi bāhyenātmanā dharmāis ca | tadyathā taimirikasya keśondukādyupacāra iti | yac ca yatra nāsti tat tatropacaryate | tad yathā bāhike gauḥ |* (Sthiramati, *Triṃśikāś-vijñapti-bhāṣya*; see Buescher 2007).

4. Two Uses of ‘I’

Can one hear an echo of the Buddhist idea in the following remark?

One of the most misleading representational techniques in our language is the use of the word ‘I’, particularly when it is used in representing immediate experience, as in ‘I can see a red patch’. It would be instructive to replace this way of speaking by another in which immediate experience would be represented without using the personal pronoun.

(Wittgenstein 1975: 88)

In another place, Wittgenstein speaks of ‘two different cases in the use of the word “I” (or “my”)', the ‘use as object’ and the ‘use as subject’ (1960: 66–7). The ‘use as object’ is the use to which it is put when we refer to ourselves as human beings, embodied entities in a public space, the use it has when, for example, one person says to another, ‘I am just going to the shops to get the paper’, or ‘I have twisted my ankle’. Having distinguished between these two uses, one strategy would be to identify one of these uses as the primary use, and analyze the other use as being in some way derivative upon the first. Indeed, it is more in keeping with Indian theory about non-literal language to speak in this way of primary and derivative uses, rather than in terms of a distinction between literal and metaphorical use. The derivative use is metonymic rather than metaphorical: that is, the term is used to refer to something else, which stands in some relation to the primary referent. Among the contemporaries of Asaṅga and Vasubandhu are Vaiśeṣika philosophers, who argue that the primary use of ‘I’ is to refer to a self (*ātman*), and that its use to refer to oneself as an embodied being, in statements like ‘I am fat’, is an act of derivative reference, that is, reference to something which stands in an ‘is the body of’ relation to the primary reference.¹²

A variant on this approach is advocated by Galen Strawson. Strawson argues that the two uses are both genuinely referential, and neither is primary, in short, that ‘I’ is not univocal. One use is to refer to what he describes as a ‘thin subject’, a thin subject being ‘an inner thing of some sort that does not and cannot exist at any given time unless it is having experience at that time’ (2008: 156; cf. 2009: 331–8). The other use is to refer to ‘the human being considered as a whole’:

¹² See the discussion under *Vaiśeṣika-sūtra* 3.2.9–14.

Are we thin subjects? In one respect, of course, we are thick subjects, human beings considered as a whole. In this respect we are, in being subjects, things that can yawn and scratch. In another respect, though, we are in being subjects of experience no more whole human beings than hands or hearts: we are—literally—inner things, thin subjects, no more things that can yawn or scratch than eyebrows or thoughts . . . — But 'What then am I?' Am I two different sort of things, a thin subject and a thick subject? This is ridiculous . . . My answer is that 'I' is not univocal. We move naturally between conceiving of ourselves primarily as a human being and primarily as some sort of inner subject (we do not of course naturally conceive of ourselves as a thin subject). Sometimes we mean to refer to the one, sometimes to the other; sometimes our semantic intention hovers between both, sometimes it embraces both.

(G. Strawson 2008: 157–8)

Elsewhere, G. Strawson (2007) is clear that the relation between the two uses is one of metonymy, and indeed that the underlying relation is one of whole to part:

I think that we do at different times successfully use 'I' to refer to different things, to human beings considered as a whole and to selves. In this respect the word 'I' is like the word 'castle'. Sometimes 'castle' is used to refer to the castle proper, sometimes it is used to refer to the ensemble of the castle and the grounds and associated buildings located within the perimeter wall, sometimes it can be taken either way. The same goes for 'I', but 'I' is perhaps even more flexible, for it can sometimes be taken to refer to both the self and the whole human being, indifferently. Our thought (our semantic intention) is often unspecific as between the two.

(G. Strawson 2007: 543)

Vasubandhu, although he does not say so here, would perhaps be content to endorse as 'conventional' the use of the first person in statements like 'I am going to the shops', a use governed by the token-reflexive rule that 'I' refers to the speaker. When 'I' is used in the expression of first person psychological judgment, however, his claim is that the reference to an inner self fails, that this use erroneously imports a subject-predicate model and imposes it upon one's inner experience. In other words, his view of this use of 'I' is that there is a combination of metonymy and error-theory. When 'I' is used metonymically to refer to the inner subject, something goes wrong, and what goes wrong is that there is nothing at the far end of the metonymic relationship for it to refer to. This, indeed, is a view which G. Strawson (2007: 543) considers and rejects:

If it turns out that the best thing to say about selves is that there are no such things, then the best thing to say about 'I' may well be that it is univocal after all, and that the apparent doubleness of reference of 'I' is just the echo in language of a metaphysical illusion. If this is right, then 'I' is not in fact used to refer to selves as distinct from human beings even when its users intend to be making some such reference and believe that they are doing so. On this view, the semantic intentions of 'I'-users sometimes incorporate a mistake about how things are. I disagree.

