

Is Consciousness Reflexively Self-Aware? A Buddhist Analysis.

Bronwyn Finnigan
School of Philosophy, RSSH, Australian National University

Correspondence: Coombs Building, Canberra, 2601, ACT, Australia
Email: bronwyn.finnigan@anu.edu.au

Abstract

This article examines contemporary Buddhist defences of the idea that consciousness is reflexively aware or self-aware. Call this the Self-Awareness Thesis. A version of this thesis was historically defended by Dignāga but rejected by Prāsaṅgika Mādhyamika Buddhists. Prāsaṅgikas historically advanced four main arguments against this thesis. In this paper I consider whether some contemporary defence of the Self-Awareness Thesis can withstand these Prāsaṅgika objections. A problem is that contemporary defenders of the Self-Awareness Thesis have subtly different accounts with different assessment criteria. I start by providing a fourfold taxonomy of these different views and then progressively show how each can withstand Prāsaṅgika objections. And I conclude by giving reasons to think that even some Prāsaṅgikas can accept some version of the Self-Awareness Thesis.

Keywords: consciousness, reflexivity, self-awareness, Buddhism, conventional truth

According to the 5-6th century Indian Buddhist philosopher, Dignāga, conscious experiences are self-aware or self-illuminating (*svasaṃvedana*). Call this the Self-Awareness Thesis. A growing number of contemporary philosophers defend its plausibility and argue both for its affinity with the continental European phenomenological tradition and its relevance to contemporary cognitive science and philosophy of mind (Coseru, 2012; MacKenzie, 2007, 2015; Siderits et al., 2011). It was historically opposed by Prāsaṅgika Mādhyamaka Buddhist philosophers, however, and their objections have been mobilised against its contemporary defence (Garfield, 2006, 2015). This article seeks to answer two questions: Can some contemporary defence of the Self Awareness Thesis withstand Prāsaṅgika Mādhyamaka critique and can some Prāsaṅgikas accept some version of the claim that conscious experiences are self-aware?

There are several factors that complicate this inquiry. Buddhist scholars discuss two historical accounts of self-awareness or *svasaṃvedana*; one ascribed to Dignāga and one to Śāntarakṣita (8th century). On neither account is self-awareness to be understood as awareness of a Self, where a Self is understood as an enduring substance that exists separate from and as the owner of events in conscious awareness. Nor is it self-conscious awareness of an 'I' that is somehow represented in each conscious experience. For both Dignāga and Śāntarakṣita, *svasaṃvedana* is the view that conscious experiences are aware *in some sense* of themselves. What this means potentially differentiates their accounts. Some follow Paul Williams (1998) in calling Dignāga's account the 'Self-Awareness' Thesis and Śāntarakṣita's the 'Reflexive-Awareness' Thesis. There are various ways to interpret these historical views, however, and subtle differences and convergences in what is defended in the name of self-awareness and reflexivity. These differences generate distinct criteria for assessment which causes confusion. This article will set aside the exegetical issue of how best to interpret Dignāga and Śāntarakṣita and will focus on what is defended in the name of self-awareness or reflexivity. It will call these accounts versions of the Self-Awareness Thesis. Given subtle differences between them, this article will begin by taxonomising the commitments of the most prominent versions before turning to assess whether they can withstand Prāsaṅgika Madhyamaka critique.

A further complication arises from the fact that historical Mādhyamikas invoke a distinction between 'two truths', conventional and ultimate, to critique the Self-Awareness Thesis. This distinction is controversial. As noted in Cowherds (2011), the Sanskrit term *satya* functions both in the sense of truth (a semantic property of statements) and reality (a truth-maker or ontological entity in reference to which statements gain their semantic properties). This causes some philosophical problems but the underlying idea is that truth is about what exists; ultimate truth is about what ultimately exists and conventional truth is about what conventionally exists. Buddhist and non-Buddhist philosophers interpret this distinction in

various ways and its plausibility and implications are fiercely contested. Abhidharma Buddhists, for instance, use it to defend nominalism. On this view, complex and conceptually structured entities (e.g. chariots, persons, tables) exist conventionally; they are conventionally real and claims about them are conventionally true. They do not exist ultimately, however, and claims that assert their existence are ultimately false. What ultimately exists are momentary token events individuated by their defining characteristics or intrinsic natures (Siderits, 2007). Madhyamaka Buddhists deny that anything ultimately exists in this sense. This matters for our purposes because Mādhyamikas take the Self-Awareness Thesis to be primarily a claim about ultimate reality and some of their objections turn on a general denial that anything can ultimately exist. Mādhyamikas nevertheless admit the use of conventional frameworks to characterize the nature of reality. Leaning on this admission, Śāntarakṣita argued that the Self-Awareness Thesis is compatible with Madhyamaka because it is merely a claim about conventional reality. Contemporary scholars make a similar move.

But what does it mean to say that the Self-Awareness Thesis is *conventionally* true? Its negative sense is clear; the explanatory target of the Self-Awareness Thesis is *not* some reductive base of momentary token events. In contemporary discussions, the claim also positively functions to locate the explanatory target as the everyday sense of phenomenal consciousness that is the typical explanandum of mainstream philosophy of mind and continental phenomenology. It also indirectly situates the thesis in dialogue with Yogācāra Buddhism. Yogācāra is traditionally understood as a form of metaphysical idealism whereby ordinary, pre-enlightened modes of conscious experiences are analysed as dependent on a more fundamental or ultimate mode of consciousness (Finnigan, 2017). Mādhyamikas typically deny the existence of an ultimate mode of consciousness. It might be reasonably thought that this denial is consistent with the claim that self-awareness correctly characterises the conventional

(viz. ordinary and pre-enlightened) mode of conscious experience that is the analysandum of Yogācāra.

