3 Dignāga and Later Developments

Summary 3.1 Dignāga on *Svasaṃvedana*: A Hotly Debated Issue. – 3.1.1 Is Self-Awareness Included in Mental Perception or Is It a Separate Type of Perception? – 3.1.2 Self-Awareness, Memory, and the Two Forms (or Aspects) of Cognition. – 3.1.3 Self-Awareness as *Pramāṇaphala*. – 3.1.4 Some Readings of Dignāga's *Svasaṃvedana*. – 3.2 Later Developments in the Issue of *Svasaṃvedana*. – 3.2.1 Dharmakīrti and the Continuation of Dignāga's Epistemological Project. – 3.2.2 Candrakīrti and Śāntideva: A Sharp Criticism of *Svasaṃvedana*. – 3.2.3 Śāntarakṣita's Contribution.

3.1 Dignāga on *Svasaṃvedana*: A Hotly Debated Issue

In Buddhism, the notion of *svasaṃvedana* is first articulated formally within an epistemological framework by Dignāga, the principal figure behind the epistemological shift in the fifth and sixth centuries AD. To better appreciate this Indian scholar's pivotal philosophical contribution, his writings will also be analyzed from the perspectives of modern philologists and philosophers who have engaged with these materials, the aim being to integrate his thought with the questioning process in the ongoing debates on the nature of self-awareness. The intention is to move beyond the comparative method that scholars have relied upon primarily thus far, in order to engage more critically and deeply with the implications of Dignāga's philosophy.

There are two main texts in which Dignāga deals with *svasaṃvedana* in detail. Both have been lost in Sanskrit, their

DOI 10.30687/978-88-6969-929-0/003

original language: the *Pramāṇasamuccaya* (henceforth *PS*) and its *-vṛtti* 'autocommentary' (henceforth *PSV*). The first chapters of these texts, dealing with perception, form the main passages devoted to the issue and will be analyzed in order to show the extent of variation in their interpretation. A comprehensive examination of Dignāga's treatment of self-awareness across these two texts would necessitate a far-reaching investigation that exceeds the limits of this study. Thus, attention will be restricted to the most contested elements of the concept of *svasaṃvedana*.

The main difficulties in this analysis stem from the selection and interpretation of the textual sources. In recent decades, modern scholars have often debated Dignāga's account of self-awareness; only in the last few years, however, have some philological studies been able to provide additional textual materials. In addition to some extant Sanskrit fragments (Hattori 1968), there are also two often problematic and divergent Tibetan translations of Dignāga's *PSV* – one by Kanakavarman and the other by Vasudhararakṣita – as well as a relatively recent hypothetical reconstruction of the original Sanskrit text undertaken by Steinkellner (2005; 2014). Steinkellner et al. reconstructed this text based on Jinendrabuddhi's (ca. eighth century AD) commentary on Dignāga's work, which frequently incorporates material from the original text (Steinkellner, Krasser, Lasic 2005). However, it must be noted that relying on this commentary carries the risk of superimposing Dharmakīrti's later, and possibly distinct, ideas onto Dignāga's thought, as Jinendrabuddhi's interpretation closely aligns with Dharmakīrti's *PV*. In light of these materials, Kellner (2010), in particular, has attempted a radical reconsideration of the issue of Dignāga's account of self-awareness, thus rebooting the contemporary debate on it and advancing a new interpretation that challenges significantly the established exegetical tradition and aligns with some modern perspectives on self-awareness. Yao (2004; 2005) is another who invites the contemporary scholarly community to consider also what Dignāga's *Nyāyamukha* has to say about this topic and disapproves of the lack of adequate attention this earlier text has received so far. A more recent tendency in scholarship – seen, for instance, in Steinkellner (2005) and Kellner (2010, 207) – has been to read the *PS* and *PSV* together as a single, unified composition in which verse and prose are interwoven rather than to treat them as two independent works. As one might expect, this approach has led to an array of intricate philological issues.

3.1.1 Is Self-Awareness Included in Mental Perception or Is It a Separate Type of Perception?

The following section engages with the philosophical foundations of self-awareness as formulated by Dignāga, examining its defining features within his epistemological system. *Svasaṃvedana* is a *pramāṇa* 'epistemically valid cognition' that is free from conceptual construction[1] and included within perceptions that are independent of the senses (SKT *indriyāni*). This category encompasses mental perceptions of external objects as well as self-awareness of desire and other mental factors,[2] and also includes conceptual cognitions.[3] Before going into the details of Dignāga's account of self-awareness, however, it is important to locate this specific type of cognition within the broader system of perception as outlined by Dignāga, a framework on which modern scholarship has yet to reach a consensus.[4]

Some scholars argue that Dignāga subsumes self-awareness within mental perception,[5] while others claim that he considers it as a separate type of perception. The former group includes Hattori (1968), Nagatomi (1980), Franco (1993; 2005), and Kellner (2010); the latter is composed of Wayman (1977-78; 1991) and Yao (2004; 2005). In a brief examination of the main arguments each side advances,

1 Dignāga, *PS* 1.3c: "***pratyakṣaṃ kalpanāpoḍhaṃ***". Please note that, here and in the following quotations, the Sanskrit text of the *PS* and *PSV* is taken from Steinkellner's (2005; 2014) hypothetical reconstruction (revised). Fully reconstructed words without attested Sanskrit fragments are in roman typeface; bold script is used for the *ślokas* of *PS* (and for the words of the *ślokas* used in the prose of the *PSV*).

2 Dignāga, *PS* 1.6ab: "***mānasaṃ cārtharāgādisvasaṃvittir akalpikā |***". The related autocommentary (Dignāga, *PSV* 1.6ab) states: "*mānasam api rūpādiviṣayālambanam avikalpakam anubhavākārapravṛttaṃ rāgādiṣu ca svasaṃvedanam indriyānapekṣatvān mānasaṃ pratyakṣam*". In Hattori's translation: "The mental [perception] which, taking a thing of color, etc., for its object, occurs in the form of immediate experience (*anubhava*) is also free from conceptual construction. The self-awareness (*sva-saṃvedana*) of desire, anger, ignorance, pleasure, pain, etc., is [also recognized as] mental perception because it is not dependent on any sense-organ" (Hattori 1968, 27; square brackets in the original). Note that both of the extant Tibetan versions (Hattori 1968, 180-1) are identical in listing instances of mental factors: "'*dod chags dang zhe sdang dang gti mug dang bde ba dang sdug bsngal la sogs pa ni*" (attachment, anger, ignorance, pleasure, pain etc.).

3 Dignāga, *PS* 1.7ab: "***kalpanāpi svasaṃvittāv iṣṭānārthe vikalpanāt |***". In Hattori's translation: "7ab. even conceptual construction, when it is brought to internal awareness, is admitted [as a type of perception]. However, with regard to the [external] object, [the conceptual construction is] not [admissible as perception], because it conceptualizes [the object]" (Hattori 1968, 27; square brackets in the original).

4 For a brief summary of the controversy, see Kellner 2010, 207 fn. 11; Coseru 2012, 254 fn. 69.

5 For a detailed study on the role of *mānasa-pratyakṣa* 'mental perception' within the whole cognizing process, in its Indian and Tibetan understandings, see Tillemans 1989.

many questions will arise regarding the nature of self-awareness and how it relates to other forms of cognition.

The earliest position on this issue was advocated by Hattori on the basis of his rendering of *PS* 1.6ab:

> there is also mental [perception, *which is of two kinds*:] awareness of an [external] object and self-awareness of [such subordinate mental activities as] desire and the like, [both of which are] free from conceptual construction. (Hattori 1968, 27; square brackets in the original; italics added)

Nagatomi (1980, 245) concurs that Dignāga subsumes self-awareness under mental perception but offers a different interpretation of this passage. According to his understanding, Dignāga would be referring not to two distinct 'kinds' of mental perception (SKT *mānasa-pratyakṣa*), but, rather, to two different 'aspects' of it: the object-cognizing aspect and the self-cognizing aspect. Nagatomi's claim is supported by two main reasons. First, if one were to read two types of perception in *PS* 1.6ab, the passage would offer no indication as to why mental perception must be acknowledged or how it differs from sense perception (SKT *indriya-pratyakṣa*). Second, it would turn out to be incongruous with Dignāga's *svasaṃvedana* theory, according to which self-awareness is considered a *pratyakṣa* 'perception' and applies to every mental event, indicating that, because of its self-luminous nature, a given cognition does not need another cognition to be cognized. These two difficulties, Nagatomi explains,

> can be removed if we understood the passage as Dignāga's explication of the first aspect, viz. the 'awareness of an object' or the object-cognizing aspect, of *mānasa-pratyakṣa* whose second *aspect* is its 'self-awareness' in the form of passion, etc. (Nagatomi 1980, 245; italics in the original)

Thus, in the former aspect, *mānasa-pratyakṣa* manifests itself in an object-form, while, in the latter aspect, it cognizes it. Nagatomi finds further support for his interpretation in Dignāga's reference (*PS* and *PSV* 1.11-12) to the two-fold appearance of every mental event, namely, the appearance of an object and that of cognition itself, which will be discussed in § 3.1.2.

Franco (1986) is perhaps the scholar most devoted to this issue, taking active part in the controversy. While it is true that he generally aligns with Hattori and Nagatomi in reading Dignāga as subsuming self-awareness under mental perception, he nevertheless articulates his own position by contrasting the opposing views of Hattori and Wayman and proposing a third interpretation that

incorporates elements of both, thus seeking a "compromise between the two" (Franco 1986, 79). Hattori interprets *svasaṃvedana* as a kind of mental perception that takes a subordinate mental activity as its object whereas Wayman considers it to be a distinct type of perception, separate from the mental one and the others. Franco's position navigates between these poles, seeking to reconcile their insights. Franco states about Dignāga:

> he did not accept *svasaṃvedana* as a fourth type of perception. The point becomes quite clear when we look at *Pramāṇasamuccaya* 1.6ab and the *Vṛtti* thereupon: in a manner which leaves no place for ambiguity Dignāga *subsumes* the self-apprehension of desire etc., under mental perception! (Franco 1986, 82; italics in the original)[6]

Franco's view represents a middle path: it incorporates Hattori's perspective by classifying *svasaṃvedana* within the domain of mental perception while, at the same time, treating it as an 'aspect of cognition' – a position that, albeit within the restricted framework of mental perception, appears to align with Wayman's interpretation of *svasaṃvedana* as a distinct type of perception.

In 1993 Franco came back to this topic, responding to Wayman's provocative footnote (Wayman 1991, 423 fn. 17) which directly expressed his criticism of Franco's reading of the passage in *PSV* 1.6ab. In the first pages of his article, Franco (1993) reflects on his understanding of the aforementioned passage against Wayman's interpretation, which holds that Dignāga posits four kinds of perception: *indriya*, *mānasa*, *svasaṃvedana* and *yogi* (Wayman 1977-78, 393). Then, addressing the issue of the Sanskrit rendering and its interpretation and making a strong case for reading only three types of perception in the *PS*, Franco demonstrates that the assumption of a four-fold perception in Dignāga's view is highly doubtful and not supported by Dignāga's own words but rather "only by Dharmakīrti's reshuffle of them" (Franco 1993, 295).

Although it appears, as Coseru (2020, 124) suggests, that Dignāga formulated *svasaṃvedana* within an epistemological framework particularly concerned with the nature and function of perception,

6 The main debate Franco recalls and relaunches revolves around the following passage of Dignāga's *PS* (1.7cd-8ab): "*bhrāntisaṃvṛtisajjñānam anumānānumānikam || smārtābhilāṣikaṃ ceti pratyakṣābhaṃ sataimiram |*". In Hattori's translation: "erroneous cognition, cognition of empirical reality, inference, its result, recollection, and desire are not true perceptions and are accompanied by obscurity (*sataimira*)" (Hattori 1968, 28). For all the references related to this philological and philosophical debate, revolving around the meaning of the word *sataimira* in this context, cf. Hattori 1968, 95-7 fn. 53; Wayman 1977-78; Franco 1986.