According to the interpretation of Vasubandhu we have reached, then, 'I' does not function as an expression of genuine reference. but is rather one of *disingenuous reference*: it is a referring expression without a referent, its use creating the false impression that there is one. That is, I suggest, the best way to understand Vasubandhu's claim that it is a 'metaphor'.

There is another possibility, though. It might be the case that the mistake which 'I'-users make is not a metaphysical one, but rather a mistake about the semantic role of 'I' itself. Perhaps what goes wrong in the use of 'I' in representing immediate experience is that speakers take themselves to be making a referential use of an expression, and in doing so mistake its true logical role. It would be as if someone thought that 'perhaps' is a referring expression, and then imagined that there must be something in the world that it designates. This view is attractive, because it does not make the argument hinge on a prior metaphysical claim about the reality of selves, which then stands in need of further, extra-linguistic, justification. Rather, once we are clear about the logical role of 'I', we see that looking for a referent is as misguided as looking for the referent of 'perhaps'. And indeed, this is what Wittgenstein seems to say about the 'use as subject' of the word 'I', its use in a sentence such as 'I have a pain'. For Wittgenstein, denying that 'I', when used as in first person psychological judgments, is a referring expression, is the only way to explain the phenomenon of immunity to error through misidentification:

[T]here is no question of recognizing a person when I say I have a toothache. To ask 'are you sure that it's you who have pains?' would be nonsensical . . . And now this way of stating our idea suggests itself: that it is as impossible that in making the statement 'I have toothache' I should have mistaken another person for myself, as it is to moan with pain by mistake, having mistaken someone else for me. . . . To say, 'I have pain' is no more a statement about a particular person than moaning is.

(1960: 66–7)

The suggestion that there is a non-referential account of the use of 'I' was developed in one direction by Anscombe (Anscombe 1975). Anscombe, however, does not distinguish two uses, and argues that the first person does not refer, even in cases like 'I have a broken arm'. Other writers have tried, following the lead of P. F. Strawson, to argue that immunity of error does not commit one to a non-referential account of 'I', and indeed to reconcile immunity with the idea that 'I' refers univocally to the embodied human being (see for example Campbell 2004; McDowell 2009). In an insightful remark about Anscombe, Campbell suggests that the best way to understand her position is as claiming that the patterns of use involving the first person do not require justification in a semantic foundation:

An alternative reaction would, of course, be to say that we ought to abandon the search for a semantic foundation for our use of the first person. There are only the patterns of use, and no explanation to be given of them. This was essentially G. E. M. Anscombe's position in her famous paper, 'The First Person', in which she claimed that the first person does not refer. This claim is generally rejected, simply because philosophers have thought that when there is a use of the first person, there is, after all, always someone around who can be brought forward as the referent. But this is an extremely superficial response to Anscombe's point. Her claim is best understood as making the point that the ascription of reference to the first person is empty or idle; it does no explanatory work.

(Campbell 2004: 18)

I have argued at length elsewhere that something along these lines is just the move made by the Mādhyamika philosopher Candrakīrti (Ganeri 2007, ch. 7). His position, I have claimed there, is that we can give a fully explanatory use-theoretic account of the role of the first person in performances of self-appropriation, an account in which it is otiose to assign a reference. Asaṅga and Vasubandhu, on the other hand, say that in the movement from a pre-attentive self-awareness to an explicit use of 'I', a transformation of some sort is involved, one which involves conceptual work (*vikalpa*), and that the use of 'I' is metaphorical or metonymic. Their view is that the use of 'I' is indeed referential, and that the use as object and the use as subject are to be understood as making reference to, on the one hand, the human being, and on the other, an inner subject of experience, this second use being derivative from the first. There is, however, no subject of experience, and so the subjective use of the first person is an error.

5. Conclusion

I have argued that the new theory of Asaṅga and Vasubandhu consists in an account of the first-person perspective. Their further claim that the first person itself, the word ‘I’, is used ‘metaphorically’ in reporting the contents of the first-person perspective, rests on a prior commitment to the non-existence of an inner subject of experience. Only this permits them to claim that its use is one of what I have called ‘disingenuous reference’. I have distinguished a different strategy, which is to begin with the observation that such reports are immune to error through misidentification, and to argue that it follows that in the proper account of the use in first person psychological judgments of the word ‘I’, the assignment of a referent is explanatorily superfluous. Some of the parts of this strategy are to be seen at work in various thinkers. Anscombe argues that ‘I’ is not a referring expression, but does not distinguish the two uses of the term. Candrakīrti explains the use of ‘I’ in a way that makes the assignment of a referent superfluous, through an appeal to the thought that its role has to do with self-appropriation (*upādāna*), but he does not base this claim on the phenomenon of immunity. The full strategy being defended here emerges only as a ‘fusion’ of components drawn from our various sources.¹³

References

- Anacker, S. (1984), *Seven Works of Vasubandhu, the Buddhist Psychological Doctor*, Religions of Asia Series (Delhi : Motilal Banarsidass).
- Anscombe, G. E. M. (1975), ‘The first person’, in Samuel Guttenplan (ed.), *Mind and Language* (Oxford: Clarendon Press).
- Buescher, H. (2007), *Sthiramati’s Triṃśīkāvijñaptibhāṣya: Critical Editions of the Sanskrit text and its Tibetan Translation*, Sitzungsberichte / Österreichische Akademie