Unfortunately, prominent Prāsaṅgika Mādhyamikas also deny the conventional status of the Self-Awareness Thesis. Thus, a simple appeal to conventionality will not sufficiently block their critique. Prāsaṅgikas challenge the Self-Awareness Thesis on multiple fronts, however, and it is sometimes difficult to distinguish arguments that target its conventional status from those that target its ultimate status. In some ways, this simplifies matters. Since Prāsaṅgikas deny both the ultimate and conventional truth of this thesis, we can simply ask whether some contemporary defence of the Self-Awareness Thesis can withstand any of their arguments. We can't take this approach to the second framing question of this article, however. We are also asking whether some Prāsaṅgika can *accept* some version of the Self-Awareness Thesis. Prāsaṅgikas offer subtly different accounts of the criteria for distinguishing conventionally true and false claims. To answer whether a Prāsaṅgika Mādhyamika can accept some version of the Self-Awareness Thesis we will need to consider their distinct views on the nature of the two truths.

This is clearly a lot to consider in one article. To limit its scope, this article will only engage some prominent contemporary defenses of the Self-Awareness Thesis and some prominent Prāsaṅgika critiques. Section 1 will provide a taxonomy of contemporary accounts of the Self-Awareness Thesis and Section 2 will explicate their distinct criteria for assessment. Sections 3 and 4 will assess these accounts against four of the most prominent Prāsaṅgika objections and will argue that they fail to successfully target contemporary views. Section 5 will argue that some contemporary versions of the Self-Awareness Thesis can provisionally satisfy some Prāsaṅgika criteria for conventional truth.

1. TAXONOMY OF CONTEMPORARY VIEWS ON SELF-AWARENESS

Dignāga famously claims that conscious experiences are self-aware. We are calling this the Self-Awareness Thesis. Dignāga articulates it as the claim that conscious experiences have “two forms [or appearances] ... the cognition [or awareness] of the object and the cognition [or awareness] of that [i.e. the cognition itself]” (PS 1.11. ab).¹ There are two components to this thesis.

First, a constitutive feature of conscious experiences is the appearance of an object. This aspect of Dignāga’s thesis is generally accepted by its historical and contemporary defenders and opponents. The idea is sometimes articulated by the claim that ordinary conscious experiences are intentional or have intentional objects or intentional content. Some care is needed with this terminology because intentionality is sometimes taken to imply a representationalist view of mental content; conscious mental states ‘intend’ an object in the sense that they are directed towards or represent some external object in mind-independent reality. Although Dignāga’s position on intentionality is complicated, representationalism is not a necessary assumption of his Self-Awareness Thesis. To see this, it is helpful to consider his view in relation to Yogācāra Buddhism. Yogācāra distinguishes phenomenology from representationalism. According to Vasubandhu’s Three Natures Theory, for instance, ordinary conscious experiences are dualistic in the sense of having an objective aspect and a subjective aspect (TSN 5). The objective aspect is the object-appearance. While there is some historical disagreement about the range of possible object-appearances it arguably includes simple sensations (e.g. of blue), spatio-temporally bounded, mid-size objects (e.g. of tables, of persons), and perhaps also less clearly demarcated ‘situations’ (e.g. of home). According to Vasubandhu, we often pre-reflectively take these object-appearances to represent real entities or properties in mind-independent reality. But this is a mistake. Properly understood, these

¹ The original translation is from Hattori (1968). See Kellner (2010) for arguments supporting the parenthetical alternatives.

objects are merely appearances in conscious awareness. A phenomenal appearance is neither a mind-independently real entity nor a representation of a mind-independently real entity. A contemporary philosopher might resist this final step but accept the implied distinction; a phenomenal appearance in conscious awareness is not necessarily a representation of an external, mind-independent object. This distinction is assumed by the first aspect of Dignāga's Self-Awareness Thesis. It is understood by its contemporary defenders as the phenomenological claim that the content of conscious experiences consists (at least in part) in a phenomenal object-appearance.

The second and more controversial aspect of Dignāga's claim is that conscious experiences are also constitutively aware *in some sense* of themselves. What this means (setting aside whether it is ultimately or conventionally true) is a matter of historical and contemporary dispute. Its contemporary defenders often use the same vocabulary to articulate the substance of this thesis but with different connotations. For instance, Dan Arnold proposes (as an interpretation of Dignāga), that to be self-aware is "to be first-personally aware" (2010: 362) which, he argues, "just is acquaintance with [our own cognitions] being contentful" (p.356). Thus understood, self-awareness does not contribute any additional phenomenal content to the experience itself but is the condition by which whatever appears appears. Matt Mackenzie also claims that self-awareness provides a "first-personal perspective" (2007: 48) but he takes this to involve the "phenomenal presentation of a subject" (2015: 282) which he once took to involve "a sense that the experience is the subject's *own*" (2007: 48) but more recently considers to merely consist of a sense in which the object appears "for a subject" (2015: 281). Christian Coseru, however, claims that self-awareness not only contributes a phenomenological sense of the object being "for me" (2012: 240), and "the experience [as being] hers" (p.242), but also provides a "vantage point", a "horizon of perceptual experience"

(p.256) and a distinct phenomenology which corresponds to the “concrete mode” in which objects are presented (p.256).

This variation obscures the target of assessment. To clarify matters, I propose a taxonomy of distinct but overlapping ideas that I take to be shared by prominent contemporary defences of the Self-Awareness Thesis. I call these ideas, respectively: Subjective Character, Modality, Structure, and Transcendence.