Franco challenges the broader interpretative framework that underlies much of contemporary exegesis:

> However, even by arguing for three against four types of perception, we are already caught in Dharmakīrti's web. For by doing so we already presuppose that Dignāga was typologyzing different types of perception. A less biased reading of Dignāga does not seem to warrant such a presupposition. Reading the text independently of Dharmakīrti, one should probably maintain that for Dignāga there is only one type of perception, that is, a cognition which is free from conceptual construction. Or better still, that Dignāga was not at all concerned with types of perception. (Franco 1993, 298)

This is also seen as an invaluable insight by other contributors to the debate. In fact, Dunne (2006, 505) follows Franco, and Kellner's specific contribution to the debate on three- *versus* four-fold perception also originates from Franco's aforementioned analysis:

> The PSV distinguishes mental perception and self-awareness in terms of its object – the former applies to external objects, the latter to mental associates – but also points out a common feature: their independence of the external sense-faculties. Agreeing with Franco's suggestion that Dignāga did not intend to provide a typology of perception in the first place, I am referring to self-awareness simply as *a form of* (mental) perception. (Kellner 2010, 207 fn. 11; italics added)

Later, Franco (2005, 632) once again draws attention to the topic by claiming that Dignāga refers to *svasaṃvedana* as a "subspecies" of *mānasa-pratyakṣa*. This time he targets the interpretation Yao gives of *PSV* 1.6ab. The conundrum is about the word *mānasaṃ*, which is found in a Sanskrit fragment but not in the Tibetan versions, and occurs towards the end of the passage: "*rāgādiṣu ca svasaṃvedanam indriyānapekṣatvān mānasaṃ pratyakṣam*" – "The self-awareness (*sva-saṃvedana*) of desire, anger, ignorance, pleasure, pain, etc., is [also recognized as] mental perception because it is not dependent on any sense-organ" (Hattori 1968, 27; square brackets in the original; see § 3.1.1 fn. 2). Franco (2005, 632-3) employs a philological argument to support his position against Yao, emphasizing that the Sanskrit

fragment, being the *lectio difficilior*, should be given preference over the Tibetan translations, as previously noted by Hattori.[7]

Yao's (2004; 2005) contribution will now be examined as he represents another prominent voice in the debate on Dignāga's *pratyakṣa*, particularly among those who argue that Dignāga regards *svasaṃvedana* as a distinct type of perception. Yao's first criticism (2005, 132), aimed at Franco and like-minded scholars, is methodological in nature. He argues that these scholars have overlooked Dignāga's early works preserved in Chinese, as well as the commentarial tradition of Dharmapāla and his Chinese counterparts. This is a sizeable omission because, in the *Nyāyamukha* (a work, as mentioned before, extant only in Chinese), Dignāga clearly describes not a three- but a four-fold type of perception – listing self-awareness separately. Thus, Yao claims:

> most contemporary scholars follow [Prajñākaragupta] to interpret Dignāga's position on the typology of perception. This reflects a general tendency among scholars of Indian Buddhism, who give Sanskrit texts a higher preference despite the fact that the Sanskrit manuscripts we have today are usually dated quite late. In the case of Dignāga, […] many scholars still prefer to study him on the basis of the Sanskrit fragments found in later commentatorial works. When these Sanskrit fragments do not agree with the Tibetan or Chinese translations, they would disregard or emend the translations accordingly without hesitation. (Yao 2004, 75)

Against Franco and other scholars, Yao (2005, 141) claims that mental perception can only be directed towards external objects and that self-awareness is the internal awareness of mental consciousness only – by itself, sense perception would not be self-cognizant. This last claim, in particular, is openly rejected by Franco (2005, 632). Moreover, Yao (2005, 141) raises another criticism against those who deny *svasaṃvedana* as a distinct type of perception, arguing that such a position stems from a misunderstanding of the relationship between self-awareness and the two-fold 'appearance' (SKT *ābhāsa*) of cognition, that is, the self-appearance of cognition and the appearance of the object. With respect to that, the first scholar Yao criticizes is Nagatomi, for, as seen previously, he understands self-awareness as the subjective aspect or the self-appearance of

7 "I do not believe that the disappearance of *mānasaṃ* was accidental. I think rather that the word *mānasaṃ* was intentionally deleted, or not copied (or not translated), from a manuscript. Some attentive copyist or reader (or translator) must have realized that Dignāga's text, as it stood, was incompatible with Dharmakīrti's statements on perception, and that by simply eliminating the word *mānasaṃ* a harmony between the two could be achieved" (Franco 2005, 633).

cognition. Yao's point in distinguishing them and clarifying their relationship is straightforward; he maintains that self-awareness inherently 'possesses' the two-fold appearance:

> The way Dignāga understands the relationship between self-cognition and the self-appearance of cognition, in brief, is that the former possesses the latter but not *vice versa*. (Yao 2005, 142)

In this section it has been shown that while some scholars argue *PS* 1.6ab presents self-awareness as a distinct type of perception, others maintain that Dignāga incorporates it within mental perception. In the end, as Franco suggests, all of this may ultimately be symptomatic of the extent to which the discussion is already shaped by Dharmakīrti's later interpretation of the topic. Although no agreement has been reached by modern scholars, this is an important point to contemplate in order to attempt locating this cognition within Dignāga's system of perception as a whole.

3.1.2 Self-awareness, Memory, and the Two Forms (or Aspects) of Cognition

To fully explore the implications of the various standpoints in this ongoing debate, it is necessary to take an additional step by outlining the distinctive features of Dignāga's understanding of *svasaṃvedana*. The main scholars discussed so far have raised the aforementioned issues for different reasons. Some focus on strictly philological aspects, concerning the availability and selection of textual materials as well as the interpretation of specific expressions. Others emphasize the role of the commentarial tradition, highlighting how later scholars and commentators may have influenced the reading of Dignāga's original account. Finally, some engage with the internal structure of *svasaṃvedana*, examining its objects, functions, components and its two aspects. In order to unpack all these controversies, it is worthwhile investigating some crucial philological points, using the most recent textual sources available, to see whether and how they could shed light on the core issue itself, that is, the nature and structure of self-awareness. Thus, the philological dimension will be involved here through the analysis of relevant passages with the aim of deepening the philosophical inquiry into the meaning of this notion, its internal constituents, and defining characteristics. This survey on Dignāga's position will begin with his demonstration of the necessity of self-awareness and its two-fold nature, and examine one of the most dense and intricate aspects of his account: the implications of his conception of *svasaṃvedana* as the result of a

pramāṇa (SKT *pramāṇaphala*) – that is, a resulting cognition, a notion that will be addressed in the course of the discussion (see § 3.1.3).

According to Dignāga, self-awareness is a *pramāṇa* free from conceptual construction, a perception that is independent of the senses and that holds mental factors as well as conceptual cognitions as its objects. The nature of the relationship between self-awareness and the subjective and objective aspects of cognition is set forth in *PS* (and *PSV*) 1.11, where Dignāga explains that cognition has two forms or appearances, namely, the object-appearance and its own appearance. Three arguments are jointly used as proof.[8] The first argument is that in order for the object-cognition and the cognition of the object-cognition to be different from each other, as they actually are, two aspects are needed. Concerning the second argument, the object-appearance has to be posited in order for memory to connect with objects that have been cognized previously. Finally, the third argument for positing the two forms of cognition is that memory retains both the object and the act of cognition itself. These three lines of reasoning do not suggest any specific stance on the ontological status of the objects of cognition, whether they exist externally or not. However, they effectively account for the internal structure of cognition, regardless of its object,[9] and provide a framework for understanding both its intentionality, which refers to its capacity to be about, represent, or cognize something, and its phenomenal feature, which pertains to how something appears subjectively.

Concerning the third line of reasoning just mentioned, in the final part of *PSV* 1.11c, Dignāga asserts that, since there is memory of both the object and its cognition, cognition has two forms or aspects, and it is also brought to awareness by itself (*svasaṃvedyatā*) (Kellner 2010, 210).[10] Hattori's translation of the same passage (*PSV* 1.11c) in its entirety reads as follows:

8 Dignāga, *PS* (and *PSV*) 1.11a-c: "*atha dvirūpaṃ* jñānam iti kathaṃ pratipādyam. (1.11ab) ***viṣayajñānatajjñānaviśeṣāt tu dvirūpatā*** | ***viṣaye*** *hi rūpādau yaj* ***jñāna****ṃ tad arthasvābhāsam eva. viṣayajñāne tu yaj* ***jñāna****ṃ tad viṣayānurūpajñānābhāsaṃ svābhāsaṃ ca. anyathā yadi viṣayānurūpam eva viṣayajñānaṃ syāt svarūpaṃ vā, jñānajñānam api viṣayajñānenāviśiṣṭaṃ syāt. na cottarottarāṇi jñānāni pūrvaviprakṛṣṭaviṣayābhāsāni syuḥ, tasyāviṣayatvāt. ataś ca siddhaṃ dvairūpyaṃ jñānasya.* (1.11c) ***smṛter uttarakālaṃ ca*** dvairūpyam iti sambandhaḥ. *yasmāc* ***cā****nubhav****ottarakālaṃ*** *viṣaya iva jñāne 'pi smṛtir utpadyate, tasmād asti dvirūpatā jñānasya svasaṃvedyatā ca*". For a translation of these passages, see Hattori 1968, 29-30.

9 Kellner (2010, 216) makes this point explicit and clear: "If Dignāga borrowed the idea of two appearances from Yogācāra, then his argumentation does not indicate this was because of its relationship to the non-existence (or non-cognizability) of external objects".

10 As Kellner (2010, 210 fn. 20) notes, here, in Sanskrit, the term *svasaṃvedya* is a passive form that seems to suggest that cognition is brought to awareness by itself – corresponding to the active *svasaṃvedana* (meaning that cognition cognizes itself).

> Some time after [we have perceived a certain object], there occurs [to our mind] the recollection of our cognition as well as the recollection of the object. So it stands that cognition is of two forms. Self-cognition is also [thus established]. (Hattori 1968, 30; square brackets in the original)[11]

Memory, therefore, not only proves the two forms of cognition but also self-awareness.

In order to see how this memory argument for *svasaṃvedana* further unfolds, the next step is to consider *PS* (and *PSV*) 1.11-12.[12] Memory requires a precondition, namely a prior experience of either objects or their cognition. An object or its cognition cannot be recalled unless it has been experienced previously. In other words, in *PS* (and *PSV*) 1.11d it is stated: "Whatsoever is recollected has been experienced before" (Hattori 1968, 110 fn. 75).[13] Therefore, either cognitions are experienced by other cognitions, just as objects are, or they are self-experiencing. If cognitions were known through other cognitions, this would lead to an infinite regress (*regressus ad infinitum*);[14] therefore, following Dignāga's reasoning, self-awareness is established.

However, despite Dignāga's whole argument, a crucial problem does not seem to have easy solutions. Kellner raises it as follows:

> If, however, Dignāga intends to establish self-awareness as an intrinsic feature of all mental states and thereby as a part of their nature, then his argument based on memory is problematic, for strictly speaking it proves self-awareness only of cognitions that are or can be remembered. But does Dignāga believe that all cognitions can be remembered? To my knowledge, it is by no means

11 As Kellner (2010, 213 fn. 31) observes, this is a delicate passage that has often led scholars (modern ones, such as Arnold, and Tibetan ones, mainly dGe-lugs) to the identification of self-awareness with the subject-aspect.

12 Dignāga, *PS* (and *PSV*) 1.11d-12b: "*kiṃ kāraṇam*. (1.11d) ***na hy asāv avibhāvite*** **||** *na hy ananubhūtārthavedanasmṛtī* rūpādi*smṛti*vat. *syād etat* – rūpādivaj jñānasyāpi *jñānāntareṇānubhavaḥ*. tad apy ayuktam, yasmāj (1.12ab1) ***jñānāntareṇānubhave 'niṣṭhā*** anavasthāiti tajjñāne ***jñānāntareṇa***+anubhūyamāne. kasmāt. (1.12b2) ***tatrāpi hi smṛtiḥ*** **|** *yena* ***hi*** *jñānena* taj *jñānam anubhūyate,* ***tatrāpy*** *uttarakālaṃ* ***smṛtir*** *dṛṣṭā. tatas tatrāpy anyena* jñānena-*ānubhave 'navasthā syāt*". For a translation of these passages, see Hattori 1968, 30.

13 Dignāga, *PSV* 1.11d: "*na hy ananubhūtārthavedanasmṛtī rūpādismṛtivat*".

14 For a thorough analysis of this argument in Dignāga (and in Dharmakīrti), see Kellner 2011b. Kellner is particularly incisive in her considerations on this point: "Self-awareness is neither *reflective* (a subsequent act of reflection directed at an earlier mental state) nor *introspective* (a look 'inside' at one's own mental realm), since both of these approaches would involve stipulating a separate higher-order mental state" (Kellner 2010, 215; italics in the original).

