The author of this thesis has granted the University of Calgary a non-exclusive license to reproduce and distribute copies of this thesis to users of the University of Calgary Archives.

Copyright remains with the author.

Theses and dissertations available in the University of Calgary Institutional Repository are solely for the purpose of private study and research. They may not be copied or reproduced, except as permitted by copyright laws, without written authority of the copyright owner. Any commercial use or re-publication is strictly prohibited.

The original Partial Copyright License attesting to these terms and signed by the author of this thesis may be found in the original print version of the thesis, held by the University of Calgary Archives.

Please contact the University of Calgary Archives for further information:
E-mail: uarc@ucalgary.ca
Telephone: (403) 220-7271
Website: http://archives.ucalgary.ca
Abstract

Parties from traditions in both East and West have claimed that the self lacks reality. I defend a version of this claim that uses some concepts and ideas drawn from an unorthodox (but argued-for) reading of the Buddhist doctrine of no-self (anattha). On this reading, the ordinary self is an illusion; its singular impression created by input from two sources. One source pertains to what I refer to as underlying witness-consciousness or awareness: this, I argue, is neither illusory nor mentally constructed. I argue that, through its very nature, awareness brings to the ‘self’ a host of non-illusory features: a subjective sense of presence, the elusiveness to its own observation, unity (disparate percepts appearing to a single point of view), endurance and invariability (non-perduring, unchanging sense of presence). The other source pertains to what I call ‘boundedness’, and it brings to the ‘self’ an ontological distinctness, such that the self is felt to be an individual, personalised entity with all the features of awareness. I argue that boundedness is not intrinsic to awareness, but comes about through desire-driven thoughts and emotions. When infused with the ‘awareness source’, their content depicts a separate self as their subject. Input from the ‘boundedness-source’ makes the self a mental construct, and the constructedness bestows on the self an illusory status. (For the self does not subjectively seem to be the mental construct it actually is: the self seems to be an unconstructed entity which thinks the thoughts rather than an entity whose existence depends on thoughts). In virtue of its two sources or ‘tiers’ — one constructed and one unconstructed — I term the phenomenon a ‘two-tiered illusion of self’. The theory is to be contrasted with standard Western theories of self-as-illusion (e.g., Hume, James, Dennett) which tend to pin the self’s illusory status on unity and endurance, by regarding these features as illusory in themselves.
Acknowledgements

In the four and a half years it took me to write this dissertation a number of people from different countries especially helped to make this possible. First of all I deeply thank my mother, Sonia Albahari, who offered me not only a home in Perth while I wrote most the thesis, but her unwavering support, confidence and interest in my project. We had numerous conversations about material in the thesis and these interchanges were instrumental in the development of my ideas. Thanks also go to my dear brothers Joseph and Benjamin for their support and feedback.

I am immensely indebted to my supervisor John Baker who continued to give me comments on thesis-chapters during the four years that I was in Perth and on leave from the program. I think it is accurate to say that without his supererogatory help I would not have finished the dissertation. His feedback and phone conversations not only significantly improved the philosophical standard of my work; they improved me as a philosopher.

A number of friends, colleagues and teachers in different countries have made it markedly easier for me to complete the thesis. In Calgary, Speranza Dolgetta was an amazing source of moral support, both through correspondence when I was in Perth (including sending me a huge box of candy and cookies to keep me motivated!) and in Calgary where her wonderful friendship helped me stay afloat and happy in numerous ways including cooking me meals during that final month when I had no time to cook. I thank my lovely roommate and friend Deanne Ivany for letting me live with her again in such a positive environment; her encouragement and advice has been a true help (thanks also to her partner Alfred!). Merlette Schnell and Renilda van Aerden played in a big role in making my life easier, including being patient, kind and efficient with the many hurdles that had to be jumped during my re-application to the program. Being able to work in a nice office has made an enormous difference to the quality of my time here and I thank Merlette for her kindness in organising this. Brian Grant was also generous enough to lend me his computer for the semester, which saved me from having to buy a laptop. I thank Greg Janzen for our stimulating conversations: from him I have learnt a
lot about the philosophy of mind. Yoshiki Kobasigawa, Helen Wirrell and Dave Mellow have also been especially wonderful since I arrived back in Calgary.

Members of the Philosophy Department from the University of Western Australia deserve special mention. Throughout the project, Michael Levine has been a source of moral and intellectual support. I thank him for showing a keen interest in my work (when he barely knew me) and for commenting on early drafts, and for twice letting me teach his Philosophy East and West course. Lee Carter and Barry Maund also went of their way to make me feel welcome in the Philosophy Department even though I was never a student there. The overall amenability from the faculty at University of Western Australia, including their willingness to let me participate in their philosophy seminars, enabled me to engage with the philosophy world in what would have otherwise been an isolating time.

I have a deep gratitude to the Buddhist community in Perth (both Theravadin and Zen) for helping me understand Buddhism on a level that I could have never had got from reading books. The weekly talks by (Theravadin) Sangha members and the meditation retreats changed not only my ideas but also my way of thinking and living. I am especially grateful to Sister Ajahn Vayama for her wise teachings and her individual encouragement on my project. I feel a special regard for Ajahn Cattamalo whose nine-day meditation retreat formed a highlight of my time in Perth. It should be noted that the Ajahns (teachers) do not necessarily agree with all my interpretations of Buddhism (I take full responsibility for the differences!). From the Zen group, I especially thank Mark Edwards for his conversations and for introducing me to the group. They gave me valuable feedback on my ideas and were a wonderful group to do retreats with (led by Sensei Ian Sweetman). Other Buddhists who especially encouraged me with the thesis were Barbara and Ainsley, Mary, Imma and Bruno, Sandra, Nick and Ramoh. Finally I would like to thank Ajahn Viradhammo whom I never met but whose recorded talks inspired me greatly.

Several friends in Perth made a big difference to the ease of the dissertation’s completion. Thanks go to my former NZ flatmate and wonderful friend Kathleen Potter. Kathleen not only offered her moral support: her sharp mind was an excellent razor against which to test many new ideas. Kathleen very kindly lent and taught me how to
use the ‘end-note’ program which considerably sped up the formation of my bibliography. Rebecca Collins, Barbara Murray, Doug Reeson and Elizabeth McCardell were other people whose friendship and intellectual conversations made a real difference to my time in Perth.

I’d like to acknowledge my support-network in New Zealand, the place from which I had initially set out to do my PhD in Calgary. First of all I thank one of my oldest friends (not in age) Jane McKessar, whose contribution to my ideas is only partially reflected in the odd footnote and whose namesake appears in a fictitious example. Since starting the thesis I have talked to Jane almost daily on email and this constant interaction has been a source of immense support. I thank Rachel and Guy for their continuing friendship and positivity. And my dear friend Catharine who had two children while I wrote the thesis and whose brother Simon pipped me at the post earlier this year (we started our Ph.Ds at the same time and were betting on who would finish first!).

Finally I would like to thank my late father, Michael Albahari, who, along with my mother, invoked my interest in Eastern philosophy as a child. My parents were always 100% behind my decision to study philosophy to an advanced level, never putting pressure on me to get a ‘real job’. My father passed away when I was working on preliminary requirements for the Ph.D: he would have been overjoyed to see me complete a thesis on this topic.
Dedication

To my parents.
Table of Contents

Approval Page .............................................................................................................. ii
Abstract ......................................................................................................................... iii
Acknowledgements ...................................................................................................... iv
Dedication ...................................................................................................................... vii
Table of Contents ........................................................................................................ viii

Introduction: The Two-Tiered Illusion of Self ................................................................ 1

Chapter One: Some Central Distinctions and The Four Noble Truths ........................... 7

   Subject and Object ...................................................................................................... 7
      ‘Subject’ .................................................................................................................. 9
      ‘Object’ ................................................................................................................ 12
      Objects and the Buddhist Doctrine of Conditioned Co-Dependence ....................... 15
      Objects and the Five Khandhas .............................................................................. 17
      The Metaphysical Status of Objects in Buddhism ................................................. 18
   Self, Sense of Self, and what is Meant by ‘Sense’ ....................................................... 19
      What is the Difference between ‘Self’ and ‘Sense of Self’? .................................... 19
      What is Meant by ‘Sense’ in ‘Sense of Self’? ......................................................... 21
      The Buddhist Perspective on the Sense of Self ..................................................... 24
   Defining ‘Person’ ....................................................................................................... 24
   The Essence of Buddhist Teaching: The Four Noble Truths ....................................... 26

Chapter Two: The Probable Buddhist Notion of Nibbanic Consciousness ....................... 38

   Introduction ................................................................................................................. 38
   Some Preliminary Considerations .............................................................................. 39
      (A) Nibbana as Intrinsically Unconditioned, Lived, Objectless Witness-Consciousness .. 42
      (B) Nibbanic Consciousness and the Khandhas ....................................................... 51
         (1) The Arahant with ‘Proximate Nibbanic Consciousness’ ................................. 52
         (2) Pre-Nibbabic Consciousness of the Ordinary Person ...................................... 54

Chapter Three: The Definition and Status of Self In Buddhism ....................................... 56

   Ownership and Identification as Reciprocal Assumptions of Self ............................. 57
      ‘The Self’ Minimally Defined in terms of Identification .......................................... 64
      Personal Ownership, The Self and Identification ................................................... 67
      The Sense of Self (Through Reciprocal Senses of Personal Ownership and Self-
      identification) Co-arises with Tanha ....................................................................... 69
   Further Features Buddhism Ascribes to the Reflexively Assumed Self ......................... 71
   What Nibbanic Consciousness Brings to the Sense of Self ......................................... 78
   The Definition and Ontological Status of Self in Buddhism ....................................... 81
   The Misportrayal of Buddhism as Endorsing a ‘Bundle Theory’ of Persons ................ 84
Chapter Four: The Reflexively Assumed Self

An East-West Convergence on Description of Self

Roles Ascribed Reflexively To the Self

Attributes Ascribed Reflexively to the Self

Role: Knower/Observer/Witness; Attribute: Mental/Aware/Conscious

Roles: Owner of Thoughts, Perceptions, Experiences, Body; Initiator of Actions – The Source of Effort and Will; Thinker or Originator of Thoughts; Seeker of Happiness over Suffering. Attribute: Bounded, viz., Ontologically Unique in Identity

Boundedness and Identification

Identification Revisited

Three Modes of Apparent Self-Identification: This-ness, Autonomy and Consistent Self-Concern

(A) The Importance of Being ‘This Very Thing’

(B) Role: Agent of Actions and Thinker of Thoughts (Autonomy)

(C) Consistent Self-Concern

Evidence for ‘Seeker of Happiness over Suffering’ as an Ascribed Feature of the Self

Ownership: Perspectival and Personal: Their Relation to Identification

Where There is ‘Mine’, There is ‘Me’ (viz., the Sense of Being a Personal Owner)

Where There is ‘Me’, There is ‘Mine’ (viz, a Sense of Personal Ownership)

Integrating Modes of Identification into a Overall Conception of Personality

Identification, Ownership, Boundedness and Tanha

Attribute: Elusiveness

Attribute: Unity (Singularity)

Attributes: Endurance and Invariability

Endurance

Invariability

Attribute: Unconstructedness

Conclusion

Chapter Five: How Do We Construe “The Self Lacks Reality”?

Introduction

What does it Mean to Deny Reality to the Self?

What is an Illusion?

The Self as Construct and Illusion

An Intersection of Agreement for Those who Deny Reality to the Self

Western Thinkers Who Deny Reality to the Self

Buddhist and the Western Accounts of ‘No-self’: Summarising the Similarity and Differences

Elimination, Reduction and Causal Roles

Conclusion

Chapter Six: Linking Problems of Consciousness with Awareness

Introduction
Introduction: The Two-Tiered Illusion of Self

In Western philosophy there exist many theories of consciousness and the self. While Eastern input into this debate is not yet mainstream, it is gaining momentum as Western interest in this tradition grows. In this project, I attempt to explicate and develop some Buddhist ideas on consciousness and the self into a substantial position in the philosophy of mind. The position will ascribe illusory status to the self as a whole, but ascribe a non-illusory status to many of ‘the self’s’ attributed features (features that I argue are intrinsic to consciousness). What makes this account unconventional is that standard Western accounts of self-as-illusion usually consider most of these attributed features (in particular, what I will refer to as unity and endurance) to have an illusory status in themselves: they are what make the self illusory.

My task will first involve extracting from Buddhist primary sources what I think is the most plausible reading of the Buddhist position on consciousness and self (or more accurately, no-self). The Buddhist school that informs my analysis is the Theravadin tradition, whose teachings are based upon texts from the Pali Canon. The Pali Canon is generally agreed by scholars to contain the earliest record of the Buddha’s own teachings (Pali being the language probably spoken in Central India during the historical Buddha’s era of around 2500 years ago). I would devote less space to this exegetical task were it not for the fact that much of what passes in the West as “Buddhist” on these matters is, in my view, substantially off-track. Even if we put aside the wider goal of defending a Buddhist account of these matters, it is still necessary to set the record straight on just what is the most likely Buddhist position on consciousness and no-self. This exercise will be as much philosophical as it is historical.

Having argued for my interpretation of the Buddhist position on consciousness and no-self (which aligns their account closely with the Advaita Vedanta tradition) I turn to the

---

1 I thank Jane McKessar for her useful feedback on earlier drafts of this introduction.
2 From John Bullitt (2002) “What is Theravada Buddhism? A Thumbnail Sketch.” On http://www.accesstoinsight.org/theravada.html. This link may be helpful to readers who are unfamiliar with Buddhism (although a more detailed account of the central teachings is offered in this project).
3 The tradition of Advaita Vedanta is generally associated with the leading figure of Sankara who lived in India at around 788-820 CE. The teachings of Advaita Vedanta (that Sankara re-interpreted) are based upon the Upanisads, of which some predate the Buddha.
task of arguing directly for a (metaphysically neutral)\textsuperscript{4} variant of this Eastern-derived account of the illusory self.

The first thing to be clear upon is that the notion of self we are concerned with in this project is a notion that describes the ordinary garden-variety self that we – or at least most of us – reflexively assume we are (a point to be argued for). Roughly speaking, the self will be a conscious subject that is a unified (non-differentiated), happiness-seeking, enduring (non-perduring), ontologically distinct ‘me’ who is an owner of experiences, thinker of thoughts and agent of actions. On our theory of “no-self”, then, a self of this description will not actually exist, although most of us will, through our very mode of living, be reflexively assuming that we are such a self. In assuming that we are such a self – a self that turns out not actually exist – we will therefore be in the grip of a deep-seated illusion. Now this illusion (and the general notion of an illusion will be spelt out in some detail during the project) will – on the proposed theory – be ‘fed into’ by what I refer to as two ‘tiers’. One ‘tier’ will be most actively responsible for the self’s illusory status. It will involve input from those mentally constructed, and hence, I will argue, illusory factors (contributed to by the content of thoughts and emotions) that make the self seem like an ontologically unique entity. The other ‘tier’, while implicated in the self’s overall illusory status, will not in itself be mentally constructed or illusory. It will involve input from ‘witness-consciousness’; that underlying factor which, I argue, makes the self seem as if it is consciously apprehending the world (regardless of the particular sensory-mental modality through which the world is apprehended).

I will argue, in accordance with the Eastern position, that witness-consciousness ‘imports’ certain features into the overall self-illusion; features that are not in themselves illusory. These features, some of which seem to qualify the self, are in fact intrinsic to

\textsuperscript{4} By “metaphysically neutral” I mean that the Eastern-derived notion of witness-consciousness that contributes to the illusion of self will be one that suspends judgement on the deeper metaphysical status of consciousness. It shall emerge from my analysis of Buddhism that such consciousness is considered, in Buddhism, to be “unconditioned”, that is, intrinsically free from any limitations pertaining to space, time, empirical/conceptual quality, and relational dependency on other things. The notion of consciousness developed in this project will be treating consciousness as metaphysically neutral on this score – as neither presupposing nor ruling out the possibility that consciousness is unconditioned. Readers may notice a similarity between this “unconditioned” consciousness alluded to in Buddhism and Kant’s depiction of the noumenal subject. In chapter two, similarities and differences between Kant’s noumenal subject and the “unconditioned consciousness” are discussed in more detail.
witness-consciousness (so I shall argue). The features (which receive a fuller description in relevant chapters) include what I refer to as ‘unity’, ‘elusiveness’ (unobservability of the subject by the subject), ‘endurance’ and ‘invariability’ (an unchanging although elusive sense of ‘being’ or ‘presence’). What makes this theory of the self-illusion differ from most Western theoretical counterparts of the ‘illusory self’ – as exemplified in accounts by Hume, James, Dennett, Flanagan and Damasio – is that the Western accounts usually pin the self’s illusory status on the unity and endurance. Such theories tend to explain these features in terms of a ‘bundle theory’, where tricks of memory and imagination employ discrete thoughts and perceptions to create the illusory impression of a unified and enduring consciousness. An illusory impression of unity and endurance is what, on these theories, makes the self an overall illusion. As it happens, this position is most commonly ascribed to Buddhism by leading scholars in the field.

What I take to be the actual Buddhist (and Advaita Vedantin) position on self-as-illusion, however, wholeheartedly rejects any such ‘bundle theory’ of unity and endurance. The trick that our mind plays on us, according to Buddhism, is one that rather results in the overall reflexive impression of being an ontologically unique or bounded self that is unified, elusive, enduring and so forth. It is our impression of being an ontological unique entity that primarily makes the self illusory. The unified, elusive, enduring (etc.) awareness forms an unconstructed, non-illusory ‘tier’ which, when combined with a ‘tier’ of mentally constructed input, works to produce the impression of such an entity. In virtue of its two contributing ‘tiers’, I describe the self-illusion as the “two-tiered illusion of self”.

The most vivid way to illustrate the two-tiered illusion of self is through the use of an analogy, which will be spelt out as the theory is developed in this project. The self that we think we are, will, as a matter of fact, have an ontological status analogous to that of a dreamt-of sound (e.g. a shrill voice) which has a non-dreamt component (e.g. the shrillness from an alarm). The dreamt-of sound, ‘the shrill voice’, is analogous to the self as construed on the Buddhist and Advaitin account: it is an overall illusion and construct. In a central sense, there is no such thing as the shrill voice and no such thing as the self (hence, ‘no-self’). Nevertheless, the shrill ringing quality – derived from the alarm-clock – is non-illusory: it is not dreamt up. There is such a thing as the shrill ringing quality. Similarly, while the self as a whole is construed as an illusion in Buddhism and Advaita Vedanta,
various features ascribed to the self on this view – those pertaining to the unified, elusive invariable, enduring and unconstructed witness-consciousness – are not considered illusory or constructed. These non-illusory aspects that most people reflexively and mistakenly attribute to the self correspond to and qualify an everyday witness-consciousness. The illusory self as depicted in Western theories, by contrast, is more akin to a dreamt-of voice whose shrill quality entirely dreamt up, with no specific input from the world outside of the dream.

Now that we have some idea of the project’s general aim – to defend a theory of the two-tiered illusion of self along the lines just described – let us briefly outline how the argument will proceed in terms of each chapter. **Chapter one** is the scene-setting chapter, where various important distinctions (for example, ‘subject and object’, ‘self’ and ‘sense of self’) are introduced. It is a chapter intended to familiarise the reader with some basic Buddhist ideas (the theory of conditioned co-dependence, the five *khandhas*, The Four Noble Truths), so that it becomes easier to grasp what may be unfamiliar concepts in chapters two and three. **Chapter one** is also intended to serve as a background that can be appealed to in justifying my interpretation of the Buddhist position on consciousness and no-self in chapters two and three. **Chapter two** aims to extract from Buddhist sources the key notion of witness-consciousness (that I term ‘nibbanic consciousness’) whose metaphysically neutral variant will relate to the concerns of this project.

**Chapter three** shows the role this notion of consciousness plays in determining the famous Buddhist doctrine of *anatta* or no-self. Once we see how Buddhism construes nibbanic consciousness (amongst other things) as contributing to the sense of self, we can then see that when Buddhism rejects the reality of a self, it is not equally denying reality to every feature of the self. It is not denying reality to those features of the self that are deemed to reflect, albeit distortedly, the character of the contributing nibbanic consciousness. Because they are distorted through their reflexive attribution to a self, such features would not, in the ordinary sense of self, parade as being overtly unconditioned.

From **chapter three** we thus get an overall idea of the self whose existence Buddhism rejects, a self that most of us are supposed to have a sense of being. In **chapter four**, we spend some time arguing that most of us do, indeed, assume ourselves to be entities of more
or less this description. During the course of the argument, we elaborate upon some further
notions, in particular those of ownership (of which there is more than one kind) and
identification. In the philosophy of mind, these notions are not all that well developed. I hope
to rectify this situation through an analysis of ownership and identification that draws upon
my understanding of how the notions are used in Buddhism. This account of identification
and ownership will prove crucial to the later theory of how the self-illusion comes about.

In chapter five we note that the existence of a self fitting this overall description has
been denied by parties from both East and West. Given the differing metaphysical
backgrounds from which the self’s reality can be denied, we may wonder whether people or
parties always mean the same thing when claiming “the self lacks reality”. The main purpose
of chapter five is to develop the concepts of illusion and construct such that they can serve as
a common way to meaningfully deny reality to the self. Thus when Buddhists claim that the
self does not exist and when Hume and Dennett claim that the self does not exist, they may
do so in a different way and for different reasons. Yet the basic content of their claim will be
the same: an entity answering to the description of ‘self’ will not exist insofar as it is an
illusion and a construct (in fact, we shall see that the self will be an illusion if and only if it is
a construct).

In Buddhism the features of unity, endurance and invariability – considered
inherently non-illusory – are ‘imported’ into the sense of self through unconstructed witness-
consciousness. In chapter six I attempt to endorse this Buddhist claim by establishing that
the concept of witness-consciousness (part of the concept of self), which I term ‘awareness’,
is to be specified with reference to concepts depicting those aspects of the self considered
non-illusory in Buddhism. Having hopefully established that a special conceptual link exists
between awareness and these aforementioned features, I argue, in chapter seven, that
awareness, as co-specified with the features of unity, endurance and so forth, is in fact not
illusory or constructed. Awareness and its features have the independent reality that
Buddhism deems them to have. I also argue, for reasons alluded to in chapters six and seven,
that this concept of co-specified awareness ought to be a central concept of consciousness in
the philosophy of mind: currently it is not.

In chapter eight, I attempt to show (with an appeal to both Buddhism and Damasio's
(1999) theory) how the overall construct and illusion of self-as-separate-entity could come
about through the process of identification (described in chapter four). This process will involve input from thoughts and emotions that seem to reflect an ongoing reflexive concern in the ‘self’s’ welfare. The content of these thoughts and emotions will reflexively depict the self as an ontologically unique entity who is engaging with the world, promoting its own well-being and happiness. The innate first-person perspective from which these thoughts and emotions are harboured, however, will not in itself be illusory but will involve the co-specified awareness (as demonstrated in chapter seven). Through this perspective (aided by identification), the unconstructed awareness and its features will silently feed into the “storyline” of a self, creating the impression of a bounded self-entity that is ontologically distinct and with all the features of awareness. The two-tiered illusion of self, whose ground is argued for in chapter eight, is described more formally in chapter nine. In glimpses beyond I outline how this theory can be utilized in a line of inquiry that may have far-reaching consequences for the philosophy of mind.
Chapter One: Some Central Distinctions and The Four Noble Truths

The aim of this chapter is (a) to make some distinctions that will be important in the discussions to come and (b) to outline some key concepts found in Buddhist and related literature.

The distinction this chapter will first explain is one that is most fundamental to this project; that between ‘subject’ and ‘object’. From this, we outline the Buddhist theory of “conditioned co-dependence”, followed by the distinction between ‘self’, ‘sense of self’ and ‘person’. We then outline the core of Buddhist teaching, known as “The Four Noble Truths”. The Four Noble Truths provide the context in which the central Buddhist conceptions of consciousness and no-self are to be understood in chapters two and three.

Subject and Object

In a short online paper entitled “Subject and Object” (1997) Mait Edey offers a useful explication of a fundamental distinction that we pretheoretically make between ‘subject’ and ‘object’. Since the subject/object distinction will be pivotal to my project, it is worth quoting Edey at some length:

For purposes of this discussion, let the term 'object' refer to anything anyone might be aware of or pay attention to. The term refers, then, not only to 'physical' objects, including whatever material processes, states, or conditions one might discriminate, but also to such 'mental' or immaterial entities or processes as pains, sensations, memories, images, dreams and daydreams, emotions, thoughts, plans, numbers, concepts, moods, desires and so on. Whatever we may think about their ontological or epistemological status in other respects, I hope we can agree at least that any of these may be objects of attention or pass in and out of awareness.

5 A published variant of this online paper appears in Journal of Consciousness Studies in 1997. Since the online version (which appears on a JCS discussion thread) is more suited to the purpose of this chapter's discussion, it is the version I shall stick with. As no information is provided about its year of appearance, I shall assume the year to be 1997.
Let the term 'subject' refer to I-who-am-aware, whatever opinion we may hold of what that T may be. I hope that, no matter what various opinions we may hold, we can all also agree that: (1) I, subject, can be aware of some object; (2) I can focus awareness in attention; and (3) I can distinguish myself from the object I attend to.

Here a crucial point should be emphasised. The distinction between subject and object, and our capacity to make that distinction, is prior to any particular opinion or theory about what either the subject or the object may be. Another way to make the point is to say that we make the same distinction, and make it the same way, regardless of what we may think we believe about the nature of the self or consciousness and their relation to the world. Yet another way to put it is to say that the distinction is not made on merely conceptual grounds. Any time you are aware of some object, or attend to some object, you won't have any trouble distinguishing it from yourself. That is, you're likely to know, immediately, without having to stop and think it over, or having to collect any evidence, which is you and which is the object. You can distinguish yourself as subject from any object whatsoever ('physical' or 'mental') any time you direct your attention to that object and realize that it is you who are aware and who pay attention, not the object. The real nature of the object and the real nature of the subject may be baffling mysteries to us, but these mysteries are no barrier whatever to knowing which is obviously which. (Edey, 1997, online).

This passage provides a solid starting point for our analysis: all of us can distinguish ourselves as observing subjects from what we observe. The distinction is cast by Edey as pre-theoretical, at least in the sense that, as explicated, the distinction does not seem to require any metaphysical commitments as to what subject and object actually are – or whether the distinction can be ultimately upheld. Such questions – including whether subject and self are the same thing – will of course arise upon more in-depth philosophical analysis. However, at this stage it will be most productive to clarify some points made by Edey in such a way that the distinction is sharpened on a pre-theoretical level. Let us first focus on what is meant by 'subject'.
'Subject'

Edey’s passage suggests two closely related aspects of what it means be a subject, pertaining to (1) its *modus operandi* and (2) the perspective-creating locus for this *modus operandi* as it stands in relation to objects. The *modus operandi* of a subject is, to put the point broadly, its realised capacity to observe, know, witness, be consciously aware. I shall use the term ‘witnessing’ synecdochically to cover these three modes of apprehension, but when I do so I am to be taken as talking only about conscious cases of such apprehension. Minimally construed, witnessing can be described as the broadest mode of conscious apprehending, subsuming all species of conscious perceiving, thinking and introspecting, whether these apprehensions are attentive or inattentive. Witnessing is for example involved in conscious perception that is visual, auditory, olfactory, proprioceptive, tactile, gustatory; and in conscious cognitive apprehension such as thinking, remembering, feeling emotions or imagining. Witnessing can be thought of as the determinable under which these more specific modes of conscious apprehending are determinates (more in chapter four). *Attentive* witnessing usually involves a focusing of the conscious awareness (through whatever sense or mental modality) on some specific object of perception, thought or introspection (for example this writing); *inattentive* witnessings involves the peripheral conscious awareness we may have of surrounding objects such as the hum of a computer.