Subjective Character: According to this idea, conscious experiences involve awareness of phenomenological elements additional to that of the object-appearance. One such element is described as the subjective character of the experience and can be understood as a phenomenological rendering of Vasubandhu’s idea that conscious experiences have both a subjective and objective aspect. There are several versions of what this consists in. For some, the object-appearance always appears to or for a subject. This is often taken to mean that there is a phenomenal aspect of ‘for-me-ness’ in every conscious experience. The subjective aspect or character is this element of ‘for-me-ness’ (Coseru, 2012; Mackenzie, 2015). Some additionally claim that the experience also involves a sense of ownership whereby the object is experienced ‘as mine’. Ownership is thus also considered to be constitutive of subjective character (Coseru, 2012; Thompson, 2011; Zahavi, 2011).

Modality: According to this idea, conscious experiences (also) involve awareness of the specific modality or way an object-appearance appears. This is often described as the mode or manner of presentation of the object, an idea that broadly corresponds with the particular modality by which it is apprehended. These modalities tend to be individuated using folk attitude descriptions; e.g. seeing, desiring, hoping, remembering. According to this view, when we experience some object we are aware both of that object and that it is (e.g.) visually seen, or desired, or remembered. It is also claimed that these distinct modalities are marked by distinct phenomenological elements in the experience. For instance, following Husserl,

Thompson (2011) argues that memories contain a phenomenal sense of ‘pastness’ that is not part of the phenomenology of visually seeing that content. And, following Jinendrabuddhi’s commentary on Dignāga, it is argued that this functions to index content to the modality of the mental state and thereby help explain how (e.g.) remembering some previously seen content is not identical to seeing that content again (Hattori, 1968: 30).

Structure: This is the idea that conscious experiences have a dualistic structure. This is arguably presupposed both by Vasubandhu’s claim that conscious experiences have both a subjective and objective aspect as well as by Subjective Character and Modality. Contemporary scholars often express this structural duality by saying that objects are presented ‘first-personally’. What this substantively amounts to is often articulated in terms of Subjective Character; for an object to be presented first-personally is for it to appear ‘for’ a subject and/or ‘as theirs’. Others offer a more deflated view, suggesting that for an object-appearance to be presented first-personally *just is* for it to appear in conscious experience. The relevant structural duality on this deflated view is between awareness and content. This represents an alternative rendering of Vasubandhu’s position. Vasubandhu articulates the duality between the subjective and objective aspects of experiences using a metaphor of grasping; the subjective aspect is the grasping of the objective aspect which is grasped, both of which are necessarily involved in each conscious experience. On a deflated reading, conscious experiences necessarily involve the apprehension of some object-appearance but without the apprehension, itself, being apprehended. By contrast, Modality assumes that the mode of apprehension does contribute some phenomenal qualities to the experience. Structure is a necessary presupposition of both approaches.

Transcendence: According to this idea, it is a condition for the possibility of conscious experience that it be self-aware. This is often described as a transcendental analysis of self-awareness. There are different ways it is understood. The most prominent accounts extrapolate

self-awareness in terms of Subjective Character and Modality and as presupposing Structure. Call this Rich Transcendence. According to this view, a conscious experience is self-aware in the sense that it necessarily involves awareness both of an object-appearance as well as of its own subjective character (a phenomenological sense in which the object-appearance appears ‘for’ a subject and/or ‘as theirs’) and/or its mode or manner of presenting the object appearance (a phenomenological dimension of the experience that corresponds to the specific modality of the experience) where this presupposes that the experience is dualistically structured. Rich Transcendence is thus an analysis which unifies Subjective Character, Modality and Structure into a Self-Awareness Thesis. It is accepted by most contemporary defenders of this thesis. Some offer a more deflated view of Transcendence. On this view, a conscious experience is self-aware in the sense that it is aware of ‘its own’ content; namely, an object-appearance. Call this Deflated Transcendence. It is a version of Transcendence because it maintains that self-awareness is a condition for the possibility of conscious experience. However, where Rich Transcendence takes ‘self-awareness’ to pick out some additional phenomenal elements that are ‘immanent’ in conscious experience, Deflated Transcendence treats self-awareness as roughly (and perhaps trivially) equivalent to the conscious awareness of an object-appearance. This is a minority view that is sometimes attributed to Dignāga, sometimes to Śāntarakṣita, but not independently defended in contemporary literature.

2. CRITERIA FOR ASSESSMENT

Is the Self-Awareness Thesis conventionally true? The various component ideas in this thesis present distinct targets of assessment.

Consider Subjective Character and Modality. Both are proposed as correct phenomenological characterisations of ordinary conscious experience. They can thus be falsified if some ordinary conscious experience is shown not to contain the phenomenological

elements deemed necessary or ubiquitous. Mackenzie (2015) assumes this assessment criterion when he attempts to refute the claim that the phenomenal quality of ownership is a necessary feature of ordinary consciousness. However, Mackenzie appeals to extraordinary cases to make his point; (i) a neuropsychological study of someone who claims to lack a sense of self-ownership, and (ii) descriptions of the experiences of those engaged in advanced meditation techniques. Given that our explanandum is ordinary, conscious experience, (ii) seems to be beyond the scope because it arguably concerns supposed experiences of an ultimate or more fundamental mode of consciousness. (i) would also be irrelevant if by ordinary we meant normal and by normal we meant statistically typical because it concerns a statistically atypical case; the authors of the study claim that it is the only reported case in neuropsychological literature. This is not how we are using the term ordinary, however. Ordinary denotes merely what is not supramundane. Mackenzie's neurophysiological study thus carries weight.