> clear from his works whether Dignāga would limit subsequent memory to certain classes of mental states, or to states that occur only under specific conditions, like wakefulness. (Kellner 2010, 215)

Moreover, Kellner continues, it is unlikely that Dignāga did not intend to present *svasaṃvedana* as an intrinsic feature of all mental states since *PS* and *PSV* 1.8cd-10 mention the self-awareness of a mental state regardless of whether it is later recalled. This potential weakness of Dignāga's argument will be brought up again later in the inquiry when Dharmakīrti's line of reasoning will be taken into account and a comparison between the two arguments will be made. For the moment, this is sufficient to reveal that, for Dignāga, the (supposed) possibility to establish self-awareness stems from epistemological issues concerning the mnemic cognitive function.

3.1.3 Self-Awareness as *Pramāṇaphala*

Another intricate and profound aspect of Dignāga's position on *svasaṃvedana* is its role as the 'result of a *pramāṇa*' (SKT *pramāṇaphala*). As mentioned above, Dignāga deals with this topic in *PS* and *PSV* 1.8cd-10.[15] In 1.8cd he says that cognition (of external objects) is a result of *pramāṇa*. In 1.9a, however, he claims that self-awareness is a result of *pramāṇa*. What is, then, the relationship between the two statements? The link between them is the particle *vā* 'or', which indicates an alternative. It introduces the second clause, namely "*svasaṃvittiḥ phalaṃ vātra*" (Dignāga, *PS* 1.9a) – "or self-awareness is the result here" (Kellner 2010, 220). To synthetize the overall point, the sentence could thus be rephrased as: the result of a *pramāṇa* is either the cognition of external objects or self-awareness. Moreover, in the *PSV* passage commenting

15 Dignāga, *PS* and *PSV* 1.8cd-10: "(1.8cd) ***savyāpārapratītatvāt pramāṇaṃ phalam eva sat*** || *na hy atra bāhyakānām iva pramāṇād arthāntaraṃ phalam. tasyaiva tu* phalabhūtasya *jñānasya viṣayākāra*tayā utpattyā ***savyāpārapratītiḥ***. *tām* upādāya *pramāṇatvam upacaryate nirvyāpāram api sat. tad yathā* phalaṃ hetvanurūpam utpadyamānaṃ heturūpaṃ gṛhṇātīty kathyate nirvyāpāram api, tadvad atrāpi. (1.9a) ***svasaṃvittiḥ phalaṃ vātra*** *dvyābhāsaṃ hi jñānam utpadyate svābhāsaṃ viṣayābhāsaṃ ca. tasyobhayābhāsasya yat svasaṃvedanaṃ tat* ***phalam***. *kiṃ kāraṇam.* (1.9b) ***tadrūpo hy arthaniścayaḥ*** | *yadā hi saviṣayaṃ jñānam arthaḥ, tadā svasaṃvedanānurūpam arthaṃ pratipadyata iṣṭam aniṣṭaṃ vā. yadā tu bāhya evārthaḥ prameyaḥ, tadā* (1.9c) ***viṣayābhāsataivāsya pramāṇaṃ*** *tadā hi jñānasvasaṃvedyam api svarūpam anapekṣyārth****ābhāsataivāsya pramāṇam****. yasmāt so 'rthaḥ* (1.9d) ***tena mīyate*** || *yathā yathā hy arthākāro jñāne pratibhāti śubhāśubhāditvena, tattadrūpaḥ sa viṣayaḥ pramīyate. evaṃ jñānasaṃvedanaṃ anekākāram upādāya tathā tathā pramāṇaprameyatvam upacaryate. nirvyāpārās tu sarvadharmāḥ. āha ca* (1.10ab) ***yadābhāsaṃ prameyaṃ tat pramāṇaphalate punaḥ*** | (1.10cd) ***grāhakākārasaṃvittyos trayaṃ nātaḥ pṛthak kṛtam*** ||".

on 1.9a, Dignāga specifies that the result is self-awareness 'with both appearances' (SKT *tasyobhayābhāsasya*), employing a *bahuvrīhi* compound that implies self-awareness possesses the very qualities it refers to: *svasaṃvedana* possesses both appearances. However, the implications of using this type of compound are not entirely unambiguous, as Kellner observes.

> [T]his could again be taken in a narrower and a wider sense: either cognition is just aware of itself as having both appearances – it is aware of itself as somehow encompassing both aspects – or it also has access to both these appearances. In other words, is the *existence of* both appearances within the scope of self-awareness, or are the appearances themselves within its scope? (Kellner 2010, 221; italics in the original)

In other terms, when getting at the internal articulation of *svasaṃvedana*, it seems that no clear indications can be found in Dignāga. Kellner appears to question the potential significance of this compound, 'having both appearances'. Is it merely a secondary, descriptive detail that adds little to our understanding of the kind of experience one can have of one's own cognition? Or is it a deliberate and meaningful specification, one that opens up a deeper interpretative layer – suggesting that self-awareness grants access to, and a deeper insight into, how each of the two appearances discloses? The relevance of that *bahuvrīhi* is still a mystery and the variety of the subsequent traditional developments of the notion of self-awareness is evidence of this ambiguity.

However, returning to the main line of reasoning in the text – which concerns the alternatives of *pramāṇaphala* being either the cognition of external objects or self-awareness – it is worth examining in detail the way this alternative is presented in *PSV* 1.9b, where the order is reversed in comparison to how it was originally stated in the *PS*. First of all, Dignāga describes the latter of the two alternatives, which will be presented here both in translation and in the original Sanskrit: "when *a cognition possessing [the form of] an object* (*saviṣayaṃ jñānam*) is itself the object to be cognized" (Hattori 1968, 29; italics added; square brackets in the original) – "*yadā hi saviṣayaṃ jñānam arthaḥ*" (Dignāga, *PSV* 1.9b) – "then one cognises the object as desirable or undesirable in conformity with self-awareness" (Kellner 2010, 222) – "*tadā svasaṃvedanānurūpam arthaṃ pratipadyata iṣṭam aniṣṭaṃ vā*" (Dignāga, *PSV* 1.9b). The text then introduces the former alternative, in which *pramāṇaphala* is the cognition of external objects, and this case is described as "when, on the other hand, *only an external thing* is [considered to be] the object" (Hattori 1968, 29; italics added; square brackets in the original) – "*yadā tu bāhyaevārthaḥ prameyaḥ*" (Dignāga, *PSV* 1.9b).

Thus, for Hattori and most scholars, the Sanskrit particle *vā* introduces two alternatives of *pramāṇaphala*: the result is the cognition of an object when the object is an 'external object'; alternatively, the result is self-awareness when the object is 'a cognition possessing the form of an object'. In modern scholarship there has been notable research around these alternatives and it has been unanimously assumed that Dignāga's reason for specifying them rests in the difference between objects of cognition, whether 'external' (in the first case) or 'internal' to the mind (in the second case). As one might expect, these two instances have been interpreted as a convenient means for Dignāga to put and keep together the Sautrāntika and Yogācāra views – the former positing external objects from which cognition arises and the latter advocating an epistemic idealism – bridging their two tenets with the notion of self-awareness.[16]

Interestingly, it is once again Kellner (2010, 225) who offers a thoroughly innovative interpretation of this particular passage. She provides a couple of philological remarks and, on the basis of those, a revealing interpretation of the particle *hi* 'because' and the compound *saviṣayam* 'with an object' in the phrase "*yadā hi saviṣayaṃ jñānam arthaḥ*" (Dignāga, *PSV* 1.9b), which opens the description of the alternative indicating self-awareness as the result. This passage could be roughly translated as: "[Self-awareness is the result] because (*hi*), when a cognition with [the form of] an object (*saviṣayaṃ*) is the object to be cognized".

Regarding this passage, the first consideration Kellner makes is that the *PSV* introduces its commentary on 1.9b with the explanatory particle *hi*, meaning 'because'. This particle introduces the reason why self-awareness is described as the result and it is present in Steinkellner's reconstruction of the Sanskrit version but was overlooked by Hattori and other scholars who have primarily relied on the Tibetan translations, where the connection is not explicitly indicated. This is a primary element to consider in the broader interpretation of this passage within scholarly discourse. Thus, according to Kellner, *PSV* 1.9b would introduce the account of self-awareness as the result using 'because' (SKT *hi*), a particle that, as noted, is absent in Hattori's and other scholars' translations. The relevance of this term will be clarified in the following pages.

This said, Kellner explains, according to Jinendrabuddhi, that another major point of that passage is the compound *saviṣayam* 'with an object', which follows immediately after the particle *hi*. Following Kellner, *saviṣayam* – 'with an object' – should not be understood as

16 Hattori (1968, 102 fn. 61), for instance, is aligned with this interpretation; for some considerations referring to the Sautrāntika and Yogācāra views, also see Yao 2005, 143; Kataoka 2016. For a radically different account, cf. Kellner 2011a.

indicating that cognition is 'endowed with' (the form of) an object, as most other interpreters suggest. Rather, it implies that what is at stake is (everything from) cognition 'extending up to and including' the object. Actually, Kellner points out that *saviṣayam* is "an indeclinable compound signifying completeness", that is, indicating 'everything up to' (see Kellner 2010, 222 fn. 58). Thus, the phrase should not be understood here as saying "when a cognition *possessing* [the form of] an object is itself the object", as Hattori (1968, 29; square bracktes in the original; italics added) translates, based on the Tibetan rendering of the compound as "*yul dang bcas pa*" 'possessing an object' (183). The entire clause should instead be translated as "when the object is [everything,] cognition *right down to* the object" (Kellner 2010, 222; italics added; square brackets in the original).

Now, bringing together both of Kellner's considerations on the aforementioned passage *PSV* 1.9b – "*yadā hi saviṣayaṃ jñānam arthaḥ, tadā svasaṃvedanānurūpam arthaṃ pratipadyata iṣṭam aniṣṭaṃ vā*" –, particularly regarding the particle *hi* and the compound *saviṣayam*, the full sentence would be translated as follows:

> [Self-awareness is the result] *because* when [everything,] cognition *right down to* the object, is the object, then one cognizes the object as desirable or undesirable in conformity with self-awareness. (Kellner 2010, 222; italics added; square brackets in the original)

As will become clear in the following discussion, this rendering – far from being a merely philological matter – significantly reshapes the content of Dignāga's account of self-awareness and its philosophical implications.[17]

The following chart displays the two exemplary scenarios that most scholars – Kellner being the notable exception – have inferred from their reading of Dignāga's *PS* and *PSV* 1.8cd-10.

When the object to be validly cognized is **just an internal object**, namely, the object-appearance		
object:	*means:*	*result:*
object-appearance	apprehending aspect	self-awareness

When the object to be validly cognized is **just an external object**		
object:	*means:*	*result:*
external object	object-appearance	cognition of the external object

17 For a useful summary on the issue, also see Arnold 2010, 348-50.

In other words, for most scholars the two paradigms of 'internalism' and 'externalism'[18] are all that is at stake here. In Kellner's words:

> Hattori and Iwata, among others, have interpreted PS and PSV 1.8cd-10 on the assumption that the status of intentional objects is the main underlying issue – is the object of valid cognition something external, or internal to the mind? (Kellner 2010, 226)

Therefore, those scholars read the particle *vā* in 1.9a – where Dignāga says that the result of a *pramāṇa* is the cognition of external objects 'or' the result is self-awareness – as presenting two alternatives: externalism in the first case, where the object is something external to the mind, and internalism in the second, where the object is the object-appearance and therefore is internal to the mind.

They then read the following passage, in 1.9c, as a shift back to the externalist perspective. From their point of view, self-awareness would bridge externalism and internalism, that is, the Sautrāntika and Yogācāra doctrines, and it would perform this function as an "intentional self-awareness", that is, "the awareness that the subjective aspect of a mental state has of its objective counterpart" (Kellner 2010, 218). Following Kellner, however, a couple of important considerations must be made. First, Dignāga offers no explicit or unequivocal support for such an understanding of self-awareness. Second, he chooses to establish *svasaṃvedana* through the memory argument, where consciousness cognizes its own 'apprehending' aspect. This implies that he assigns greater relevance to a facet of consciousness's reflexivity that is distinct from the awareness of the apprehended aspect by the apprehending one. In light of this, it seems unlikely that Dignāga conceives of *svasaṃvedana* as the awareness that the subjective aspect of a mental framework has of its objective aspect.