Now a subject does not present itself, in relation to observed objects, as something disembodied – and hence not as witnessing from no point of view. As Thomas Nagel (1986) makes clear, whenever the world is viewed (with its objects), the viewpoint will always seem to be from *somewhere*. This view from somewhere pertains to the second aspect of what it is to be a subject. What seems to create a discernible viewpoint for witnessing – as it would seem to the witness – is a psycho-physical instantiation in space and time (which might turn out to be a complex process). This extends not only to the viewpoint witnessing takes within a waking, material world (with its physical body) but to how things appear during dreams and such like: it is not as if the witnessing suddenly loses all sense of occupying a spatio-temporal perspective. *A subject S is hence a specific perspectival locus for witnessing, in relation to which objects are witnessed.* This locus implicitly *appears* – whether dreaming or waking – to be psycho-physically constrained in space and time, offering a specific viewpoint from which the world, dreamt or actual, is witnessed.
We might now wonder whether the *modus operandi* of a subject, namely the witnessing, could ever *come apart* from the specific (apparently spatio-temporal) perspective from which it seems to view the world. Could there be any such thing as pure witnessing with no definite perspectival locus — and hence no observed objects which would seem to necessitate the spatio-temporal perspective of a *subject* from which the objects are viewed? If witnessing could occur without a definite perspective, (that is, without a subject) then the term 'subject', as we have defined it here, may refer to something that is not as psychologically basic as it pre-theoretically seems to be. In everyday experience, filled as it is with observable sensory or mental objects (whose observation demands that witnessing assume a specific perspective), one is likely to get the impression that witnessing and a specific spatio-temporal perspective must always *go together* as a psychologically basic unit.

It just so happens, however, that Buddhism regards witnessing *sans* a spatio-temporally limited perspective (and without any objects that are witnessed) to be a genuine psychological possibility which may (a) be realised through advanced meditative states and (b) is said to reflect the real, unconditioned nature of witness-consciousness (without a subject as I’ve defined it). To avoid begging questions for or against this possibility, I refrain from formulating my definition of ‘subject’ or ‘witnessing’ in a manner that would either rule out or presuppose the psychological possibility of this scenario. While I presume the scenario of ‘pure witnessing’, as I shall call it, to be quite unimaginable for most people (insofar as we would be hard or psychologically impossible to imagine what, if anything, it would be like), we can at least gain some conceptual handle on what the scenario of objectless, perspectiveless witnessing might involve. It would imply witnessing *sans* any input from the five senses: no sights, sounds, smells, tastes, tactile and proprioceptive sensations. There would furthermore be no cognitively sensed objects: no thoughts, concepts, observable emotions or desires. Without any objects witnessed, attentively or inattentively, including objects pertaining to one’s bodily and mental state in space and/or time, there may well be no cues that could lend witnessing the implicit and ubiquitous impression of occupying a specific spatio-temporal perspective.

Why is it relevant to mention — at this early stage especially — this possible scenario of pure witnessing? What does it matter to our pre-theoretical definition of ‘subject’ whether its *modus operandi* of witnessing can or cannot occur without the perspectival confines that
are necessary for a subject? It matters because not mentioning it may reinforce various assumptions about the subject – assumptions in line with ordinary object-filled experience – that Buddhist philosophy, as I understand it, would fundamentally reject. The assumption I mainly have in mind is that witnessing implies a subject; or to frame it in more detail, that the outlook of witnessing must always be confined to a particular (seemingly spatio-temporal) perspective. Since the aim of this project is to first articulate (before defending a neutral variant of) the Buddhist position on consciousness and no-self, it is important, from the outset, to set the stage for thinking about consciousness in a way that is accessible to – and does not prematurely rule out – the Buddhist position on consciousness. As it will emerge in chapter two, witness-consciousness, in Buddhism, is implicitly construed as something that is not (and does not necessarily seem to be) limited to or by any spatio-temporal (and hence psycho-physical) perspective. Witness-consciousness is regarded as a metaphysically basic ‘substance’ (for want of a better word), completely unconditioned by what Kant refers to as the parameters of time, space, quality and relation. At a later stage, we will be saying more about what this involves: the aim for now is to make sure no questions are begged by supposing that whenever there is witnessing, there must be a delimiting subject that is witnessing. If the Buddhist position (as I understand it) is correct, then there can be pure subjectless witnessing; that is, witnessing without a subject, viz., without the confines or appearance of a spatio-temporal (hence psycho-physical) perspective.

Allowing that the Buddhist position might be correct also calls into question a further common assumption; the assumption that witnessing must be categorically an activity or state of a subject (which is an individual). Such an assumption is at odds with (my understanding of) the Buddhist notion of witness-consciousness which is to be construed as metaphysically basic and not as a state or activity of anything else (and so perhaps as more like an individual). While this project will not include a fully-fledged argument for the reality of such metaphysically basic consciousness, it is hoped that nothing said in this project will rule out the possibility that witness-consciousness could be metaphysically basic in this way. While the surface-grammar will make it sound like I talk about witnessing as an activity or state, this surface-grammar should not be taken to indicate any deep metaphysical reality, just as talk of the ‘setting sun’ does not imply that the sun literally sets. I leave open the possibility that witnessing – everyday witnessing – could, in its essence, be metaphysically
basic in the manner suggested by my understanding of Buddhism, and hence not, in any essential way, a state or activity of a subject. On such a scenario, those psycho-physical factors that appear to confine witnessing to a particular spatio-temporal perspective (thus helping generate the impression that witnessing implies a subject) would be as unessential to the nature of witnessing, as clouds, whose formation can appear to confine the blue sky to a circular shape, are unessential to the nature of the blue sky. More will be said about the Buddhist notion of witness-consciousness in chapter two.

When trying to conceive of pure witnessing, the question will probably have arisen: would there be something it is like to experience this unusual mode of witnessing, should it be psychologically possible? Or would it be more akin to how we might conceive of a deep coma, with an absence of any phenomenal inner life whatsoever? This has direct bearing on the question of whether witnessing, as opposed to the objects witnessed, has its own non-intentional phenomenal feel, a feel that it may bring to ordinary conscious experience. It is fairly uncontroversial that there is something it is like to experience an object via this or that sense modality; for example that there is something it is like to see a red apple – or simply to see, simpliciter. The overwhelming hold that discernible sensory and mental objects tend to have upon our field of awareness makes it hard to introspectively decide whether there is more to the experience than just these objects. Later in the thesis, I shall argue that there is indeed a non-intentional phenomenal quality that witnessing qua witnessing brings to a conscious experience – even if it can never occur in a subjectless and hence objectless fashion.

'Object'

We turn now to the term, object. In the following, I will expand upon Edey's useful stipulation. Edey stipulates that the term 'object' refer to "anything anyone might be aware of or pay attention to" – regardless of whether it be external or internal to the (witnessing) subject's mind. I will, more precisely, define an object as anything that a subject can, in

---

6 Barry Dainton (2002) uses the word "tangible" to convey the idea of a phenomenal dimension to witnessing or 'awareness' (a synonymous term in this context).
7 Greg Janzen argues (in his forthcoming Ph.D dissertation) that with regards to vision, the what-it-is-like-ness is due to simply the seeing, rather than to what is seen (e.g.colour).
principle, pay attention to. This definition does not imply that something fails to be an object if inattentively witnessed, but it does mean that a subject must, in principle, be able to focus their attention on it. For example, two seconds ago, I was inattentively witnessing the hum of this computer. At this current point in time, however, I (as a subject) focus my attention on the sound, demonstrating that the hum of the computer is — and always was — an object. Thus anything that can in principle be attended to by a subject — and so implicitly viewed as separate from the subject qua witnessing — qualifies as an object on our analysis: the term is very general, cutting across a traditional distinction between individuals, properties and relations. This general designation of ‘object’ — as something that a subject can in principle attend to — will be central to this project and neutral with regard to debates on representation. The term ‘object’ will, for example, remain neutral on the question of whether an observed object (for example a cat) reduces to a Lockean idea that partially resembles the feline object in terms of its primary but not secondary qualities, or whether the observed cat is a mind-independent object whose presentation is mediated via sensory peculiarities, or whether it is some phenomenon constructed by the Kantian categories of space and time in response to unknowable noumena, or whether …(and we can go on at some length). The point of central relevance is that the object is anything that can be attended to by a witnessing subject, whether through channels that are mental, visual, auditory, gustatory, olfactory, tactile, proprioceptive, or pertaining to those channels, to the very qualia associated with each sensory or mental modality. Our analysis allows object-terms to adverbially modify those terms denoting perceptual experience, if that is called for. It is possible for colour-subjectivists to analyse colour in this way, as in “she perceived the chair blue-ly”. In our context, the “perceiving blue-ly” still qualifies as an object since the blueness is a visual experience to which a subject can attend. The witnessing is not itself modified by any such experience denoted by the term ‘blue-ly’ since the same quality of witnessing can quickly be focused elsewhere and non-blue-ly to another sense modality and its deliverance. (The sense-independent quality of witnessing will be argued for at length later in the dissertation).

I say in principle and not in practice, since there will be various things (for example the entire empirical universe, sub-atomic particles, or bat-qualia) where limitations of cognitive power (e.g., on the part of a human subject) will prevent them from being objects of a single attentive purview. While such items may in practice escape the focus of attention (for example by humans) they can in principle be attended to (by a subject with the necessary cognitive powers). That is enough to render them objects on my definition.
While ‘object’ is defined in relation to a subject, we are not stipulating that an object be occurrently experienced by a subject in order to qualify as an object. It is enough that an object be the kind of thing that can in principle be (experientially) attended to by a subject, whether that thing is ‘internal’ (such as a thought or mood) or ‘external’ (such as an apple-tree). This includes the physical world and its properties (occurrently observed or unobserved) as well as the world of thoughts, concepts, ideas, dreams, memories and pains – and hence it fits with Edey’s intuitively drawn definition. Does it include the subject, viz., the witnessing perspective? In other words, can a subject qua witnessing perspective ever, in principle, be the attended-to object of an experience by a subject – whether the observed ‘object-subject’ be reflexively one’s own witnessing perspective or the witnessing perspective of another? We can immediately note that the subject does not seem to be something that can ever be attended to – whether it be someone else’s witnessing perspective or one’s own. With regards to other subjects, it is pretty much agreed that even having direct access to another subject’s ‘internal objects’ such as their thoughts and pains is an impossibility – perhaps conceptually as well as empirically. So it seems even more unlikely – and I would say incoherent – that the very point of view from which another subject views the world could, qua that point of view, become the object of another subject’s attentive witnessing (a first-person witnessing perspective experiencing another person’s first-person perspective: the mind boggles!)\(^9\). Can one’s own witnessing perspective (viz., the subject) be the reflexive object of one’s attentive witnessing? Again, it is very hard to see how this could in principle be possible. The main reason – to be elaborated upon later – is that the witnessing subject (and for that matter, witnessing) seems systematically elusive to reflexive observation: just as an eye cannot directly see itself, witnessing – and the witnessing subject – cannot seem to attend to its own witnessing, or to the subject’s viewpoint from which the witnessing ensues. Whether this is a benign point of logic (as Ryle suggests)\(^10\) or there is more to it (as will be argued later), it does not seem that the subject can be its own object of attention in this reflexive sense either. Thus it would seem, on reflection, that the

\(^9\) The best one can hope for, if one is a physicalist, is some kind of identity, reduction or elimination of the witnessing perspective to brain-process, meaning that when the relevant brain-process is observed by a third party, then so is the witnessing perspective – should it exist. But note: the observed object would no longer be a subject qua subject, but merely (at the most) a subject qua brain-process. The subject-object distinction would not be obviously transgressed.

subject/object distinction is not obviously collapsible in this way, and so unless proved otherwise, it will remain an axiom in this thesis.

**Objects and the Buddhist Doctrine of Conditioned Co-Dependence**

I will now say more about the notion of objects as it relates to Buddhist cosmology. In doing this, we introduce ideas that will aid in our later analysis of how to understand the concept expressed by the word *nibbana* in Pali (*nirvana* in Sanskrit). *Nibbana* is depicted as being *unconditioned* – a notion that will make little sense unless we first understand what it means, in Buddhism, for something to be *conditioned*. That is the purpose of this section. **Note that the exercise here is not to defend, but to elaborate Buddhist thought on this matter** – although much of it should come across as fairly uncontentious. With this proviso in place, I shall avoid the tedium of repeating “according to Buddhism” after every little phrase: this can, for purposes of the discussion, be taken as a given.

In Buddhist cosmology, the world of objects – as we’ve outlined it above – amounts to what is known as ‘conditioned existence’. Each object within conditioned existence is characterised by a co-dependence upon other conditioned objects in a manner that is both *synchronic* and *diachronic*.11 *Synchronic co-dependence* means that a given object cannot, at a given time, exist without a concurrently supporting network of other conditioned objects. ‘Other conditioned objects’ can include the internal structure and functionality of an object, in other words, objects that make up a more complex object. Our living bodies, for instance, depend upon the mutual functioning of innumerable internal objects such as organs, DNA, cells and atoms. Supporting objects also occur outside of the conditioned object; our living bodies depending upon such factors as breathable air, temperature, gravity and sustenance.

On Buddhist cosmology, there is no object whose existence does not synchronically depend upon other specifiable conditions, such that the withdrawing of those conditions (themselves objects) would not bring about the destruction of that object. There is, in other words, no such thing as an indestructible object, whether this be a quantum particle or a depressed

---

11 The division of conditioned co-dependency into explicit ‘synchronic’ and ‘diachronic’ dimensions is my own. However, both are implied in an abstract formulation of the doctrine, expressed in the *suttas* (and quoted in Harvey), hence “That being, this comes to be; from the arising of that, this arises; that being absent, this is not; from the cessation of that, this ceases” (1990, 54). On this, Harvey comments: “This states the principle of conditionality, that all things, mental and physical, arise and exist due to the presence of certain conditions, and cease once their conditions are removed: nothing (except *Nibbana*) is independent” (1990, 54).
mood. At any moment in time, the synchronic existence of a conditioned object is made actual by the mutual manifesting of other favourable conditions. And each of these favourable conditions is itself dependent upon other conditions, suggesting a vast network of interconnected objects. That is partly what it means for an object to be ‘conditioned’. (On C.B. Martin’s (1996) system of dispositional ontology, this would be known as the “mutual manifestation of reciprocal disposition partners”\textsuperscript{12}.

As well as synchronic co-dependency, Buddhism postulates diachronic co-dependency (or conditionality) of objects. This means that whether on a microscopic level (external to a subject’s usual perspective) or a psychological level (accessible to a subject’s perspective) objects are constantly being created, changed and destroyed in a non-random, law-like fashion. One micro-moment of objects will condition or influence the next micro-moment according to laws of nature that Buddhism says operate at a number of different levels (mechanical, biological, psychological and so forth)\textsuperscript{13}. While some of these laws (e.g., mechanical) apply more often to objects of the external physical world, Buddhism places particular emphasis on the train of objects that appear directly (even if inattentively) to a subject’s perspective, and which, relative to this perspective, pass in and out of that subject’s conscious awareness (or witnessing). It is the world of objects as \textit{occurrently} (not just dispositionally) viewed by a subject, whether attentively or inattentively. I shall refer to these objects as ‘objects of awareness’ or simply ‘awareness-objects’.

Objects of awareness can include many items from the general set of objects (remember, objects can \textit{in principle} be attended to by a subject) and hence include features of the surrounding world, one’s physical body, sensations, perceptions, thoughts, volitions and emotions. They comprise the totality of objects that are either at one time, or over time, experienced by a subject. Buddhism maintains that these objects, when divided into their smallest components, arise and pass away from our conscious awareness in much the same way that Hume described when he tried to introspect the self; with an almost “inconceivable rapidity” (1739, 162). The nature and flow of these objects in relation to one’s awareness are of course subject to the spectrum of laws, such as mechanical/biological (stubbed toe leads to physical pain) or psychological (yearning for absence of pain leads to mental suffering).

\textsuperscript{12} See Armstrong, Martin and Place (1996) especially 71-87 and 184-191.
\textsuperscript{13} A good explication of these laws can be found in Bhikkhu Payutto’s book \textit{Good, Evil and Beyond: Kamma in the Buddha’s Teachings} (1993, 1-2).
Buddhism maintains that there is not a single object of awareness – or indeed object in general – be this thought, perception, situation, emotion or atom, whose existence has not been lawfully brought about or influenced by previous objects (where objects include states), whether from the preceding micro-moment or fifty years ago. There is no such thing as an object without a historical precedent: Buddhism does not accept that there can be something from nothing. Even the big bang is considered, in Buddhist cosmology, to be one of countless conditioned phases of the universe: conditioned existence is regarded as having no known beginning. On a more local note, as we said before, the subject need not be explicitly knowledgeable or cognisant about either the presence or kind of object that is being experienced – in particular, their status as arising and passing away from the purview of witnessing.

**Objects and the Five Khandhas**

There are many ways of categorizing objects, when the term ‘object’ is defined in the general sense I have introduced above. Buddhist literature suggests a particular categorization, (pertaining to what I have termed ‘objects of awareness’), which will be important in the discussions to come. The categories in question are, in Pali, called *khandhas* (Sanskrit: *skandhas*). These *khandhas*, which specify types under which awareness-objects can be tokened, are: ‘form’ (physical objects including bodies), ‘feelings’ (pleasant, unpleasant and neutral sensory or cognitive qualia), ‘perceptions’ (recognition of objects and qualia), ‘mental fabrications’ (thoughts, desires, personality traits, intentions), and the object-oriented ‘consciousness’ (awareness of objects specific to each sensory or mental modality). In more detail, while ‘form’ pertains to anything overtly physical that can be observed, the body is alluded to in particular, with its sensory and cognitive receptors. ‘Feeling’ pertains to the quale or raw feel of experiences that are associated with sensory or cognitive modalities, such as the redness of red, the aroma of coffee, the sadness of an emotion or all of these combined. Feelings are further parsed into three useful categories: pleasant, unpleasant and neutral. Pleasant feelings are those associated with physically or mentally enjoyable

---

14 See for example Harvey (1990, 32) who writes “The Buddhist view, in fact, is that there is no known beginning to the cycle of rebirths and the world...However far back in time one goes, there must have been a prior cause for whatever beings existed at that time”.

15 Descriptions of the five *khandhas* can be found in most Buddhist texts, including Harvey (1990, 49-50) and Gethin (1998, 31-32).
experience; unpleasant feelings pertain to physical and mental discomfort or suffering; neutral feelings are all those which are neither pleasant nor unpleasant (such as the background hum of a computer). When it comes to explaining the ‘Four Noble Truths’ we shall see that this parsing of feelings into pleasant, unpleasant and neutral categories is highly relevant. ‘Perception’ refers to the recognition of objects as particular objects, such as books, accidents, numbers and colours. It requires the functioning of memory, such that newly arisen input – mental or sensory – can be brought under a pre-existing concept. ‘Mental formations’ or ‘fabrications’ involve cognitive entities and events such as memories, volitions or intentions, thoughts, emotions, desires, attachments and character traits. Finally, ‘consciousness’ describes witnessing as it is directed toward (and co-dependent on) specific objects observed by the sensory/mental modalities. So there is eye-consciousness, ear-consciousness, nose-consciousness, mind-consciousness etc.

In its context as one of the five aggregates of conditioned existence, this object-oriented and hence relational notion of consciousness is not what I’m talking about when I refer to subjective ‘witnessing’ and there will be little discussion of it in this thesis. However, it is also true that without witnessing – in its non-relational generic sense – there could not be any object-directed and perceptual modes of this knowing or witnessing. This will become clearer as the thesis proceeds.

The Metaphysical Status of Objects in Buddhism

Having provided some (but by no means a complete) overview of how objects (in my stipulated sense) are regarded in Buddhism, we finish with some general points about the metaphysical status pertaining to their inherently conditioned, and hence impermanent nature. To do this, it will be useful to bring to bear on the exposition a fourfold categorization of modes of conditioned existence. These modes can, I think, be found in Kant, but I want to use them without any commitment to Kant’s theory about these modes. These modes in effect delineate key dimensions along which observable objects can be co-dependent: the four are space, time, quality and relation. To be conditioned by space is to take up room in space, with position and size. Physical objects such as the body will, on the Buddhist analysis, be conditioned by space. To be conditioned by time is to be either destructible and hence impermanent, or indestructable and existing forever within time (as opposed to
existing outside of time). Since each of the five *khandhas*—mental or physical—are impermanent through their synchronic and diachronic co-dependence, they will be conditioned by time. To be conditioned by **quality** means that an item will be describable through parameters that are empirically or conceptually delimited. Any quality that can be ascribed to an object will be definable in terms of conceptual and empirical qualities (including qualia) that can be the focus of attention. In the context of human observation, these will include observable empirical qualities that pertain to or can be known through the five sense modalities (discernible objects with colours, shapes, sounds, tastes, tactile and proprioceptive qualities) as well as qualities and concepts that pertain to mental deliverances such as emotions, thoughts and mathematical concepts. Obviously all tokenings of the *khandhas*, being observable through sensory and/or mental modalities, will be conditioned by quality. To be conditioned by **relation** is to not exist independently but to rely upon other objects for existence. The conditioned synchronic and diachronic co-dependence of objects in relation to other objects (according to Buddhism) will ensure that no object can exist in isolation; hence all objects, on Buddhist cosmology, will be conditioned by relation.

**Self, Sense of Self, and what is Meant by ‘Sense’**

**What is the Difference between ‘Self’ and ‘Sense of Self’?**

While the notion of self that concerns Buddhism will be properly outlined in chapter three, it is important to first get a clear grasp of the general distinction between ‘self’ and ‘sense of self’, since these are liable to be (and have been) confused. For the current purposes of explaining this distinction—as it will relate to the angle of this project—we will assume that ‘self’ means something like Descartes’ *res cogitans*: a thinking subject with unbroken endurance in its identity during a lifetime. It will hence be more than *just* what we have meant by ‘subject’. Now if ‘self’ refers to an enduring *res cogitans*, then one’s ‘sense of self’ will refer, at the very least, to a felt, deep-seated, *belief or assumption* that one is a self-entity of this description. For the sake of the current argument, we shall assume that this deep-seated belief or assumption of being an enduring *res cogitans* is common amongst humankind. We should note that such a belief or assumption need not be reflectively obvious to the person who has it: just as one may be unable to articulate specific grammatical rules
one has mastered, one may be relatively incognisant of the kind of entity one assumes one is 
(more on this soon).

Now the main reason for making a distinction between the self and the sense of self is 
that it enables us to articulate and explore a possibility that will be of central importance to 
Buddhism and this project in general. The possibility is this: while the sense of self widely 
exists, meaning that most people assume themselves to be a self-entity (a possibility that of 
course needs arguing), the self itself does not exist – meaning that there is no such entity that 
most people assume themselves to be (a possibility that will also need arguing!). As it will 
transpire, the enduring res cognitans that illustrates our current example is not dissimilar in 
description to the self-entity that Buddhism rejects. We can at this stage hazard a guess that 
Buddhist cosmology will not allow for a distinct entity (including a perspectively-bound 
subject) that endures unceasingly over time. Now, the rejection of such a self’s existence is 
quite compatible with the fact that most people unwittingly assume the existence of this self 
by taking themselves (qua a subject) to be such an entity (viz., an enduring res cognitans). 
While the deep-seated and common assumption of self will on this scenario turn out to be a 
false one, it will nonetheless be true that the assumption of self, viz., the sense of self, will 
exist. Should this be the case, then the correct thing to say will be that while the sense of self 
is real (for example, non-illusory), the self is not real (for example, illusory)\(^\text{16}\). I underscore 
this point, so as to avoid falling into the easy trap of saying or thinking that the sense of self 
is unreal or illusory, when actually meaning that the self is unreal or illusory. Now if the self 
were on the other hand, contra Buddhism, to be real, then it would simply mean that the 
sense of self is none other than self sensed.\(^\text{17}\) In this case, both the self and the sense of self 
would be real (for example, non-illusory). Distinguishing the self from the sense of self, then, 
gives us logical room to investigate whether a possibly ubiquitous sense of self is explained 
by, or amounts to, an actual self – as it would subjectively seem to – or whether, as 
Buddhism maintains, it is in fact explained by, and hence amounts to, cognitive conjurers 
that trick us into believing we are such a self.

\(^{16}\) The notion of ‘real’, in context of the self, will be properly analysed in chapter five. 
\(^{17}\) Obviously this would render the self a different kind of entity from most items, where in most cases, it is 
clear that to have a sense of an item (for example, a sense of danger) will not convey the impression that the 
sense of that item is none other than the item sensed. The self is exceptional in that it does not 
phenomenologically seem different from the sense of self – a claim that shall be argued for in chapter five. For 
now we can accept it provisionally.
It is helpful to give a parallel example that may be more familiar to some. Let us suppose that hard determinism is correct and that there is no such thing as free-will. That is, we are supposing that it is not the case that given a situation where we seem to exercise agency, we could have chosen to do otherwise. Every action is fully determined by factors of which none pertain to an agent's freedom to act otherwise. Free-will does not exist. Yet it is surely possible, and indeed likely, that many people harbour a deep-seated sense/belief/assumption/feeling that given an identical situation, they could have chosen to act otherwise. This assumption of being a free agent, of having free-will, is very real – despite the fact that free-will does not, on this scenario, exist. So while the sense or assumption of free-will exists, free-will does not exist: the deep-seated assumption turns out to be a mistaken one. The hard determinists will attempt to explain the common belief in free-will not in terms of free-will – which would subjectively seem to explain it – but in terms of cognitive and psychological factors that do not include free-will. Similarly, while the sense of self may exist (as it does on our above example) the self whose reality is assumed may not exist. If it does not exist, then the sense of self's reality will have to be explained by factors that do not include reference to the self in their ontology.

What is Meant by 'Sense' in 'Sense of Self'?

Now one may wonder at the choice of terms used to describe this deep subjective allegiance to the self's existence. While I have chosen the term 'sense' to be primary, my usage of other terms such as 'belief', 'assumption' and 'feeling' is meant to convey that the term 'sense' in this context is more complex than in some other contexts. The reason for allocating the word 'sense' as primary is that the turn of phrase 'sense of self' is already in vogue and while lacking ideal precision, it captures the general gist very well. What, then, do we mean by 'sense' in this context? Well, let us distinguish it first from that associated with the five sensory organs, as put by the Merriam-Webster online dictionary (2004): "specialized animal function or mechanism (as sight, hearing, smell, taste, or touch) basically involving a stimulus and a sense organ". This is not the kind of sense alluded to in 'sense of self'. But the same dictionary offers another definition that is more to the point: "A definite but often vague awareness or impression <felt a sense of insecurity> <a sense of danger>". One immediate thing to come across from this definition is that 'sense' in this context is not a
success-term; to have a sense of X does not imply that X exists. For example if one has a sense of danger, then there need not be danger that is sensed. ‘Sense’ in this context is thus more akin to a subjective impression than a perception, which well suits the case of the self whose existence is in question.

As a kind of “a definite but often vague awareness or impression”, the term ‘sense’ as applied to ‘self’ has a further advantage. It manages to convey a subjective experience; that there is, in Nagel’s famous phrase, *something it is like*, from the first-person perspective, to have or undergo a general conscious impression of X. It hence rules out non-conscious machine-sensing that might be ascribed to a thermostat. And it invites the question: just what does comprise this impression – or sense – when it seems to indicate an X such as a self, danger, happiness or free-will? In the case of the self, we have already mentioned and questioned the most obvious option: that it is the *self* that comprises the sense of self. Given that ‘sense’ is not a success-term, however, it could also be the case that other cognitive factors contribute to it instead, factors which perhaps include what we have already designated as ‘the subject’, viz., a witnessing perspective that appears psycho-physically restricted. The virtue of this notion of ‘sense’, then, is that while conveying the idea of a first-person experience, it does not commit this experience to a particular genealogy but leaves the question open to inquiry.

Now we’ve also used the term ‘belief’ to portray something significant about the sense of self. This term indicates that the sense of self conveys not only a subjective impression, but a propositional attitude with truth-bearing content that is, perhaps implicitly, assented to by the subject. The content of this belief, *<there is a self>* , can thus be true or false, depending on whether there *is* such a self. If the self does not exist, then we can say that the assented-to belief is *false* ; if the self does exist, then it will be described as a *true* belief. Now although we will incorporate this doxastic dimension into what is meant by ‘sense of self’, it is misleading to refer to one’s allegiance to the self’s existence as merely a *belief that* the self exists or even a *belief in* the self’s existence. ‘Belief that’ carries connotations of (a) involving content that forms an obvious object (from the subject’s perspective) in relation to the subject holding the belief, for example: “I believe that it is sunny today” and (b) being easily revisable if objectively proved false. Should it turn out to rain, then I can quickly revise my belief to match the newly discovered fact. ‘Beliefs in X’
are more difficult to revise, since they often involve a greater element of faith, along with a vested interest in the existence of believed-in items such as God or a friend’s integrity. Still, they carry connotations that the believed-in item forms an obvious object (from the subject’s perspective) in relation to the subject who holds the belief. The belief that pertains to a ‘sense of self’, as it will be used in this thesis, is far closer to home. The entity that is believed to exist does not obviously seem, from the subject’s perspective, to be a separate object from the subject holding the belief, but seems, rather, to be integral to the very subject holding it. In this way, the belief in the self’s existence can be thought of as a reflexive belief, seeming, from the first-person perspective, to issue from and qualify the subject that harbours it. When we harbour a sense of self, then, the entity that we believe that we are is, minimally, the subject that harbours that belief. Being this axiomatic, it will not be the kind of belief that is easily revised if objectively proved false.