Structure is the view that conscious experiences have a dualistic structure. It is presupposed by Subjective Character and Modality. It might seem not to be independently assessable. For instance, if we treat Subjective Character and Modality as correct and exhaustive in accounting for the phenomenological elements in conscious experiences additional to the object-appearance, then falsifying these views falsifies Structure because there would then be no 'secondary' phenomenological elements to serve as relata for a dualism. If this is right, it could helpfully simplify the assessment task; we need merely assess the truth of Subjective Character and Modality. There is a deflated version of Structure, however, whereby the relevant dualism is to be understood simply in terms of a distinction between awareness and content. This would survive the falsification of Subjective Character and Modality; insofar as there is conscious awareness of some object-appearance there is a structural duality between awareness and content. Deflated Structure can be falsified in at least two ways, however. First, one might argue that it simply amounts to the claim that conscious experiences are intentional;

they necessarily involve ‘awareness of’ some object-appearance. This claim can be falsified, in principle, if there is some conscious experience that lacks intentional content. Second, deflated Structure was earlier articulated in terms of a metaphor of grasping. As we will later discuss, this metaphor was traditionally analysed in terms of an act-patient relation and Prāsaṅgikas reject its presupposition by the Self-Awareness Thesis as internally inconsistent and regressive (Arnold, 2017). Deflated Structure might thus be falsified if, in fact, shown to be inconsistent and regressive.

Transcendence is presented as a transcendental analysis of conscious experiences. According to Rich Transcendence, a conscious experience would not be possible unless Subjective Character and/or Modality and Structure were the case. They are defining characteristics and hence necessary for all conscious experiences. According to Deflated Transcendence, a conscious experience would not be possible unless a deflated sense of Structure were the case; namely, awareness of an object-appearance. Neither Rich Transcendence nor Deflated Transcendence can be directly falsified by considering the phenomenological content of conscious experiences. This is because they are metaphysical claims about the ontological status of the experience. It might turn out that all conscious experiences do, in fact, have some sort of subjective character and modality but this would not thereby prove that these are *necessary* features. Transcendence makes this strong claim; conscious experiences *would not be possible* unless they were in some sense self-aware. How could this modal claim be falsified? If it cannot be falsified, it cannot be established as substantively true as opposed to a mere posit. What criteria for evaluating its truth could there be?

The Self-Awareness Thesis is often evaluated in terms of whether it is a necessary condition for the possibility of *other* related phenomena. This is a weaker and potentially more empirically grounded metaphysical claim. Prominent candidates of relevantly related

phenomena include (i) episodic memory, (ii) a substantive conception of self, and (iii) an immediate and indubitable acquaintance with phenomenal content. If these phenomena can be explained in terms that does not presuppose self-awareness in cases where it is posited as a necessary condition, this counts against the plausibility of self-awareness. It does not thereby falsify it; self-awareness might still be a necessary condition for conscious experience or a phenomenally rich sense of subjective character even if not for these additional phenomena. But if no positive argument can be given to establish the truth of this stronger claim, these counter-arguments provide reasons for doubt.

3. PRĀSAṄGIKA MADHYAMAKA OBJECTIONS

Prāsaṅgika Mādhyamikas deny the ultimate and conventional status of the Self-Awareness Thesis. In what follows I will consider some of the most prominent objections against this thesis, ultimate or conventional, and assess whether they hit any of the above contemporary targets.

No ultimate existents argument: Madhyamaka is a philosophical school that follows Nāgārjuna's reasoning in *Mūlamadhyamakakārikā*. It takes its philosophical target to be committed to the theses: (i) all things exist in dependence on causes and conditions, (ii) something is ultimately real only if it possesses an intrinsic nature, (iii) intrinsic nature is an essential property that secures numerical identity and independent existence. Through a series of reductio arguments, Nāgārjuna denies that anything can satisfy these criteria without contradiction. Most take this to imply that nothing satisfies the criterion specified in (ii) and there is thus no ultimate reality, so conceived. Prāsaṅgikas directly charge defenders of the Self-Awareness Thesis with falsely assuming an ultimate reality of inherent existents. For instance, as we will discuss, Dignāga appeals to a relationship between perceptual experience and episodic memory to help substantiate his thesis. In Mipham's commentary, Candrakīrti

responds by arguing that this supposes that past experiences and present recollections (i) are causally related, and (ii) ultimately exist and so have an intrinsic nature. But, according to Nāgārjuna, (i) and (ii) are inconsistent (Mipham, 2002: 247). This line of reasoning also informs some minor arguments to the effect that what satisfies (ii) does not exist and non-existing things do not cause anything. Mipham argues, for instance, that an intrinsically existing perception cannot cause an intrinsically existing recollection because a non-existent cause cannot cause a non-existent effect.

Since contemporary scholars defend the conventional status rather than the ultimate status of the Self-Awareness Thesis it would seem that they avoid this line of objection. However, the 14th century Tibetan Prāsaṅgika, Tsongkhapa, argues against this move. Tsongkhapa believes that underlying the appeal to conventionality is an attempt to define the general or essential nature of conscious experience, the truth of which, he believes, can only be established if one (fallaciously) posits essential properties that secure numerical identity and independent existence (Garfield, 2006). Contemporary defences of the Self-Awareness Thesis assume Transcendence and so do, indeed, attempt to define the general character of conscious experiences. And, as mentioned earlier, there is an unanswered question about how best to assess the truth of Transcendence. There is no reason to think, however, that it must be grounded in an essential property that secures both the numerical identity (and thus distinctness) of individual conscious experiences as well as their independent existence. Contemporary defences of the Self-Awareness thesis can thus avoid the no ultimate existents argument.

Unproven thesis argument: Candrakīrti objects that the Self-Awareness Thesis is unproved. Mipham elaborates that there is nothing that *could* prove its existence and that it is absurd to claim it exists if wholly unknown (2002: 245). This objection can be directed against Transcendence; if there is no way to prove the truth of this claim, if there is no way to know

whether self-awareness is indeed a necessary condition for conscious experiences, then it can be denied.