Kellner's innovative reading of Steinkellner's impressive reconstruction (based on Jinendrabuddhi)[19] shows a different possibility: instead of externalism and internalism, that *vā* particle marks a shift from the 'particular' case of having just external objects of cognition to the 'general' case of having all kind of objects, internal (object-appearances) and external to the mind. Then, it follows that in 1.9c, the text returns to the 'particular' case – namely, the former one. For Kellner, what is at stake in 1.8cd-10 is not the type of intentional object the mental event engages with (as shown in the

18 "The externalist theory assumes that some extra-mental, material object causes a perception that has its form. By contrast, the internalist theory assumes that perception, as well as all other cognitive activity, takes place solely within the mind, and that nothing else is to be experienced by cognition" (Kellner 2011a, 294).

19 For Jinendrabuddhi's commentary on 1.8cd-10 of the *PS* and *PSV* cf., in particular, Steinkellner, Krasser, Lasic 2005, 58-77.

chart above), nor is this what the passage aims to clarify. Rather, the actual underlying problem is the following:

> Self-awareness is the result because, owing to its providing access to how objects of valid cognition appear subjectively, it allows for a comprehensive conception of the result, applicable to intentional objects as well as mental associates, and also applicable regardless of whether externalism or internalism are advocated. (Kellner 2010, 226)[20]

Thus, the main point that Dignāga seems to be making in this passage is crucial, in that it might actually suggest an account of self-awareness as non-conceptual access to 'how' all kind of things appear to the mind, relegating the issue of internalism or externalism to secondary importance. In addition to the possible differences in the internal structure of self-awareness, namely its various internal articulations concerning cognitive means and objects, what emerges is an account of *svasaṃvedana* that is universally applicable to any form of cognition. The key issue then is the development of a model of self-awareness that can accommodate all types of experiences. In fact, recalling *PSV* 1.9b on the basis of Kellner's work, with the compound *saviṣayam* (suggesting completeness), Dignāga is understood to be addressing the totality of objects: self-awareness is the result with respect to every kind of object, both external and internal to the mind. He would then explain that this is the case 'because' – hence the relevance of the Sanskrit particle *hi*, highlighted by Kellner and previously discussed – the determination of objects as desirable or unattractive conforms to the self-awareness that pertains to them. In other words, Kellner states:

> Dignāga's argument can be explicated as claiming that intentional objects (as well as mental associates) are determined as desirable or undesirable depending on how they appear in the mind, and it is this *how* of appearing that is accessed through self-awareness. From this perspective, the feature of self-awareness that grants it its status as the result is its access to the way things subjectively present themselves in the mind. (Kellner 2010, 222-3; italics in the original)

20 As for Dignāga's commitment to either internalism or externalism, Kellner (2016, 226) points out that in the *PS* "the existence of two alternative accounts, externalism and internalism, is presupposed without any further elaboration; no commitment is made to one or the other as superior. This is, as is well known, different in the *Ālambanaparīkṣā*, where Dignāga points out fundamental problems of externalism and presents internalism as the superior account of what counts as the 'object support' (*ālambana*) of a mental state".

In its universal applicability, then, self-awareness boils down to the immediate access to how things subjectively and experientially manifest, being either external or mental. If Dignāga borrowed the notion of the two appearances from the Yogācāra discourse on cognition, that might have been in order to introduce an internal and phenomenal structure that could fit 'any' mental state (Kellner 2010, 216). Moreover, coming back to the 'mysterious' *bahuvrīhi* compound in *PSV* 1.9a, *tasyobhayābhāsasya*, stating that self-awareness is the awareness a mental state has of both its appearances, it is important to bear in mind that this model of self-awareness is a form of access to both the apprehending aspect and the object-appearance of cognition, even though the nature of this access is not further explained in Dignāga's work. Therefore, based on this, one could argue for certain variations in the internal articulation of self-awareness – which, in any case, as *pramāṇaphala*, remains the access one always has to experience – including the possibility for a cognition to access the subjective aspect (for instance, in the memory function) and the objective aspect.

It is therefore time to consider another remarkable aspect of the internal articulation of *svasaṃvedana*, in light of the distinctive features of Dignāga's thought concerning *prameya*, *pramāṇa*, and *phala*. These three aspects of a mental event – the object of epistemically valid cognition, the epistemically valid cognition, and the correspondent cognitive result – are not separate from one another, despite the tendency to distinguish them metaphorically (*PS* and *PSV* 1.10). On the basis of this statement, Dunne stresses the absence of a causal structure in Dignāga's account of self-awareness:

> That reflexive awareness is noncausal follows from its simultaneity with its object, namely, the awareness that is reflexively perceived itself. Indeed, what can be most confusing about reflexive awareness is the notion that it is a cognition distinct from its object. This distinction is clearly the case for all forms of perception, including mental perception (*mānasa-pratyakṣa*), for in all cases the object (*grāhya*) of perception is its cause […]. In contrast, what Dignāga first identifies as the three aspects of an awareness – namely, reflexive awareness, the objective aspect (*grāhyākāra*), and the subjective aspect (*grāhakākāra*) – are all ultimately identical and hence simultaneous. The notion that reflexive awareness is cognizing the subjective- and objective-aspects is merely a way of conceptualizing the process of knowing (see the *locus classicus* in *PS* 1.10 […]). (Dunne 2004, 276-7 fn. 93)

One more piece needs to be added to the overall picture of what Dignāga's account of *svasaṃvedana* might be. Despite the common belief that a cognition performs the activity of apprehending its object,

for Dignāga – following a Sautrāntika position – it does not perform 'any' action at all (*PS* and *PSV* 1.8cd). Rather, it is a 'result', an effect that is similar to its cause, resembling the apprehended object by bearing its aspect or image. Along this line, it is interesting to quote the important paradigm shift Matilal (1986a, 112-13) suggests, from mental 'acts' to mental 'episodes': whatever happens in one's mind is something that takes place within oneself and is not something that one enacts. This would actually shape a model whose "demands are minimal" (Matilal 1986a, 113): there is no need for any actors or agents, and mental episodes may be treated as effects brought about by certain conditions.

Thus, as has been shown, in *PS* and *PSV* 1.8cd-10 Dignāga presents *svasaṃvedana* as *pramāṇaphala*. For decades, it has been thought that its primary role, as conveyed by this passage, was to bridge the gap between externalism and internalism, and that self-awareness's nature consisted in the intentional cognizance of the objective aspect by the subjective one. However, more recently, mostly thanks to Kellner's contribution, it has been argued that its role as *pramāṇaphala* is rather that of providing access to 'any' experience whatsoever, to the 'how' of appearing of any possible object. Thus, while it may indeed serve to bridge the gap between externalism and internalism, it does so in a fundamentally different way than previously suggested – namely, by indicating self-awareness as the very means of accessing the subjective cognition of any object, regardless of whether objects are conceived as internal or external to cognition, and encompassing mental factors.

3.1.4 Some Readings of Dignāga's *Svasaṃvedana*

The complexity of the textual materials and their exegesis has contributed to the diversity of interpretations proposed by modern scholars. Now that the main elements have been critically examined, it is possible to explore some of the most significant contemporary interpretations of Dignāga's standpoint, particularly in relation to current debates in the philosophy of mind. This, in turn, allows for a deeper appreciation of how scholars of religious studies and the philosophy of mind attempt to decipher and integrate the contributions of the Indian Buddhist tradition into a dialogue with modern and contemporary thought and concepts.

Given the extensive discussion of Kellner's contribution in the preceding section, it seems fitting to return briefly to her position. In her commentary on *PS* and *PSV* 1.12, she draws on the infinite regress argument to rule out any interpretation of Dignāga's *svasaṃvedana* as either "*reflective* (a subsequent act of reflection directed at an earlier mental state)" or "*introspective* (a look 'inside' at one's own

mental realm)" (Kellner 2010, 215; italics in the original). Arguing that self-awareness is made possible by any separate higher-order mental state would simply go against the above-mentioned argument in Dignāga's text.

Nevertheless, there are a few interpreters who suggest a reading of Dignāga's account of *svasaṃvedana* that implies a second-order mental event. For instance, Matilal (1986a, 152) seems to read self-awareness along these lines. Another example is Yao (2005, 159), who assumes that Dignāga would support the higher-order theory (HOT) perspective, which argues that a mental state becomes conscious only when it is the object of a higher-order state. Yao acknowledges that Dignāga gets close to the position held by the phenomenological tradition, in that he appears to regard self-awareness primarily as immediate access to one's own mental state as they present themeselves, rather than as a retrospective glance at it occurring at a later moment. However, Yao ultimately concludes that it is possible to "*safely* assume that Dignāga would support the HOT theory" (Yao 2005, 159; italics added), that is, a higher-order thought theory, according to which a mental state becomes conscious when it is taken as the object of a higher-order thought. According to Yao, since Dignāga posits a three-fold cognition – comprising self-awareness itself, self-appearance as subject, and object-appearance as object – it follows that he adopts a HOT perspective. Within that framework, self-awareness, as a form of 'perception', is a non-conceptual type of cognition; yet, being a 'mental' perception, it is "of the nature of mental" (Yao 2005, 159) and does not rely on sensory organs. It is precisely for this reason – insofar as the kind of perception at stake is distinct from that which is tied to the senses – that Yao does not associate Dignāga's account with the strand of philosophy of mind known as higher-order perception (HOP) theory.

Some scholars label Dignāga's self-awareness as 'introspection' (Stcherbatsky [1932] 1962, 12; Hattori 1968, 95 fn. 50). Notably, Wayman applies the same label to *svasaṃvedana* as it is presented in the works of both Dignāga and Dharmakīrti, portraying it as a kind of mirror, "the only faculty of the system that can be credited with the role of figuring out the system as given" (Wayman 1991, 429). One more instance of a recent reading that frames Dignāga's self-awareness in terms of introspection is Ganeri's:

> This is perhaps the rationale behind Dignāga's claim that, when one is introspectively thinking about one of one's own thoughts, the subjective aspect of the latter becomes the objectual aspect of the former. That is, the objectual aspect of a second order thought = the subjective aspect of its first order thought. (Ganeri 1999, 470-1)

Kellner's precise understanding of Dignāga's self-awareness as a "*mode* of awareness" (Kellner 2010, 227; italics in the original) that provides access to how things present themselves subjectively is quite an exception in the landscape of modern scholarly exegesis and is a thoughtful insight on the issue. It seems that Kellner aims to highlight the structure or function of self-awareness in opening the phenomenality of experience to one's own immediate knowledge of it, being a constitutive disclosing part of one's cognitions but in such a way that thinking of the quality of subjective experience as separate from one's own access to it would be deeply misleading. Could self-awareness be thought of in terms of the immediate and lively 'transparency' of the quality of experience, whatever this might be? In light of the same textual and philological work shared with Kellner, Arnold (2010, 349) ends up interpreting self-awareness not as a "mode" but rather as a "quality" of the mental, a feature or nuance that is there regardless of what the cognition's content might be. Compared to Kellner's stance, despite an overall proximity in their readings of Dignāga's work, Arnold's approach arguably situates Dignāga's standpoint on a slightly different level of philosophical discourse. Rather than focusing on 'access to' experience, as Kellner does, Arnold emphasizes the 'features of' experience itself. Kellner's insight on Dignāga's contribution appears as a way to conceive self-awareness as a 'modality' that is constitutive of experience and that is there as an experience-disclosing factor, an access to it, no matter what the actual qualities of the accessed experience turn out to be. On the other hand, in Arnold's understanding, Dignāga's contribution can be seen as addressing the status of awareness or cognition in its qualitative dimension. His focus is on the fact that there is always a way in which experience unfolds for someone, rather than on the very access one has to it. It is interesting to examine the implications of claiming that *svasaṃvedana* is the result from such an angle:

> [It] amounts, on my reading, to the point that it is only as first personally cognized that anything is epistemically accessible to us at all. I take it that Dignāga advances, in this way, a case for thinking that *epistemic* idealism (the view that we are only immediately acquainted with mental items) represents the only reasonable epistemological position, even for those reluctant to

> give up reference to external objects. (Arnold 2010, 349; italics in the original)[21]

Arnold (2005a, 35) identifies the role of *svasaṃvedana* as being the epistemological foundation or the basic model of any possible cognition, the only certainly immediate cognition, and in focusing on its immediacy he is very close to the position held by the phenomenological tradition (see Part Three). He argues that cognition is characteristically contentful, and for Dignāga, as interpreted by Arnold, this is "just to make a phenomenological point (not an ontological or metaphysical one): that cognition seems to be *of* things" (Arnold 2010, 353; italics in the original).