With the sense of self there is hence, from the subject’s perspective, no obviously felt gap between the believed-in item – the self – and the subject who holds the belief. (‘Felt’ in this context is another way of conveying ‘sense’, viz., a general conscious impression). The subject believes, in other words, that they are a self. Importantly, the subject believes that they are a self not by subscribing to some abstract intellectual treatise about the existence of a self – which would once again forge a felt gap between belief-holder and the state of affairs believed – but through deeply assuming the self’s existence. Hence my inclusion of the word ‘assumption’, suggesting something that is ‘lived’ and taken for granted. The nature or existence of such an assumption, like mastered grammatical rules, need not be obvious qua that assumption to the subject who harbours it. Indeed, since a widespread assumption of self will, if the assumption exists, colour the very way that we as subjects approach and think about the world, it will be difficult to gain enough cognitive distance to introspectively discern its qualities. The assumed self will not after all come across as an observable object. We shall have to resort to other methods to get a picture of this self that we (may) have a sense of being – a task to which chapter four is dedicated. Like an assumption of free-will, any subscription to the self’s existence will be implicated in the very way that we approach the world; in our patterns of behaviour, thoughts, desires, emotions and motivations. Someone who assumes she is a freely-willed agent, for instance, need not reflectively believe that she is free or that she even harbours a belief of this kind. Yet her belief qua assumption
can be betrayed through becoming racked with guilt at making a bad decision. Unless she assumed she could have chosen otherwise, she would not, arguably, feel guilty. Similarly, someone who harbours a belief that they are a self – (let us say for now) an enduring thinking entity – need not reflectively know that they identify as such an entity. The assumption can shine forth in a display of nerves before giving a speech. The nerves suggest an implicit belief on the part of the apprehensive subject; a reflexive belief that the very same – and hence enduring – entity will give the speech and possibly make a fool of himself. It is a belief to which he need not, and probably will not, overtly subscribe.

The Buddhist Perspective on the Sense of Self

I have mentioned that from a Buddhist perspective, the distinction between self and sense of self will be a crucial one to make. From our reflections on the kind of thing that the sense of self is in relation to the self, a matter of further importance emerges. It is that the notion of self that concerns Buddhism will not be an obscure metaphysical concept that is the exclusive domain of some intellectuals educated in traditional Western ideas. Since a fear of death or non-existence is so common, an immortal soul-entity for instance is unlikely to be the kind of thing that most people could have a sense of being. Buddhism purports to be a highly practical system of thought: the self that it appears primarily concerned with is depicted as an entity we commonly assume ourselves to be. The notion of self alluded to in Buddhism will hence aim to depict a very ordinary kind of thing: an entity whose existence is reflexively subscribed to by almost every Joe Bloggs on the planet. In chapter three, we will attempt to discern from Buddhist writings just what this supposed self amounts to; in chapter four we will aim to show that this kind of self-entity is indeed something that most people – Buddhist or not – will identify themselves as being. In other words, chapter four will be devoted to arguing that most people do indeed have a sense of self of a description very similar to that alluded to in Buddhism. In chapter five, we discuss what it means to deny existence to this self that we arguably have a sense of being.

Defining ‘Person’

In a context where the existence of an ordinary self is being questioned, we need a neutral non-question-begging term that enables us to refer to people or individuals in a manner that
does not presuppose anything about selves or their sense. While our term ‘subject’ can often be used for this end, it has its limitations. Even in its usual mode of being a psycho-physically restricted witnessing perspective, we noted before that a subject does not, qua subject, seem directly observable from either the first or third-person perspectives. Now, when giving a story about the sense of self’s ontology, the notion of subject will be highly relevant, perhaps pertaining to a component of the self, for instance. However, it will also be useful to have a term which designates something that is, on the whole, observable from both third-person and first-person perspectives; that is largely an object in other words. The term ‘person’ will be reserved for this purpose. ‘Person’ will thus, in this thesis, be a term of art that designates a mostly observable psycho-physical entity but with the capacity for subjective witnessing along with an array of sensory and cognitive faculties. While a person need not harbour a sense of self, our analysis will suggest that most persons of the human variety do. The term ‘person’ will remain neutral on just how the psycho-physical components come together; whether for instance, they can be composed around a genuine self-entity (a “brain-pearl” as Dennett 1991 puts it) or whether they amount to only a bundle of perduring psycho-physical processes that contributes a perspective to the witnessing. We shall see that Buddhism holds a variant of the latter position: a person will be identifiable in virtue of law-like relations that hold between tokenings of the five aggregates or khandhas. It is also important to note that the term ‘person’, as we are using it here, may have little to do with what makes someone a ‘person’ in the ethically relevant sense. It will not, for example, build in such restrictions as a capacity for reflective self-awareness or intelligence or, of course, a sense of self. While my discussion will allude to humans as the prototypical person, there is no reason why the arguments should not, in some cases, also apply to non-human animals who are subjects of experience (such as chimpanzees, cats, rats and fish) and who may even harbour a sense of self in the way that I define it. The main point of relevance is that a person be a mostly observable functioning psycho-physical unit with a witnessing perspective on the world, whether or not they are human or have a sense of self or indeed whether or not persons incorporate selves, that term having the

---

18 I say mostly observable because a person incorporates a subject which seems unobservable.

19 Daniel Dennett, *Consciousness Explained* (1991, 430) is not a fan of the selfy-homunculus or ‘brain-pearl’.
kind of metaphysical baggage which has been attached to it by philosophers like Descartes et al.

**The Essence of Buddhist Teaching: The Four Noble Truths**

We shall now outline the core ideas in Buddhism known as “The Four Noble Truths”. This outline – which will allude to such phenomena as *kamma* and rebirth (to be described) – may seem out of place in an investigation of the kind of issues I will be examining. However, it will provide what I believe to be a necessary background against which to understand my interpretation of the Buddhist position on consciousness and no-self in chapters two and three. Some of the concepts introduced in chapters two and three (of which a few recur in later chapters) are quite novel to Western philosophy and are in danger of being misunderstood without some familiarity with the context in which they occur.

I have mentioned that Buddhism is an essentially practical body of teaching. By this, I mean that the focus of Buddhist teaching is not on expounding complex theoretical doctrine for the sake of it, but on eliminating suffering and dissatisfaction. Any elaborate-sounding cosmological or ontological claim, such as that pertaining to the conditioned co-dependence of objects, or to *kamma*, rebirth and the unconditioned *nibbana* (terms that will be explained) is only there in context of the Buddha’s practically oriented teaching on how to eliminate suffering and dissatisfaction. This includes, no less, the Buddha’s teachings on no-self (*anatta*). The essence and emphasis of the Buddha’s teachings on suffering etc., is captured in the following sutta concerning the Four Noble Truths:

---

20 I use the turn of phrase “what the Buddha said” or “the Buddha’s teachings” when explicating what the historical Buddha was alleged to have taught. By this I don’t mean to take a stance on whether he literally said those things – it is just easier to follow this convention. I will also refer to the individual Noble Truths with capital letters to indicate that they are known by that name in Buddhism, rather than to presuppose that they are indeed true.

21 Suttas, drawn from the Pali Canon, are the closest we have to original teachings by the historical Buddha. Many of the Pali words have similar-sounding Sanskrit translations, e.g., *nibbana* (in Pali) is *nirvana* (in Sanskrit); *kamma* (in Pali) is *karma* (in Sanskrit).

22 Since my familiarity with the Four Noble Truths stems mainly from attending or listening to tapes of ‘*dhamma*’ talks (over several years) by a variety of senior monastics in the Buddhist tradition, I will not refer to secondary texts to endorse each exegetical point unless the context demands it. My general impression has been that senior monastics, being practitioners who try to live their lives according to the Four Noble Truths, are often the best authorities on expounding it (which is not to say that there are not differences between their interpretations of the Truths, some of which I may not agree with). Readers wishing to gain another overview on the Four Noble Truths can consult Buddhist sources such as Gethin (1998, 59-84), Harvey (1990, 47-72) and Venerable Ajahn Sumedho (1992).
That both I and you have had to travel and trudge through this long round is owing to our not discovering, not penetrating four truths. What four? They are: The Noble Truth of Suffering [Dukkha], The Noble Truth of the Origin of Suffering [Dukkha], The Noble Truth of the Cessation of Suffering [Dukkha], and The Noble Truth of the Way Leading to the Cessation of Suffering [Dukkha]. [From the Digha Nikaya, Sutta 16, quoted in Venerable Ajahn Sumedho, (1992, 8)].

Expanded further, the Four Noble Truths – using key Pali terms – are as follows:

1) *Dukkha* (approx. suffering, stress, unsatisfactoriness, unfulfilment) exists.
2) The origin of *dukkha* is *tanha* (approx. craving, thirst, attachment, emotional investment).
3) The cessation of *dukkha* lies in the cessation of *tanha*. The cessation of *dukkha* involves *nibbana*.
4) There is a path to the cessation of *dukkha* (and hence to *nibbana*): the Noble Eightfold Path, which involves the practice of insight-wisdom (*panna*), meditation (*samadhi*) and virtue (*sila*).

Throughout this project, I will primarily use the Pali term ‘*dukkha*’ (rhyming with ‘book-ah’) for what has most commonly been translated as ‘suffering’ and the Pali term ‘*tanha*’ for what has most commonly been translated as ‘craving’. While ‘*nibbana*’ has no common translation, it has sometimes been alluded to as ‘enlightenment’ or ‘awakening’. The reason for preserving these Pali terms is that there is no single-word English equivalent that captures their meaning. While my preference is to avoid non-English terminology when possible, earlier drafts replete with the words ‘suffering’ and ‘craving’ were misunderstood by its reader, John Baker – and for good reason. The English words turned out, on reflection, to be too crude and extreme to properly convey the Buddhist teachings. For instance, Baker rightly remarked that ‘craving’ in modern English carries connotations of rabid nicotine withdrawal symptoms! In outlining the Four Noble Truths, we shall see that the Pali terms are rather more subtle and layered than this. Note that during this outline, the same remarks will apply as to when I outlined the Buddhist notion of conditioned co-dependency: we can assume that

---

23 The reader should not get the impression that I am a Pali scholar! My remarks on the common translation of these select words stem from familiarity with the Buddhist teachings, together with John Baker’s feedback upon my initial employment of these translations.
the Four Noble Truths are being *described* rather than *defended*. It is important to keep this in mind since, when the context demands it, wider Buddhist cosmology will be invoked.

**The First Noble Truth**, then, is that *dukkha* exists. *Dukkha* alludes to the *khandha* of unpleasant feeling, whether mental or physical, overt or covert. On a physical level, it includes all overt suffering such as that associated with sickness, injury and nicotine withdrawal. On a mental level, it includes any overt suffering from pangs of guilt and disappointment, to major depression, to not wanting a physical pain that is present, to craving a cigarette that is not present. But as well as that, *dukkha* pertains to the more covert unpleasantness that can occur even while one is “having a good time”. An example of this is needing to go to the bathroom whilst listening to an amazing concert. Even if the discomfort is not attended to but only peripherally witnessed, it is still *dukkha*, a *khandha* of unpleasant feeling. There is also a kind of *dukkha* referred to in Buddhism as the *dukkha* of change or impermanence. Has anyone ever noticed a “this won’t last” thread of anxiety running through the back of a wonderful experience? This feeling of unease or anxiety can mar, even if just slightly, an enjoyable time. A further example of *dukkha* during pleasure is the feeling that although one is enjoying an experience, one is not, somehow, gaining the amount of enjoyment that one feels one *ought* to be gaining from it. This can typically happen when visiting a place of ‘breathtaking’ scenery, such as the Rocky Mountains. I once heard a Buddhist monk remark (in a recorded talk) that looking at the Rockies was almost painful, as if his mind could not get its rightful fill of hedons! This *dukkha* is a feeling of unfulfilment, a sensing that one is not as full of happiness and joy as one might be. A related but more active kind of *dukkha*, also very common, is what I shall term ‘doer-*dukkha*’. Doer-*dukkha* pertains to anxiety at the thought of not getting the best possible outcome from one’s actions. An example of this is the student who, while listening to a conference paper, suddenly thinks of a ‘brilliant’ question. If only he could ask it, everyone in the room will see how intelligent he is! But the joyful anticipation is also veined with anxiety: will it after all be a brilliant question or could he have misunderstood the speaker? This is doer-*dukkha*. *Dukkha* thus pertains to not only the out-and-out suffering of an experience that is on the whole mentally
or physically painful, but to the threads of unease, physical pain, dissatisfaction, anxiety or unfulfilment that can occur in an experience that is on the whole enjoyable\textsuperscript{24}.

As well as pertaining directly to physical and mental states of persons, 'dukkha' is commonly used as a term in Buddhism to apply to all of conditioned existence, hence tokenings of the five khandhas: physical form, feeling, perception, mental formations and object-oriented consciousness. But why should the world and all its objects be depicted as unsatisfactory or suffering? My computer sure gives me dukkha, but how can it be dukkha? Indeed, how can the khandha of pleasant feeling be dukkha? I interpret this as not an intrinsic, but a relational ascription, which means that any object, if clung to as a source of happiness, has the potential to elicit in that person dukkha. Put another way, it means that no object has the potential to elicit in someone a level of satisfaction or fulfilment that is free from potential dukkha. In this way, all conditioned objects are deemed in Buddhism to be unsatisfactory as a source of dukkha-free happiness. To see what is meant by this, we need to turn to the Second Noble Truth.

The Second Noble Truth states that the origin of dukkha is tanha. Basically, 'tanha' conveys an emotional attachment to objects being a particular way, such that (a) one’s happiness or absence of mental dukkha seems to depend upon a favoured configuration of objects (whether imagined or actual) such that (b) if they did not configure the way one would like, then there would be mental dukkha of a particular emotional nuance (for example, that of anger, sadness, anxiety, disappointment or guilt). We first discuss tanha in its relation to mental dukkha, before turning to tanha in relation to physical dukkha.

In its relation to mental dukkha, the presence of attachment or tanha may not be obvious to the person who harbours it, which is why the term ‘craving’, implying an intense occurrent desire, has inappropriate connotations. For instance, someone who is attached to the good life and is getting mostly what he desires – a nice house, good health, car, job and girlfriend – may not overtly feel the tug of his emotional investment in these things. Yet if any of these items were to be lost or threatened, resulting in his mental suffering (through

\textsuperscript{24}It may be pointed out that sometimes threads of pain or suffering can actually enhance the overall hedonic tone of an experience, as the pain of a mountain-climber might add to his overall elation at having climbed the mountain. This is not a problem with our account of dukkha, however, since the Buddhist is not committed to denying that dukkha can play this kind of role.
negative emotion), his attachment to them would suddenly become obvious. He would not necessarily become any more attached to the items than he was before – it is just that the changing circumstances would reveal his existing attachments or tanha towards them, showing that he had let his happiness depend on them. The Second Noble Truth asserts that if mental dukkha such as this is to arise, there has to be the presence of tanha. If the man were not, in the first place, attached to his house, girlfriend, job etc – if he lacked tanha or emotional investment with reference to these items – then he would not, according to Buddhism, feel any negative emotion at their loss.

Buddhism mentions three classes of item or situation towards which one may harbour tanha: ‘objects of sense-desire’, ‘existence’ and ‘non-existence’. To have tanha for objects of sense desire is to be attached to the khandha of pleasant feeling that may arise from objects associated with the five sense-organs or the mind. It involves emotional investment in such things as food (most especially, the pleasant taste sensations), a satisfying career (the pleasant sensations associated with mental fulfilment), and a relationship (nice emotional sensations of romance). Tanha with respect to pleasant sensory objects can lead to dukkha if the desired object is threatened, denied or withdrawn. The reason that objects of sense desire are sometimes classed as being ‘dukkha’, even though they feel pleasant, is that their conditioned co-dependence means that they are not immune to change and decay. To emotionally invest one’s happiness in these impermanent objects is to hence guarantee the possibility of dukkha. For chances are that either the object of pleasant sensation will not be acquired (leading perhaps to anger or disappointment) or it will change or disappear (leading perhaps to grief) or the pleasant sensations will no longer arise in connection with the associated objects: one may get bored, for instance. No conditioned thing, as an object of tanha, is exempt from the possibility of eliciting dukkha through change, no matter how pleasant their sensations at a particular time. This renders them unsatisfactory as a source of happiness that is completely free of potential dukkha.

To have tanha for ‘existence’ means, in Buddhist parlance, that one emotionally invests in being a particular kind of person, for example, as a person who is successful and accepted by one’s peers. The student who burns to ask a conference-question out of a hope to be seen as intelligent is acting out of tanha for existence. Tanha in this context feeds on praise, fortune, love and fame, nurturing what is colloquially known as an ego. Tanha for
existence is not exempt from the possibility of causing dukkha, since the changing world will
guarantee that no-one, no matter how intelligent, is immune from rejection, disappointment
or making a fool of oneself. Doer-dukkha is often motivated by tanha for existence.

On the Buddhist picture, tanha for ‘non-existence’ arises in connection with being
rejected too often or making a fool of oneself: one yearns to not exist. When the student’s
question backfires and he wishes that “the earth would open up and swallow him” there is
temporary tanha for non-existence – obviously implying dukkha. In the most extreme cases,
this can manifest as an urge for suicide, while less extreme examples can show through in
patterns of self-criticism (e.g. “I am not worthy of this praise”). A more general tanha for
non-existence can pertain to not only one’s ‘self’, but to the khandha of unpleasant feeling
associated with sensory or mental objects. Just as one can emotionally invest in the presence
of pleasant sensations, one may emotionally invest in the absence of unpleasant sensations.
To do the latter is to risk suffering the dukkha of wishing an existing situation were
otherwise, or that of hoping that an unpleasant situation will not arise. Again, Buddhism
claims that while certain objects are prone to elicit dukkha, there must also be a vested
emotional interest in their non-existence for that dukkha to have a mental dimension. Hence
tanha is required if mental dukkha is to occur. If one did not care at all about the presence of
any situation, including physical pain, but were perfectly at ease with whatever happened,
without tanha for its non-existence, then there could not be that layer of negative emotion
which arises with wishing that a situation were otherwise.

The Buddhist claim, then, is that tanha is necessary for the arising of dukkha. We
have so far mentioned, with some plausibility, how tanha could be at the root of mental
dukkha. But on the Buddhist account, tanha is also implicated in physical dukkha. Suppose
that I bang my ankle on the furniture, causing intense physical pain. While I might be blamed
for being careless, where is the emotional investment in this story? What can it mean for
tanha to be deemed necessary for the pain of physical injury, disease, old-age, sickness and
death, once we factor out all the mental suffering? It is here that we must refer to the
Buddhist doctrine of kamma (sanskrit: karma), rebirth and samsara. Since the idea here is to
obtain a general picture (also necessary for understanding the Third Noble Truth) I will not
elaborate upon these concepts beyond what is necessary for such a picture.
Basically, Buddhist cosmology maintains that the moral quality of a person’s action (where actions can include thoughts) is determined by the amount of *tanha* which helps motivate that action. So long as *tanha* is at all present in a person’s mindset, the person’s thoughts and actions which spring from that mindset will be *kammically potent*\(^{25}\). A *kammically* potent action is one which generates future consequences for that person (where a person is analysed in Buddhism as being constituted by witnessing together with tokenings the five causally conditioned *khandhas*). The less *tanha* in the mindset, the better the kammic consequences (‘better’ meaning a yield of pleasant mental or physical sensations), the more *tanha* in the mindset, the worse the consequences (‘worse’ meaning a yield of unpleasant mental or physical sensations). Even when a person dies, they cannot escape the kammic consequences of their former actions – where actions include thoughts on the deathbed that spring from a mind with *tanha* – which leads to rebirth into another body (which is actually a reconfiguration of *khandhas*)\(^{26}\). The overall conditions of a person’s rebirth – whether favourable or unfavourable – will depend on how much wholesome or unwholesome *kamma* has been accrued. But even in a ‘favourable’ rebirth (such as affluent circumstance), to simply have a body and mind in a physical world will involve physical *dukkha* at some time or other, such as that from banging one’s ankle on the furniture. This baseline of physical *dukkha* is thus largely due to the *kamma* of simply having a physical body, and that *kamma* is the result of action springing from *tanha* of a previous lifetime. And so long as one dies with a mindset harbouring *tanha*, as most of us will according to Buddhism, physical *dukkha* will arise in the next birth. Which raises the question: is it possible to escape from all *dukkha* and indeed, from *samsara*, the potentially endless cycle of birth, death and *dukkha*? This pertains to the Third Noble Truth.

Before moving onto this, I will note that it is mostly *mental dukkha* in its relation to *tanha* that will be of relevance to this project. Aside from exegetical purposes, none of my

\(^{25}\) In the spectrum of ethical theories, Buddhism fits most closely with virtue ethics. The moral value of an action depends upon the state of mind (virtuous or vicious) from which the action springs.

\(^{26}\) Note that on the Buddhist position, it is not some essential enduring soul-substance is reborn into another body. While a person is living (in the conventional sense) their ‘identity’ is largely preserved by a process of rapid-fire *khandha*-tokenings conditioning each other from one moment to the next (similar to how Hume (1739) described mental contents as rapidly arising and passing away, one lot influencing the next). When the physical body dies, the four remaining mental *khandha*-tokens continue to condition each other (in accordance with the law of *kamma*) causally influencing the particulars of a new physical body (viz., *khandha*-tokenings of the type *form*) that they come to be associated with.
arguments will depend upon the adoption of any controversial Buddhist cosmology. The particular interest to be pursued with reference to the Second Noble Truth is the implied relation posited between tanha and a sense of self – related to the teachings on anatta (no-self). For as we shall see in chapter three, Buddhism regards tanha to be not only the main cause of dukkha, but to co-arise in the mind with a sense of self; a point that will be discussed in some detail. A sense of self will hence also, by implication, be at the root of dukkha. The notion of tanha in relation the sense of self will play a major role in later arguments.

The Third Noble Truth states that the cessation of dukkha lies in the cessation of tanha. The cessation of dukkha involves nibbana (which we shall see is not to be construed as out-and-out annihilation). Given that tanha is at the root of dukkha, the Third Noble Truth functions like a modus tollens on the Second Noble Truth: root out tanha, and dukkha will disappear. In a Western context, this will generally be considered a novel approach to eliminating dukkha. Usually, the thing we seek to do to avoid suffering is to change the external conditions that we perceive to be causing the suffering or discomfort. Buddhism does not deny that this can serve as a temporary expedient to stop dukkha – but it is only temporary. So long as our mindset harbours tanha – emotional investment in things being a particular way – then sooner or later the conditioned cards will fall in a way that is out of line with our preference, resulting, once again, in dukkha. On the Buddhist picture, the only way to get rid of dukkha once and for all, so that it stops arising altogether, is to get rid of tanha, the mindset needed for dukkha to arise.

The elimination of tanha and hence dukkha is considered by Buddhism to be a real psychological possibility, forming the ultimate goal of Buddhist practice. The term used to

---

27 The Stoics seem a notable exception, however. Of the Stoics, William James wrote: "The first care of diplomatists and monarchs and all who wish to rule or influence is, accordingly, to find out their victim's strongest principle of self-regard, so as to make that the fulcrum of all appeals. But if a man has given up those things which are subject to foreign fate, and ceased to regard them as parts of himself at all, we are powerless over him. The Stoic receipt for contentment was to dispossess yourself in advance of all that was out of your own power, - then fortune's shocks might rain down unfelt. Epictetus exhorts us, by thus narrowing and at the same time solidifying our Self to make it invulnerable: 'I must die; well, but must I die groaning too? I will speak what appears to be right, and if the despot says, then I will put you to death, I will reply, 'When did I ever tell you that I was immortal? You will do your part and I mine; it is yours to kill and mine to die intrepid; yours to banish, mine to depart untroubled.' How do we act in a voyage? We choose the pilot, the sailors, the hour. Afterwards comes a storm. What have I to care for? My part is performed. This matter belongs to the pilot. But the ship is sinking; what then have I to do? That which alone I can do - submit to being drowned without fear, without clamor or accusing of God, but as one who knows that what is born must likewise die.'"(1890, 311-312).
designate this goal—the cessation of *tanha* and *dukkha*—is ‘*nibbana*’—a term, that we shall see, has more than one nuance. A person who attains *nibbana* is termed an *Arahant*. Now Buddhists distinguish between *nibbana* while the *Arahant* is living, and *nibbana* when the *Arahant* dies, a distinction that directly bears upon the difference between mental and physical *dukkha*. While the *Arahant* is alive, the elimination of *tanha* only involves the cessation of mental *dukkha*. There is still physical *dukkha* for the living *Arahant*, due partly to the *kamma* of past *tanha* which earned him or her a physical rebirth. Lacking *tanha*, hence any emotional aversion to the physical *dukkha*, the *Arahant* cannot harbour any mental *dukkha*—that of wishing the situation were otherwise. He or she may hence feel the most excruciating physical pain, but will not mind a bit, meaning, that they would become no happier (or less happy) if the pain were to cease. This is what it means to harbour no emotional investment or *tanha* in any states of affairs. I am sure most people will agree that such a state, were it to be possible, would be quite extraordinary. Now an *Arahant*, lacking *tanha*, will generate no more *kamma* and hence, no conditions for future rebirth into the cycle of *samsara*. So when the *Arahant* dies—a event referred to as *parinibbana*—there will be no continued and conditioned flux of *khandhas* into another psycho-physical existence. No more perspective will be generated for the witnessing (should there be witnessing): there will be no more subject in relation to objects. With no rebirth, there can be no *dukkha*, mental or physical. In this context, we can appreciate the meaning of the Third Noble Truth: the complete cessation of *dukkha* (mental and physical) must require the cessation of *tanha*.

I have mentioned that the term ‘*nibbana*’ has more than one nuance. In the discussion above, it pertains to the mind where *tanha* and (mental and/or physical) *dukkha* has ceased or else to the event of *tanha* and mental *dukkha* ceasing. But ‘*nibbana*’ does not merely denote the cessation of *tanha* and *dukkha*: it denotes something positive as well, since, as we shall see in chapter two, it has something to do with a *mind* that has been liberated from *tanha*. The liberated mind of the *Arahant* is the positive thing that ‘*nibbana*’ denotes: it is what Buddhism refers to as ‘unconditioned’. So far, we have just talked about how Buddhism regards the conditioned world, pertaining to the five kinds of *khandha*. The conditioned and

---

28 When it comes to explaining *nibbana*, I have been greatly helped by two secondary sources: Thanissaro Bhikkhu’s *The Mind Like Fire Unbound* (1993a) — which contains many primary sources from the suttas — and Harvey’s *The Selfless Mind* (1995). The details of *nibbana* will be elaborated in chapter two.

29 In chapter two we provide evidence from the suttas to suggest that there would be witnessing for the *Arahant* who dies.
observable world of khandha-objects, with their co-dependency, is limited by the
aforementioned modes of space, time, quality and relation. But Buddhist cosmology also
alludes to that which is unconditioned by these modes of existence and is not observable as
an object of witnessing (such as that pertaining to perception, thought, or introspection). The
mind of the Arahant is deemed unconditioned and unobservable in this manner. Given that
the modus operandi of a subject, namely witnessing, does not seem observable either (as an
object of attention) we may wonder if witnessing per se is deemed relevant to the Arahant’s
liberated mind. In chapters two and three, we shall consult Buddhist suttas which suggest that
this is most likely to be the case. Moreover, we shall see that there is reportedly something it
is like for the Arahant to experience a liberated and unconditioned mind: it is not as if the
goal of Buddhist practice is to become a zombie. This suggests that the witnessing dimension
of nibbana – which I shall term ‘nibbanic consciousness’ – has its own phenomenal feel that
is not derived from any observable objects. The ‘something it is like’ tone of nibbanic
consciousness is said, in fact, to involve the most immense peace and happiness that can ever
be felt, completely untainted by the presence or possibility of mental dukkha.