This is a strong objection against Dignāga if his Self-Awareness Thesis is best interpreted as committed to Transcendence. Dignāga was an epistemological empiricist. Accordingly, for a metaphysical claim to be warranted it must be epistemically established. If there is no legitimate epistemic way to establish the existence of some object, claims of its existence can be denied. According to Dignāga, there are only two legitimate epistemic means of establishing the truth of a claim: perception (*pratyakṣa*) and inference (*anumāna*). Transcendence cannot be established perceptually because it is a modal claim about the possibility of conscious experience rather than (simply) a phenomenological claim about its nature. It also cannot be established inferentially because according to Dignāga's abductive view of inferential reasoning we would need some perceptual evidence to serve as the supporting reason for the inference (Katsura, 1983). It would seem, however, that no relevant perceptual evidence can be provided. Since there is no other legitimate epistemic way to establish the truth of this thesis, the Self-Awareness Thesis can be denied.

This argument employs Dignāga's own conception of logic to defeat the Self-Awareness Thesis. It has very strong implications. In particular, it would seem to falsify all modal claims. Its contemporary defenders need not accept it. They could insist that Transcendence can be true even if unverified or unverifiable by experience. To avoid arbitrariness, they would need *some* supporting reasons but these need not be experiential. What kind of reasons could they be? They could be comparative reasons grounded in a study of alternative accounts and assessed against various theoretical virtues. Following Thompson (2011), for instance, one might argue that the Self-Awareness Thesis better explains the nature and possibility of conscious experience than its nearest competitor. The main historical and contemporary competitors are higher order or other-illuminating theories. To establish that self-awareness

better explains the nature and possibility of conscious experiences, we would need to assess the respective merits and deficiencies of these competing theories against standards of consistency, both internal and in relation to other theoretical commitments.

Regress argument: Candrakīrti accuses the Self-Awareness Thesis of being internally inconsistent and regressive. Dignāga accuses higher-order theories of being internally inconsistent and regressive. Similar accusations are levelled against contemporary versions of higher-order theories. Unfortunately, these accusations do not cancel each other out. Since it is beyond the scope of this paper to assess higher-order theories, I will engage with this issue hypothetically. If the Self-Awareness Thesis can withstand this critique and higher-order theories not, there is comparative reason to prefer the former to the latter. What then is Candrakīrti's argument and does it successfully refute contemporary defences of self-awareness? According to Arnold (2017), Candrakīrti understands the structural dualism implied by the self-awareness thesis to be an act-patient relation; there is an act of apprehending (grasping) some content which is apprehended (grasped). If the act of apprehending is also apprehended, then either (i) the act of apprehending apprehends itself, or (ii) there is some further act of apprehending by which the initial act of apprehending is apprehended. According to Mipham, Candrakīrti thinks (i) is as absurd as the idea of a sword cutting itself, a finger touching itself, or an eye seeing itself (2002: 245) and (ii) is regressive.

Is this a genuine dilemma? A tempting response is to deny (ii), accept (i) and argue that it begs the question to deem it absurd. But this is too quick. On Arnold's analysis, Candrakīrti's charge of absurdity is grounded in a Sanskrit grammatical distinction between acts and patients according to which agents and patients of acts are necessarily distinct; nothing can be both the agent and patient of the same act. If apprehension is an act, it is grammatically distinct from what it apprehends. Arnold thinks this grammatical distinction should be challenged. However, grammatical issues are often difficult to untangle from logical issues, which can be grounds for

refutation. A similar grammatical concern arises for the English term ‘awareness’ and the related expression ‘awareness of’ in contemporary articulations of the Self-Awareness Thesis. All defenders of the Self-Awareness Thesis accept that conscious experiences are aware both of some object-appearance as well as in some sense of themselves. As we have seen, some interpret the underlying structural dualism as a general distinction between awareness and content. And many characterise self-awareness in phenomenological terms. One might argue that this combination of views gives rise to a dilemma. On the one hand, it could be argued that phenomenological articulations of self-awareness just add more content to the experience rather than substantiating what it could mean for an experience to be ‘aware of itself’ without reference to its phenomenal content. On the other hand, positing an ‘awareness-of’ relation without phenomenal content as referent seems grammatically (if not logically) absurd. The Self-Awareness Thesis thus assumes either that conscious experiences necessarily involve (i) ‘awareness-of- x ’ where x is empty, which is absurd, or (ii) awareness of the subjective and objective dimensions of content but needs some additional act of awareness for there to be awareness of this awareness, which is regressive. The resulting dilemma parallels that originally identified by Candrakīrti. Neither position is acceptable, so the Self-Awareness Thesis can be denied.

This dilemma, while challenging, can be resisted. One approach is to argue that it conflates two distinct interpretations of Structure that were identified in our taxonomy and which, if treated independently, can slip between its horns.

First, we identified a deflated sense of Structure according to which conscious experiences presuppose a structural dualism between awareness and content but without the awareness, itself, forming part of its content. How is the conscious experience ‘self’ aware? Following Arnold (2010), one might argue that it is aware of ‘its own’ content; namely, its object-appearance. How then is self-awareness distinct from conscious awareness or just plain

consciousness? According to Arnold, there is no difference. Dignāga's Self-Awareness Thesis is just a straightforward analysis of conscious experiences; they are of some phenomenal content and are aware of it. Arnold (2017) further argues that Śāntarakṣita defends this deflated sense of self-awareness in his famous claim that self-awareness marks conscious experiences as "distinct from inanimate things" like rocks and tables and stars. If so, then Dignāga, Śāntarakṣita, and Self-Awareness Theories that presuppose deflated Structure (and deflated Transcendence) can avoid the above dilemma. Conscious experiences involve awareness of phenomenal content, are 'self'-aware in the sense that they are aware of 'their own' content, and do not need additional states or acts of awareness for this to be so.