If cognition's being contentful is known by the subject, this can function as a common ground for both Sautrāntika and Yogācāra tenets: any cognition is "only as first-personally known" (Arnold 2010, 353). On his reading of Dignāga, Arnold further insists on this unique feature of self-awareness, noting that the object of a cognition "is how the cognition seems to us" (Arnold 2009, 141), that is, the way in which experience is given and presented to the subject, from a first-person perspective. In contemporary terms, the identity between the 'intentional' content – what a cognition is about – and the 'phenomenal' quality – how it is presented to the experiencer – constitutes the most salient feature of self-awareness, to the extent that this minimal cognitive disclosure has often been chosen as what guarantees and establishes knowledge itself.[22]

The possibility of interpreting self-awareness in line with the classical phenomenological view is also explored in Coseru's

21 The term "Epistemic Idealism" is used by Arnold to refer to Dunne's understanding, that is, to the view that "All Entities are Mental" (Dunne 2004, 59), meaning "idealism […] only with regard to what we know" (Arnold 2008, 15). This is also compatible with an ontological or metaphysical commitment to existent external objects. Thus, the difference between Sautrāntika and Yogācāra would lie only in their metaphysical stances and not in their epistemological perspective, since both hold that one is immediately acquainted only with mental things, with the contents of one's own awareness.

22 Also see Arnold 2009, 141; for the equivalence of intentionality and phenomenality in the Yogācāra view, see Ram-Prasad 2007, 70. With regard to phenomenality and intentionality, Ganeri (1999, 471-2) offers a different perspective by arguing against the possible claim that Dignāga's notion of the subjective aspect refers to the phenomenological quality of experience, to 'how it feels' to the experiencer. For Ganeri, while the phenomenological quality of an experience does not give any information as to what its intentional content is, Dignāga's subjective aspect involves also the specific intentional content of experience. Thus, for him (Ganeri 2012, 171), the apprehending aspect "is better described with the help of the idea of a mode of presentation as a constituent of intentional content: the subject-aspect is an intentional mental state's mode of presentation of its own object-directed intentionality".

remarkable contribution.[23] In his view, Dignāga's explanation closely aligns with Husserl's concept of 'noematic content', referring to the 'perceived as such', which phenomenological reduction aims to reveal (Coseru 2012, 237). Nevertheless, whereas Coseru emphasizes the intentionality dimension of Dignāga's *svasaṃvedana*, MacKenzie, in dialogue with Coseru and inspired by readings of Dharmakīrti and his commentator Śākyabuddhi, chooses to set intentionality aside, or more precisely, to subsume it under the experience's phenomenal availability, in terms of the various subjective and objective contents. Thus, he offers a perspective which considers self-awareness as a form of "unmediated acquaintance of consciousness with its own subjective and objective contents" (MacKenzie 2015b, 42). For MacKenzie (2017), in line with the aforementioned accounts, Dignāga's self-awareness is not a distinct higher-order cognition but rather an intrinsic feature of first-order consciousness itself: a "primitive" and "direct acquaintance" one has with one's own experiences. This means that, according to MacKenzie, this leads to a same-order model of self-awareness,[24] without requiring any higher-order structure. In recent years, Coseru (2020) has revisited Dignāga's contribution, stating:

> Dignāga's dual-aspect theory of mind can in large measure be understoodas an attempt to bring the debate about the primacy of either "intentionalityover phenomenality" or of "phenomenality over intentionality" to a resolution. (128)

In Coseru's interpretation, Dignāga's perspective holds that conscious experience entails a subjective aspect, marked by an inherent openness to what is given, and an objective aspect, which determines the specific content toward which the mental state is directed. In addition to these two aspects, every conscious cognitive episode is characterized by a reflexive awareness of itself (*svasaṃvedana*), which constitutes its distinctive mode of presentation. Through Coseru's lens, Dignāga's notion of self-awareness encompasses several key dimensions: the "self-intimating aspect" of cognition, or its for-me-ness; the "modality-specific nature of conscious apprehension", according to which remembering is not the same as perceiving, and

23 For a deeper exploration of the phenomenological dimension of self-awareness, based on experiential reports, see Metzinger 2020; 2024. See also Coseru 2020 for an innovative interpretation of Dharmakīrti's account and the classical Indian Buddhist debate on self-awareness, advocating for the efficacy of phenomenal experience as a vehicle for self-knowledge.

24 Cf. the first-order representational theories (Dretske 1995; Tye 1995), which argue for a type of reflexiveness that is implicit in the first-order consciousness itself.

both differ from conceiving; and, finally, the distinct "phenomenal character" that marks each mental state (Coseru 2020, 133).

This brief overview of the main interpretations of Dignāga's account of *svasaṃvedana*, while reaching no determinate conclusions, ends with a further question, adding one more piece to the puzzle or, perhaps, providing a way to reformulate some of the main issues just mentioned. In several of the key contributions examined thus far, *svasaṃvedana* appears to have emerged as a foundational concept – one that accounts for the very possibility of experience itself.[25] However, in attributing such a central role to it, one may wonder whether this involves a metaphysically overreaching claim, potentially at odds with the core spirit of Buddhist philosophy and some of its fundamental tenets, such as no-self (SKT *anātman*) and dependent origination (SKT *pratītyasamutpāda*). Should *svasaṃvedana* be understood not as a metaphysical principle, but as a doctrinal and soteriological concept that aims to articulate and clarify specific features of experience within the broader context of Buddhist thought?

By addressing this question, one targets the ontological status of self-awareness and the role it plays in a wider philosophical tenet or view. Actually, this is a question that runs through all the studies on the different traditional accounts of self-awareness. Williams (1983, 329) introduces it in terms of "the problem for the philosopher" and some years later Kellner (2010, 227) advances her thoughts on it referring to the specific case of Dignāga's view, where *svasaṃvedana* seems to be there to explain certain features of consciousness – a view shared by Cozort (1998, 157). The bottom line is the alleged experiential base of experience, meaning the threat of subjectivity, deeply linked to that mineness.[26] Commenting on Dignāga's account of self-awareness, Ganeri argues along this line of thought by recalling Nāgārjuna's view and saying:

> as soon as one allows the idea that experience has a base at all, the game is lost. That stronger claim leads immediately to the Mādhyamika View, that there is neither base nor place. (Ganeri 2012, 167)

25 Cf. Coseru 2020, 134 for a discussion on Dignāga's central concern, which is presented as "pragmatic rather than soteriological" and framed around the guiding question: "what relevant criteria must be in place for our mundane cognitions to succeed?". Self-awareness thus emerges as the necessary 'condition' for the Buddhist practitioner's attainment of self-knowledge.

26 For a lucid and brief description of the 'whole' of one's own experience and its difference from any other person's in relation to the Buddhist notion of self-awareness, see Mookerjee 1935, 328-30.

However, Dignāga's stance on the subjectivity of experience is, for Ganeri, deeply aware of this risk:

> Dignāga goes to the philosophical heart of the matter. As soon as one postulates a base for experience distinct from the experience itself, whatever it may be, nothing can block its subsequent identification with self, its identification, in other words, with the *place* of experience, a site of experiential ownership. The only way to defend a No Place View, therefore, is to base each item within the stream of experience *in itself*. That is the fundamental point of transition from an Abhidharma to a Yogācāra View. (Ganeri 2012, 167; italics in the original)

As such, Dignāga's brief remarks on *svasaṃvedana*, its first formulation within the Indian Buddhist epistemological context, give rise to a vast range of issues and interpretations that the Indo-Tibetan tradition continues to grapple with.

3.2 Later Developments in the Issue of *Svasaṃvedana*

The scope of this research is primarily focused on the Tibetan contribution to the global philosophical discourse on self-awareness. However, an overview of the main Indian Buddhist developments of this idea is inevitable in order to contextualize the debate that later unfolds in the Tibetan setting. The initial stages of this discourse in the pre-Dignāga period have been examined, followed by a close reading of Dignāga's challenging exposition. Here the core difficulties surrounding the subject came to the fore. The next step will be to investigate the transmission of this legacy to Tibet, along with its many complexities. This requires a prior outlining of the essential aspects of Indian Buddhist thought in the post-Dignāga phase as similar ideas will re-emerge in the Tibetan context and will also resonate with contemporary philosophical discussions on the topic.

3.2.1 Dharmakīrti and the Continuation of Dignāga's Epistemological Project

After Dignāga, Dharmakīrti (*Nyāyabindu* 1.10 and *Pramāṇavārttika* – henceforth *PV* – 3.485-503) continues the epistemological discourse on *svasaṃvedana* that had begun. His *PV* is ostensibly promoted as a commentary on Dignāga's *PS*, but it would be misleading to take these thinkers as exemplifying a unified body of thought. Here, the main philosophical aspects of Dharmakīrti's account of self-awareness will be presented in an attempt to grasp its unique

features while also linking them to some of their interpretations in modern scholarship.

Dignāga's only argument for self-awareness is that from memory, implying the infinite regress. In Dharmakīrti, on the other hand, the discourse on self-awareness in the *PV* is much wider in scope, but the sections in which he provides many of the important arguments have not yet been studied carefully. However, in order to appreciate the different philosophical implications of Dignāga's and Dharmakīrti's arguments for self-awareness from memory, it is necessary first to recall the elements of Dignāga's line of reasoning (in *PS* and *PSV* 1.11d-12). To begin with, memory requires past experience,[27] and one can remember either past objects or their cognitions.[28] Dignāga's argument, then, is that cognitions must also have been experienced prior to their subsequent recollection: either they have been experienced by a separate cognition or by the same cognition, and since the former would lead to an infinite regress, the latter is asserted and self-awareness is thus established.

Is the problem of infinite regress truly compelling in Dignāga's argument? Kellner argues (2011b, 417; 2010, 215) it is not: in order to be such, an infinite regress must entail that 'every' single cognition that cognizes a cognition is remembered. This would imply that "all cognitions are remembered – which, however, is wildly implausible" (417). Alternatively, one could take this as implying merely a 'possibility' of memory: every cognition might possibly be remembered. However, in doing this, one would be claiming arbitrarily, as a matter of principle, that some cognitions cannot be remembered – a position that would, in turn, raise further philosophical difficulties.

Now, turning to an analysis of Dharmakīrti's argument for self-awareness from memory, the *PV* contains only a brief reference to Dignāga's reasoning: "from memory, too, self-awareness is established" (Kellner 2011b, 419).[29] The closest argument to that of Dignāga's infinite regress is, instead, in *Pramāṇaviścaya* (henceforth *PVis*) 1 (40.11-41.13 *ad* 54cd). Unpacking *PVis* 1.54cd,[30] these are the main passages according to Kellner:

27 Dignāga does not explain why and how experience (of objects or object-cognitions) is a precondition for memory; on the other hand, Dharmakīrti (*PV* 3.179) for his part makes it clear: if there were memories of cognitions without any previous experience of them, it would follow that cognizing one's own past cognitions would be the same as cognizing others' cognitions.

28 Interestingly, Tsong-kha-pa Blo-bzang-grags-pa (1970, 26.1-8.9) presents the memory argument in a different way: subsequent memory establishes that the earlier cognition 'includes' self-awareness. On this point, cf. Kellner 2011b, 419.

29 Dharmakīrti, *PV* 3.485a-b1: "*smṛter apy ātmavit siddhā jñānasya*".

30 For the Sanskrit passage and the translation, see Kellner 2011b, 420.

> (1) we perceive objects not because they exist, but because their perception exists. In other words, the existence of an object's perception is a precondition for the object's being perceived. (2) If the existence of the perception of the object is unknown, it cannot serve as the basis for subsequent forms of behaviour that presupposes existence. (Kellner 2011b, 420-1)

This said, (1) and (2) are the premises from which Dharmakīrti concludes in (3) that if perception is unestablished, the perceived object is also unestablished.