The relevance of the Third Noble Truth to this project’s inquiry – insofar as it seeks
to elucidate the Buddhist position – lies in the relation between what I term ‘nibbanic
consciousness’ and the Buddhist position on the self and sense of self. I shall suggest in
chapters two and three that one cannot get a proper picture of how Buddhism regards the
ontology of (no)self – and the sense of self – unless one also gets a picture of what is meant
by nibbana, in view of how nibbanic consciousness could relate to ordinary witnessing. For
once one acknowledges that the witnessing dimension to nibbanic consciousness must be
unconditioned by time, then this immediately suggests that the witnessing dimension to
nibbanic consciousness cannot be something that suddenly pops into existence when a person
becomes an Arahant. The witnessing must – somehow – be present in a way that is
unaffected by time. This invites a number of questions: could the timeless witnessing of
nibbanic consciousness be somehow contributing – in collaboration with tanha – to the
ordinary sense of self prior to the attainment of nibbana qua event? Could it in fact be the
very witnessing that is present in ordinary conscious states? Could this ever-present
witnessing be instrumental to the mind’s liberation from the co-arising tanha and sense of
self – as required for the event of nibbana? Radical as it may sound, I shall present the
Buddhist position on consciousness and no-self as answering a resounding “yes” to these questions.

I say ‘radical’ not only because the position outlined is radical by Western standards (I know of no Western philosopher who defends it) but because it radically departs from how the Buddhist position on anatta or no-self is usually presented. The usual scholarly tactic on no-self has been to ignore nibbanic consciousness, concentrating only on what the Buddha allegedly taught about the self and its denied reality. While not disagreeing that Buddhism denies the self reality, the story does not stop there. By stopping the story there, many scholars convey the impression that on the Buddhist position, the sense of self is only contributed to by conditioned reality (none of which contains a self). While Buddhism does indeed suggest that (a) a self fails to underpin the sense of self and (b) conditioned khandhas contribute to the sense of self, it also suggests that (c) unconditioned reality, that of nibbanic consciousness, is crucial for a sense of self – and indeed for the mental khandhas – to arise. When (c) is excluded from the ‘Buddhist’ analysis of self, as it usually is, the resultant picture becomes curiously nihilistic – quite out of keeping, I shall suggest, with the Third Noble Truth. Subsequently – as I argue in chapters two and three – it is often a distorted picture of Buddhist ‘no-self’ and consciousness that makes its way into the Western academic arena. The aim of chapters two and three is to put this picture right. Not until the picture is clear can it be properly defended. While arguing for the reality of nibbanic consciousness is too ambitious for this project, I do aim to show that a metaphysically neutral variant of the Buddhist position on sense of self, no-self and witness-consciousness is highly plausible.

The Fourth Noble Truth states: there is a path to the cessation of dukkha (and hence to nibbana): the Noble Eightfold Path, which involves the practice of ‘insight-wisdom’ (panna), ‘meditation’ (samadhi) and ‘virtue’ (sila). Being a practical tradition, Buddhism does not make lofty reference to a mode of flawlessly happy consciousness without prescribing a system of training that will purportedly effect this transformation. Now this project will refer to the Buddhist transformative goal in terms of cessation to both the sense of self and tanha. Describing the goal of Buddhist practice in this way will set the stage for talking about the practice in a manner that is directed and precise. One can be then in a position to describe the
practice in terms of dismantling the sense of self as well as ‘uncovering’ nibbanic consciousness. Of course the sense of self in its relation to tanha and witness-consciousness will have to be accounted for before we can do this, a task which – after space-constraints forced me to edit out the chapter on this dismantling process – turns out to occupy the whole thesis. Not until one has said enough about how the sense of self could be ‘put together’, can one suggest, without sounding vague, how it could possibly be taken apart. The methods alluded to in the Fourth Noble Truth – ‘meditation’, ‘insight-wisdom’ and ‘virtue’ – are methods that Buddhism deems relevant to the dismantling of the sense of self, a process which, philosophically speaking, is very interesting. I mention the Fourth Noble Truth so as to indicate how the role of such practice would relate to the central Buddhist schema, and to give some idea of the scope of this project (which can be viewed as setting the stage for this most interesting inquiry).

We now have sufficient background against which to understand and interpret the Buddhist position on consciousness and no-self. To this we now turn.
Chapter Two: The Probable Buddhist Notion of Nibbanic Consciousness

Introduction

The goal of this chapter is to elucidate the Buddhist notion of *nibbana* as a mode of witness-consciousness, which I shall term ‘nibbanic consciousness’. In chapter three I attempt to show how this notion of nibbanic consciousness, amongst other things, plays a central role in understanding the famous Buddhist doctrine of ‘no-self’. It will transpire that the witness-consciousness pertaining to *nibbana* is a metaphysically robust version of the witnessing described in chapter one. As something unconditioned, it is not to be construed as the object-oriented and conditioned ‘consciousness’ from the five *khandhas*, nor is it to be construed (at least intrinsically) as an activity or state of a subject. Gaining a clear idea of what Buddhism has to say about witness-consciousness and no-self is necessary before we can carry out this project’s aim of defending a (metaphysically neutral) variant of the Buddhist position on consciousness and no-self.

Readers may wonder why two chapters are devoted to what might, on the face of it, appear like a simple exegetical task. Why not just quote a few suttas and secondary sources and have done with it? The problem is that, from a scholastic point of view, extracting notions about no-self and nibbanic witness-consciousness from Buddhist suttas is not like reading *Goldilocks and the Three Bears*. This is not because the Buddha was being deliberately obscure, but simply because the practical orientation of his teachings did not always present an ideal format for those interested in its theoretical implications. Careful inference is therefore required on behalf of those who seek such clarifications. Peter Harvey (1995) and Thanissaro Bhikkhu (1993a) have done some fine work in articulating the probable Buddhist notion of *nibbana*. While the purposes of my project incline me to present the notion of *nibbana* from a different angle to these scholars, their work on ‘*nibbana*’ certainly informs my analysis and can be considered as a complement to what I offer. Let me state from the outset that my account of *nibbana* should not be considered as complete: the focus of this project forces me to exclude many important facets (such as its relation to compassion, wisdom and more broadly, the conditioned *khandhas*). In this and the next chapter, I shall mainly be using Thanissaro Bhikkhu’s (1993a), (1993b), (2003) translations
of the relevant Buddhist suttas on *nibbana* and no-self (*anatta*). While not a Pali scholar
myself, his translations strike me as being the most intelligent and least clumsy of those I’ve
encountered. Readers who doubt this are of course free to consult alternative translations (or
the original Pali if they are a scholar).

When it comes to articulating the Buddhist notion of ‘no-self’ (*anatta*) – at least as an
ontological doctrine – matters are less straightforward. The problem, as I see it, has not been
so much in showing that Buddhist cosmology does not support a self, but in adequately
defining the self that Buddhism purportedly rejects. No scholar that I know of has, in my
opinion, provided an adequate definition of the self Buddhism rejects – and the notion of self is *not*
clearly spelt out in the suttas. I hence spend most of chapter three seeking to define the
self that Buddhism is primarily concerned with. I shall suggest that a major reason the notion
of ‘self’ is so often misconstrued is that scholars focus too exclusively on the *anatta* suttas
without considering their probable relation to the *nibbana* suttas in context of the Four Noble
Truths. Once we do take into account the *nibbana* suttas (the reason for interpreting them
first in this current chapter), a very different picture emerges, one which has vital
implications for how Buddhism actually denies the existence of self – and hence for this
project as a whole. Another matter often overlooked is that in the suttas on *anatta*, the
Buddha primarily intended ‘the self’ to be construed not as a mere theoretical postulate, but
as a self that we, theoreticians or not, commonly assume ourselves to be.\(^3^0\) When this is taken
into account, important constraints are placed on how the self is to be defined in Buddhism. It
must be the kind of entity which most people could plausibly have a *sense* of being. Too
often ‘the self’ in Buddhism is portrayed as something similar to *nibbana* – as if we assume
we are eternally happy, for instance.

Let us now turn to this chapter’s current focus: ‘nibbanic consciousness’. Our
analysis of nibbanic consciousness will inform discussion in the following chapter (three).

**Some Preliminary Considerations**

When explaining the Third Noble Truth (in chapter one) it was intimated that
‘*nibbana*’, the goal of Buddhist practice, can be characterised in several ways. There is

\(^3^0\) I say ‘primarily’ because the Buddha also warns of the dangers of dogmatically clinging to various elaborate
theoretical doctrines of self. It will become clear, however, that the Buddha’s main imperative behind the suttas
on *anatta* (no-self) is for people to let go of an ordinary, reflexive sense of self.
nibbana as the *event* of cessation, denoting the key moment when *tanha* is wholly eliminated from a person’s mindset, implying the cessation of (usually just mental) *dukkha*. ‘*Nibbana*’ in this context denotes the transition of a person with *tanha* to a *tanha*-free *Arahant*. ‘*Nibbana*’ can also refer to a mind of a living Arahant in terms of it lacking *tanha* and mental *dukkha*. There are further characterisations of *nibbana*, where reference to an ‘unconditioned mind’ comes into play. I shall henceforth use the term ‘*nibbanic consciousness*’ to refer to the unconditioned mind of the Arahant. We shall see, in this chapter, that nibbanic consciousness is most likely to involve an unconditioned *witness*-consciousness, where the notion of ‘unconditioned’ contrasts with the notion of ‘conditioned’ outlined in the previous chapter. We shall later see (in chapter three) that Buddhism would regard ordinary everyday witnessing (of the kind described in chapter one) to be identical, in essence, to this unconditioned witness-consciousness.

Before consulting the suttas for evidence of such an interpretation, a note on the semantic category of ‘nibbanic consciousness’ is in order. While this will involve jumping a few guns on the characterisation of nibbanic consciousness as timeless and objectless, it should also defuse some misconceptions that might otherwise cloud one’s approach to the notion of nibbanic consciousness. Near the beginning of chapter one, we mentioned that the consciousness Buddhism regards to be metaphysically ultimate – the consciousness we have now termed ‘*nibbanic*’ – cannot, in Buddhism, be considered to be intrinsically a state or activity (of a subject or person), as the ordinary concept of consciousness or witnessing would initially seem to demand. We should not for instance be misled by the convention of referring to a subject or person as “coming to experience nibbanic consciousness at time t” anymore than we should be misled by talk of “the sun rising or setting”. The timeless nibbanic consciousness he or she “partakes in” is not, intrinsically, a mere *state or activity* of his (qua subject’s) mind or consciousness that somehow comes into existence (like a state of anger) when conditions, or the lack of them, are ripe. It is rather the very *nature* of mind that is being unveiled, a mode of unconditioned reality that does not *intrinsically* have anything to do with the mental or physical states (viz., configurations of *khandhas*) which delimit the conditioned parameters of a subject or person.

Any appearance of a state-change in a subject or person from “not partaking” to “partaking” in nibbanic consciousness should be properly analysed as a kind of Cambridge
change, where a change in the conditions external to nibbanic consciousness (for example, the cessation of tanha and awareness-objects) will result in a change of relational predicates that can be ascribed to the term ‘nibbanic consciousness’. Thus we can correctly say that at time \( t \), nibbanic consciousness stood in relation \( x \) to tanha and awareness-objects (where relation \( x \) means tanha and awareness-objects are present) while at \( t_1 \), nibbanic consciousness stands in relation \( y \) to tanha and awareness-objects (where relation \( y \) means tanha and awareness-objects are not present). Only in this extrinsic and relational context can nibbanic consciousness be designated a ‘changeable state’. The occasion of tanha and other khandhas disappearing from the purview of a living ‘subject’s’ witnessing (hence the temporary dissolution of a subject) – or else the khandhas being altogether destroyed upon death of the Arahant (the permanent dissolution of a subject) – can be thought of as analogous to the disappearance of clouds revealing a shining sun that was there all along. The sun is not brought into existence by the clouds’ disappearance at time \( t \), any more than nibbanic consciousness is brought into existence by the disappearance of tanha and other awareness-objects at time \( t \). So to think of the subjectless and objectless nibbanic consciousness as intrinsically a state or activity of the Arahant (qua subject or person) – as if it were just another co-dependent, time-bound khandha like anger – is deeply mistaken. It will transpire that the effect of the khandhas – in particular, those pertaining to tanha – is not to create conditions that may inhibit the manifestation of nibbanic consciousness, but to create conditions that may partially obscure it, as clouds obscure the sun. This will apply, on our Buddhist analysis, not only to the consciousness of Arahants, but to everyday witnessing-consciousness of Joe Bloggs which is ‘covered’ by configurations of the five khandhas (which include tanha) and which convention would label as a ‘state’ or ‘activity’.

Buddhism regards nibbana to involve a mode of consciousness that is unconditioned and yet lived, in that there is something it is like to ‘undergo’ nibbanic consciousness. In the following sections, we will (a) look to some of the suttas that are suggestive of this picture and (b) provide the most probable story of how nibbanic consciousness will relate to ordinary witnessing (as defined in chapter one). With regards to (b), we may have gathered that although the intrinsically unconditioned nibbanic consciousness is not something that comes into being upon its direct unmediated experience (by an Arahant), there is still a clear sense in which its ‘timeless ever-presence’ (a notion that needs unpacking) will not normally be
obvious. If the Buddha had regarded nibbana as being already obvious to us, then what would be the point of the Four Noble Truths? In view of this, I will introduce further terms that are related to the primary concept of the subjectless and objectless 'pure nibbanic consciousness', namely; 'proximate nibbanic consciousness' (nibbanic witnessing by the tanha-free Arahant who occurrently witnesses khandhas) and 'pre-nibbanic consciousness' (nibbanic witnessing by tanha-bound subjects who may occurrently witness khandhas). The difference between these modes of nibbanic witnessing must be understood as extrinsic - to do with the obscuring khandhas in a given conscious moment - rather than intrinsic - to do with the nibbanic witness-consciousness.

I make the usual reminder that the goal here is to explicate rather than defend the reality of nibbanic consciousness and its cognates. While the practice-oriented suttas are not always explicit with the theoretical details, it is hoped that the account offered in this and the next chapter will be, in the most parsimonious way, in keeping with Buddhist cosmology in its relation to the Four Noble Truths. For ease of reference, I shall refer to the suttas with ascending Roman numerals (S.I, S.II...) – their Pali origin, translator and hosting text to be indicated in footnotes. Another point to note is that I cite only that part of the sutta which is relevant to the inquiry. Readers should not assume that each quoted sutta is complete.

(A) Nibbana as Intrinsically Unconditioned, Lived, Objectless Witness-Consciousness

Let us first consider evidence that nibbana involves objectless (and subjectless) witness-consciousness. I believe it is most fitting to interpret these supporting suttas as depicting a mode of objectless and subjectless consciousness as it appears to the first-person perspective of the Arahant, rather than as metaphysical proclamations of the objective or intrinsic non-dependence of consciousness on the khandhas (which would render such consciousness non-physical). By referring to the ‘first-person perspective’ of an Arahant in this context, I do not mean to suggest that the perspective or viewpoint of an Arahant experiencing nibbanic consciousness appears spatio-temporally delimited in the manner that I have specified for the notion of a subject. Rather, I mean to suggest that there is something it

31 I say that those with 'pre-nibbanic consciousness', according to Buddhism, 'may' occurrently witness khandhas since Buddhism considers it possible for non-Arahants (hence people with tanha) to experience episodes of objectless consciousness. Because there is still tanha in their mindsets, this is not considered to be the 'pure' nibbanic consciousness.
is like for nibbanic consciousness to be ‘undergone’. The following sutta seems to depict nibbana from the what-it-is-like viewpoint (and I shall soon point to evidence that this is so) as opposed to characterising it in more objective terms (although I shall soon suggest that Buddhism collapses this distinction when it comes to the ‘ultimate’ depiction of nibbanic consciousness). To endorse the claim about nibbana’s ‘objective’ characterisation we appeal to other suttas. For now, let us consider the following:

(S.I) Consciousness without feature, without end
   luminous all around:
   Here water, earth, fire, & wind have no footing.
   Here long & short
   coarse & fine
   fair & foul
   name & form
   are all brought to an end.
   With the stopping
   of [the activity of] consciousness,
   each is here brought to an end.

In this sutta, “water, fire, earth and wind” pertain specifically to the khandha of physical form. The other descriptions (“long and short” etc.) are covered by the terms “name and form” (nama-rupa) which denote all five psycho/physical khandhas. The consciousness that is “without feature, without end” signifies a mode where all the usual conscious activity, involving the awareness of khandhas, has ceased. This means that from the first-person perspective, there will be no cognizance of any khandha, mental or physical, to present a sense of limitation (“feature”) – and hence a confined perspective of a subject – to the consciousness. The depiction of consciousness as “luminous” is, furthermore, suggestive of a lived dimension, despite the absence of apprehended objects. The following sutta (S.II), which refers to nibbanic consciousness as the “sphere to be realised”, reinforces the idea that there is something it is like to undergo such objectless consciousness. It provides an alternative parsing of the sensory-mental objects of which the nibbanic consciousness is free:

32 From Thanissaro Bhikkhu’s *The Mind Like Fire Unbound*, (1993a, 29-30). The sutta is from *Digha Nikaya*, section 11. All the suttas mentioned in this chapter are translated by Thanissaro Bhikkhu and appear in *The Mind Like Fire Unbound*. 
Monks, that sphere should be realised where the eye (vision) stops and the perception (mental noting) of form fades. That sphere is to be realized where the ear stops and the perception of sound fades...where the nose stops and the perception of aroma fades...where the tongue stops and the perception of flavor fades...where the body stops and the perception of tactile sensation fades...where the intellect stops and the perception of idea/phenomenon fades: That sphere should be realized.

Why should we interpret “consciousness” or the more metaphorical “sphere to be realised” as ‘witnessing’? In light of what was said in chapter one, there are two pertinent reasons to suppose that the term ‘witnessing’, (normally the \textit{modus operandi} of a subject), is well-suited to capture what is meant by “consciousness” in these suttas. We have noted that witnessing does not come across as being object-like, either from the third-person or, as is relevant here, the first-person perspective. If witnessing \textit{qua} witnessing could be attended to as an object from the first-person perspective, then its observed limitations would immediately render it unsuited to capture what is supposed to seem \textit{free} from such limitation. As that which can observe or attend to sensory-mental limitation, witnessing does not in \textit{itself} seem limited by these observable object-qualities. Another reason for interpreting “consciousness” in these suttas as ‘witnessing’ is that witnessing, as noted in chapter one, does not logically require the co-presence of objects that are witnessed. While most of us will be unable to imagine what objectless witnessing is like, it is at least possible to \textit{conceive} of it without contradiction. Since the above suttas allude to a mode of consciousness that is free from observed objects, ‘witnessing’ seems well placed to capture what is meant here. Unless there is evidence to the contrary, we shall henceforth treat the “consciousness” of \textit{nibbana} as a mode of \textit{witnessing}. Now, there is more to nibbanic consciousness than just witnessing – or to put the point another way, the witnessing aspect to \textit{nibbana} is metaphysically significant. Because I prefer to reserve the term ‘witnessing’ for the specifically \textit{apprehending} aspect (regardless of metaphysics) I shall continue to refer to the unconditioned \textit{nibbana} as simply ‘nibbanic consciousness’ (of which we can interpret the consciousness as involving witnessing). We will now consider these metaphysical dimensions.

An instructive way to approach the concept of nibbanic consciousness is to note some parallels that can be drawn between Kant’s (1787) notion of the “noumenal subject” and that of ‘nibbanic consciousness’. Both are portrayed as being unconditioned by space, time,

\footnote{Thanissaro Bhikkhu (1993a, 32) \textit{Samyutta Nikaya}, chp XXXV,section 116.}
quality and relation – although Buddhist suttas do not employ these four Kantian terms to describe the manner in which nibbanic consciousness is unconditioned\textsuperscript{34}. To see that these terms are indeed aptly applied, we will soon consider a sutta, which, unlike (S.I) and (S.II), is explicit on the \textit{objective} or \textit{intrinsic} non-dependence of nibbanic consciousness on the psycho-physical \textit{khandhas}, suggested in that sutta by repetition of the word “the”. Before introducing it, it is important to say something about the epistemic conditions under which Buddhism would, in this context, apply the word ‘objective’, ‘intrinsic’, or some similar variant. What is said here is extrapolated from my general understanding of Buddhism, rather than based upon specific secondary sources. I believe this extrapolation to make the most epistemological sense of Buddhism’s metaphysical claim that \textit{nibbana} is, as the following sutta suggests, “the ultimate”. It shall emerge from this that the notion of nibbanic consciousness will, despite any parallels, not be entirely congruent with Kant’s “noumenal subject”, since the former, unlike that latter, is claimed to be humanly knowable as it is in itself.

Unlike what is customary in traditional Western science, Buddhism does not consider the objective or intrinsic character of nibbanic consciousness to be \textit{revealed} and \textit{properly understood} through third-person investigation and analysis, such as through studying the brain or behaviour of an Arahant, or by reading these and other words. Buddhism, as I see it, maintains that the objective, unconditioned nature of nibbanic consciousness can only be revealed and known \textit{fully as it is in itself} through its direct, unmediated first-person experience. Let us first talk about the revelation aspect. That the full intrinsic character of nibbanic consciousness would not be revealed through third-person observation is not an unreasonable supposition, given its objectless witnessing nature. Third-person observation requires mediation by conditioned sense-modalities that automatically present their data as an \textit{object}. And witnessing, as already noted, does not \textit{seem} like something that can, \textit{qua} witnessing, ever be an \textit{object} of observation. It hence stands to reason that direct objectless experience will be required for the full intrinsic character of witnessing to be properly revealed. Whether being the recipient of such \textit{revelation} is either required for, or amounts to, a \textit{complete knowledge or understanding} of the character of witnessing is a harder contention to defend. Buddhism holds that any knowledge that is obtained from analysing reports of

\textsuperscript{34} The description of nibbanic consciousness in these particular Kantian terms is my own.
such revelation – such as in this project – will be considered a fragmented and inferior means to understanding *nibbana*.

On this picture (as I understand it), Buddhism will maintain that only through a direct unmediated experience will the unconditioned, *khandha*-independent, *objective* reality of *nibbanic* consciousness be laid bare and understood. There will furthermore be no room for subjective variation between different revelations of its experience. If subjective variation between certain *Arahants’* experience of *nibbanic* consciousness were possible, then that would automatically render the experiences to be mediated by conditions pertaining to *khandhas* of each *Arahant*, implying that the experiences would not *in fact* be that pertaining to objectless consciousness. Moreover, it is claimed that the experience revealed through an unmediated appearing-to or ‘uncovering’ of *nibbanic* consciousness in its objectless mode will not differ – either ‘numerically’ or ‘qualitatively’ – from the intrinsic unconditioned reality of *nibbanic* consciousness itself. When it comes to the first-person experience ‘of’ *nibbanic* consciousness, I understand Buddhism to maintain no gap between appearance and reality, subjective and objective, *knowing and being*. The first-person experience ‘of’ *nibbanic* consciousness effectively *is* *nibbana*. As I interpret the position, no ‘aspect’ of *nibbanic* consciousness can remain ‘hidden’ from the purview of the *Arahant*. With this in mind, we can now consider the following *sutta*:

(S.III) The unfashioned, the end,  
the effluent-less, the true, the beyond,  
the subtle, the very-hard-to-see,  
the ageless, permanence, the undecaying,  
the featureless, non-differentiation,  
peace, the deathless,  
the exquisite, bliss, solace,  
the exhaustion of craving,  
the wonderful, the marvellous,  
the secure, security,  
*nibbana*,

---

35 What makes the different ‘instances’ ‘numerically different’ is extrinsic to *nibbanic* consciousness, due to the *khandhas* as ‘coverings’. There is no intrinsic difference with respect to the ‘underlying’ *nibbanic consciousness* that is experienced on these different ‘instances’ (if there was a difference, then *nibbanic consciousness* would, like the *khandhas*, be conditioned).

36 I place ‘numerical’ and ‘qualitative’ in scare-quotes because strictly speaking, these terms are befitting of conditioned, rather than unconditioned reality. Being undifferentiated, unconditioned reality (in a Buddhist context) offers no contrast class that could give these terms a domain of meaningful application.
the unafflicted, the passionless, the pure,
release, non-attachment,
the island, shelter, harbor, refuge,
the ultimate.\(^{37}\)

(S.III) encapsulates much of what is central about the notion of nibbanic consciousness. It is fitting to begin by talking about how it depicts nibbanic consciousness as **unconditioned by quality**. While (S.I) and (S.II) relay the idea that nibbanic consciousness does not appear qualitatively restricted from a subjective or first-person perspective (suggesting objectless consciousness), (S.III), in virtue of the words ‘the featureless, non-differentiation’, suggests this to be objectively true of nibbanic consciousness in general. This means that nibbanic consciousness will elude any ‘external’ characterisation by delimiting quality, rendering it unobservable from a third-person perspective. In the previous chapter, we stipulated that ‘quality’ in a Buddhist context pertains to anything that can be characterised with reference to sensory or mental observation – which covers all of *conditioned* existence. It hence pertains to all the observable empirical qualities that characterise or can be known through the five sense modalities (discernible objects with colours, shapes, sounds, tastes, tactile and proprioceptive qualities) as well as qualities and concepts which characterise mental deliverances such as emotions, perceptions, thoughts and mathematical concepts. Now, in Buddhism, “differentiation” is defined by limitations of the “six spheres of contact”, which amounts to those deliverances of the five senses and mind:

(S.IV) However far the six spheres of contact go, that is how far differentiation goes. However far differentiation goes, that is how far the six spheres of contact go. With the remainderless fading and stopping of the six spheres of contact, there comes to be the stopping, the allaying of differentiation.\(^{38}\)

Nibbanic consciousness, then, is both ‘objectively’ and ‘subjectively’ undifferentiated by any qualities that are discernible by the six senses or intellect. As a consequence, nibbanic consciousness could never be discovered by traditional scientific methods that rely upon data from quality-sensitive instruments. In denoting something unconditioned by quality, ‘ nibbanic consciousness’ seems notably akin to Kant’s ‘noumenal subject’, namely, as “not a

\(^{37}\) Thanissaro Bhikkhu, (1993a, 14). *Samyutta Nikaya* chp XLIII, sec 1-44.

real whole but a simple” (1787, A404). I interpret Kant to mean here that the noumenal subject, as a simple, is undifferentiated by any empirical or conceptual determination. But while Kant believes the deliveries of sense and reason to exhaust the scope of human knowledge and experience, Buddhists do not consider this true with respect to nibbanic consciousness. The idea that nibbanic consciousness can be known as it is in itself — namely, through unmediated first-person experience (notably not as a separate object of experience) — sets it apart from Kant’s humanly unknowable noumenal subject. This ‘phenomenal’ (for want of a better word) dimension to nibbanic consciousness is furthermore depicted in (S.III) as extremely positive in hedonic tone: “exquisite, bliss, solace, wonderful, marvellous”. Since nibbanic consciousness intrinsically lacks any objects and their limitations, both mental and physical dukkha will be absent.39

There is a further consequence of importance here. Anything pertaining to the “allaying of differentiation” carries connotations of an absolute unity — absolute in the sense that there is no observable differentiation to suggest a lack of unity. Experienced as it is in itself sans khandhas, with no sense of a subject/object division, nibbanic consciousness cannot therefore seem like a disunified or fragmented witnessing but must carry a sense of absolute intrinsic unity. In later chapters, outside the context of nibbanic consciousness, we shall elaborate on the notion of unity and argue that ‘disunified witnessing’ is a contradiction in terms. Witnessing will be shown to be intrinsically unified — rendering it yet more suited to capture what is meant by the ‘consciousness’ of ‘nibbanic consciousness’.

Let us consider evidence that nibbanic consciousness is considered as unconditioned by space. This can be found in a simple implication: if something is unconditioned by quality in virtue of being beyond the “six spheres of contact”, then it cannot possibly be conditioned

---

39 One may wonder at whether a contradiction lurks in the notion of an untainted happiness that is not an object of consciousness but is intrinsic to objectless nibbanic witnessing. While there is not the space to address this concern in detail, an analogy may suffice in showing that there is not obviously a contradiction. Suppose we compare witness-consciousness to light (as has often been done). The light, which has the potential to illuminate objects, cannot illuminate itself. Similarly, the witnessing, which has the potential to notice objects, cannot witness itself (as an object of attention). Let us now suppose that the light is spectrally coloured (as through a prism). The coloured light cannot illuminate its own hue any more than it can illuminate its own luminosity; the hue is simply part and parcel of the illuminating quality that sheds light on other objects. Similarly, it seems we can imagine without contradiction that ‘nibbanic happiness’ (as I shall call it) ‘colours’ the pure nibbanic witness-consciousness with a ‘hedonic hue’ that, like the pure witnessing, cannot be attended to as an object. This will not rule out the possibility that the nibbanic happiness/consciousness has its own intrinsic phenomenal feel — it is just that such happiness/consciousness will escape its own attentive purview and so will never be an object.
by space, since spatial limitation bespeaks of discernible qualities that can be known through these spheres. That is, being conditioned by space involves an object’s restriction by spatial parameters (such as position and shape), which can be verified through sensory observation. In Buddhism, the *khandha* of physical form, with its elements of “earth, fire, water and wind” (pertaining to density, heat, fluidity and airflow) is allocated the ‘space-filling’ niche. Since this renders the spatial physical *khandhas* describable with reference (at the very least) to those sensory-mental modalities of sight, touch and perception, they must also be conditioned by quality. Nibbanic consciousness, lacking any such discernible quality, both objectively and subjectively, cannot therefore be conditioned by space.