Most contemporary defenders of the Self-Awareness Thesis accept some form of Rich Transcendence whereby self-awareness involves awareness of Subjective Character and/or Modality in addition to some object-appearance. And most also hold that the relevant structural dualism is internal to or constitutive of the content of the experience. Object-appearances appear for-me, as-mine, and are presented in a certain manner or mode. Why think this shows that the conscious state is aware 'of itself' rather than simply aware of some additional phenomenal content? In reply, one might argue that this question unnecessarily insists on a sharp distinction between awareness and phenomenal content. Consider Modality. On this view, conscious experiences always take the form of some particular mode of apprehending object-appearances, where particular modes are individuated (at least in part) in terms of phenomenal qualities. The paradigm case is episodic memory. On Thompson's analysis, episodic memory phenomenologically contains an object-appearance as well as a sense of pastness that is phenomenologically constitutive of the experience. Episodic memories are self-aware, one might argue, in the sense of being aware of the phenomenal qualities that are distinctive of 'their own' particular mode of apprehension. This view also avoids the above dilemma. Conscious experiences involve awareness of (a distinct kind of) phenomenal content,

are ‘self’-aware in the sense that this element of content is distinctive of their own particular kind of experience, and do not need additional states or acts of awareness for this to be so.

4. THE CASE OF EPISODIC MEMORY

Most contemporary defenders of the Self-Awareness Thesis argue that it is a necessary condition for some other phenomena. Dignāga and Dharmakīrti both argue that it is a necessary condition for episodic memory. These arguments are receiving a contemporary revival (Ganeri, 2018; Kriegel, 2019; Thompson, 2011). According to Dignāga, the recollection of some previously seen object involves recollecting that it was perceived. This presupposes that the original perception contained both the object and the *perceiving* of it (PS 1.11d). Dharmakīrti goes further and argues that if such were not the case, recollection would be simply another instance of visually perceiving that object. Since there is clearly a difference between recollection and sense perception, the latter must be self-aware (Kellner, 2011).

Prāsaṅgika Mādhyamikas are often presented as responding to these arguments by defending a higher-order theory of consciousness as alternative explanation. Śāntideva, for instance, responds by claiming “memory comes from connection with another experience, like rat’s poison” (BCA 9:23). Garfield (2006) explicates the simile as follows: just as a hibernating bear, which is bitten by a rat while sleeping and wakes to experience the pain caused by infection (poison), can ‘know’ (or infer) that it was bitten despite having not experienced the bite, so too one can develop a cognitive state directed at a past perceptual episode even if one was not also aware that one was perceiving at the time of that perceptual episode. Most agree that this case is quite strange. Several have pointed out that the simile is inapt as it is analogous to the claim that one can have an episodic memory of a mental state that was never consciously experienced (Bernier, 2015; Thompson, 2011). Most set the example aside and simply take Śāntideva to be directly opposing the Self-Awareness Thesis by expressing commitment to a

higher order theory of consciousness. While this would not disprove the Self-Awareness Thesis, its viability as an alternative explanation would undermine the rational support derived from the claim that self-awareness is a necessary condition for episodic memory.

There is a growing literature on this issue. One point often overlooked is that it is not obvious that Śāntideva *can* consistently express a commitment to a higher order theory of consciousness. This is because, traditionally understood, the Prāsaṅgika criteria for conventional reality disallows the positive defence of theses and explanations. As mentioned, Nāgārjuna employs reductio arguments to deny that anything can both be dependently originated and ultimately exist with an intrinsic nature. Most take this to imply that nothing ultimately exists. Candrakīrti goes further and insists that Nāgārjuna's reasoning does not establish a positive thesis as the result of a valid argument. If it did, he thinks it would thereby assert an ultimate truth about the ultimate nature of reality and result in the same inconsistencies as Nāgārjuna's opponents. To be consistent, Candrakīrti thinks that Mādhyamikas are methodologically constrained to using reductios (*prasaṅga*) against their opponent's theses but without having a thesis of their own. The label 'Prāsaṅgika' represents this philosophical position. Tibetan philosophers fiercely divide over how it is best understood (Dreyfus & McClintock, 2003). On a traditional interpretation, it entails that a Mādhyamika neither defends a position on ultimate reality nor appeals to such a position to analyse and adjudicate competing claims about conventional reality (Tillemans, 2016). This does not foreclose the *holding* of any positive views about conventional reality. However, it was thought to methodologically constrain Mādhyamikas to simply accepting the common-sense views of the person on the street without subjecting it to rational or epistemic analysis. The term for this methodological position is *lokaprasiddha*; accepting what 'the world acknowledges' to be the case. It can be understood as both a methodological constraint and a criterion for conventional truth.

If a Prāsaṅgika must adhere to *lokaprasiddha* as traditionally understood it follows that they cannot rationally defend a higher-order explanation of episodic memory in opposition to the Self-Awareness Thesis. This is because they cannot defend or explain anything. Prāsaṅgikas accept and practice according to the widely accepted standards and language of ordinary, everyday folk but do not engage in deeper philosophical questioning about justificatory status or grounds. According to Tillemans (2016), this conception of Prāsaṅgika Madhyamaka is best understood as a form of Quietism, tasking itself with pointing out the inconsistencies that plague philosophical theorising without embroiling itself in those problems. If this is right, one might argue that Śāntideva was inconsistent when he offered an alternative explanation to undermine the claim that the Self-Awareness Thesis is necessary for the possibility of episodic memory. But this is too hasty. While *lokaprasiddha* disallows philosophical defence or explanation, it permits assertions so long as they simply express beliefs about (some) folk ontology. Śāntideva's assertion "memory comes from some other experience" might be understood in just such terms; it is pretty banal and few would disagree. Unfortunately, this claim is also consistent with the view that episodic memory can only be explained by the Self-Awareness Thesis. A philosophical position is not refuted by a refusal to philosophise. As a result, the Self-Awareness Thesis can withstand this response.