The main problem or weakness of this argument, as Kellner (2011b, 421) notes, resides in a slight but crucial shift in the reasoning: from saying, with premise (1), that perception needs to 'exist' in order for its object to be cognized, to stating, in premise (2), that perception needs to be 'known' in order to exercise its function. Interestingly, for Dharmakīrti, the regress would turn out to be as follows: when the object-cognition is perceived by another cognition, that second-order perception is not established and therefore it needs another perception to follow and establish it. A person then needs to wait for the end of this chain of perceptions without perceiving anything because so long as one member of the chain is not established, none are established. Since there is no end to the sequence of perceptions, therefore, one would be unable to perceive any objects.

The similarities between Dignāga's and Dharmakīrti's arguments lie in the shared requirement that any cognition must itself be cognized – whether this occurs through a subsequent act of cognition or through self-awareness. Since the former leads to a vicious infinite regress, the latter is (allegedly) established. Nevertheless, to talk about only 'one general' regress argument within the Indian epistemological project would overlook the significant differences between the two arguments, particularly in their premises – premises whose weakness has just been pointed out. For Dignāga, the 'subsequent memory' of the cognition of an object is what is at stake, and the memory of all the higher-order cognitions leads to infinite regress; Dharmakīrti, on the other hand, focuses his argument on the 'cognition of an object', and, since a cognition has to be cognized, this is what actually leads to infinite regress.

Since, by contrast with Dignāga, Dharmakīrti offers additional arguments in support of self-awareness,[31] it is worth examining the so-called *sahopalambhaniyama* argument, widely regarded by subsequent philosophers as the definitive rationale for *svasaṃvedana* (Taber 2010). It consists of

> the argument that the most salient fact about objects of awareness is the constraint (*niyama*) that they be known only together with the apprehension thereof (*sahopalambha*). (Arnold 2010, 327)

In an effort to understand this line of reasoning in dialogue with the modern philosophical tradition, Taber (2010, 292-3) proposes reading this as a specific instance of the traditional principle of the 'identity of indiscernibles' according to which "two things having exactly the same properties are identical". In this context, Taber argues, Dharmakīrti maintains that identity – or more precisely, non-difference – is not established through the sharing of all properties but rather by the sharing of a single decisive one: the awareness of a cognition is not something over and above the cognition itself since there is no difference between the two in terms of either content or temporal structure.

Arnold, by contrast, proposes interpreting the issue through the lens of Sacks's (2005, 444) notion of 'situated thoughts'. For him, thoughts are essentially embedded in the context of an experiencing subject's perspective and situation.[32] In these terms, Dharmakīrti's point will turn out to be that anything one might say about the nature of an experience will inevitably involve some form of 'it-is-like' for its subject. One's thoughts about experience are always already phenomenologically embedded in the very thing one is trying to explain: awareness has an ineradicable first-person character. However, leaving aside this particularly modern take on the *sahopalambhaniyama* argument, it states that there can be no awareness of a cognition without awareness of an object-form: in

31 It is worth mentioning another argument Dharmakīrti (*PV* 3.448-59) offers, a reasoning that leads to the recognition that self-awareness is important for establishing subjectivity and privacy, and for avoiding another person's access (that is, a yogi's direct access) to one's own mental state. In Moriyama's (2010, 261) words: "when we compare it to Dignāga's argument, we notice an interesting change of focus from the temporal sequence of cognitions in a single mind-stream to the co-occurrence of cognitions in the same moment in multiple mindstreams – what Dharmakīrti tries to avoid with self-awareness is […] the absurdity that one person can experience another's mental states".

32 For an attempt to analyze the *sahopalambhaniyama* argument by focusing on its logical and philosophical features, assessing its strengths and weaknesses as an argument, its formal validity, and the defensibility of its premises, see Taber 2020.

apprehending itself, a cognition must apprehend itself as possessing a form (*ākāra*), that is, as having some content.[33] What is at stake is

> the two parts of the equivalence which is the *sahopalambhaniyama* of cognition and object – no perception of an object without perception of awareness and no perception of awareness without perception of an object-form. (Taber 2010, 293)

Delving deeper:

> Dharmakīrti presents the decisive point in *Pramāṇavārttika* [3.]335: Whenever an object is apprehended as having some form, awareness will be apprehended as well; but conversely, whenever awareness of a particular form is apprehended, an object having that form is apprehended. From this we can confidently conclude that the form of the object and awareness are actually in some sense the same. (292)[34]

In this regard, in *PVis* (1.54ab), Dharmakīrti makes the straightforward statement: "Blue and its cognition are not different because they are necessarily perceived together" (Kellner 2011b, 419).[35]

In examining the main philosophical implications of this argument, Arnold (2010, 357) finds it fruitful to engage with several contemporary views in order to establish a dialogue on the topic of self-awareness. In particular, drawing on insights from the Buddhist epistemological tradition, he seeks to challenge, expand, and reconsider certain modern philosophical perspectives. Through this approach, he further aims to clarify the distinctive contribution of the Buddhist tradition to ongoing philosophical debates. In doing so, he finds it convenient to recall the "governing disjunction" formulated by Bilgrami (2006, 28), an interpretive framework through which theories of self-awareness in the philosophy of mind are analyzed as falling into two distinct models, with the claim that only one of these can be correct: either the 'perceptual' model or the 'constitutive' one. According to the former approach, self-awareness of intentional states is characterized as direct or immediate, akin to how colors or sounds are perceived; such an understanding of *svasaṃvedana* would see it as a distinctive kind of ordinary perceptual awareness, only with a specific object or content, that is, a mental state. According to the latter model, self-knowledge would instead be constitutive of intentional states themselves, in such a way that they would depend

33 See also Matilal 1986b.

34 For other occurrences of the same argument in Dharmakīrti's works, cf. Taber 2010.

35 Dharmakīrti, *PVis* 1.54ab: "*sahopalambhaniyamād abhedo nīlataddhiyoḥ |*".

fundamentally on the self-awareness one holds of them. One's very concept of a mental state would then imply that mental states are inherently tied to one's capacity to be aware of them rather than existing independently of that awareness.

In other words, in the perceptual model, cognition is treated analogously to the perception of external objects. According to Bilgrami's perspective, these exist independently of whether they are perceived. A house's existing, for instance, is not constituted by someone's perception of it; similarly, mental states are taken to be independent of one's awareness of them. On this view, the awareness of a mental state is a separate act, added to the state itself. By contrast, the constitutive model maintains that a mental state is not truly a mental state unless it is accompanied by self-awareness. Here, awareness is not external to cognition but internal to its very structure: to have a mental state just is to be aware of it. In this sense, mental states lack the kind of independence from one's epistemic access that external objects possess. Under this model, therefore, cognition is not merely something observed; it is something that, in occurring, conceptually entails awareness of itself.

Bilgrami argues that these two views are mutually exclusive – either self-awareness is perceptual or it is constitutive. Both cannot be correct simultaneously. However, Arnold (2010, 357) suggests that Buddhist thinkers like Dignāga and Dharmakīrti significantly blur this disjunction. It is in this light that Arnold's decision to draw on Bilgrami's work becomes clear as a means of illuminating the distinctive contribution of Dignāga and, above all, Dharmakīrti. In fact, they treat self-awareness as perceptual (by calling it *pratyakṣa*), but their arguments – especially Dharmakīrti's *sahopalambhaniyama* – point to a constitutive understanding where cognition and awareness of cognition are inseparable. Drawing on Dharmakīrti's and Dignāga's accounts, Arnold argues for a constitutive model of self-awareness that is nevertheless perceptual in nature. This position is framed within Buddhist idealism, in which perceptual knowledge and self-awareness would not turn out to be different, and nothing is considered independent of the cognition through which it is apprehended.[36]

Moreover, a key philosophical implication of Dharmakīrti's account often emphasized by modern exegesis is the phenomenological dimension of self-awareness. This can be characterized as follows:

36 See Arnold 2010, 328. Thus, as Arnold (2010, 356) and Yiannopoulos (2012, 156-7) note, Bilgrami's theory of the "governing disjunction" seems to be problematic: it begs the question by asserting a priori that perception is the cognition of an objectively-existing external world.

> Dharmakīrti [...] seems to conflate two distinct concepts in his use of the term *svasaṃvedana*: following Dignāga, as a primitive non-objectual source of a sense of ownership (i.e. the implicit *de se*); but also as a cognition's immediate presentation to itself of its intentional content (i.e. the transparency of content). The second concept is in reference to the grounds of one's access to the contents of one's own mental states. (Ganeri 2012, 169 fn. 8)

In this way, two sides of the account of self-awareness emerge: a sense of minimal ownership upon the experience itself,[37] on the one hand, and an aspect of intentionality, on the other, *via* an immediate presentation of consciousness to itself. In this regard, the Yogācāra context in which Dharmakīrti operates is particularly well-suited for articulating a phenomenological account of consciousness without requiring any metaphysical commitments. On the one hand, the immediate acquaintance that consciousness has with its unfolding cognitive events does not necessarily entail a belief that these states belong to a self; nor, on the other hand, does this immediate access to experience necessarily imply any specific ontological stance regarding the existence of external objects.

Among the most influential contemporary studies on Dharmakīrti's theory of intentionality is that of Dreyfus (2007), who argues that, within Dharmakīrti's epistemological project, the reflexive nature of cognition is intrinsically linked to the very structure of mental activity. Within this process, the aspects of consciousness play a crucial role. If subjective and objective aspects do not exist separately, and a mental state encompasses them both in such a way that the subjective aspect 'beholds' the objective one, "self cognition is nothing over and above this beholding" (Dreyfus 2007, 202).[38] There is no subject-object duality between the two, they are not separate, and there is an intimate contact between them: a mental state directly experiences itself, has an immediate access to itself, and self-awareness is nothing but the self-revealing feature of a mental event. Dreyfus also suggests that the aspects of consciousness which are representative of the objects play a crucial role as intermediaries between the external world and the mind. In fact, these aspects possess a double nature, partaking of both domains – being the forms of the objects cognized by the consciousness and the "aspected

37 This account of self-awareness implies that all one knows is something that 'seems' or 'appears' to them; thus, the resemblance with Kant's transcendental unity of apperception is made clear by many scholars (Arnold 2005b; Dreyfus 1997, 398 ff.). In fact, both stances emphasize that immediate cognitive acquaintance is never with things-in-themselves but only with things-as-they-appear. This way of appearing follows a necessary structure which entails, first and foremost, a perspectival unity of experience.

38 Cf. Dreyfus, Thompson 2007, 103-4.

consciousness itself" (Dreyfus 2007, 201). The *ākāra* is the mark of the object in the mind, such that

> [t]o be aware of an object means to have a mental state that has the form of this object and is cognizant of this form. The aspect is the form or epistemic factor that allows us to distinguish mental episodes. (336)

It is with respect to this that Dreyfus describes Dharmakīrti as a "representationalist" (336),[39] thereby seemingly aligning him with a Sautrāntika metaphysical framework. However, this represents only one aspect of a far more complex issue, one that requires considering Dharmakīrti's metaphysical position in light of his ability to navigate between Sautrāntika and Yogācāra ontologies. Although this shift in perspective remains the subject of ongoing debate, its significance and implications continue to be actively explored (Dunne 2004; Dreyfus 2007; Kellner 2011a).

Nevertheless, fully-fledged intentionality arises only when conceptions are present to discriminate and determine what objects are. Perception, without this ability, may thus be described as exhibiting "phenomenal intentionality", to be distinguished from "cognitive" (or conceptual) intentionality (Dreyfus 2007, 109). In that respect, Dreyfus points out a problem that Dharmakīrti raises and leaves unsolved – a problematic legacy for the later tradition to deal with. For Dharmakīrti, valid cognition boils down to perception and inference, and they are distinguished both by their modes of apprehension and by their objects. (Real) specifically characterized phenomena (SKT *svalakṣaṇāni*; TIB *rang mtshan*), or particulars, are the only possible objects for perceptions; (unreal) generally characterized phenomena (SKT *sāmānyalakṣaṇāni*; TIB *spyi mtshan*), universals or conceptual constructs, are the only ones for inferences. Since they necessarily hold different objects, perception and inference lack any epistemological point of contact. Nevertheless, fully-fledged intentionality requires both perception and conceptuality. Therefore, the problem of the need for a bridge between the perceptual and conceptual components of intentionality in self-awareness lies at the core of Dharmakīrti's account of *svasaṃvedana*. Eventually, the very core of his extensive epistemological project intersects with the teleological and soteriological dimensions of his discourse insofar as the epistemological framework is intended to describe the cognitive approach to a reality that the Buddhist practitioner must explore and understand by advancing through a series of stages that lead to full-fledged knowledge and ultimate liberation. A gap or inconsistency

39 Cf. MacKenzie 2007, 47.

within the epistemological system would therefore compromise the very trajectory of liberation it is meant to support.[40]

In fact, one crucial aspect of the whole issue of self-awareness lies in its ultimate function considered both from an epistemological and a soteriological standpoint. In Wayman's (1997) reading, the need for a separate and specific type of *pratyakṣa* that is *svasaṃvedana* – a need shared, according to him, by both Dignāga (although, as previously noted, this remains a matter of debate) and Dharmakīrti – is linked to the fundamental necessity of a faculty that can play the role of figuring out the whole system.