We can note the similarity between nibbanic consciousness and Kant’s spatially unconditioned noumenal subject, despite any variation in the background metaphysics. For Kant, human “intuitions” of space and spatial qualities originate not from the mind-independent world (as would naively seem) but from the noumenal subject – that intrinsically unknowable entity which must be postulated if law-like conscious experience is to be possible. Kant maintains that this entity from which space and its qualities originate cannot in itself be conditioned by space, but must rather present a unity with respect to it (1787, A 404). In the following chapter, we consider how nibbanic consciousness – even if extrinsically ‘covered’ by the co-presence of spatio-temporal *khandhas* pertaining to a sense of self – will also, on our analysis of Buddhism, present a unity with respect to them. In later chapters, and outside the context of nibbanic consciousness, we will argue that witnessing *qua* awareness brings to ordinary conscious experience a unity that is not only inferred (in the manner of Kant’s noumenal subject) but is directly experienced.

We now consider evidence that suggests nibbanic consciousness to be *unconditioned by time*. Kant’s noumenal subject is “the unconditioned unity in the plurality in time” in that it is “not numerically different at different times but one and the very same subject” (1787, A 404). In context of Kant’s metaphysics, this means that the noumenal subject is outside of time, rather than continuing forever in time. Since Kant regards time, like space, to originate from the noumenal subject, he sees the noumenal subject as exempt from temporal plurality: it presents a unity by not being numerically different at different times (and so not perduring). Like Kant’s noumenal subject, the depiction of nibbanic consciousness as “the ageless, permanence, undecaying, deathless” is to be understood as eternal in the sense of
being outside of time, rather than as existing forever within time. It presents a unity with respect to time by being, in Kant’s phrase, “not numerically different at different times”. With the microscopically fleeting and conditioned co-dependent khandhas on one hand, and unconditioned reality on the other, Buddhist cosmology has no room for a hybrid entity that will unchangingly exist forever in time. If something is unconditioned, then it must be outside of time altogether.

Unlike Kant’s noumenal subject, however, the unifying presence of timeless nibbanic consciousness can, it is claimed, be experienced as it is in itself. This invites a potential objection: when experienced fully as it is in itself by a living Arahant, nibbanic consciousness will be experienced at a time t, for example, at 4pm. Does this not mean it is conditioned by time? The short answer is “no”. As we indicated earlier, what makes it appear like a temporal event has everything to do with the dispersing and re-grouping of khandhas that are conditioned by time. The nibbanic consciousness would be conditioned by time if it were brought into existence upon that moment, which, of course, it is not. With this in mind, nibbanic consciousness might be described as being ‘timelessly ever-present’. The timeless aspect suggests that nibbanic consciousness, in its undifferentiated (hence non-plural) unity, is intrinsically unaffected by any time-bound configurations of khandha. The ever-presence suggests that at any given khandha-moment (or ‘k-moment’ for short), the unified nibbanic consciousness will be present – even if not obviously so. And in the absence of any k-moment (or where there is no passing of time) nibbanic consciousness will be present. We shall see in chapter three that the timeless dimension to nibbanic consciousness, even when ‘covered’, has important implications when it comes to articulating how Buddhism construes the sense of self. We will also see (in later chapters) that even outside the context of nibbanic consciousness, witnessing will be noted for its ‘subjective sense of presence’.

We finally look at whether nibbanic consciousness is considered unconditioned by relation. To be unconditioned by relation, on Kant’s account, is to be “not inherent [in something else] but self-subsistent” (1787, A 404), which would imply, as with the noumenal subject, an unreliance on anything else for existence. The unreliance of nibbanic

What if, for the sake of argument, all living beings were to be destroyed? Would there still be nibbanic consciousness? Given its timeless non-dependence upon any khandha, mental or physical, the answer must be ‘yes’. It is just that the mind of the Arahant, when freed from tanha and other ‘covering’ khandhas, is uniquely conducive to fully experiencing its intrinsic nature (nibbanic consciousness) as it is in itself.
consciousness upon the *khandhas* can be inferred from the fact of its being unconditioned by time. Since nibbanic consciousness is not time-bound, it cannot depend upon any time-bound entity, viz, the *khandhas*. If it did depend upon the *khandhas*, then it would, like the decaying *khandhas*, be itself subject to decay. But nibbanic consciousness is portrayed as unequivocally exempt from those laws of conditioned co-dependence (both synchronic and diachronic) that govern the five *khandhas*. Thanissaro Bhikkhu has noted that the complete freedom of the Arahant’s mind from dependency upon any object renders it indescribable both “from the outside”, and from the perspective of the Arahant who experiences it (1993, 28-29). There are simply no quality-restricted criteria by which it could be positively described. Although cast as a relatively ‘positive characterisation’, our discussion has mainly focused on the restrictions to which nibbanic consciousness is not bound. The only truly positive aspect has been in the depiction of an immensely peaceful and happy witnessing. However, even this much seems portrayed, in the overall context of nibbanic consciousness, as being quite beyond classification, description and imagination; a theme that is repeated in the suttas, for example:

(S.V) When all phenomena [viz., the conditioned *khandhas*] are done away with
All means of speaking are done away with as well.

Nibbanic consciousness hence emerges every bit as unconditioned as the Kantian noumenal subject – but yet with an experiential dimension.

**(B) Nibbanic Consciousness and the Khandhas**

We now ask how the Buddhist analysis might plausibly relate the intrinsically timeless, objectless nibbanic consciousness, to conscious ‘states’ where there is the co-presence of *khandhas*. For this purpose, two important groups of person will be considered: (1) the Arahant (in whom *tanha* has ceased) and (2) the ordinary person with *tanha*.

---

41 Nibbanic consciousness would hence not be physical, since it would not depend in any way upon the body or brain. Why then, its association with properly functioning brains (as opposed to say tables or dead brains) or, in Buddhist terms, the psycho-physical *khandhas*, for example those of an Arahant? Any role played by the *khandhas* must be understood in terms of their holding in the right relations to allow for the various degrees of revealing (rather than creating) the intrinsic nibbanic consciousness. The *khandhas* hence serve to cover, in various degrees, what is timelessly everpresent.

(1) The Arahant with ‘Proximate Nibbanic Consciousness’

We have seen that for nibbanic consciousness to be experienced fully as it is in itself—where knowing and being coincide—there cannot be the co-presence of experienced khandhas. Nibbanic consciousness as it is in itself will be objectless. Upon the Arahant’s death, known as parinibbana, the psycho-physical khandhas will not reconfigure into another birth. The nibbanic consciousness formerly associated with khandhas of the Arahant will fully know itself as it is in itself, free from the confines of space, time, quality and relation. Now during their psycho-physical lifetime, the mind or consciousness of an Arahant will come into regular contact with khandha-objects as they engage with the world. As Peter Harvey points out, this implies that the objectless nibbana as objectless “is not experienced by the Arahat all the time” (1995, 197). However, Harvey also notes that “he or she can repeatedly re-experience it before entering it for a final time at death ... The Arahat’s full experience of nibbana, as a state [sic] in which the personality-factors [khandhas] stop, might be seen as his ‘participating in’ this timeless reality” (1995, 192).

What emerges from this picture, then, is the following. While the tanha-free Arahant has ready access to timeless, objectless nibbanic consciousness, insofar as their khandhas can temporarily ‘stop’ to allow for its full experience, there are plenty of k-moments where khandhas are present to their field of consciousness. In these k-moments of an Arahant, where nibbanic consciousness can be thought of as at least partially ‘covered’, how would nibbanic consciousness reveal itself? What comes across from the accounts of reported Arahants is that even while interacting in the world, their minds, never far from full nibbanic consciousness, are replete with aspects whose source can be most parsimoniously attributed to nibbanic consciousness.43 Having realised the goal of the Noble Eightfold Path—eliminating tanha and experiencing nibbana—they can never revert to their former state of tanha and of not knowing what nibbanic consciousness is fully like. (Perhaps in a similar fashion, one who has clearly ‘seen through’ a magic trick can never be duped by it again). Realising this ultimate Buddhist goal is said to involve the deepest understanding of conditioned phenomena, meaning that the mind, fully knowing the impermanence of khandhas, is never again drawn to seek happiness from them (which is another way of saying

---

43 A detailed depiction of the Arahant’s general mindset, supported by suttas, can be found in Thanissaro Bhikkhu, (1993a, 95-120). Rahula (1996, 43) also gives a nice depiction.
that \textit{tanha} ceases). The mind of the Arahant fully comprehends the conditioned \textit{khandhas} as \textit{anicca}, \textit{dukkha} and \textit{anatta} – that is, as impermanent, as an unsatisfactory source of true happiness, and as not pertaining to a self (more on this later). What is termed as ‘ignorance’ (\textit{avidya}) about conditioned existence is hence said to be lost – resulting in an increase in ‘wisdom’ (\textit{panna}) with respect to this factor. The Arahant also apprehends the unconditioned nature of nibbanic consciousness as not \textit{anicca} and not \textit{dukkha} (but still \textit{anatta} since there is no place for a self – more on this in chapter three). Upon the first such direct apprehension (not merely glimpse) of unconditioned reality, \textit{nibbana} – the complete cessation of \textit{tanha} – is realised.

On all reports, the mind of a living Arahant is not only free from \textit{tanha}, mental \textit{dukkha} and ignorance about conditioned reality, but is radiantly and constantly happy in the knowledge of its real nature as unconditioned (indeed, the untainted happiness is deemed as inherent to the mind’s very nature). Such a mind’s reflexive knowledge of its real unconditioned nature is never forgotten whilst engaging with objects of the world. The Arahant’s sense of witness-consciousness is said to be very powerful, present-centred and constant, with their awareness never ‘lost’ in the imagined time-zones of thought. Buddhism maintains that being regularly lost in thought feeds a mindset with \textit{tanha}, which, they claim, is pivotal to maintaining a sense of self. We shall soon see that on the Buddhist analysis, the witnessing of a usual \textit{tanha}-ridden person, while sourced in nibbanic consciousness, does \textit{not} reflexively know its real nature as unconditioned – because it assumes, reflexively, that it is a self. Arahants, lacking the co-arising \textit{tanha}, are said to be without a sense of self. At this stage we can surmise that a pull of \textit{tanha} will strongly condition the arising of mental \textit{khandhas} – hardly conducive to ready partakings in objectless consciousness. Freed from \textit{tanha}, however, the mind of the Arahant will be naturally ‘with’ whatever is happening in the present – not compulsively looking elsewhere (for example to past or future scenarios) for gratification. This may considerably free up their mind to ‘partake in episodes’ of objectless consciousness. Because of its ready ‘proximity’ to full nibbanic consciousness, I shall henceforth refer to the Arahant’s witness-consciousness with presence of \textit{khandhas} as ‘proximate nibbanic consciousness’.\footnote{Since editing out the final chapter (on how meditation could work to undermine sense of self), there will not be many more opportunities to mention proximate nibbanic consciousness. The notion nevertheless serves as an}
(2) Pre-Nibbanic Consciousness of the Ordinary Person

While ‘proximate nibbanic consciousness’ denotes the witness-consciousness of an Arahant who is observing objects, I allocate the term ‘pre-nibbanic consciousness’ to denote the nibbanic witness-consciousness of the person whose mind harbours tanha — no matter how much or how little. On the Buddhist position, there can be no intrinsic difference between the ‘pure’ nibbanic witness-consciousness of an Arahant and the pre-nibbanic witness-consciousness of a criminal. The difference between ‘pure’, ‘proximate’ and ‘pre’ nibbanic consciousness is only extrinsic — to do with the ‘covering’ of khandhas and/or tanha. A criminal has the potential to attain nibbana — a potential to be realised not by his “watering the seed” of a latent nibbanic consciousness so that it “blossoms into” fully-fledged version, but by his eradicating tanha so that conditions become ripe for an “uncovering” of the timelessly ever-present nibbanic consciousness, intrinsic to his mind. Our inquiry will now turn to the extent to which, on this Buddhist position, the underlying objectless nibbanic consciousness would be revealed in a mind with the coverings of tanha. How much, if at all, would pre-nibbanic consciousness, with the co-presence of tanha, resemble the tanha-free objectless nibbanic consciousness? Obviously the resemblance cannot be veridical — otherwise there would be no need for the Four Noble Truths to help one’s mind to realise nibbana.

From reflections on this chapter we can note an immediate similarity: everyday witnessing will resemble that of nibbanic consciousness. Indeed, not only has this been assumed immediately above (viz., ‘nibbanic witness-consciousness’), but this chapter has also provided reason for supposing that the ‘consciousness’ of nibbanic consciousness is best cast as ‘witnessing’. Hence I propose that the modus operandi of the subject — the broadest mode of conscious apprehending — will be most parsimoniously cast, on the Buddhist position, as the witnessing inherent to pre-nibbanic consciousness. While Buddhism does not make any notion of witnessing explicit — since the focus of practice is on discerning the nature of khandha-objects as anicca, dukkha, anatta — witnessing as ‘that which knows’ objects as anicca, dukkha and anatta is assumed in the very possibility of such practice. For

important contrast-class against which to understand ‘pre-nibbanic consciousness’, a notion that is needed, in turn, to understand the Buddhist position on consciousness and no-self.
there must be that which can progressively know the conditioned nature of khandhas and of course, know itself reflexively as the unconditioned nibbanic consciousness.

Resuming our cloud-covered sun analogy, if the sun is analogous to the pure objectless nibbanic consciousness, the witnessing inherent to pre-nibbanic consciousness may be compared to the brightness of clouds from the sun behind them. We may now wonder whether Buddhism would regard any further aspects of nibbanic consciousness to be “brightening the clouds”. For instance, does Buddhism regard the ordinary conscious state of Joe Bloggs to harbour any indications – even if somewhat muted or distorted – of timeless everpresence, non-dependence upon objects, unchangingness, non-differentiation or unity (whether synchronic or diachronic), a non-intentional phenomenal feel, or intrinsic happiness? To determine this, we must now consider what Buddhism has to say about the self: the topic of chapter three.
Chapter Three: The Definition and Status of Self in Buddhism

In Buddhism, the sense of self is considered to co-arise with tanha such that whenever one of these factors is present in the mind, then so is the other. Given that is it the goal of the ‘Noble Eightfold Path’ of practice to eliminate tanha, then it must also be its goal to eliminate the sense of self – a point made explicit in Buddhism. The self that we have a sense of being is considered, moreover, not to exist. As I see it, the most challenging task is not that of showing that Buddhism rejects the existence of a self – which we shall see comes across fairly clearly in the suttas – but that of extracting a clear and plausible definition of the self whose existence Buddhism rejects. The self is generally depicted, in Buddhism, as an entity that most people have a sense of being. While more theoretical concepts of self are also alluded to in the suttas (and the existence of such selves implicitly rejected) these concepts of self are peripheral to the main concerns of Buddhism and will not be discussed in this chapter. In order to extract from Buddhist sources the relevant definition of self (namely, that of the entity we have a sense of being), we shall appeal to the well-known suttas on anatta or not-self, as well as take into consideration the suttas on nibbana – a task that will occupy most of this chapter. Seeing which features of nibbana ‘show through’ in what I have termed ‘pre-nibbanic consciousness’ will be essential to understanding how Buddhism construes the sense of self and hence, the self.

As mentioned in chapter two, I know of no scholars who I believe to have achieved this in a satisfactory manner. Because the Buddhist definition of self has usually been misrepresented in Western literature, the central Buddhist doctrine of anatta or no-self – which relies on the definition of self – has also been misconstrued in Western literature. For these reasons I base my analysis mainly on primary sources. The goal of my discussion, in other words, is as much historical as philosophical: I want to present a definition of the self which most closely fits with primary Buddhist sources, and then show how the Buddhist doctrine of anatta or no-self ought to be construed, given the motivations behind this definition of self in Buddhism. This construal will significantly depart from how the

45 As translated by Thanissaro Bhikkhu.
‘Buddhist’ definition of self and anatta has been popularised in the West – as my critique of secondary sources, at the end of this chapter, will reflect.

**Ownership and Identification as Reciprocal Assumptions of Self**

Central to understanding what ‘self’ means in Buddhism are the notions of identification and ownership. These notions are implicitly appealed to in the suttas below and will need considerable unpacking, a task that will run into chapter four (and so will require some patience on behalf of the reader). For now, we can take the term ‘identification’, on the Buddhist position, to pertain to a subject’s adopted sense of self, involving the reflexive and implicit assumption, on the part of the subject, that various psycho-physical attributes (comprised mainly of tokenings from the five khandhas) are in some way assimilated to itself\(^46\). Ownership can be regarded, in Buddhism, as reflecting a broad mode of such identification, where the subject implicitly takes on the mantle of ‘me-as-owner’. This assumed identity as ‘owner’ is evidenced through the tacit regarding of certain other khandhas to be ‘mine’ or to ‘belong’ to ‘me’. We shall see that Buddhism regards the ownership mode of identification to be central to the very notion of self, such that: if there is a self (or ‘me’), then there is what belongs to a self (‘mine’), and if there is what belongs to a self (‘mine’), then there is a self (‘me-as-owner’). From the suttas below, it can be discerned that a sense of ‘me’ and ‘mine’ is sought to be eliminated through the Noble Eightfold Path of practice. In such a scenario, the subject would no longer identify with (tokenings of) khandhas as ‘me’ (for instance, would no longer reflexively assume them to be a part of itself), thus eliminating the psychological platform of a ‘me’ from which other khandhas can be regarded as ‘mine’. On Buddhist cosmology, when the sense of self is eliminated, then so is rebirth into samsara, with its endless round of dukkha. We should be reminded of our definition of subject:

A subject S is a specific perspectival locus for witnessing, in relation to which objects are witnessed. This locus implicitly appears – whether dreaming or waking – to be psycho-physically constrained in space and time, offering a specific viewpoint from which the world, dreamt or actual, is witnessed.

\(^{46}\) In this project, the term ‘itself’ is to be understood colloquially and is not meant to convey the reality of a self.
It is fundamental to the following discussion that this notion of subject does not involve appeal to ownership or identification and so is not the same as the notion of self depicted in Buddhism. On the Buddhist position, we are to understand that the (witnessing) subject, from its own psycho-physical perspective, makes the assumption of being a self through its assuming various khandhas to be ‘me’ (hence integrated with its existence) or ‘mine’. This assumption of self is regarded, in Buddhism, to be deeply mistaken. Let us then consult some of the relevant suttas (whose themes are repeated in other suttas). We continue referring to them with Roman numerals, with only the relevant parts to each sutta quoted:

(S.VI) Bhikkhus, there being a self, would there be what belongs to my self?"
"Yes, venerable sir."
"Or there being what belongs to a self, would there be my self?"
"Yes, venerable sir."  

(S.VII) “Just as a dog, tied by a leash to a post or stake, keeps running around and circling around that very post or stake; in the same way, an un instructed, run-of-the-mill person -- is not well-versed or disciplined in their Dhamma -- assumes [khandhas (form, feeling, perception, mental formations, consciousness)] to be the self, or the self as possessing [khandhas], or [khandhas] as in the self, or the self as in [khandhas].

He keeps running around and circling around that very form...that very feeling...that very perception...those very fabrications...that very consciousness. He is not set loose from form, not set loose from feeling...from perception...from fabrications...not set loose from consciousness. He is not set loose from birth, aging, & death; from sorrows, lamentations, pains, distresses, & despairs. He is not set loose, I tell you, from suffering & stress.

[The mode to aspire is where one] doesn't assume [khandhas (form, feeling, perception, mental formations, consciousness)] to be the self, or the self as possessing [khandhas], or [khandhas] as in the self, or the self as in [khandhas].

He doesn't run around or circle around that very form...that very feeling...that very perception...those very fabrications...that very consciousness. He is set loose from form, set loose from feeling...from perception...from fabrications...set loose from.

consciousness. He is set loose from birth, aging, & death; from sorrows, lamentations, pains, distresses, & despairs. He is set loose, I tell you, from suffering & stress".  

(S.VIII) [On the Arahant] “In the same way, a monk investigates form, however far form may go. He surveys feeling...perception...fabrications...consciousness, however far consciousness may go. As he is investigating form...feeling...perception...fabrications...consciousness, however far consciousness may go, any thoughts of 'me' or 'mine' or 'I am' do not occur to him”.  

(S.IX) ‘...How do you construe this, monks: If a person were to gather or burn or do as he likes with the grass, twigs, branches & leaves here in Jeta’s Grove, would the thought occur to you, “It’s us that this person is gathering, burning, or doing with as he likes”?’

‘No, sir. Why is that? Because those things are not our self and do not pertain to our self.’

‘Even so, monks, whatever is not yours: Let go of it. Your letting go of it will be for your long-term happiness and benefit. And what is not yours? Form (body) is not yours...Feeling is not yours...Perception...Mental Processes...Consciousness is not yours. Let go of it. Your letting go of it will be for your long-term happiness & benefit.’  

It is clear from this selection of suttas that the assumption or sense of a self, at least in relation to the khandhas, is something that Buddhism considers, like tanha, to lead to dukkha and, as with tanha, something to be got rid of if dukkha is to be eliminated. The relation between the sense of self and tanha can thus be taken to be a very close one – a point I examine later in this section. We can also gather that the self Buddhism concerns itself with here is not something lofty and esoteric, but is something that a “run-of-the-mill” person, harbouring tanha, will assume themselves to be. The phrase “run-of-the-mill person [who is] not well-versed or disciplined in their dhamma” thus applies to anyone who is not an Arahant; to anyone who still harbours tanha in their mindset. Such a person, harbouring tanha and a sense of self toward the khandhas, will be reborn into samsara and dukkha. On the Buddhist picture almost all of us will be such a person.

48 Samyutta Nikaya chp XXII sec 99 translated by Thanissaro Bhikkhu (2003 online).
49 Samyutta Nikaya chp XXXV sec 205 translated by Thanissaro Bhikkhu (2003, online).
Another point to come across is that the self that a "run-of-the-mill" person assumes he is, is some kind of owner. To be an owner, as (S.VI) indicates, is to stand in a relation of 'belongingness' to something else. From (S.VIII) we can gather that a sense of belongingness involves thoughts of "me", "mine" or "I am"; while (S.IX) alludes to ownership through an injunction not to regard khandhas as "yours" (hence not with thoughts of ownership) and (S.VII) alludes to the owner-self as "possessing" the khandhas. What, then, does it mean to be an owner, such that one possesses the khandhas? Whatever it means, it will have to be distinguished from two other prevalent kinds of ownership that I shall for convenience term 'perspectival ownership' and 'possessive ownership'. We outline these contrast-terms before saying more about the notion of ownership relevant to the suttas.

For a subject to own something in a perspectival sense is for that thing, an object, to appear to the subject seemingly in a way that it can appear to no other subject. All 'private' phenomena such as thoughts, intentions, perceptions and sensations — at least as they appear from a first-person perspective — will be perspectivally owned by a subject: i.e., I view them from my perspective. In this sense, we commonly speak of such things as 'my toothache, my thought about Canada, my perception of the sea, my intention to move'. When it comes to objects (such as trees) that are agreed to be external to the 'inner life' of a person, what will be perspectivally owned is not the tree (unless one is a solipsist) but the specific manner through which the tree or other object appears to an observing, psycho-physical subject, namely in the associated visual, auditory, tactile and mental qualia. In relation to objects that appear to a subject in this special way, and which can hence be called 'mine' by a subject, the subject may be termed a 'perspectival owner'.

For a subject to own something in a possessive sense is, roughly, for that object to be regarded as theirs by right of social convention, actual or ideal. Ownership of clothes, houses and even people can fall in this category. In relation to possessively owned objects, the

---

51 I say "seemingly" in case it logically turns out that one person can access another person's experience, such as through clairvoyance. This turning out to be the case will not impinge on our definition, since what matters is only that objects appear a particular way from a subject's first-person perspective through virtue of occurring, as Philip Gerrans puts it "within her mind, not the mind of someone else" (2002, 37). This manner of occurring is one that, as it happens, would seem to be accessed only by the subject to whom the object occurs. Perhaps, Ronald de Sousa suggests, what I refer to here as 'perspectival ownership' can be divided further into 'titularity', 'privileged access' and 'incorrigibility' (2002, online). If so, then it may be possible to treat 'perspectival ownership' as an umbrella term for these closely related subsets. The important matter, for our purposes, is that it broadly contrasts in the aforementioned way with the Buddhist notion of 'ownership'.

subject can be termed a ‘**possessive owner**’. It seems plausible to suppose that a subject’s ownership of their body involves, at the very least, both perspectival and possessive ownership. For on one hand, bodily existence and movement can be known to a subject by way of perspectively owned perceptions and such sensations that cannot, *ipso facto*, be had by any other subject. In this context, the subject may speak of ‘my body’ in a perspectival sense. On the other hand, one’s body can be regarded as an object that will one day die, or whose parts can be lost or donated whilst alive. In a context where one may donate blood as they donate clothes, the subject can be described as having or renouncing possessive ownership of it (or of part of it).

The notion of ownership alluded to in the Buddhist suttas is neither perspectival nor possessive – but is nonetheless implicated in the prevalent attitude one has towards their body and mind (the body being composed of *khandhas* of physical form). Importantly, the Arahant is depicted as lacking any sense of this kind of ownership toward such *khandhas*, which immediately distinguishes it from the other two sorts of ownership. Since Arahants still interact with the world, for instance, they could not lose the impression of perspectively owning objects: objects, such as physical *dukkha*, will still appear to them in a way that cannot be accessed by other subjects. Nor will they suddenly cease to recognise such social conventions as those that dictate possessive ownership of their robe and bowl. While neither perspectival nor possessive, this third kind of ownership, a sense of which Arahants are depicted to lack, is clearly portrayed in the suttas as a sense of ownership had by *most people* towards various *khandhas* – in particular, as it happens, those “very” *khandhas* that are perspectivally or possessively owned by them. Perhaps it a kind of ownership-sense that a person could not locate with introspective ease unless they had experienced its loss. When we put the point this way, we can immediately point to (although not yet clearly define) a notion of ownership that is alluded to in contexts of Western psychology and which seems to answer to what is talked about in the suttas. It is a sense of ownership or ‘myness’, which, while ubiquitous, may on occasion be lost or suspended. Unlike with the Arahants, the loss of ownership-sense in such cases does not seem to affect their mindset in a global manner. Rather it involves the person regarding a subset of *khandhas*, formerly felt as theirs, to ‘not

---

52 As we shall make clear in chapter four, the Arahant, unlike persons with a sense of self, will not **identify with** the perspectival or possessive owners of their *khandhas*. 
belong’ to them. I shall henceforth refer to this kind of ownership – which still requires further elucidation – as ‘personal ownership’ whose subject identifies as a ‘personal owner’. A sense of personal ownership with respect to selective body-parts or mental contents, then, is reportedly found lacking in some cases of depersonalisation and anosognosia. Anosognosia is a mental affliction where the patient has a physical deficit (such as paralysis) but does not, on some cognitive level, recognise that he has it. As the philosopher Nikolinakos reports, this can elicit a denial of (personal) ownership towards the affected body part:

When such patients deny a deficit, they also tend to show indifference toward the affected body part or deny ownership of it and justify such denial with confabulations, e.g., ‘the limb was left behind by another patient’, ‘the limb belongs to the examiner’. They may also ignore the request to move the affected limb, or they may respond to the request by saying ‘here you are’ without being able to acknowledge that no movement has taken place....Patients who combine anosognosia and limb paralysis present a graded reflexive consciousness when they acknowledge the condition of paralysis in parts of the affected limb but deny it with respect to other parts of the same limb, e.g., the patient who denied ownership of the left hand but did not deny paralysis of the left arm and elbow...Such patients may persist in their denial of ownership of the specific body part in spite of the visual evidence, which may even be acknowledged explicitly. They may even acknowledge the oddity of their statements that it is paradoxical to hold that the forearm and hand, which they do not consider as belonging to them, are attached to their elbow. (2004, 323, 324).