5. REVISITING MADHYAMAKA CRITERIA FOR CONVENTIONAL TRUTH

There are notorious difficulties with *lokaprasiddha* as a criterion for conventional truth. It was heavily criticised even by those sympathetic to Madhyamaka. The 8th century Indian Buddhist philosopher Kamalaśīla, for instance, complained that it absurdly entailed that every belief would be true simply because its content was acknowledged to be the case, including the 'mistaken' views that the Buddha claimed to be at the root of suffering (Tillemans, 2016). This approach flattens out conventional truth to mere belief without offering any intelligent means

of adjudicating competing beliefs. This is especially problematic when extended to moral judgments (Finnigan, 2015). But even in contexts where one can reasonably predict consensus, it unduly restricts the possibilities of rational inquiry. Not only does it follow that the Self-Awareness Thesis is conventionally false, by its lights, but so too is every competitive theory. This is because *all* theorisation is excluded from the domain of the conventional. It follows that no penetrating analysis in the philosophy of mind will count as either ultimately or conventionally true. This is a depressing outcome.

There is reason to think that not all Prāsaṅgikas adhere to such a strict criterion for conventional truth, however. Arnold (2005), for instance, provides textual evidence to suggest Candrakīrti admitted ‘defining characteristics’ within the scope of *lokaprasiddha* and so Arnold thinks he *should* permit the activity of defining the conventional nature of ordinary conscious experience. This would not, of itself, establish that the Self-Awareness Thesis is the correct definition. Tsongkhapa follows Candrakīrti in denying its conventional truth. But *must* Tsongkhapa, or a Prāsaṅgika more generally, accept this conclusion? Tillemans (2016), for instance, identifies Tsongkhapa as the main representative of a group of Mādhyamikas that seek to preserve the possibility of rationally analysing conventional truth without reintroducing the notion of intrinsic nature. There are at least two classes of rational norms that they could reasonably admit; logical norms and widely accepted epistemic standards (Finnigan, 2018). Nāgārjuna adhered to logical norms of coherence and consistency to refute his opponent’s theses. Since Prāsaṅgikas follow Nāgārjuna in his reasoning, coherence and consistency should be acceptable norms for the rational analysis of conventional beliefs. In addition, Candrakīrti identified four epistemic means of establishing the truth of ontological claims that, in his view, are widely accepted by the ‘people of the world’. They are: empirical observation, inference, reliable testimony and analogical similarity (Ruegg, 1981). Rather than restricting

lokaprasiddha to the assertion of widely-held folk ontological beliefs, it might be extended to permit their rational assessment against these widely-held epistemic standards.

How should these rational norms function as criteria for assessing accounts of the Self-Awareness Thesis? One might straightforwardly argue that an account is conventionally true if it satisfies some of the above norms. But this standard for conventional truth is too weak. It would follow that (e.g.) mere analogical similarity would be sufficient. This seems wrong. Defenders of the Self-Awareness Thesis often appeal to the analogy of a lamp illuminating itself when illuminating other things to support their position. But this is thought merely to add plausibility by contributing illustrative clarity rather than sufficiently establishing truth. Should we then remove analogical similarity from the list? Not necessarily. We might instead interpret the above epistemic standards as an ordered list with a decreasing scale of evidential support corresponding to a degreed analysis of conventional truth. Analogical similarity would thus provide some plausibility to an account of the nature of conscious experience, but not as much as reliable testimony or abductive inference (I will return to this), both of which gain their authority in relation to empirical observation.

If we adopt this strategy, does some account of the Self-Awareness Thesis count as conventionally plausible to some degree according to this epistemic scale? Prāsaṅgikas unanimously deny that self-awareness is analogically similar to a lamp illuminating itself. But even if it were, it would only constitute a very weak form of evidence. What about reliable testimony? This criterion is relevant in contexts where empirical observation is not practically accessible but is available in principle. It is because the assertions of some authority figure are grounded in valid inference and/or substantiated by empirical evidence that we (do or ought to) treat them as authoritative. Our argument is a principled argument, so I will set this aside.

That leaves inference and empirical observation. Subjective Character and Modality might satisfy the empirical observation criterion. The same cannot be said for Transcendence

which, as we have discussed, is a modal claim and thus not directly verifiable by experience. If the conventional criterion of inference is abductive (akin to that of Dignāga as discussed earlier) then it also cannot be established inferentially. However, Prāsaṅgikas can also appeal to inferences that draw on rational norms of consistency and coherence. If the arguments in this paper have been successful, they demonstrate that contemporary defences of the Self-Awareness Thesis are internally consistent (or at least, that they withstand prominent objections against their consistency). Are they also coherent explanations of ordinary conscious experiences? If we interpret coherence to imply consistency with a wider network of conventionally true beliefs that satisfy the above epistemic standards, then the answer is arguably ‘yes’. This does not yet establish that the Self-Awareness Thesis is conventionally true. This is because there may be other internally consistent and coherent explanations of ordinary consciousness. Higher-order theories of consciousness, for instance, might also satisfy these criteria (if shown to withstand Dignāga’s regress objection). The Self-Awareness Thesis and Higher-Order theories are mutually exclusive and so jointly inconsistent. But since consistency is a rational norm we are deeming admissible to Prāsaṅgika for the assessment of conventional truth claims, it can justify their comparative assessment. Undertaking this task is beyond the scope of what can be achieved here. But on the assumption that some version of the Self-Awareness Thesis can succeed in such comparative assessment, it can withstand Prāsaṅgika critique and warrantably count as conventionally true.

CONCLUSION

This article critically examined contemporary accounts of the Self-Awareness Thesis and considered whether they can withstand Prāsaṅgika critique and count as conventionally true. It provided a taxonomy of four distinct but overlapping views that are shared by prominent contemporary defences of this thesis; Subjective Character, Modality, Structure and

Transcendence. It also demonstrated their distinct assessment criteria. It then showed how these views can withstand four of the most prominent Prāsaṅgika objections; the no ultimate existents argument, the unproven thesis argument, the regress argument, and positing a higher-order theory of consciousness as alternative explanation. It then considered whether Prāsaṅgika Mādhyamikas can admit *any* positive analysis of ordinary consciousness. This required considering Prāsaṅgika criteria for distinguishing conventional truth and falsity and concluded that contemporary defences of the Self-Awareness Thesis can satisfy some such criteria. This is only a provisional conclusion, however. To sufficiently establish its conventional truth requires comparative assessment against competing theories. But that is a task for another article.