> Inference (*anumāna*) could not do it, because associated with 'delusion' (*bhrānti*), even though being the best of the faculties so tainted. 'Perception of the yogin' (*yogi-pratyakṣa*) cannot do it, because it was acknowledged to be unmixed[41] with the guru's precepts. But 'introspection' [*svasaṃvedana*] can entertain as object disparate features or images, as though in a mirror; therefore, could feasibly have the entire system as its object (the *svalakṣaṇa*). (Wayman 1997, 429)

How is it possible that self-awareness has the whole system of cognition (as formulated by Dharmakīrti) as its object, that is, as a *svalakṣaṇa*? Dunne (2006) provides a detailed explanation of this point on the basis of Dharmakīrti's theory of self-awareness and yogic perception (as it is articulated, for instance, in *PVis* 1.28-31). This is the type of knowing that should lead the practitioner through the teleological framework of liberation. Through the sequential development of the three types of insight, based respectively on hearing (or studying), thinking and meditating (*śrutamaya-*, *cintāmaya-*, and *bhāvanāmayajñāna*), one should start studying the concepts and discourse of the Noble Truths and deepen one's meditation on them. At some point, thanks to the third insight, one finally has a non-conceptual, vivid, perceptual cognition of those concepts. However, at this point a question immediately arises: how is it possible to shift from a conceptual cognition to a non-conceptual one (*PVis* 1.7ab and *PV* 3.287)? Moreover, how can one 'perceive'

40 For valuable contributions to the soteriological aspect of Dharmakīrti's thought, cf. Dreyfus 1997; Steinkellner 1999; Wayman 1991; Dunne 2004; 2006. As for Dignāga's and Dharmakīrti's accounts of self-awareness, see Yiannopoulos's (2017, 156 ff.) insistence upon the fact that dualistically structured cognitions are not epistemically reliable whereas non-dual self-awareness is.

41 Wayman (1997, 424) explains that the expression refers to the yogi who has gone beyond the insight or cognition induced by 'hearing' and 'thinking' and has reached the insight induced by 'contemplating' or meditating (the three insights being, in Sanskrit, *śrutamaya-*, *cintāmaya-*, and *bhāvanāmayajñāna*). For a better understanding of the way Dharmakīrti unpacks Dignāga's expression (*PS* 1.6cd), cf. Dunne 2006, 505 ff.

something that, being a conceptual construct (*sāmānyalakṣaṇa*), is unreal? Concepts are "causally inefficacious": universals are "incapable of the causal activity" necessary to generate a perceptual image or *ākāra* (Dunne 2006, 510-11). These are crucial questions. According to Dunne, it is here that *svasaṃvedana* intervenes as the solution for both difficulties: as far as it is known through self-awareness as a mental event, any cognition – even a conceptual one – is a particular and thus a real element. Self-awareness would constitute the solution, therefore, acting as the central pillar of the entire cognitive system, with significance on both the epistemological and soteriological levels (Dunne 2006, 512 fn. 41). Following his interpretation, since in *apoha*, the process of 'exclusion' that forms the concept, the negation is not ontologically distinct from the *ākāra* in conceptual cognition, that cognition, which is a concept, *qua* mental event is actually a particular; and it is known as a particular mental event by self-awareness itself. Dunne elaborates on this explanation by using the example of recognizing an individual through the act of making a conceptual judgment:

> The judgment, being conceptual, has a vague – not vivid – image that results in part from the *apoha* process; in short, the image is vague in that it is not a phenomenally clear depiction of the object that it represents. Nevertheless, even though the judgment's image is vague as a representation, it is nevertheless an image. In other words, the judgment does contain some type of phenomenal content. And as a mental event, that phenomenal content is a real mental particular that can be known in its nature as a mental event *through reflexive awareness* (*svasaṃvitti*). In relation to that reflexive awareness, however, the content no longer appears to stand for something else; that is, it is no longer conceptual. (Dunne 2006, 512-13; italics added)[42]

This is the role of *svasaṃvedana* that seems to be claimed and argued for by Dharmakīrti, according to Wayman's and Dunne's

42 See Dharmakīrti, *PV* 3.287: "*śabdārthagrāhi yad yatra taj jñānaṃ tatra kalpanā | svarūpaṃ ca na śabdārthastatrādhyakṣam ato 'khilam ||*". In Dunne's translation: "A cognition that apprehends a linguistic object (*artha*) is a conceptual cognition of that [object] which it is cognizing. The actual nature [of any cognition *qua* mental event] is not a linguistic object; therefore, any [awareness of awareness itself] is direct [and hence non-conceptual]" (Dunne 2006, 512; square brackets in the original).

interpretations.[43] The centrality of self-awareness in Dharmakīrti's soteriological and epistemological framework ends up being stressed by some subsequent strands of the tradition. On the basis of Śākyabuddhi's (ca. 660-720) reflections upon Dharmakīrti's project, for instance, conventional perceptions and inferences are considered as instruments of knowledge and action in life, but are left behind once one has reached the ultimate instrument of knowledge, that is, the non-dual self-awareness. In his commentary on *PV* 3.212-13, where awareness is described as being in itself undifferentiated, while, owing to confusion, its appearances are perceived as dualistic, Śākyabuddhi (*Pramāṇavārttikaṭīkā* vol. mdo 'grel nye 203*v*) remarks:

> even if external objects do not exist, awareness nevertheless arises with that cognitive image [i.e., with an image that appears to be external]. (Dunne 2004, 406; square brackets in the original)[44]

He then adds: "However, dualistic awareness is not real" (Dunne 2004, 406).[45] In fact, the pairing of non-dualism and non-ignorance constitutes a line of reasoning that runs as a consistent thread throughout the entire philosophical discourse that has developed around *svasaṃvedana*. As will be shown (§ 6.3), this theme reappears in the Tibetan tradition as well, particularly at the delicate intersection between the soteriological and epistemological dimensions of the broader issue of self-awareness.[46]

43 Dreyfus (1997, 413), instead, although acknowledging the link between soteriology and epistemology, is extremely cautious – as opposed to Dunne – when talking about the problem of how conceptual understanding could be deepened and intensified and thus gradually made clearer until it is so vivid as to be non-conceptual. He (1997, 413) claims, "Dharmakīrti does not discuss this […]. How can conceptual cognitions, which are mistaken, become […] undistorted, merely by becoming vivid?". Moreover, Dreyfus (1997, 414) argues that Dharmakīrti explains yogic perception in relation to inference without linking it to apperception. While Dharmakīrti discusses the clear nature of mind in a soteriological context, he does not explicitly associate it with apperception. Therefore, the idea that yogic perception is apperceptive likely stems from sources other than Dharmakīrti's texts. Although there may be a connection between wisdom and self-awareness, it does not directly correspond to the type of *svasaṃvedana* Dharmakīrti describes. This suggests that the link between yogic perception and apperception emerged later, even if it remains compatible with Dharmakīrtian thought.

44 Śākyabuddhi, *Pramāṇavārttikaṭīkā* vol. mdo 'grel nye 203*v*2: "*phyi rol gyi don med na yang shes pa nyid de'i rnam par skye pa*".

45 Śākyabuddhi, *Pramāṇavārttikaṭīkā* vol. mdo 'grel nye 203*v*2: "'*on kyang ngo bo nyid gnyis pa can gyi shes pa de yang bden pa ma yin par*". Cf. Yiannopoulos 2017, 157 ff.

46 It must be remembered here that the Tibetan tradition is divided on this issue between those who support epistemology as meaningful to the soteriological discourse and those who do not (Dreyfus 1997, 439-40). For a lucid consideration on the tight relationship between the Buddhist epistemological project and its philosophy of liberation, and on its biased interpretation by modern scholarship, see Steinkellner 1982.

Dharmakīrti's continuation of Dignāga's epistemological project adds precious elements to the formulation of the notion of *svasaṃvedana*. Reflection on the implications of the *sahopalambhaniyama* argument and the crucial role of *ākāra*s has demonstrated that the two key aspects of Dharmakīrti's view of cognition are the phenomenality and intentionality of consciousness. Nevertheless, as has been shown, Dharmakīrti's entire system hinges upon a problematic point that cuts across both the epistemological and soteriological dimensions, with self-awareness potentially emerging as the crucial factor in the search for a resolution.

3.2.2 Candrakīrti and Śāntideva: A Sharp Criticism of *Svasaṃvedana*

The notion of self-awareness has been the object of sharp criticism from numerous Indian Mādhyamikas commonly associated with the Prāsaṅgika school, such as Candrakīrti and Śāntideva, who commented shortly after the epistemological enterprise of Dignāga and Dharmakīrti. It is important to mention their contributions to the broader discourse, as they also constitute the main landmarks for the subsequent critiques developed within the dGe-lugs system in Tibet. This study will return to these figures when examining the views of Tsong-kha-pa and Mi-pham rNam-rgyal-rgya-mtsho (ch. 5), highlighting how the dGe-lugs tradition, following Tsong-kha-pa, interprets Candrakīrti and Śāntideva as rejecting self-awareness entirely, while Mi-pham, drawing on Śāntarakṣita's legacy, understands their critiques as denying only its ultimate existence. For now, the focus will be on the contributions of Candrakīrti and Śāntideva.

In *Prasannapadā* 7.8-12, Candrakīrti follows and retraces Nāgārjuna's work, but it is in *MAv* 6.72-77 that, analyzing and criticizing the notion of self-awareness, he (*MAv* 6.75 [TIB 6.76]) poses a crucial question that threatens the very foundations of Yogācāra metaphysics. In this passage, he questions the knowability and, by extension, the very existence of the other-dependent (SKT *paratantra-svabhāva*; TIB *gzhan dbang*), whose nature consists in being not-different from mind. Thus, his whole criticism ultimately addresses the existence and knowability of the mental, which lies at the core of Yogācāra metaphysics.[47] Even before that criticism, in Candrakīrti's *Madhyamakāvatāra*, three more attacks are found

47 See Candrakīrti, *MAv* 6.75c [TIB 6.76c]. This critique is framed against the Yogācāra doctrine of the three natures (SKT *trisvabhāva*), which holds that only the other-dependent is really existent. Candrakīrti wants to know: if the *paratantra-svabhāva* exists as empty of both subject and object, then who is aware of its existence? For the whole argument's implicit passages, see Garfield 2006, 205.

on the doctrine of self-awareness. The first targets the memory argument for self-awareness, showing that it begs the question (Candrakīrti, *MAv* 6.73 [TIB 6.74]).[48] The point, as formulated by Cozort, is as follows:

> Candrakīrti argues that self-consciousness is not necessary for the function it is imputed to serve – facilitating later memory of the subjective aspect of experience – because the memory of a previous consciousness and that previous consciousness are not inherently different. (Cozort 1998, 163)

The second point raised by Candrakīrti rejects the possibility of any self-cognizing cognition from the conventional standpoint of everyday experience (Candrakīrti, *MAv* 6.74 [TIB 6.75]). The third points out its incoherence due to entailing an identity of agent, action and object (Candrakīrti, *MAv* 6.75c [TIB 6.76c]).[49]

For his part, Śāntideva (*Bodhicaryāvatāra* – henceforth *BCA* –, 9.17-26) gathers together various elements that have previously been used in arguments for and against *svasaṃvedana*. In criticizing a Yogācāra opponent, he recalls the blade of a sword that cannot cut itself (Śāntideva, *BCA* 9.17) and criticizes the metaphor of the lamp (Śāntideva, *BCA* 9.18-22). He then attacks the memory argument by stating that memory can be established by virtue of a 'connection to having experienced something else' (TIB "*gzhan myong ba dang 'brel ba las* |", Śāntideva, *BCA* 9.23 [TIB 9.24]). The arguments used by Śāntideva are rather obscure and, particularly with respect to the example he gives for explaining memory without *svasaṃvedana*, modern scholars do not always agree in their interpretations of the main point he makes. The example offered involves a hibernating bear that, despite being bitten by a rodent, remains asleep. When it awakens in the spring, the bear feels the pain caused by the resulting

48 As formulated by Cozort (1998, 163), the point is as follows: "Candrakīrti argues that self-consciousness is not necessary for the function it is imputed to serve – facilitating later memory of the subjective aspect of experience – because the memory of a previous consciousness and that previous consciousness are not inherently different". As Arnold points out, "The main thing that Candrakīrti's critique in *Madhyamakāvatāra* adds that is not in the *Prasannapadā* is a refutation of the memory argument for *svasaṃvitti*" (Arnold 2005b, 106 fn. 46).