Now someone with a paralysed arm but without anosognosia will usually associate the visual cues of it as attached to their body – a body that seems perspectivally owned because of such cues – with a sense of personal ownership or ‘my-ness’ towards the body and its paralysed arm. In the above example, however, it does not seem as if the perspectival cues of seeing one’s paralysed arm as attached to the body are being associated with the kind of personal my-ness that is usually felt towards one’s arm, paralysed or not. That the patient is aware of the visual cues (and so has some sense of perspectively owning the arm) is evident through the fact that they can be puzzled by a mismatch between the visual evidence and the lack of felt ownership.

When it comes to the perceptions, sensations and thoughts themselves (as opposed to external body parts inferred through such sensations etc.), a decoupling of a felt personal ownership from perspectival ownership towards them can be yet more evident. Cases of
depersonalisation provide good evidence that while some sensations, perceptions or thoughts present a unique perspectival aspect to the subject, the subject feels as if the very same mental objects do not belong to them. It is this kind of ‘personal’ ownership or ‘belongingness’, then – whose felt loss is evidenced through cases of anosognosia and depersonalisation – that seems relevant in the above suttas. It is a kind of ownership that seems taken for granted by most people, but whose feeling can yet be lost upon occasion. (We will discuss the difference between global versus partial loss of the sense of personal ownership in later chapters).

Can more be said about this notion of personal ownership? It is notable that Western philosophy does not seem to have formalised a distinction between perspectival and personal ownership, the two sometimes being conflated despite their differences. Perhaps this is unsurprising since the two usually co-occur; the suttas themselves suggesting that where there is perspectival ownership towards the mind and body, there is almost always a sense of personal ownership towards them. While the concept of personal ownership does not seem that well articulated in Western philosophy, the suttas, when interpreted correctly, may have something illuminating to offer. The key to their account is that personal ownership (unlike perspectival or possessive ownership) is to be analysed in terms of identification. In what follows, I will first say more about what is meant by ‘identification’, suggesting that the ordinary notion of self that concerns Buddhism is one that is minimally defined in terms of identification. I will then provide evidence from the suttas which analyses the notion of personal ownership in terms of the self and hence, in terms of identification.

53 Thus Tim Bayne writes: “Patients suffering from depersonalisation complain that their sensations and perceptions no longer feel as though they belong to them”. (2004, 222-223). Later in this project, I will, during the course of my arguments, be referring to cases of depersonalisation.

54 Bayne (2004) notes for instance that some authors have argued for a distinction between two kinds of ownership, namely (1) “a bare sense of being the subject of an experience” and (2) “the sense of being its author or agent” (2004, 222). Bayne illustrates this with reference to a jingle running through one’s head, where: “one experiences oneself as the subject of the auditory experience, but not its author” (ibid). Since one can clearly have a sense of myness without agency toward the jingle, Bayne wisely reserves the term ‘ownership’ for the notion of subjecthood as opposed to agency (in chapter three we further justify this difference between ownership and agency). While improving on these authors, however, Bayne does not say enough about subjecthood to forge a clear distinction between perspectival and personal ownership. Later, it transpires that he must mean ‘personal ownership’; since he suggests that a subject can, through such experiences as depersonalisation, lose the sense of ownership toward their thoughts and experiences. His analysis fails to distinguish this from the legitimate sense in which such thoughts etc., will remain (and will seem to remain) owned, namely, the perspectival sense.
Identification, as we mentioned earlier, involves a reflexive assumption on the part of the subject that various (psycho-physical) attributes are in some way assimilated with itself, the subject. It sometimes useful to draw upon a further distinction between, on the one hand, the terms ‘identify(ing) with’ and ‘identify(ing) as’, which describe what the subject implicitly does (by assuming itself to be assimilated with the attributes), and, on the other hand, the term ‘identified with’ (in this context), which describes the state of affairs implicitly assumed by the subject, namely, some level of assimilation between the subject and the psycho-physical attributes. Hence we can say that by identifying with or identifying as various attributes, the subject assumes itself to be identified with them in some integrative manner.

We shall now say more about the distinction between ‘identifies with’ and ‘identifies as’, before further elaborating on the phrase ‘identified with’.

For purposes of clarity I shall throughout this project adopt the following distinction between the phrases ‘identifies with’ and ‘identifies as’. It should be seen as a useful stipulation rather than as an argument for such a distinction per se. When a subject identifies with X, we shall stipulate that this implies (in our given context) the reality of the X that the subject identifies with (such as the khandhas). When a subject identifies as X, on the other hand, then this does not imply the reality of X: X, in our given context, may or may not exist (such as the self). ‘Identifies with X’ hence implies the reality of X, while ‘identifies as X’ does not imply the reality of X.

Now let us consider the phrase ‘identified with’ as it occurs in the above context. When a subject identifies either with X or as X, where X is some psychological or physical aggregate(s), this means that X is reflexively assumed, by the subject, to be in some way identified with itself. The term ‘identified with’ is to be construed either as X being assimilated with the whole subject, or as X being assimilated with a part of the subject, or as the subject being assimilated with X (more on this soon). It is important to note that the fact that a subject identifies with any X does not imply that X really is identified with the subject in any way, anymore than a philosopher’s ‘arguing for Y’ implies that the argument for Y succeeds.

In chapter four, when we say more about the concept of identification, it will be seen that the subject identifies with an item through appropriating the idea of that item to their perspective.
The relation that Buddhism posits between selfhood and identification is elaborated upon in (SVII), where the "run-of-the-mill person...assumes [khandhas (form, feeling, perception, mental formations, consciousness)] to be the self, or the self as possessing [khandhas], or [khandhas] as in the self, or the self as in [khandhas]". While a fuller description of what such identification could amount to will be postponed until chapter four, we can still get some initial idea of what is meant here.

Suppose that a "person" – which I think is most usefully analysed as a subject – "assumes" various khandhas to "be the self". Since I take this to mean that a general assimilation between the subject as a whole and a set of khandhas is being presupposed here (on the part of the subject that identifies with them) it makes most sense to construe the set of khandhas broadly, so as to include tokenings from both mental and physical quarters. The set will amount to the collection of khandhas, which, at any given time, is assumed by the subject to be 'me' (viz., a self) as a whole. Hence it will include many elements from the body (with its sensations etc.), the mental faculties, and from the personality (which incorporates ideas of one's past and anticipated future). I shall refer to this collection as the 'body-mind'. Expressed first-personally, I take the implicit content of this assumption to be something like: "I (qua subject) am assimilated with this body-mind, this body-mind is what I am". Through this assumption, the subject will implicitly and reflexively feel as if the body-mind is assimilated to itself in such a manner that the body-mind qua subject is implicitly assumed to be a singular 'me' or 'self'. We express this implicit assumption by saying that a subject, through identifying with the body-mind, identifies itself as a singular, integrated self or me. The 'self', as something that the subject identifies itself as, designates a subject to which the body-mind is somehow assimilated. 'Sense of self' (in this context) designates the feeling/assumption: "I (qua subject) am this body-mind, this body-mind is what I am".56

When a subject takes on this assumption of a general bodily-mental identity, it will often be evidenced through a particular concern taken in 'bodily-mental me', for example when out shopping for new clothes that nicely fit 'me' (where 'me' includes the overall self-conception). This now relates to that part of the sutta that speaks of "the self as possessing khandhas". The sense of a <self qua subject-that-is-body-mind> (held on the part of the

---

56 Of course, as mentioned in chapter one, the subject will probably not be reflectively aware of taking on this identity; it will most likely be simply assumed as true in the relevant context.
subject) will give rise to such assumptions as personally possessing other khandhas in relation to this assumed identity as a body-minded self. Suppose one buys a pair of trousers that, from both a bodily and mental perspective ‘fit’ very well – a fact that one delights in. Not only are the new trousers possessively owned (in the way I have defined the notion of possessive ownership): there is also a sense of personal ownership or my-ness toward the trousers possessed – and it is this kind of possession that the sutta is alluding to. ‘The self’ in this context designates me as personal owner of (viz., ‘possessing’) the trousers. One identifies as their personal owner – a self – through identifying with the body-mind that ‘fits’ them (mentally and physically) and with the possessive owner (‘possessive’ in the sense that I have defined as opposed to in the sutta) who has bought them.

Suppose that the possessive trouser-owner now decides that their messy hair (in particular) is quite out of keeping with their smart image. The kind of attitude towards the messy hair in particular, may be borne of assuming “khandhas... in the self”. The subject does not merely view the hair as any old hair, but views the hair as a part of ‘me’, that is, as an integral part of the overall self (i.e. as a part of the subject qua body-mind) that ‘I’ want to change. After the new hairdo, we can suppose that one now reflects: “Finally I fit into the corporation; I am properly a part of it”. This is assuming the “self as in the khandhas”, where the subject is identifying itself as a part of some other set of khandha-tokenings. The khandhas, in this case, are those whose tokens make up the corporation.

While this account of identification requires further elaboration – a task postponed until chapter four – we will have gained some initial idea of what is meant by ‘identification’ in relation to (SVII). The kind of self that the subject identifies itself as being, will, in this context, minimally pick out a subject that is in some way identified with khandhas, whether through being assimilated with them as a general self (khandhas assumed to be self) or as a part of the self (the khandhas assumed to be in the self) or as a part of the khandhas (the self assumed to be in the khandhas). As we mentioned earlier, however, we cannot assume from the subject’s identifying with the khandhas as a self, that the subject is in fact identified in any way with the khandhas as a self (or as part of a self); it is just that it seems to the subject (implicitly) as if it is identified with the khandhas as a self. When we consider (SVII) in context of the other suttas mentioned above, which make it clear that Buddhism is indeed concerned with the ordinary self we assume ourselves to be, it also apparent that Buddhism is
offering a minimal definition for selfhood in terms of identification. On this specification, the 'self' will minimally designate a subject who, through the process of identification, has assimilated various khandhas to itself as part of its perceived identity.

**Personal Ownership, The Self and Identification**

In the following very important sutta, the self, which we can now surmise must involve identification, is further defined in terms of **personal ownership** such that whenever there is a self or 'me' (implying identification), then there must be what is personally mine, viz., what is **personally owned**; and conversely, whenever there is what is **personally owned** or mine, there must be a me, viz., a **self** or personal owner (implying identification). This sutta links the notion of personal ownership to a self and hence, to the notion of identification:

(SVI) "Bhikkhus, there being a self, would there be what belongs to my self?"
"Yes, venerable sir."
"Or there being what belongs to a self, would there be my self?"
"Yes, venerable sir."

In this sutta, the self is co-defined in terms of an ownership relation with 'what belongs to my self'. The primary candidates for 'what belongs to my self' are the khandhas. It is natural to interpret the second half of (SVI), first of all, to be saying that if something were to belong to a self – to be personally possessed or owned by a self in other words – then the self would automatically assume the role of **personal owner** in relation to the owned item. The personal owner-self that the subject identifies itself as being, would, given the analysis of (SVII), amount to the subject as assimilated with all those khandhas that at a time seem to reflect to the self's role as personal owner of the personally owned item. They would hence, ipso facto, include (a) the collection of psycho-physical constituents that a subject identifies with globally as **me** (viz., as a self) and (b) those items that are specifically brought to bear upon

---

57 We are not entitled to conclude from (SVII) alone, that the key notion of self that concerns Buddhism is to be defined in terms of a subject's identification with the khandhas, and hence to be defined as the ordinary "run-of-the-mill" self a subject assumes itself to be (through these channels). On its own, (SVII) is quite consistent with a reading such that each subject is in fact an eternal metaphysical self which has nothing in essence to do with the khandhas, but which the subject, through ignorantly identifying with the khandhas as me (or as part of me), mistakenly assumes them to be (or be part of) this eternal self. It is not until we consider (SVII) in conjunction with other suttas such as (SVI), that it becomes clear that the self that concerns Buddhism is indeed that ordinary self which a subject assumes itself to be.
the global self’s particular relation to the owned item(s) in question\textsuperscript{58}. (SVI) suggests, then, that if there is a personal ‘mine’, then there must be a personal ‘me’; or put another way: if there is personal ownership of X, then the subject must be no less than a personal owner of X – which, as (SVI) makes clear, must amount to a self. As we saw from (SVII), a self involves identification, hence, the subject will identify as a personal owner, through identifying with those khandhas that would (should the self be real) reflect its status of a personal owner.

The first half of (SVI) seals the co-definition of ‘self’ with ‘personal owner’ (and hence ‘personal ownership’) stating that if there is a self, then that self must stand in a relation of personal ownership or myness to some X. We have already established (according to the second half of the sutta) that standing in a genuine (rather than merely apparent) relation of personal ownership to some X will guarantee that the subject not only identifies as, but is identified with – in a manner that it is identical to – the personal owner of X, which is the self. Hence, we can infer from the sutta that (a) whenever there is personal ownership towards some X, then there is a self (who is personal owner) and (b) whenever there is a self, then there is a personal owner (via the personal ownership relation). The self, we have inferred from (SVII) and (SVI), will involve identification on the part of the subject such that its assumption of being identical to a personal owner will in fact reflect a personal owner through the subject’s assimilation with the relevant khandhas. The self, whose existence implies personal ownership, will hence also imply the fact that there is identification (on part of subject) as a personal owner in such a manner that the self really is a personal owner. I shall henceforth refer to the a subject’s genuine identity with the self \textit{qua} any specified role (such as personal owner) as ‘self-identification’ or ‘self-identity’. Of course, whether a subject is identical to a such a self is a question at issue.

The upshot of this analysis, then, is that personal ownership in Buddhism delimits the self to be, at the very least, a personal owner, where a personal owner involves a subject whose assimilation with various khandhas reflects its overall self-identity with a personal owner. The subject, through self-identification, is at the very least a personal owner – and so is hence not a mere subject, as stipulated in our definition above. To have a sense of self, then, is (minimally) for a subject to reflexively feel as if it is identical to an owner-self.

\textsuperscript{58} For example, if the personally owned item (‘what belongs to a self’) were some trousers, then the personal owner (viz., the self) would be likely to include the subject’s assimilation with all those bodily-mental khandhas which reflect the specific impression that the subject is a ‘me-\textit{qua}-mind-body that owns the trousers’.
Should the subject turn out to be an owner-self, the sense of self must pertain to the owner-self sensed. Should the subject turn out not to be an owner-self, then the sense of self will not pertain to the owner-self sensed. When one has a sense of personal ownership towards X, then one has a sense of myness towards it, entailing a reciprocal sense of self-identification with the personal owner of X. We express this neutrally (in a way that doesn’t presuppose the self’s actual existence) by saying that one identifies as the personal owner of X, namely as a self.

Applying this analysis to the anosognosic, we can now say that when he feels that his arm does not belong to him, this means that he fails to identify as the personal owner of the arm. Or when he feels that the arm is not a part of who he is, this means that he fails to identify with that part of the body as a part of himself. In chapter four, we further explore the relation between assumed self-identification and the roles of perspectival and possessive ownership.

The Sense of Self (Through Reciprocal Senses of Personal Ownership and Self-identification) Co-arises with Tanha

It is clear that Buddhism regards a sense of a self to be something that leads to dukkha. The ideal mindset, as suggested in (SVIII), is one where a person regards the khandhas in such a way that “thoughts of 'me' or 'mine' or 'I am' do not occur to him”, such that they no longer play a role in his attitudes, motivations or choices. Given our analysis so far, we can surmise that a lack of 'mine'-thoughts amounts, specifically, to the lack of a sense of personal ownership toward the khandhas, and the reciprocal lack of 'me' or 'I am'-thoughts pertains to the subject’s no longer identifying with any of the khandhas as me or as a part of me. Having renounced a sense of self-identification and personal ownership toward the khandhas, where one no longer regards them in this manner to be me or mine, one is also “set loose” from dukkha (mental dukkha whilst physically alive; mental and physical dukkha when the body dies). Now we already know from the Third Noble Truth that Buddhism regards a mindset with tanha to lead to dukkha, and the cessation of tanha to lead to the cessation of dukkha. Given that a parallel story is true with the sense of self in relation to dukkha – in virtue of the subject’s reciprocal senses of personal ownership and self-identification – we can surmise that the relation between tanha and the sense of self will be very close. The two will, at the very least, co-arise together.
I am unsure as to whether the Buddha intended the relation between tanha and sense of self to be conceptually or empirically construed – or whether this empirical/conceptual distinction would even be considered relevant to the context of eliminating dukkha. For purposes of this project, however, it is best to treat the proposed relation as an empirical claim such that whenever there is a sense of self, there is tanha, and whenever there is tanha, there is a sense of self; where the aspects of self-sense most relevant to tanha are those pertaining to the subject’s implicit assumption of being the personal owner of various khandhas. Analysing the relation in this way opens up the possibility of its empirical investigation – a possibility we shall investigate later in the project. I anticipate that if the Buddhist analysis proves accurate – with the senses of self-identification and personal ownership co-arising with tanha and hence potential mental dukkha – then it will contribute substantially to the debate and discussion of ownership and identification in the West. For what it will provide is an effective empirical means of testing whether a person is identifying as a personal owner – and hence as a self – and if so, what items they are harbouring a sense of personal ownership or my-ness towards.

Suppose we limit our pool of potential items (towards which one might have a sense of personal ownership) to Xs that are perspectivally and/or possessively owned. The Buddhist claim would therefore be that there will also be a sense of personal ownership or self-identification with respect to item X iff one has tanha with respect to X (where the bi-conditional is regarded as indicative of empirical co-dependence). Determining if one has tanha with respect to X might be achieved in a number of ways, including imagining: “If X were to be lost or gained, would my happiness or (mental) suffering alter?” If the answer is “yes”, then one would likely harbour tanha and hence a sense of personal myness or self-identification with regards to X. If the answer is “no”, then one may not harbour tanha with respect to that X. In the case of the anosognosic with the foreign-feeling arm, the Buddhist analysis would predict that he harbours little or no tanha towards it, and is therefore emotionally indifferent to its fate, feeling no negative emotions if it were to be harmed. It is interesting to therefore note that indifference is mentioned by Nikolinakos (above) as a common feature of anosognosia, alongside the lack of ownership-feelings. Should the anosognosic suddenly become emotionally concerned again about the arm’s welfare, then Buddhism would predict a return of tanha towards it and hence a return of his sense of
personal ownership towards the arm – with the reciprocal sense of self-identification as its personal owner – and probably a elimination of the anosognosia with respect to that limb.

Where X as tanha-object is neither possessively nor perspectivally owned, such as the situation in Iraq, the right thing to say will probably be that one identifies as the personal owner of their tanha-feelings with regard to X. There may well be further notions of non-personal ownership upon which a sense of personal ownership might piggyback. Which ever way we decide to parse the details, it should be evident enough that the Buddhist analysis, if correct, will be of significant value to the Western discussion on ownership. It will also, if correct, offer a solid way of testing whether, or to what degree, the sense of self (evidenced by the subject’s reciprocal sense of self-identification and personal ownership towards khandhas) is being eroded through practice of the Noble Eightfold Path. It hence provides us with a potential objective means for testing whether the methods advocated by Buddhism, such as meditation, are in fact doing what they are claimed to do, namely, eroding the sense of self-identification. When put in this light, the practice of meditation sounds considerably less ethereal.

**Further Features Buddhism Ascribes to the Reflexively Assumed Self**

We have spent some time elaborating (although by no means completely) the Buddhist notion of self as centrally a personal owner – of which we can now take as a given, its attitude of reciprocal ownership towards the khandhas. We now consider some further aspects that Buddhism would attribute to the self we assume we are. We can immediately note that built into the notion of ‘personal owner’ is the central but implicit aspect of a witnessing subject who is identified with various khandhas at various times. ‘Witnessing subject’ can thus be incorporated into the definition of self that is implied by Buddhism. To have a sense of being a personal owner is for the witnessing subject to identify itself as a personal owner, through identifying with a collection of khandhas that help underpin the subject’s feelings of being a personal owner of the objects in question (which doesn’t entail that the subject is actually identified with the requisite khandhas). Now we may recall from chapter one that a feature of a subject – as opposed to objects – is that a subject, unlike objects, cannot, in principle, be the focus of attention (reflexive or otherwise). (This also applies to the subject’s modus operandi of witnessing). It stands to reason, therefore, that if a
subject identifies with various khandhas at a time, such that the khandhas seem assimilated to the subject, that the khandhas will not overtly appear as objects from the perspective of the witnessing subject. If the khandhas did overtly appear to the subject as objects – as able to be attended to – then how could they possibly seem identified with the witnessing subject in the first place? In view of this, it makes sense to suppose, on the Buddhist analysis, that a subject’s identifying itself with various khandhas at a time will be felt as a unity: the khandhas feeling somehow integral to – at one with – the subject who identifies with them as a self. We can hence ascribe the notion of ‘synchronic unity’ to that self Buddhism supposes we have a sense of being.

The subject’s assumed unified identity with the khandhas will, on our analysis of Buddhism, exert a curious effect upon the way that the khandhas (viz., objects of awareness) are apprehended – or more fittingly, upon the way they fail to be apprehended. We have noted that Buddhism regards the objects of the world – parsed into the psycho-physical khandhas – to be conditioned and hence impermanent (anicca pronounced ‘annie-cha’). Now if various khandhas, through being identified with by a subject, fail to be overtly noticed by the subject as objects of witnessing, then we can infer that they will also fail to be overtly noticed as anicca – as obviously coming and going from the subject’s field of witnessing. It will be as if the subject, assuming the khandhas to be self, becomes somehow change-blind to their coming and going. The skewed perception of impermanence will also imply, on the Buddhist analysis, that their status as causing dukkha when emotionally invested in (or as not being conducive to dukkha-free happiness) will fail to be apprehended with clarity.

Buddhism maintains that if one were to clearly apprehend the khandhas as anicca and therefore as dukkha, then there would be no motivation to emotionally invest in them. That we do emotionally invest in the khandhas, according to Buddhism, reflects our lack of clear apprehension of their nature as anicca and dukkha.

This inference about the shortfall in apprehending various khandhas as anicca and dukkha – due to a sense of self-identification – is borne out in the oft-cited Anattalakkhana sutta (S.X below), where the Buddha urges his followers to regard the khandhas as anicca, dukkha and anatta. To fully regard the khandhas as anatta, we can now surmise, is to regard them as not pertaining to a self, and hence, without the reciprocal senses of self-identification and personal ownership. While it is these aspects to the self-sense that Buddhism regards as
being linked to *tanha* and *dukkha* – and hence to be wisely got rid of – there is *more* to the Buddhist notion of self than just a personal witness-owner. The following sutta, besides bearing out the above inference, contains further central clues as to how Buddhism construes the self that we have a sense of being:

(S.X) I have heard that on one occasion the Master was staying at Varanasi, in the Game Refuge at Isipatana. There he addressed the group of five monks:

'Physical form, monks, is not the self. If physical form were the self, this body would not lend itself to dis-ease. One could get physical form to be like this and not be like that. But precisely because physical form is not the self, it lends itself to dis-ease. And one cannot get physical form to be like this and not be like that.

'Feeling is not the self... Perception is not the self... Mental processes are not the self...

'Consciousness is not the self. If consciousness were the self, this consciousness would not lend itself to dis-ease. One could get consciousness to be like this and not be like that. But precisely because consciousness is not the self, it lends itself to dis-ease. And one cannot get consciousness to be like this and not be like that.

'How do you construe thus, monks -- Is physical form constant or inconstant?' -- 'Inconstant, Lord.' -- 'And whatever is inconstant: Is it easeful or stressful?' -- 'Stressful, Lord.' -- 'And is it right to assume with regard to whatever is inconstant, stressful, subject to change, that "This is mine. This is my self. This is what I am"?' -- 'No, Lord.'

'... Is feeling constant or inconstant?... Is perception constant or inconstant?... Are mental processes constant or inconstant?...

'Is consciousness constant or inconstant?' -- 'Inconstant, Lord.' -- 'And whatever is inconstant: Is it easeful or stressful?' -- 'Stressful, Lord.' -- 'And is it right to assume with regard to whatever is inconstant, stressful, subject to change, that "This is mine. This is my self. This is what I am"?' -- 'No, Lord.'

'Thus, monks, any physical form whatsoever that is past, future, or present; internal or external; blatant or subtle, common or sublime, far or near: every physical form -- is to be seen as it actually is with right discernment as: "This is not mine. This is not my self. This is not what I am."

'Any feeling whatsoever... Any perception whatsoever... Any mental processes whatsoever...

'Any consciousness whatsoever that is past, future, or present; internal or external; blatant or subtle, common or sublime, far or near: every consciousness -- is to be seen as it actually is with right discernment as: "This is not mine. This is not my self. This is not what I am."'\(^59\)

---

\(^59\) In Thanissaro Bhikkhu (1993a, 79-80). *Samyutta Nikaya* chp XXII sec 59. A similar theme is repeated in suttas throughout the *Khandha Vagga* (*Samyutta* XXII-XXXIV).
From the imperative tone of this sutta, we can surmise that the Buddha would not be urging people to perceive the *khandhas* as impermanent (*anicca*), *dukkha* and without self (*anatta*) unless there was an existing tendency to regard them as otherwise. Our central question now is what, given the Buddhist picture, is most likely to account for such a tendency? From (S.X) we can gather that the tendency to not view *khandhas* as *anicca* and *dukkha* has to do with viewing them in terms of a self: the sense of self, through identification, is somehow obscuring our perception of the *khandhas* as *anicca* and *dukkha*. Could there be something about their identification as a self that makes it seem as if the *khandhas* lack *anicca* and *dukkha*? Indeed, there could something about the assumed self that itself seems to lack *anicca* and *dukkha*? The sutta provides a strong hint that the commonly assumed self does, in some important way, lack these attributes.

"Dis-ease" (meaning a lack of perfect ease) and "stress" are Thanissaro Bhikkhu's translations of *dukkha*. When the Buddha says of each of the *khandhas* that they are not a self because they lend themselves to *dukkha*, the implication, spelt out above and noted by Harvey (1995, 46), is that if some X were a self, then that X would not lend itself to *dukkha*. Not-*dukkha* can hence be surmised as being attributed, in some way, to the self that we assume we are (and we shall see that how 'in some way' gets spelt out is crucial to the analysis). A parallel story can be applied to *anicca*. The sutta suggests that it is not right to regard what is inconstant, stressful (*dukkha*) or subject to change as a self. The implication, also noted by Harvey (1995, 46), is that if some X were to be rightfully assumed as a self, then that X would not be inconstant or subject to change – not *anicca*, in other words. The implication is also, crucially, that most of us do take ourselves to be an entity that somehow lacks in *anicca* and *dukkha*; an assumption of self which is distorting our perception of those *khandhas* that are viewed as "mine", "myself" and "what I am". Finally, as Harvey also notes, the opening passage of (S.X.) suggests that if any of the *khandhas* were a self, then they would be under one's control, and hence not lend themselves to *dukkha* (on the assumption that we seek to avoid *dukkha*). Harvey suggests (with backing from further suttas) that the notion of control in this context be construed reflexively, implying that the controlled self is under its own control, making the self a controller or agent (1995, 49-50).
Now when using this derived information such as 'not dukkha' and 'not anicca' to help define that self-entity which Buddhism says we assume we are, we need to be careful. The most common step for scholars (a mistake that I confess to making in earlier drafts) has been to confidently ascribe the positive counterpart of these attributes to the self-entity as if this is the kind of entity Buddhism supposes we assume we are. Hence Harvey concludes his otherwise insightful analysis of the Anattalakkhana and other suttas with the statement:

It can thus be seen that the Self-ideal which early Buddhism worked with was of an unconditioned, permanent, totally happy 'I', which is self-aware, in total control of itself, a truly autonomous agent, with an inherent substantial essence, the true nature of an individual person (1995, 51).

If the reader thinks this sounds a little like nibbanic consciousness, he or she is not far wrong. Harvey notes the explicit similarity with nibbana, stating the only differences to be that the self, unlike nibbana, carries the identity of 'I am' or 'this I am', as well that of being a controller. Buddhism does not construe the nibbanic element as the “agent-controller of action, as the Self is seen to be” (1995, 53). Nibbana is also not construed as an individual, bounded entity; a corollary of it not being a locus of self-identity.

It is not my intention to criticise the general gist of Harvey’s comparison, since I think it is, generally speaking, on the right track. It is to Harvey’s credit that he makes the comparison between the assumed self and nibbana; a comparison that few scholars have acknowledged (and one we shall soon follow up). Yet it clearly needs refinement, since the self we assume we are, even if in some sense lacking anicca and dukkha, is clearly not an unconditioned, permanent, totally happy ‘I’. We only need reflect upon the amount of misery in the world, perhaps in our own minds, to see that an assumption of one’s self-nature as perfectly happy is not common amongst humanity. We only need reflect upon the widespread fear of death and non-existence to see that an assumption of being unconditioned or permanent is not ubiquitous. This does not rule out the possibility that somehow, underneath all the dukkha, the nature of the mind is that of perfect, unconditioned, nibbanic happiness. But if the nature of mind is perfect in this way, then it is not glaringly obvious. Assuming the Buddha was intelligent, it is therefore unlikely that he would have intended the commonly assumed ‘self’ to be construed as sounding so blatantly similar to nibbana. Indeed, the
Buddha would not have delineated a Noble Eightfold Path to the cessation to *dukkha* if he had thought that our assumed self-nature was already perfectly happy, as Harvey’s analysis would have us believe.