¹³ In his contribution to this volume, Matthew MacKenzie argues that the performativist model which I find in Candrakīrti needs to be supplemented rather with embodied and enactivist elements. Such a move “fuses” the Buddhist theory with ideas drawn from recent phenomenological literature, a theme of many of the contributions to this volume. I have tried instead to ‘fuse’ the Buddhist theory with elements taken from the recent analytical tradition, and the existence of both such possibilities suggests to me that Indian theory might well serve to create the intellectual space for a *rapprochement* between those hitherto separated strands of Western thought.

- der Wissenschaften, Philosophisch-Historische Klasse (Wien : Verlag der Österreichischen Akademie der Wissenschaften).
- (2008), *The Inception of Yogācāra-Vijñānavāda* (Wien: Verlag der Österreichischen Akademie der Wissenschaften).
- Campbell, J. (2004), 'What is it to Know what "I" Refers to?' *The Monist* 87.2: 206–18.
- Dreyfus, G. and Thompson, E. (2007), 'Asian Perspectives: Indian theories of mind', in Morris Moscovitch, Evan Thompson, and Philip David Zelazo (eds.), *The Cambridge Handbook of Consciousness* (Cambridge: Cambridge University Press).
- Galloway, B. (1978), 'Vijñāna, samjñā and manas', *The Middle Way* 53.2: 72–5.
- (1980), 'A Yogācāra Analysis of the Mind, Based on the Vijñāna Section of Vasubandhu's *Pañcaskhandaprakaraṇa* with Guṇaprabha's Commentary', *Journal of the International Association for Buddhist Studies* 3.2: 7–20.
- Ganeri, J. (1999), 'Self-Intimation, Memory and Personal Identity', *Journal of Indian Philosophy* 27.5: 469–83.
- (2007), *The Concealed Art of the Soul: Theories of Self and Practices of Truth in Indian Ethics and Epistemology* (Oxford: Clarendon Press).
- (forthcoming), *Mind's Own Nature: The Reconciliation of Naturalism with The first person perspective*, ch. 7.
- Kramer, J. (2008), 'On Sthiramati's *Pañcaskandhakavibhāṣā*: a Preliminary Survey', *Nagoya Studies in Indian Culture and Buddhism: Sambhāṣā*. 27: 149–72.
- McDowell, J. (2009), 'Referring to Oneself', in *The Engaged Intellect* (Cambridge, MA: Harvard University Press).
- Perry, J. (1998), 'Myself and I', in Marcelo Stamm (ed.), *Philosophie in Synthetischer Absicht* (Stuttgart: Klett-Cotta).
- Potter, K. H. (2003), *Buddhist Philosophy from 350 to 600 A.D.*, *Encyclopedia of Indian Philosophies*, vol. 9 (Delhi: Motilal Banarsidass Publishers).
- Recanati, F. (2007), *Perspectival Thought: A Plea for Moderate Relativism* (New York: Oxford University Press).
- Sartre, J. (1957), *The Transcendence of the Ego: An Existentialist Theory of Consciousness* (New York: Noonday Press).
- Schmithausen, L. (1987), *Ālayavijñāna: On the Origin and the Early Development of a Central Concept of Yogācāra Philosophy*, *Studia Philologica Buddhica*, 2 v. (Tokyo: International Institute for Buddhist Studies).
- Shoemaker, S. (1984), 'Self-Reference and Self-Awareness', in *Identity, Cause, and Mind* (Cambridge: Cambridge University Press).
- (1996), *The First-Person Perspective and Other Essays* (Cambridge: Cambridge University Press).

- Strawson, G. (2007), 'Selves', in Brian P. McLaughlin and Ansgar Beckermann (eds.), *The Oxford Handbook of Philosophy of Mind* (Oxford: Oxford University Press).
- (2008), *Real Materialism and Other Essays* (Oxford: Clarendon Press).
- (2009), *Selves: An Essay in Revisionary Metaphysics* (Oxford: Clarendon Press).
- Strawson, P. F. (1966), *The Bounds of Sense: An Essay on Kant's 'Critique of Pure Reason'* (London: Methuen).
- Waldron, W. S. (2002), 'Buddhist Steps to an Ecology of Mind: Thinking about "Thoughts Without a Thinker"', *Eastern Buddhist* 34.1: 1–52.
- Wittgenstein, L. (1960), *Preliminary Studies for the 'Philosophical investigations', Generally Known as the Blue and Brown Books* (Oxford: Blackwell).
- (1975), *Philosophical Remarks* (Chicago: University of Chicago Press).
- Zahavi, D. (1999), *Self-Awareness and Alterity: A Phenomenological Investigation* (Evanston, IL: Northwestern University Press).
- (2005), *Subjectivity and Selfhood: Investigating the First-Person Perspective* (Cambridge, MA: MIT Press).