49 In this regard, see Arnold 2005b, 91-2: "It is important to note, though, the extent to which Candrakīrti's version of that argument is informed particularly by Sanskritic grammatical analyses; [...] Candrakīrti takes Dignāga's *svasaṃvitti* as an action – that is, as some kind or episode of cognition that will admit of the sort of agent-instrument-object analysis that can necessarily be given for anything involving a verb". This is precisely the point that Śāntarakṣita denies. See Nurboo 2022 for a thought-provoking comparison between Dignāga's and Candrakīrti's accounts of self-awareness, arguing that, despite their methodological differences, their fundamental epistemic agenda remains identical.

infection and, through that sensation, recalls the experience of the bite; actually, the bite was not experienced at the time it occurred, only later. To mention just a few interpretations of this line of reasoning, Cozort writes:

> According to Śāntideva, self-consciousness is unnecessary because the earlier object and the consciousness that experienced it are relatedly remembered. For instance, when one remembers having seen a patch of blue, one does so by first of all remembering the patch of blue and then remembering the eye consciousness that saw the patch of blue. It seems that this is *not* a case of merely *inferring* that an eye consciousness must have been present, since that would not actually be a memory of a previous awareness, but of *experiencing newly* what was previously experienced, if it was at all, only in a subliminal way. (Cozort 1998, 169; italics added)

While, for Cozort, it is not a matter of inference, for Williams, inference itself is what is at stake. He comments on the example as follows:

> In the case of the poisoning the person is not remembering that something experienced occurred. Rather he is *inferring* the occurrence of something not actually experienced at the time, something which, nevertheless, accompanied an experience of being bitten. What Śāntideva appears to want to say, therefore, is that in a memory of seeing blue I remember blue but *infer* seeing blue, although at the time there was only the experience of blue, not a concurrent experience of seeing blue (i.e., self-awareness as well). In remembering seeing blue I remember blue but *infer* (if I wish) seeing blue. (Williams 1998, 153; italics added)

Garfield offers yet another interpretation, reading the example of the bear as positing a cognitive state directed toward the past. He grounds this interpretation in the Tibetan commentary by rGyal-tshab Dar-ma-rin-chen (1973), one of Tsongkhapa's principal disciples (1364-1432), and explains it as follows:

> 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. (Garfield 2006, 210; italics added)

The diversity of interpretations is evidence of the obscurity of the example itself (cf. Thompson 2011; Jinpa 2002, 128) while the nature of the specific link between experience and memory emerges as a recurrent point of debate for the subsequent Madhyamaka tradition, both in India and Tibet (Kellner 2010, 215 fn. 35).

3.2.3 Śāntarakṣita's Contribution

The discussion now turns to a final figure within the Indian Buddhist landscape, Śāntarakṣita, whose contribution significantly shaped subsequent developments on the topic of self-awareness and who would later serve as a key reference point for Mi-pham rNam-rgyal-rgya-mtsho's elaboration in the context of the Tibetan reception of this notion. In the eighth century, Śāntarakṣita, together with his disciple, Kamalaśīla,[50] developed what Williams (1998) defines as a theory of "reflexive" self-awareness; it is considered a "more simple, fundamental" (Yao 2005, 149) version of self-awareness, of which Mi-pham rNam-rgyal-rgya-mtsho was to be the main Tibetan heir. Some of Śāntarakṣita's relevant verses will be examined, as they later (ch. 5) serve as a key reference point for understanding Mi-pham's contribution in the context of the Tibetan reception of the notion of self-awareness.[51]

Śāntarakṣita takes as a given that there is a clear distinction between two classes of entities, the sentient and the non-sentient or matter, and that this distinction pertains only to the ordinary level of discourse. The important criterion for this differentiation is self-awareness itself, the absence of which renders something non-sentient. Śāntarakṣita's position, as found in his *Madhyamakālaṃkāra* (henceforth *MAl*) (16), is as follows:

> Consciousness arises as the opposite
> of the nature of what is matter;
> whatever belongs to the nature of non-matter,
> that is consciousness [aware] of itself.[52]

50 Although the Tibetan tradition considers Śāntarakṣita a Yogācāra Sākāravādin, Ichigō (1989, 177-9) suggests that this classification may be misleading or at least inaccurate. While Śāntarakṣita does adopt certain elements of the Sākāravāda position, considering cognition as aspectual, and cognitive appearances as real, other aspects of his philosophy align more closely with the Alīkākāravāda view, which holds that such appearances are deceptive or illusory. For this reason, it would be more appropriate to classify him simply as a member of the Yogācāra school. See Funayama 2007 for a detailed study on the difficulty of classifying Śāntarakṣita and Kamalaśīla in relation to the *satyākāravāda/sākārajñānavāda* and *alīkākāravāda/nirākārajñānavāda* categories.

51 For further studies, see Ichigō 1985; 1989. Cf. Blumenthal 2004 for a detailed explanation of how Śāntarakṣita's stance is in dialogue with Dharmakīrti's.

52 Śāntarakṣita, *MAl* 16: "*rnam shes bem po'i rang bzhin las || bzlog pa rab tu skye ba ste || bem min rang bzhin gang yin pa || de 'di bdag nyid shes pa yin ||*". Cf. Blumenthal's translation: "Consciousness is produced in the opposite way from that which is of an inanimate nature. That which is not the nature of being inanimate is the self-knowledge of this [consciousness]" (Blumenthal 2004, 83; square brackets in the original).

Self-aware cognition is the very quality which defines sentience. That which is conscious must be self-aware, by definition.[53]

Next, Śāntarakṣita rejects the idea that reflexive awareness entails any identity among agent, object and action, as these components are absent from its structure. According to him, it has a unitary nature and should be understood as a basic, inherent feature of sentience. In saying so, he refutes the notion that consciousness should be conceived as a form of action. In fact (Śāntarakṣita, *MAl* 17):

> In the nature of what is single and partless,
> a threefold nature cannot be sustained;
> the self-awareness of that
> does not [exist] as the entities of action and doer.[54]

Before moving on to the next stanza, Śāntarakṣita's autocommentary argues that since consciousness does not rely on anything external to itself in order to be illuminated, it must possess a nature that is self-illuminating. This leads into the next major point (Śāntarakṣita, *MAl* 18):

> Therefore, this [consciousness],
> being the nature of consciousness, is able to cognize itself.
> With regard to the nature of other objects,
> how can it come to cognize it?[55]

Since consciousness is by nature the opposite of non-sentience, it is not possible for consciousness to contact it. Thus, in knowing an object, consciousness must be apprehending itself in the form of an object, that is, it must be cognizing an aspect of itself. Arnold reads this point in terms of "ontological parsimony":

53 For a speculation on the reasons why Śāntarakṣita takes this position, see Arnold 2005b, 95.

54 Śāntarakṣita, *MAl* 17: "*gcig pu cha med rang bzhin la || gsum gyi rang bzhin mi 'thad phyir || de yi rang gi rig pa ni || bya dang byed pa'i dngos por min ||*". Cf. Blumenthal's translation: "Self-cognizing cognition is not an entity which [exists as] agent and action [with its object] because it would be incorrect for consciousness, which is of a single, partless nature, to be three (i.e., knower, knowing, and known)" (Blumenthal 2004, 86; square brackets in the original).

55 Śāntarakṣita, *MAl* 18: "*de'i phyir 'di ni shes pa yi || rang bzhin yin pas bdag shes rung || don gyi rang bzhin gzhan dag la || de yis ji ltar shes par 'gyur ||*". Cf. Blumenthal's translation: "Therefore, this [consciousness] is capable of self-consciousness (*bdag shes*) since this is the nature of consciousness. How [though] could that cognize the nature of objects from which it is distinct?" (Blumenthal 2004, 88; square brackets in the original).

> insofar as cognition is constitutively distinct from putatively material objects, it makes more sense for the direct objects of cognition to be of the same nature. (Arnold 2005b, 96)

However, with the distinction between the sentient and non-sentient, representationalism – described by Garfield (2006, 212) as "curiously Cartesian in character" – emerges as its direct consequence.[56] Since a state of consciousness is immaterial and cognizant by nature, its content must be as well. This epistemological theory closely aligns with the Yogācāra position, with the key difference that Śāntarakṣita – who does not regard the mind as ultimately existent – confines the validity of this view to the conventional level only.[57] Suganuma offers a short but useful summary of the view Śāntarakṣita holds of self-awareness:

> Briefly speaking, cognition is defined as 'the no-unconscious' (*ajaḍa*), not as the unconscious (*jaḍa*), only because the cognition is 'self-cognisability', in the concrete, 'self-revelation' (*ātmaprakāśa*). Śāntarakṣita explains it more concretely. A cognition does not depend upon any other cogniser (*anyadvedaka*: *shes byed gzhan*) for cognition of its own form; and yet this all is not uncognised; this is what is meant by 'self-cognition' (*svasaṃvid*: *rang rig pa*). According to him, cognition is the illumination of its own by its nature. In order to illuminate itself, cognition never needs any other things than itself. (Suganuma 1963, 69-70; Author's adjustments for the transliterations from Tibetan)

For Śāntarakṣita, cognition is devoid of the cognized and the cognizer, is one and impartite. Śāntarakṣita's account of self-awareness may be seen as bearing resemblance to Kant's notion of the transcendental unity of apperception:

> Candrakīrti's arguments fail to undermine the understanding of *svasaṃvitti* that can be developed following Śāntarakṣita – with the invulnerability of Śāntarakṣita's view now expressed as a function of its being a basically *transcendental* idea. More precisely, while Candrakīrti's critique targets the view on which *svasaṃvitti* is considered a particular kind of *intentional* cognition (considered, that is, to *display* intentionality), Śāntarakṣita's is more like the

56 For thoughtful considerations of the reasons why the 'given' embraced by Śāntarakṣita and Kamalaśīla differs from the 'given' that is rejected by the dGe-lugs scholars, see McClintock 2003.

57 For a parallel analysis of the exposition of this theory as it appears in Śāntarakṣita's *Tattvasaṃgraha*, see Ichigō 1989, 173-7. For an interpretation of Śāntarakṣita from a naturalist perspective, see MacKenzie 2019.

> view that *svasaṃvitti* is *itself* 'intentionality'. (Arnold 2005b, 78; italics in the original)

Svasaṃvitti or *svasaṃvedana* would then denote the intentional and constitutive structure that characterizes any cognition[58] and would be able to defend itself against Candrakīrti's criticism that agent, action, and object would become identical. The whole issue will return (§ 5.2) in the context of the Tibetan philosophical framework, where the way Śāntarakṣita's landmark assertions influenced the subsequent tradition will be made clear. In conclusion, Śāntarakṣita's view of self-awareness presents it as a 'presupposition' of experience itself, rather than as a philosophical 'explanation' of specific cognitive functions (as in Dignāga's case, for instance, in relation to memory). For Śāntarakṣita, it is 'obvious' that there are sentient and non-sentient entities, and that self-awareness is what makes the former sentient. This line of reasoning represents a fundamentally different approach from the view that self-awareness exists solely to enable specific mental functions.

In this way, the rich and nuanced spectrum of Indian Buddhist reflections on the topic of self-awareness in the post-Dignāga phase of its development laid the ground for the subsequent Tibetan discourse, where various standpoints would intersect, evolve, and continue to shape the philosophical conversation on the reflexivity of mind.

58 See MacKenzie's (2017) useful identification of the "transcendental" role of *svasaṃvedana* as the distinguishing feature or nature of consciousness, and its "epistemic" role as the most basic and secure means of knowledge.