ACKNOWLEDGEMENTS

I would like to thank Dan Arnold, Evan Thompson, the ANU Centre for Consciousness Work in Progress Group, and the editor of this special issue for helpful comments on an earlier draft of this paper.

ABBREVIATIONS

BCA *Bodhicaryāvatāra* of Śāntideva in Wallace & Wallace
PS *Pramāṇasamuccaya* of Dignāga in Hattori
TSN *Trisvabhāvanirdeśa* of Vasubandhu in Anacker

REFERENCES

- Anacker, S. (1998). *Seven Works of Vasubandhu*. Delhi: Motilal Banarsidass.
- Arnold, D. (2005). *Buddhas, Brahmins and Belief*. New York: Columbia University Press
- Arnold, D. (2010) Self-awareness (*svasaṃvitti*) and Related Doctrines of Buddhists Following Dignāga: Philosophical Characterizations of Some of the Main Issues, *Journal of Indian Philosophy* 38: 323-378
- Arnold, D. (2017) Should Buddhists deny that we are conscious? Revisiting Mādhyamika critiques of self-awareness, Unpublished article.
- Bernier, P. (2015) Dignāga on Reflexive Awareness *Philosophy East & West*, 65, 1, 125–156

- Coseru, C. (2012). *Perceiving Reality: Consciousness, Intentionality and Cognition in Buddhist Philosophy*. Oxford: Oxford University Press
- Cowherds, The (2011). *Moonshadows: Conventional Truth and Buddhist Philosophy*. Oxford University Press.
- Dreyfus, G. & McClintock, S. L. (Eds.) (2003). *The Svātantrika-Prāsaṅgika Distinction*. Boston: Wisdom Publications
- Finnigan, B. (2015). Madhyamaka Buddhist Meta-Ethics: The justificatory grounds of moral judgments, *Philosophy East and West*, 65, 3, 765-785
- Finnigan, B. (2017). Buddhist Idealism, in Goldschmidt, T. & Pearce, K. (Eds.). *Idealism: New Essays in Metaphysics*. Oxford: Oxford University Press
- Finnigan, B. (2018). Madhyamaka Ethics, in Cozort, D. & Shields, M. (Eds.). *Oxford Handbook of Buddhist Ethics*. Oxford: Oxford University Press
- Ganeri, J. (2018). Mental Time Travel and Attention, *Australasian Philosophical Review*, 1, 4, 353-373
- Garfield, J.L. (2006). The Conventional Status of Reflexive Awareness: What's at Stake in the Tibetan Debate, *Philosophy East and West*, 56, 201-228
- Garfield, J.L. (2015). *Engaging Buddhism: Why it matters to philosophy*. Oxford University Press
- Hattori, M. (1968). *Dignaga, On Perception*. Cambridge, Mass: Harvard University Press
- Katsura, S. (1983). Dignāga on Trairūpya, *Journal of Indian and Buddhist Studies* 32, 1, 15-21.
- Kellner, B. (2010). Self-Awareness (*svasaṃvedana*) in Dignāga's *Pramāṇasamuccaya* and –*vṛtti*: A Close Reading, *Journal of Indian Philosophy*, 38, 203-231
- Kellner, B. (2011). Self-awareness (*svasaṃvedana*) and Infinite Regresses: A Comparison of Arguments by Dignāga and Dharmakīrti, *Journal of Indian Philosophy*, 39, 411-426
- Kriegel, U. (2019). Dignāga's Argument for the Awareness Principle: An Analytic Refinement, *Philosophy East & West*, forthcoming
- Mackenzie, M. (2007). The Illumination of Consciousness: Approaches to Self-Awareness in the Indian and Western Traditions, *Philosophy East and West*, 57, 1, 40-62
- Mackenzie, M. (2015). Reflexivity, Subjectivity, and the Constructed Self: A Buddhist Model, *Asian Philosophy*, 25, 3, 275-292
- Padmakara Translation Group (Trans.) (2002). *Introduction to the Middle Way: Candrakīrti's Madhyamakavatara with commentary by Jamgön Mipham*. Boston & London: Shambala

Ruegg, D.S. (1981). *The Literature of the Madhyamaka School of Philosophy in India*. Wiesbaden: Otto Harrassowitz.

Siderits, M. (2007). *Buddhism as Philosophy*. Aldershot, Burlington, Ashgate.

Tillemans, T. (2016). *How Do Mādhyamikas Think?* Boston: Wisdom Publications

Thompson, E., (2011). Self, No Self? Memory and Reflexive Awareness, in Siderits, M., Thompson, E., & Zahavi, D. (Eds.). *Self, No Self?* Oxford: Oxford University Press, 157-175

Siderits, M., Thompson, E., & Zahavi, D. (Eds.). (2011). *Self, No Self?* Oxford: Oxford University Press

Wallace, V.A. & Wallace, B.A. (1997). *A Guide to the Bodhisattva Way of Life*. Ithaca, New York: Snow Lion Publications

Westerhoff, J. (2009). *Nāgārjuna's Madhyamaka: A Philosophical Introduction*, Oxford: Oxford University Press

Williams, P. (1998). *The Reflexive Nature of Awareness*, Delhi: Motilal Banarsidass

Zahavi, D. (2011). The Experiential Self: Objections and Clarifications, in Siderits, M., Thompson, E., & Zahavi, D. (Eds.). *Self, No Self?* Oxford: Oxford University Press, 56-78