Now this is not to say that the Buddha did not, in other passages, explicitly caution against propounding notions of a self that *did* closely match the description of Harvey’s passage. But we need to understand the context in which such cautions were made. Around the time of the Buddha, it was apparently common for scholars and pandits to expound exotic theoretical views of a self (some of them similar to the Christian idea of an immortal soul), views that the Buddha regarded as leading to *dukkha* if dogmatically adhered to. The Buddha’s quest for pandits and scholars to renounce their theoretical views of the self should not distract us from the main force of his suttas on *anatta*; to lose that sense of self which almost all of us – theoreticians or not – are said to hold onto. Our current aim is to get a plausible picture of how Buddhism would construe this ‘commonly assumed’ self, a picture that does not make it sound obviously like *nibbana*, but which may yet include features that, given Buddhist psychology, are most parsimoniously explained by nibbanic consciousness.

From (S.X), we have clues that the assumed self is not *dukkha*, not *anicca*. I think that the most intelligible way to construe the self, on the Buddhist account, as being ‘not *dukkha*’ is, perhaps paradoxically, through looking to *tanha*. *Tanha* involves a misdirected urge to seek perfect happiness and avoid *dukkha* – misdirected, since the urge is sought to be satisfied through seeking ideal configurations of the impermanent *khandhas* (which leads to *dukkha*!). Misdirectedness aside, on the (Buddhist) assumption that the urge to seek happiness and avoid *dukkha* is itself universal amongst humankind, it makes the most sense to construe the ‘not-*dukkha*’ element of self as simply *that aspect which seeks happiness and avoids *dukkha***. When combined, as it perhaps must be, with the assumption that this urge can be properly satisfied through the *khandhas* (an assumption borne through not properly apprehending their impermanence) it is only natural to suppose that the subject will fail to properly construe the *khandhas* as conducive to *dukkha* – hence prompting the Buddha’s imperative in (S.X) to correct this bias by regarding the *khandhas* as *dukkha*.

Thanissaro Bhikkhu’s translation of (S.X) conveys *anicca* or impermanence as having two aspects to it, or at least as being describable in two ways, namely through *inconstance* and *being subject to change*. I take ‘inconstance’ to mean a lack of numerical
continuity, such that any conditioned object or its fleeting parts are prone to decay. I take ‘subject to change’ to mean that even if a conditioned object (such as a person) is conventionally designated to exist numerically over a period of time, its qualities, in virtue of its fleeting parts, will change or vary from one moment to the next. Now (S.X) hints that the self we assume we are is somehow construed as lacking in anicca and hence, as lacking in inconstance and changeability. And I suggest that there is a way of construing the self in this manner, which avoids an implausible casting as death-defyingly constant and unchangeable. The suggestion is as follows: from the viewpoint of the apparent self (which is actually the viewpoint of the witnessing subject which assumes itself to be a self) it is as if, from one conscious moment to the next, the self’s numerical existence is not gappy or perduring (meaning with different temporal parts). From one conscious moment to the next, it seems as if the self endures, rather than perdures. The qualitative invariably can be construed in a similar way: from the viewpoint of the apparent self, it is as if there is something essential about the self – its very me-making quality – which does not, from one conscious moment to the next, change in quality. It stays exactly the same. This can seem to hold despite the changing flux of khandhas that are over time being identified with as ‘me’ or as ‘a part of me’. For it can be supposed that such identification will seem, from the outset, to be effected by not a mere subject but by a qualitatively invariable khandha-independent self.

Now we are not of course claiming at this stage that this is indeed how we construe our selves: this argument is left to chapter four. The current claim is rather that such a construal makes good sense of how Buddhism would cast the self that we assume we are. On one hand, the construal manages to avoid the obvious implausibility of the self seeming to defy death or non-existence, while on the other hand it bestows the assumed self with enough permanence to plausibly interfere with one’s perception of identified-with khandhas as impermanent (namely, as inconstant and variable). For if we assume ourselves to be an entity which endures and does not change (from one conscious moment to the next), then it is natural to suppose that while the subject assumes various khandhas to be assimilated with this self-entity, namely as ‘me’ or ‘part of who I am’, those khandhas will, from the subject’s perspective, lose their appearance as impermanent. Fear of death may be explained as an oscillation between viewing the body as ‘me’ and viewing it as an object that will one day
perish. Buddhism makes no claim that the things a subject identifies with as ‘me’ will not vary from one moment to the next.

**What Nibbanic Consciousness Brings to the Sense of Self**

The self, on the most plausible analysis of Buddhism, thus far includes features of being a personal owner who seeks happiness and avoids dukkha and is permanent, viz., enduring in an invariable manner from one moment to the next. The sense of self (or assumption of self) on the Buddhist analysis is accordingly our subjective impression of being such an entity that includes these features. Our question is now: what, given Buddhist metaphysics, is most likely to account for this impression of endurance, invariability and a happiness-seeking urge, which comes through in the sense of self? The effect of *tanha*, along with the subject’s sense of identification with *khandhas* as a me-qua-personal owner, will no doubt exert some influence on this impression. Indeed, the urge to seek happiness (and avoid *dukkha*) from the *khandhas* is what it *means* to have *tanha*. Yet, on the Buddhist picture, we might still reasonably ask: from where does our very intuition or belief in an ideal of *dukkha*-free happiness arise? While we perpetually seek it from the *khandhas*, our very seeking it seems to suggest that we harbour a belief that such happiness is possible to attain. What accounts for this belief, given that the *khandhas*, on the Buddhist picture, can *never deliver* this ideal of *dukkha* free happiness? The subject’s identifying with *khandhas* as a unified *me* will also, on the Buddhist picture, reinforce the impression of the *self’s* (moment-to-moment) endurance and invariability, by making it seem as if the identified-with *khandhas*, as integral to this unobservable ‘me’, are lacking their native inconstancy and variability. Yet one can still reasonably ask, *on the Buddhist picture*: what accounts for the very impression of such endurance and invariability? Being inherently impermanent, the *khandhas* seem intrinsically unsuited to deliver such an impression. It may be highly significant that Hume, with his empirically based ontology of impermanence, admitted defeat when it came to explaining the impression of what he called “uninterruptedness” and “invariability” of the *self*.60

Unlike Hume’s system, Buddhism offers a ready answer to these questions. Note firstly that our claim here is not that answers cannot possibly be found in an ontology of

---

60 Thus Hume writes: ‘But all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought or consciousness. I cannot discover any theory, which gives me satisfaction on this head’ (1740, 175).
impermanent khandhas. The claim is rather that, against the background of Buddhist metaphysics, the most economical and obvious answer will involve an appeal to nibbanic consciousness. From chapter two it was determined that nibbanic consciousness, being timeless, cannot be anicca (unenduring and changeable); and, partaking in the nature of dukkha-free happiness, cannot be dukkha (where ‘dukkha’ can be understood as disposed to cause suffering and dissatisfaction). We have already identified the witnessing aspect, in Buddhist metaphysics, with pre-nibbanic consciousness. Pre-nibbanic consciousness, we can recall, is nibbanic consciousness as it appears through the filter of tanha and (we can now say) the sense of self. It therefore makes the most parsimonious sense to attribute the aspects of endurance and invariability, which seem inherent to a self, to the witnessing, viz., pre-nibbanic consciousness. On this picture, the nibbanic witnessing, through the subject, brings to the sense of self not only an impression of witnessing, but also that of endurance and invariability. We can surmise that the impression of endurance and invariability will be distorted enough by identification with time-bound khandhas to make it appear as if the self is enduring in time rather than timelessly. The nibbanic witnessing will also bring to the sense of self a deep-seated belief that a perfect, dukkha-free happiness is attainable. When distorted by tanha, it will manifest as an urge to seek such happiness from the khandhas, an urge that is perpetually frustrated (hence conducive to dukkha).

We might wonder if nibbanic consciousness will bring any further aspects to the sense of self on the Buddhist analysis. Earlier we mentioned synchronic unity; that feature whereby when a subject identifies with various khandhas as a part of itself, the khandhas are somehow felt to be unified with the identifying subject (and hence not viewed as overt objects). The resultant impression of self will thus reflexively convey, at a given time, a unified subject-entity. Now we can ask, as we did in the case of apparent permanence: just what, on the Buddhist picture, will best account for this impression of a unified self? The identified-with khandhas will likely reinforce an impression of me-ness in the unified entity. The impermanent, differentiated khandhas do not however seem to offer an obvious explanation for the impression of unity itself (and we can note that Hume admitted a failure to explain synchronic unity which he termed ‘simplicity’). Nibbanic consciousness, on the other hand, has already been pegged (in chapter two) as undifferentiated and hence unified, providing Buddhism with an obvious answer to the origin of the unity-impression.
‘synchronic’ unity will, on this picture, be inherent to the very witnessing which seems, like an in(di)visibility cloak, to enfold its unity around the identified-with khandhas.

We may finally wonder at whether the ‘unconditioned by relation’ aspect of nibbanic consciousness will reveal itself at the pre-nibbanic level. Is there any way in which the self we assume we are, on the Buddhist picture, would appear as somehow apart from, and not dependent upon, the khandhas? Indeed there is. Even with its assumption as self, the witnessing subject will, in some important sense, retain an impression of being independent of the witnessed objects (the khandhas), in a manner that reflects the basic subject-object division. This impression will not be lost during identification. That is because when the subject identifies itself with khandhas as ‘me’ or as ‘a part of me’ their statuses as separate, impermanent objects will, for reasons already explained, become effectively invisible to the subject. It will therefore appear to the subject (who assumes it is a self) that the self is, ontologically speaking, quite independent of any such objects. Now we are not saying that should the subject-self turn out to be real, it would in fact ontologically depend upon the khandhas that it identifies itself with. To the contrary (and this point shall be made clearer in chapter four), it may well seem, to the subject, as if its essential selfhood stands ontologically prior to any objects (viz., the khandhas) that it may assimilate to itself. It may hence seem to the subject as if a pre-existing self identifies itself with various khandhas, rather than as if the khandhas are helping to construct the essential self upon each new identification. On this picture – suitably attributed to Buddhism – identification will seem, from the viewpoint of the subject, to evidence rather than to construct an owner-self. While the self will not appear to be unconditioned by relation in an objective sense (hence as immortal or non-physical), it will nevertheless seem, so long as it exists, that its subjective or ‘internal’ existence is not constructed by any objects of witnessing, overt or covert. I shall henceforth refer to this feature, implicitly originating from nibbanic consciousness on the Buddhist picture, as unconstructedness.

Peter Harvey alludes to evidence from the suttas that directly support some kind of ‘unconditioned’ ascription to the self (1995, 46). However, as with the other attributes, Harvey does not temper this with the kind of modification needed if we are to plausibly assume ourselves to be such an entity.
The Definition and Ontological Status of Self in Buddhism

We are almost ready to offer a definition of that self which Buddhism claims most people to have a sense of being. Before we do this, two further points need mentioning. The first point pertains to a corollary of ‘identification as a personal owner’, which will later prove central to our analysis of the self. When the subject assumes itself to be a personal owner-self, through identification with khandhas, it is assuming itself to be a unified entity with a boundary. While boundedness (where ‘the self’ ends and ‘the other’ begins) will not be directly observable by the subject qua ‘personal owner’, it will be clearly evidenced through the fact that some khandhas are identified with as ‘me’, while other khandhas are viewed as ‘not me’. Boundedness underscores a self/other distinction, casting the self as an ontologically distinct entity. The relationship between boundedness and self-identification will be spelt out further in chapter four. The second point pertains to the (non-intentional) phenomenal dimension that is deemed inherent to the witnessing of nibbanic consciousness. Does it come across in pre-nibbanic witness-consciousness, through the ordinary sense of self? Since I can find no explicit evidence from the suttas that it does, I shall for now leave it out of the proposed definition. In a chapter four, however, it will be argued that a non-intentional phenomenal dimension to witnessing can be easily inferred from the definition of self that I have extracted from Buddhism. Without further ado, here is that definition:

A self is defined as a bounded, happiness-seeking/dukkha-avoiding witnessing subject that is a personal owner and controlling agent, and which is unified and unconstructed, with moment-to-moment endurance and invariability.

This is the entity that Buddhism supposes that we – or most of us – assume we are. Is such an entity held to exist in Buddhism? The assumption or sense of being such an entity is certainly held to exist: it is after all that thing, along with co-arising tanha, that is deemed to perpetuate dukkha – and something be got rid of, along with tanha, if dukkha is to cease. But the entity itself – the self that we think of as me – has no place in Buddhist ontology. This is perhaps most clearly expressed in the following oft-cited verses, where ‘dhammas’ refers to all of reality, both conditioned and unconditioned:

(S.XI) ‘All conditioned things are anicca (inconstant, changeable)’
When one sees this with discernment
And grows disenchanted with dukkha,
This is the path to purity.

‘All conditioned things are dukkha (stressful)’
When one sees with discernment
And grows disenchanted with dukkha
This is the path to purity.

‘All dhammas are anatta (not-self)’
When one sees with discernment
And grows disenchanted with dukkha
This is the path to purity.\(^{62}\)

The first two verses allude only to conditioned things, namely the khandhas, and we should by now have enough background in Buddhism to interpret their gist. Of current relevance is the third verse: “All dhammas are anatta”. Now that we finally know what Buddhism means by ‘self’, we can inject some meaning into the inference that Buddhist ontology, while having room for that which is not anicca or dukkha (namely nibbanic consciousness) does not allow for the hybrid entity of self. To see all dhammas as anatta thus means seeing (including reflexively seeing) everything as not self – the khandhas and nibbanic consciousness alike. It is not as if losing the propensity to see the conditioned khandhas as (personally) ‘me’ or ‘mine’ implies a fall-back position of seeing nibbana as ‘me’ or ‘mine’. In Buddhism, the impression of me or mine with regards to anything at all not only leads to dukkha, but involves a fundamental misapprehension of reality, since there is nothing in the entire cosmological system, conditioned or unconditioned, which answers to ‘me’ or ‘mine’.

It is of utmost importance to understand that the Buddhist rejection of this self-entity from their ontology does not imply their equal denial of reality to each and every feature attributed to the self. From the account we have presented, it should be fairly obvious that to the extent they are not contributed to by identification and tanha, those features of witnessing, unity, endurance, invariability, unconstructedness and the sought-after happiness-ideal, will not, in themselves, be considered to lack reality. We have argued that it is most parsimonious, on the Buddhist picture, to suppose that these features are imported into the sense of self by the ever-present nibbanic consciousness. The features will only lack reality

\(^{62}\) Thanissaro Bhikkhu, (1993b). From the Dhammapada nos. 5, 6, 7 of chapter XX (or verses 277, 278 and 279). I insert the Pali words in lieu of his English translations.
to the extent that they are unwittingly and reflexively attributed to a self and are somewhat distorted or muted (in their impression) through this process of attribution (viz., identification). The lack of the self’s reality will pertain to the subject’s purported status of being a bounded personal owner and agent whose independent existence seems to be reflected through the ubiquitous impressions of ‘me’ and ‘mine’ (including ‘me’ as the agent of actions). On the Buddhist account, moreover, whenever the subject identifies as a bounded, personal owner-self, there must be tanha in the subject’s mindset; in a slogan: tanha is present if and only if a sense of self is present. The upshot of this Buddhist picture is that the sense of self, while not grounded in a real self, is contributed to by elements of which some have reality (being sourced in nibbanic consciousness) while some do not (being distorted by tanha).

Given this picture, the purpose of the Noble Eightfold Path of Practice can be understood, in Buddhism, to selectively whittle away those tendencies that have the intrinsic nature of mind – nibbanic consciousness – somehow duped into assuming it is a bounded self-entity. What are whittled away will not be the contributing sense of unity or permanence (etc.), but only those tanha-involving tendencies towards identifying with khandhas as ‘me’ and ‘mine’, tendencies that have the subject believing it is a bounded personal owner and agent. Only when those tendencies have been completely eradicated will it be known with full discernment that “All conditioned things are anicca”, “All conditioned things are dukkha” and that “All dhammas are anatta”.

This completes our analysis of how Buddhism defines the ordinary sense of self and how its ontology, while having no room for that self we have a sense of being, still acknowledges the contribution nibbanic consciousness must make to the sense of self. It is fitting to conclude this section with some passages from Shankara (who lived at around 788-820 CE), the ‘sage’ or ‘Arahan’ historically associated with the tradition of Advaita Vedanta. I have argued elsewhere (Albahari, 2002) that by “Atman” and “Self” Shankara means something like ‘nibbanic consciousness’, which partially comes across in the following passages. This apparently striking similarity between Advaita Vedanta (AV) and Buddhism has been very helpful in my interpretation of the Buddhist suttas, since AV makes plain the metaphor of nibbanic witness-consciousness as ‘covered’ – a metaphor that I have,
indeed, borrowed from this tradition. Advaita Vedanta will be alluded to again in later chapters. Here, then, are the passages:

That Reality [Atman] sees everything by its own light. No one sees it. It gives intelligence to the mind and the intellect, but no one gives it light...Its nature is eternal [timeless] consciousness... It never ceases to experience infinite joy...It is the knower of the activities of the mind and of the individual man. It is the witness of all the actions of the body, the sense organs and the vital energy. It seems to be identified with all of these, just as fire appears identified with an iron ball. But it neither acts nor is subject to the slightest change. The Atman is birthless and deathless. It neither grows nor decays. It is unchangeable, eternal. It does not dissolve when the body dissolves. Does the ether cease to exist when the jar that enclosed it is broken? (1968, 62-63).

From the reference to the ‘iron ball’ we can surmise that AV maintains that the pure nature of nibbanic consciousness – Atman – is, through ‘ignorance’ of how things are, ‘covered’ by tanha and self-identification, but is yet partially revealed through the witnessing. The aim of spiritual practice, as in Buddhism, is to free pure witness-consciousness from this “bondage”:

Man is in bondage because he mistakes what is non-Atman for his real Self. This is caused by ignorance. Hence follows the misery of birth and death. Through ignorance, man identifies the Atman with the body, taking the perishable for the real...He becomes enmeshed in the things of the senses like a caterpillar in the threads of its cocoon...Wrapped in its five coverings, beginning with the physical...the [pure] Atman remains hidden, as the water of a pond is hidden by a veil of scum. When the scum is removed, the pure water is clearly seen. It takes away a man’s thirst [tanha], cools him immediately and makes him happy. When all five coverings are removed, the pure Atman is revealed...as unending, unalloyed bliss...The Atman dwells within, free from attachment and beyond all action. A man must separate his Atman from every object of experience, as a stalk of grass is separated from its enveloping sheath (1968, 64, 67).

The Misportrayal of Buddhism as Endorsing a ‘Bundle Theory’ of Persons

I have already suggested that no scholars I know of have portrayed the Buddhist notion of anatta – at least in its ontological formulation – in an adequate way. This is partly because the notion of self alluded to in Buddhism has been poorly defined (for example, Harvey, op. cit.) and partly because scholars have tended to focus on the anatta suttas without
considering their relation to the suttas on nibbana in the broader context of the Four Noble Truths. And as I have argued elsewhere, an almost religious urge by many scholars to separate Buddhism from Upanisadic teachings – a separation that I do not think is warranted – has contributed in a major way to this trend. The upshot is that many scholars – enough to make this the dominant position – have depicted the Buddha as not only rejecting the existence of a self, but also as rejecting key features attributed to the self such as unity, endurance, invariability and even that which knows experience (hence the witnessing). With the rejection of these features, the scholars depict the Buddha as holding that a person’s reality consists of nothing more than the five impermanent, dukkha-ridden khandhas (translated as ‘aggregates of conditioned existence’). Hence it is not uncommon to see deep parallels being drawn between Buddhist and Humean ontology of persons-as-thoroughly-impermanent. Here are three such examples from contemporary literature:

[Hume concludes] that ‘[persons] are nothing but a bundle or collection of different perceptions which succeed each other with an inconceivable rapidity, and are in perpetual flux and movement.’...Buddhist writers typically make the same point by analysing a person into the ‘Five Aggregates’ [khandhas]. Since a person is nothing more than the sum of these five aggregates, and since soul, in the sense of a permanent unchanging subject of consciousness [viz. the self], cannot be identified with one or more of the five, soul cannot exist...It seems clear that Hume and the Buddhists say the same thing for the same reasons: both analyse the ‘soul’ [viz. sense of self] into a series of events or processes, and do so because this is what experience reveals. (A.H. Lesser, 1979, 58).

Moment by moment new experiences happen and are gone. It is a rapidly shifting stream of momentary mental occurrences. Furthermore, the shiftiness includes the perceiver as much as the perceptions. There is no experiencer, just as Hume noticed, who remains constant to receive experiences, no landing platform for experiences...Suffering arises quite naturally and then grows as the mind seeks to

---

63 See Albahari (2002). I have argued that many scholars, such as those cited in this section, confuse a common ‘eternalist’ misreading of Upanisadic teachings on Atman, a misreading prevalent around the time of the Buddha, with the Upanisadic teachings themselves. When interpreted correctly – for example as Shankara interprets them (1968) – then it becomes hard to defend any grounds upon which to separate the Upanisadic notion of Atman from that of nibbanic consciousness.

64 Some notable exceptions to this stance include Peter Harvey (1995), Thanissaro Bhikkhu (1993a), (1993b), Christian Lindtner (1999), Jonathan Shear (1996), Arthur Deikman (1996) and Robert Forman (1998). These authors do not depict Buddhism as construing a person’s mind to be thoroughly impermanent – although they do not link nibbanic consciousness to anatta in the manner carried out here.

65 Be reminded that I use the term ‘person’ as I have defined it in chapter one and that this notion of person is relevant to the current discussion.
avoid its natural grounding in impermanence and lack of self (Varela, Thomson and Rosch, 1991, 60-61).

Although the Buddha cites various characteristics that something must have if it is to be considered a self, the most important is that of permanence and identity over time. But when we look to our experience, there is nothing but impermanence: our bodies, feelings, and thoughts are forever coming and going. In this sense the Buddha is in complete agreement with Hume: where there is diversity there can be no identity. (James Giles, 1993, 186).

Further examples, while not making explicit comparisons with Hume’s ‘bundle-theory’ of persons, reveal that this ‘Humean’ interpretation of Buddhism is widespread. Amongst the following scholars to be quoted, Rahula, Kalupahana and Gethin are well known; their work often consulted by other scholars or put on reading lists for courses in Buddhism:

Buddhist thought presents these five aggregates as an exhaustive analysis of the individual. They are the world for any given being – there is nothing else besides (Rupert Gethin, 1998, 136).

The whole human personality, according to Buddhism, is nothing more than the effectively functional psycho-physical organism. The whole endeavour of the Buddha and Buddhism is to make one realise one’s own personality and existence in terms of these unenduring and dependently arisen factors...(H.S. Prasad, 2000, 139, note 3).

What we call a ‘being’, or an ‘individual’, or ‘I’, according to Buddhist philosophy, is only a combination of ever-changing physical and mental forces or energies, which may be divided into five groups or aggregates (pancakkhandha) (Sri Walpola Rahula, 1996, 20).

Therefore, he [the Buddha] undertook the task of redefining the concept of man. According to him, this was merely a “bundle of perceptions” (sankharapunja) or a group of aggregates (khandha), not discrete and discontinuous, but connected and continuous by way of causality, a ‘bundle’ (kaya) which, for the sake of convenience, is designated by such names as Sariputa and Moggallana. (David Kalupahana, 1976, 39, note 4).

The following exchange between the monk Nagasena and the King Menanda, or a shorter sutta along the same vein, is sometimes cited in support of this interpretation:

(S.XII)...The Venerable Nagasena replied: "I am known as Nagasena, your majesty, and that is what my fellow monks call me. But though my parents may have given me
such a name, it is only a generally understood term, a practical designation. There is no question of a permanent individual implied in the use of the word."

King Menander was perplexed. "How can you declare," he said, "that there is no permanent, individual identity implied in the use of the name, 'Nagasena'?"

Nagasena asked, "How did you come here -- on foot, or in a vehicle?"

"In a chariot."

"Then tell me, what is the chariot? Is the pole the chariot?"

"No," replied the King.

"Or the axle, wheels, frame, reins, yoke, spokes, or whip?"

"None of these things is the chariot," the King replied.

"Taking the separate parts all together and laying them down on the ground, side by side -- is this the chariot?"

"No again," said the King.

"Then is the chariot something separate from all its parts?"

"No, your reverence."

"Then for all my asking," concluded Nagasena, "I can find no chariot. The word 'chariot' is a mere sound. You know what the word chariot means. And it is just the same with me. It is on account of the various components of my being that I am known by the practical designation 'Nagasena'." 66

The purpose of this passage is to point out the basis upon which one thing is to be distinguished from another, allowing us, in practical contexts, to designate various objects with the use of conventional names. It points out that in the case of persons, we must not be led to believe any "permanent individual identity" (viz., a self-entity) lurks behind the use of such proper names as 'Nagasena' -- any more than a chariot-essence lurks behind use of the term 'chariot'. It is on account of the khandha-components, rather than some permanent individual essence, that we are able to designate, for practical purposes, objects such as chariots and persons by their conventional names. Now we should note that nibbanic consciousness, being undifferentiated, would play no specific role in the designation of one thing from another, and so there would be no need to allude to it in the context of (S.XII). Yet if scholars of Buddhism have failed to properly acknowledge the role that nibbanic consciousness does play in our overall constitution as persons (through the witnessing aspect), then it is easy to misinterpret (S.XII) as being about the holus-bolus nature of persons, reading into it the idea that we are nothing but the five conditioned aggregates. But

66 I have only been able to locate this longer version of the sutta on the following website:
http://www.trinity.edu/madeau/Asian%20Religions/Lecture%20Notes/Mahayana%20Buddhism/Buddha%20Nagasena.htm. According to Harvey (1995, 35-36), this is a well-known passage, which appears in the Milindapanha. 25. A extract from this version appears in the Samyutta Nikaya chp 1 sec 135 and has been quoted by Kalupahana (1976, 39).
if we look carefully, (S.XII) is not saying this at all: it is merely pointing out that the conditioned khandhas, as opposed to individual essences, are what allow us to designate persons in a practical context. The misinterpretation of (S.XII) typifies the kind of trap scholars fall into when, upon determining the Buddhist account of a person’s nature, they ignore the suttas on nibbana. When portrayed as endorsing a ‘bundle theory’ of persons, Buddhism is sorely misrepresented in the West.

Conclusion

In this and the previous chapter, I have attempted to elucidate the Buddhist position on a particular notion of consciousness, relating this notion to its well-known doctrine of no-self (anatta). In summary, this kind of consciousness, not to be confused with that from the five conditioned khandhas, is of an intrinsic unconditioned, witnessing nature that I have termed ‘nibbanic consciousness’. By its nature nibbanic consciousness, on the Buddhist position, is intrinsic to all witnessing, but its nature as unconditioned, we can surmise, will not be obvious to most people. While the ‘coverings’ of tanha and identification are the most effective in obscuring its unconditioned nature, this nature is nevertheless partially revealed (on my interpretation of Buddhism) through the ordinary sense of self, in the features of witnessing, unity, relative permanence, unconstructedness, and in the ideal of a happiness that is sought after. I have allocated the term ‘pre-nibbanic consciousness’ to describe, on the Buddhist position, the witnessing and its various features as partially revealed through the sense of self. Pre-nibbanic consciousness is hence nibbanic (witness)-consciousness that reflexively assumes it is a self – or, to put it another way, nibbanic consciousness as ‘covered’ or ‘duped’. For given Buddhist ontology, this assumption on the part of nibbanic consciousness will be a mistaken one, one that must be eliminated if dukkha is to cease. The elimination of this assumption from a person’s mindset will mean that the innate nibbanic witnessing, no longer assuming it is a self, will reflexively know its nature as ‘uncovered’ nibbanic witness-consciousness, and hence, without any sense of ‘me’ or ‘mine’. Lacking this extrinsic ‘covering’, it will no longer be pre-nibbanic consciousness, but proximate nibbanic consciousness (when objects are apprehended) or pure nibbanic consciousness (when objects are not apprehended). Pure nibbanic consciousness is intrinsic to all forms of
witnessing – ‘pre’ and ‘proximate’ – the difference lying in the extrinsic ‘coverings’ of khandha-formations, such as those pertaining to tanha and identification.

Now we may be wondering whether such a metaphysically loaded notion of ‘witnessing’ as nibbanic consciousness – or the sense of self as contributed to by nibbanic consciousness – will have an application in Western philosophy. With its dominant physicalist paradigm, Western philosophy is not likely to embrace anything so radical as unconditioned nibbanic consciousness (which could not possibly be identical to anything physical). In chapter one I promised that the arguments of this project would not, aside from exegetical purposes (such as to clarify the Buddhist position), rely upon the adoption of what is at least highly contentious about Buddhist metaphysics. What I propose, therefore, is that we begin our inquiry by treating the Buddhist notions of ‘witness-consciousness’, ‘self’ and ‘sense of self’ as metaphysically and ontologically neutral, with a view to advancing independent arguments for some aspects of the Buddhist position on consciousness and no-self. I suggest, for a start, that we limit the concept of witness-consciousness to the notion that was set out in chapter one and re-iterated in this chapter under the definition of ‘subject’. This definition does not require that we treat the witnessing as nibbanic. Rather, it leaves open the possibility that witnessing might be nibbanic, through its lack of explicit commitment of witnessing to such categories as ‘state’ or ‘activity’. We have seen that nibbanic witness-consciousness cannot be regarded in any intrinsic sense to be a state or activity of a person, any more than the sun can be regarded as literally rising or setting. The common impression of witness-consciousness as a state or activity of a person would be regarded, in Buddhism, to be a mistaken impression. The implication of such a possibility for this project is that we shall not be letting any pre-established grammatical categories dictate our inquiry into the nature of witnessing. To suppose that ordinary witnessing must be intrinsically a state, for instance, would be to beg the question against the very possibility that ordinary witnessing may in fact be pre-nibbanic consciousness, and hence, identical in its essence to unconditioned nibbanic consciousness.

The metaphysically neutral notion of witnessing as set out in chapter one can be imported into the Buddhist definition of the self, namely, that self which Buddhism alleges we have a sense of being. We shall hence, upon starting our inquiry, leave open the question as to whether the witnessing, which appears in the definition of self, is in fact pre-nibbanic
witnessing or whether it is something more mundane. Upon starting our inquiry we leave open the question as to whether unity, endurance and invariability (for example), perhaps 'piggybacking' on the witnessing, are real aspects attributed to a self, or whether they, perhaps along with the other features, lack any reality whatsoever. We leave open the question as to whether this self as a whole exists, or whether it does not exist. We leave open the question as to the features deemed 'pre-nibbanic' in Buddhism (unity, endurance etc) do in fact 'piggyback' upon the witnessing in some conceptual sense – whether real or not. And we leave open the question of whether we commonly do, in fact, assume ourselves to be a self-entity of the description implied by Buddhism. This final question will be tackled in our next chapter – and the other questions (apart from the first), in the chapters that follow.

The aim of this project is hence to show that the Buddhist account of consciousness and no-self, when treated with relative metaphysical neutrality, is a highly plausible account that deserves independent recognition in the philosophy of mind. I hope to show that witness-consciousness is conceptually tied to the features regarded as 'pre-nibbanic' in Buddhism (although we shall not regard them as pre-nibbanic) – and that this unified, enduring, invariable (etc) witness-consciousness is real in and of itself, mistakenly identifying itself as a self-entity. I will argue that Buddhism is right to suppose that tanha and identification, given their relation to unified, enduring (etc.,) witness-consciousness, are what undermine the reality of the self that we assume we are. Before we can progress with such an aim, however, we need demonstrate that we commonly do, indeed, take ourselves to be an entity of the description that Buddhism alludes to as 'the self'. To this chapter, we now turn.
Chapter Four: The Reflexively Assumed Self

An East-West Convergence on Description of Self

Our goal so far has been to clarify the notion of self that has been implicated in the Buddhist suttas. It is a self that Buddhism supposes most people to have a sense of being and, furthermore, a self that as a whole has no reality on the Buddhist metaphysical system (although a subset of its ascribed features, sourced in nibbanic consciousness, do have reality). In this chapter, we investigate whether most subjects plausibly do, in fact, reflexively identify themselves as such selves. We hence look to whether we have a sense of being the kind of entity alluded to in Buddhism. If we do indeed harbour this sense of self, as this chapter will argue, then we can in later chapters address the ontological question: are we really the self-entity that we assume we are? If the self is real, then this sense of self will reflect a real self that is sensed. If the self lacks reality, then our sense of self will not reflect an actual self and will have to be explained by factors that do not appeal to the self in their ontology.

To recapitulate, the self alluded to in Buddhism is as follows:

A bounded, happiness-seeking/dukkha-avoiding, witnessing subject that is a personal owner and controlling agent, and which is unified and unconstructed, with moment-to-moment endurance and invariability.

How do we go about discovering whether we assume ourselves to be such an entity? Such an assumption is not likely to be introspectively obvious. As we mentioned in chapter three, the purported self conveys the impression of being essentially subject-like rather than object-like: it will therefore seem, as a whole, to elude introspective attention. We must therefore use other techniques, such as looking to common patterns of behaviour, linguistic practices, phenomenal descriptions, emotions, motivations and philosophical puzzles that may arise. A further highly useful source of evidence, although this will be indirect evidence, will lie in passages from Western philosophers who have sought, for various reasons, to extract notions of the self we assume we are. Suppose that the emerging notion of ‘ordinary’ self that
Western philosophers are concerned with coincides, for the most part, with the notion that concerns Buddhism. This will serve as strong initial (although indirect) evidence that Buddhism is indeed alluding to an assumption of self that is, as they claim, common to humanity, rather than something which turns out to be peculiar to their tradition.

Of the philosophers I shall quote, some are explicitly setting out to defend their description of self, while others are just assuming it. Yet others are describing the self as part of an attempt to defend or refute claims about its existence. This is all fine – so long as we keep in mind that such further ontological questions are not the current concern. While these philosophers emphasise different aspects of the self, I will suggest that an overall picture of 'self' will emerge that includes and/or expands upon the description that I gleaned from the Buddhist literature. From these quotations, I shall therefore extract a list of key, defining features ascribed to the self. Following this, I will outline each feature, arguing that each is commonly ascribed to that entity which most of us take ourselves to be identical to. Some of these arguments will be developed further in later chapters. Here then are the selected quotations beginning, fittingly, with the philosopher who made famous the notion of 'res cogitans':

Rene Descartes:

I am a thinking (conscious) thing, that is, a being who doubts, affirms, denies, knows a few objects, and is ignorant of many, -- [who loves, hates], wills, refuses, who imagines likewise, and perceives; for, as I before remarked, although the things which I perceive or imagine are perhaps nothing at all apart from me [and in themselves], I am nevertheless assured that those modes of consciousness which I call perceptions and imaginations, in as far only as they are modes of consciousness, exist in me. (1641, Meditation III, 42).

...And certainly the idea I have of the human mind in so far as it is a thinking thing, and not extended in length, breadth, and depth, and participating in none of the properties of body, is incomparably more distinct than the idea of any corporeal object..., (1641, Meditation IV, 63).

David Hume:

But self or person is not any one impression, but that to which our several impressions and ideas are supposed to have a reference. If any impression gives rise to the idea of self, that impression must continue invariably the same, through the whole course of our lives; since self is supposed to exist after that manner. ... There is
properly no *simplicity* in [the mind] at one time, nor *identity* in different, whatever natural propension we may have to imagine that simplicity and identity. ... What then gives us so great a propension to ascribe an identity to these successive perceptions, and to suppose ourselves possessed of an invariable and uninterrupted existence through the whole course of our lives? (1739, 161-162, 163).

**John Locke:**

*Personal Identity in Change of Substances.* That this is so, we have some kind of evidence in our very bodies, all whose particles, whilst vitally united to this same thinking conscious self, so that we feel when they are touched, and are affected by, and conscious of good or harm that happens to them, are a part of ourselves, i.e. of our thinking conscious self. Thus, the limbs of his body are to everyone a part of himself; he sympathises and is concerned for them. Cut off a hand, and thereby separate it from that consciousness he had of its heat, cold, and other affections, and it is then no longer a part of that which is himself, any more than the remotest part of matter. Thus, we see the substance whereof personal self consisted at one time may be varied at another, without the change of personal identity; there being no question about the same person, though the limbs which but now were a part of it be cut off... (1690, 213).

*Person,* as I take it, is the name for this self..... It is a forensic term, appropriating actions and their merit, and so belongs only to intelligent agents, capable of a law, and happiness, and misery. This personality extends itself beyond present existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to itself past actions, just upon the same ground and for the same reason as it does the present. All which is founded in a concern for happiness, the unavoidable concomitant of consciousness; that which is conscious of pleasure and pain desiring that the self that is conscious should be happy. And therefore whatever past actions it cannot reconcile or appropriate to that present self by consciousness [memory], it can be no more concerned in than if they had never been done... (1690, 220).

**William James:**

If the stream [of consciousness] as a whole is identified with the Self far more than any outward thing, a *certain portion of the stream abstracted from the rest* is so identified in an altogether peculiar degree, and is felt by all men as a sort of innermost centre within the circle, of sanctuary within the citadel, constituted by the subjective life as a whole. Compared with this element of the stream, the other parts, even of the subjective life, seem transient external possessions, of which each in turn can be disowned, whilst that which disowns them remains. Now, *what is this self of all the other selves?*

Probably all men would describe it in much the same way up to a certain point. They would call it the *active* element in all consciousness; saying that whatever qualities a
man's feelings may possess, or whatever content his thought may include, there is a spiritual something in him which seems to go out to meet these qualities and contents, whilst they seem to come in to be received by it. It is what welcomes or rejects. It presides over the perception of sensations, and by giving or withholding its assent it influences the movements they tend to arouse. It is the home of interest, not the pleasant or the painful, not even pleasure or pain, as such, but that within us to which pleasure and pain, the pleasant and the painful, speak. It is the source of effort and attention, and the place from which appear to emanate the fiats of the will... Being more incessantly there than any other single element of the mental life, the other elements end by seeming to accrete round it and to belong to it. It becomes opposed to them as the permanent is opposed to the changing and inconstant. (1890, 297-298).

Gilbert Ryle:

When a child, like Kim, having no theoretical commitments or equipment first asks himself ‘Who or What am I?’, he does not ask it from a desire to know his own surname, age, sex, nationality, or position in the form. He knows all his ordinary personalia. He feels that there is something else in the background for which his ‘I’ stands, a something which has still to be described after all his ordinary personalia have been listed. He also feels, very vaguely, that whatever it is that his ‘I’ stands for, it is something very important and quite unique, unique in the sense that neither it, nor anything like it, belongs to anyone else. There could only be one of it. (1966, 31).

Daniel Dennett:

...for ourselves, it seems, consciousness is precisely what distinguishes us from mere "automata". Mere bodily "reflexes" are "automatic" and mechanical; they may involve circuits in the brain, but do not require any intervention by the conscious mind. It is very natural to think of our own bodies as mere hand puppets of sorts that "we" control "from inside"... There are notorious problems with this idea, but that does not prevent it from seeming somehow right: unless there is a conscious mind behind the deed, there is no real agent in charge. When we think of our minds this way, we seem to discover the "inner me," the "real me." This real me is not my brain; it is what owns my brain ("the self and its brain"). (1991, 32).

Richard Baron:

The self which this paper rejects is not the individual organism... It is the self as one who thinks and acts... there is not really a knowing self, observing the world and learning. There is only an organism with its brain being etched by its encounters with the rest of the world. (2000, sec 2.0) ... Denying the reality of self means... not accepting that there is anything special about a particular individual: someone else with the same properties would be just as good... If you look at a person and see only the organism, characterised by its properties, you will not notice if another qualitatively identical person is substituted... there would be nothing special about the individual as such. (2000, sec 4.0) ... We must distinguish between the feeling that
one is a self and the fact (or rather fiction) that one is a self. We can feel many things to be the case even though they are not the case. (2000, sec 4.1) ... the sense of self is taken by the person feeling it to prove that there is a real, enduring, singular self... (2000, sec 6.1).

John Canfield:

What is our very idea or conception of the I'? (1990, 19) ... Something that is not the body but owns it; something that perceives, thinks, and wills; something that persists over time. This is the core conception of self virtually everyone has or presupposes. By saying we presuppose it, I mean that our way of talking about ourselves and our beliefs about ourselves imply this view; or at least, we would normally take it to be true, on reflection, that there is such an implication. (1990, 20).

Galen Strawson:

What, then, is the ordinary, human sense of self, in so far as we can generalise about it? (1997, 407) ... I propose that the mental self is ordinarily conceived or experienced as: (1) a thing, in some robust sense (2) a mental thing, in some sense (3,4) a single thing that is single both synchronically considered and diachronically considered] (5) ontically distinct from all other things (6) a subject of experience, a conscious feeler and thinker (7) an agent (8) a thing that has a certain character or personality. (1997, 407-408).

Peter Strawson:

...we have not only the question: Why are one’s states of consciousness ascribed to anything at all? We have also the question: Why are they ascribed to the very same thing as certain corporeal characteristics, a certain physical situation, etc? It is not to be supposed that the answers to these questions will be independent of one other (1959, 89-90).

Owen Flanagan:

...an “I” that stands behind all conscious experience [and]... constitutes the core of the self, our conscious control centre, the source of all action plans, and the agent for whom all experiences accrue before being filed for future reference or discarded... (1992, 177) ... the mind’s “I” is an illusion. ... The illusion is that there are two things: one side, a self, an ego, an “I,” that organizes experience, originates action, and accounts for our unchanging identity as persons and, on the other side, the stream of experience (1992, 178).

Jonathan Shear:

---

67 The paper by Baron is unpublished and only available online (a fact that I verified by e-mail). The paper shall be referred to again in a later section of this chapter.
In the West, philosophers such as Descartes, Hume, Kant and Sartre have been utterly unable to come to anything even approximating agreement as to whether there is, or even possibly could be, anything in our experience which could fulfil our commonsensical conception of self as the unitary conscious subject of one’s thought and experience (1996, 359).

Antonio Damasio:

Besides those images [of what you perceive externally] there is also this other presence that signifies you, as observer of the things imaged, owner of the things imaged, potential actor on the things imaged. There is a presence of you in a particular relationship with some object. ...The presence is quiet and subtle, and something little more than a “hint half guessed,” a “gift half understood,” to borrow words from T.S. Eliot... (1999, 10). In all the kinds of self we can consider one notion always commands centre stage: the notion of a bounded, single individual that changes ever so gently across time, but, somehow, seems to stay the same... (1999, 134). Whether we like it or not, the human mind is constantly being split, like a house divided, between the part that stands for the known and the part that stands for the knower. (1999, 191).

It should be evident from these passages that a split between a knowing and conscious subject and the known object (in Mait Edey’s sense) is absolutely basic to this notion of ordinary self. Metaphorically speaking, the subject might be thought of as the grain of sand upon which the self-pearl is grown. The subject becomes characterised as ‘the self’ through a collection of roles (for example ‘observer’, ‘owner’, ‘actor’) and attributes (for example, ‘conscious’, ‘bounded’) that ‘bind’ themselves to the subject. Some of these roles and attributes seem intrinsic to subjectivity itself; others, arguably, are additional. I would like to suggest there is something like a common idea of the self that can be seen in the above writers. In the several sections that follow, I will try to outline this idea with a view to showing that this common idea describes the self that we assume we are.

Roles Ascribed Reflexively To the Self:

In this common idea of the self, we reflexively take the self to be a subject which is a:

- Knower/Observer/Witness of experience, source of attending
- Owner of thoughts, perceptions, experiences, body, personality
- Initiator or agent of actions – the source of effort and will
- Thinker or originator of thoughts
• Seeker of happiness over suffering.

Attributes Ascribed Reflexively to the Self:

In the common idea of the self, we also reflexively take the self to be a **subject** which is:

- Conscious/mental/aware at least for most of the time
- Bounded, viz., ontologically unique in its identity
- Elusive to observation and attention ("hint half guessed"), vague
- Unified/singular/simple
- Uninterrupted (hence enduring) and unchanging (hence invariable) in its essence (responsible for our identity over time).
- Unconstructed by – and standing apart from – observable objects in the ‘stream’ of consciousness

Each of the phrases for these roles and attributes will receive explanation and amplification in the following sections.

The striking similarity to emerge between the notion of self alluded to by Western philosophers and the Buddhist notion of self adds independent credibility to the Buddhist claim that this is indeed a self whose existence is assumed by most people. Features such as ‘thinker of thoughts’ and ‘elusive to observation’ are notable additions. With this Western philosophical endorsement, we shall now provide further arguments for the claim that most people reflexively assume themselves to be an entity with these roles and attributes. Where necessary, we elaborate upon what these features involve, or point to further distinctions such as that between perspectival and personal ownership (of which both are ascribed to the self).

The roles and attributes are not to be considered independently from each other, since some of them seem to be inextricably linked – for example the roles ‘knower/observer/witness’ and the attributes ‘conscious/mental/aware’. Where linked, we consider them in relation to each other. Importantly, it should be borne in mind that in arguing that most people assume themselves to be entities with each of the above-mentioned facets, we are *not* arguing or supposing that any such facet has independent reality.
Role: Knower/Observer/Witness; Attribute: Mental/Aware/Conscious

The role ‘knower/observer/witness’ would appear to best characterise that which makes the subject of experience subjective and hence something with a mental or conscious inner life. As we noted in chapter one, the *modus operandi* of the subject would appear to be that of witnessing, this being the broadest mode of conscious apprehending, attentive and inattentive.\(^{68}\) What helps contribute to the impression that the witnessing subject is something separate from objects (and of the category ‘individual’) is that fact that when presented with objects, witnessing appears to emanate from a particular psycho-physical perspective of a body and mind. The subject seems to be the locus of our first-person perspective on the world. Let us review our definition:

A subject S is a specific perspectival locus for witnessing, in relation to which objects are witnessed. This locus implicitly appears — whether dreaming or waking — to be psycho-physically constrained in space and time, offering a specific viewpoint from which the world, dreamt or actual, is witnessed.

From its outset, this project has supposed that a subject/object distinction, as outlined by Edey (1997, online) and elaborated in chapter one, is assumed as axiomatic by anyone with an inner conscious life who interacts with the world and its objects. To recapitulate a section of the passage from Edey:

You can distinguish yourself as subject from any object whatsoever (‘physical’ or ‘mental’) any time you direct your attention to that object and realize that it is you who are aware and who pay attention, not the object. The real nature of the object and the real nature of the subject may be baffling mysteries to us, but these mysteries are no barrier whatsoever to knowing which is obviously which. (1997, online).

As this passage (and the fuller version in chapter one) suggests, the fact that we pre-theoretically distinguish ourselves as conscious, observing or witnessing subjects in opposition to observed internal or external objects seems obvious enough, and is indeed taken for granted by many authors of the above quotations. It is an assumption that underpins the way we use language. When we use the first-person pronoun ‘I’ in everyday contexts, for instance, there are many contexts in which this ‘I’ does not seem to refer to any object in

\(^{68}\) Rather than repeating what was said in chapter one, I will assume the reader can refer back to it if necessary.
our purview, but to ourselves as the subject that observes those objects. If pushed for further argument that we make this distinction, we might suggest that the very notion of first-person memory involves an appeal to witnessing. (We are not presuming here that the notion of ‘first-person memory’ picks out anything real in the world – rather we are appealing to what seems, ordinarily, to be the case). When we recall experiencing object Y, for instance, our memory is not considered genuine unless we really did experience Y. In order to have experienced Y, then Y will have had to have stood in what John Perry (1998) would call an “agent-relative role” by appearing, in either a direct or mediated fashion, to a perspective that is relative to the standpoint of a conscious subject in a given space and time. For Y to appear this way, such that it can be consciously recalled later, it will have to be consciously (as opposed to merely unconsciously or sub-consciously) apprehended at the time. As I see it, the only way to make sense of ‘consciously apprehend Y’ – in such a way that Y can be consciously and explicitly recalled later – is to construe that apprehending as involving a form of conscious mental recording – in other words, as involving conscious observing or witnessing. That we can consciously recall an experience seems to crucially depend upon the fact that we, as conscious subjects, witnessed it to begin with. It notably does not seem to matter if the recalled experience pertained to an object that was, at the time, seen, heard, tasted, smelt, felt, imagined or remembered. What matters to this argument is not the specific perceptual or introspective mode through which the objects are witnessed, but the fact that they are consciously witnessed, whatever the mode. The witnessing that pertains to ‘the subject’ can be regarded as a generic ‘knowing’ or ‘apprehending’ which seems common to all conscious perception and introspection. As I mentioned in chapter one, witnessing can be considered as the determinable under which the specific perceptual and introspective modes, so long as they are conscious, are determinates.

The word ‘conscious’ (as opposed to ‘unconscious’ or ‘subconscious’) has been mentioned several times. By ‘a conscious experience’, we mean in Nagel’s (1986) phrase, that there is something it is like to have that experience – or, perhaps, to be a thing which has that experience. However we construe it, an inner phenomenal life seems central to what it means to be a subject whose modus operandi is witnessing. That is, the subject of the above characterisation, which seems at the core of the self, is not merely a lifeless thing, like a rock, but is a mental or conscious thing with a subjective inner life. Regarding this subjective inner
life, we might further ask: is there something it is like to be a subject or self per se, such that the witnessing qua self seems imbued with its own non-intentional feel, whose character is yet a part of any given experience? Or does the subjectivity of a given conscious experience seem exhausted by the phenomenal character of the objects of witnessing, including the phenomenal characters of the different perceptual or introspective modes such as that of seeing or imagining?

Reflection upon the nature of conscious experience suggests the former to be true. We naturally forge – and allude to – a subject/object distinction (or as this chapter will suggest, a self-other distinction) on the basis of first-person experience, that is to say, on the basis of how things seem. Why would we allude to such a distinction if our experience seemed exhausted by its objects? It will transpire (during this chapter) that inherent to harbouring a sense of self will be the reflexive harbouring of an impression that the self (namely, oneself) is a subject-like, bounded entity that seems to stand apart from its objects. As such a subject-like entity, the self will not seem like a mere abstract inference, but will seem like something whose existence and character is gleaned, even if only elusively, through first-person experience, such that there seems something it is like to be a subject qua self. This something-it-is-like-ness about the self that we have a sense of being, will, being subject-like, seem to glean none of its intrinsic selfy character from observable objects (since the subject seems inherently un-object-like). The self will seem to stand apart from its objects, with its own non-intentional phenomenal character. That we implicitly ascribe this phenomenal dimension to the self-as-subject of experience – which does not seem reducible to qualia that characterise objects of experience – should become clearer as we proceed with our analysis.

Roles: Owner of Thoughts, Perceptions, Experiences, Body; Initiator of Actions – The Source of Effort and Will; Thinker or Originator of Thoughts; Seeker of Happiness over Suffering. Attribute: Bounded, viz., Ontologically Unique in Identity.

I consider the roles ‘owner/thinker/initiator/happiness-seeker’ together with the attribute of boundedness (hence ‘ontological uniqueness’) because the roles, at least on a certain interpretation, provide strong evidence that we attribute boundedness to the self. Once we see how the roles ‘thinker/owner/actor’ can play out, we can properly appreciate how the self appears as a bounded, unique entity that is distinct from all other things. First we will
describe boundedness in general, showing how a sense of self-other boundedness is evidenced by identification in general. Then we show how a sense of boundedness is evidenced through common modes of identification in which these roles become apparent.

**Boundedness and Identification.**

We first explain what is meant here by 'boundedness'. We mentioned in chapter three that boundedness is implied by self-identification. By 'boundedness', we mean that the self is an ontologically distinct entity – not merely *qua* an impersonal point of view who observes objects (as implied by the bare notion of 'subject'), but a personalised 'thickened' thing with a numerically unique identity who observes objects. As Ryle puts it:

> He also feels, very vaguely, that whatever it is that his 'I' stands for, it is something very important and quite unique, unique in the sense that neither it, nor anything like it, belongs to anyone else. There could only be one of it (1966, 31).

Boundedness is hence that feature by which the self *qua* self is identical to something with a unique identity that is separate from all other things. To have a *sense* of being a bounded entity is to have a sense of *being* a unique thing that is separate from all other things. To have a sense of being a bounded thing is to *identify as* a uniquely separate thing. We shall see that the best evidence for a subject identifying *as* a uniquely separate thing, is that of the subject consistently identifying *with* select items, in such a manner that those items are tacitly assumed to be assimilated to the subject (recall our stipulation that 'identifies *with* X' implies X, while ‘identifies *as* X’ does not imply X).

When identifying with items, the boundary sensed between 'self' and 'other' will not be something that seems directly observable like just another object. This point will become clearer when we look at further aspects of the self we assume we are, in particular, that of 'elusiveness'. The self, being subject-like, will seem to elude its own observation in such a manner that anything identified with as part of the self will seem to inherit this observational elusiveness. The self's boundary – where the self ends and 'other' begins – will hence also seem somewhat elusive. Yet the very fact that the (apparent) self perceives various objects as *not* being identified with as 'me' (or as a 'part of me') will, given that some items *are* being identified with as 'me', be *suggestive* of a boundary between a thickened or personalised subject and the observed world. The 'entity within this boundary', a personalised self, will
thus be felt "very vaguely" to be something "quite unique" as if "there could only be one of it".

In this section, I carry out a promise made in chapter three; to elaborate further upon the notion of identification, identification being the strongest evidence for a sense of the self’s boundedness as an ontologically unique entity. In doing this, I shall no longer use the terminology of Buddhism (for example, that of ‘the khandhas’). As an upshot of this departure, the details of this chapter’s account may at some points vary from how the notion of identification has been articulated in context of the Buddhist suttas. Nevertheless, the core ideas behind the notion of identification, as elaborated in this chapter, are based firmly upon my interpretation of how the notion is employed in Buddhism.

Strong evidence for a sense of boundedness can be found not only through the fact that identification is something we commonly do, but also through three common modes of self-identification, referred to in the literature and/or this chapter as: ‘this-ness’, ‘thinker/agent’ and ‘consistent self-concern’. Our assumed roles as ‘thinker of thoughts’ and ‘agent of actions’ are therefore demonstrated in this exercise. I then relate this analysis to the most major identity-role to characterise the self: the owner of experiences. I argue that the Buddha was right to suppose that (a) personal ownership and self-identification stand in reciprocal relations to one other such that there is a self if and only if there is a personal owner and (b) that most people identify themselves as a personal owner of experience and hence, as a self. In arguing for this, it will also become evident that ‘perspectival owner’ is a feature central to the self we assume we are. The section is finished by raising the question of whether (as Buddhism would hold) the sense of the self’s boundedness could, contrary to appearances, actually amount to (as opposed to just being evidenced by) identification with various items, through the assumed roles of self-identity. We also consider preliminary evidence for the Buddhist claim that the reciprocal senses of self-identification and personal ownership co-arise with tanha, such that whenever one phenomenon is present in the mind, then so is the other.

Before we continue, a note on use of the term ‘subject’. Appeal to this term is unavoidable when explaining the notion of identification. Readers should not take this to mean that a subject has independent reality; only that identification is to be analysed in terms
that appeal to a subject. Should the subject turn out to lack reality, then so, too, will identification.

**Identification Revisited**

“Arnie identifies with his fit body”.

“How are you Julie?” “We’re doing great. We’re pregnant!”

“The blue hair and body-piercings are part of who I am. I feel incomplete without them”.

Identification seems to occur when a subject ‘assimilates’ various things to its perspective, so that they don’t overtly appear as things separate to that subject. From the subject’s point of view, it is as if the identified-with items become assimilated to the subject (this fit body, this blue hair), or the subject becomes part of a larger item (that doting couple where the pronoun ‘I’ entirely disappears from conversation with the outside world). And when the union of subject and assimilated thing, identified with as ‘me’ or ‘we’, confronts the world as ‘self’ or ‘us’, this world is invariably thrown into relief as ‘other’. This sense of separation between ‘self’ and ‘other’ is evidence for a feeling of boundedness that is had by the identifying subject. Identified-with items may include gender, race, character traits, preferences, aversions, bodily attributes, social status, profession, possessions, other people, political affiliation, sports-teams and countries. It may be as simple as a perspectival owner of one’s perception or as complex as a Marxist manifesto. Now without specifying how identification exactly works, it should be clear that while it is correct to say that a subject identifies itself with or as something, qua that thing, for example a football team, it is really only the idea of that thing which could get appropriated, in any literal sense, to the subject’s perspective. I shall henceforth use the term ‘appropriates’ to convey what a subject actually does upon identifying itself with or as an item: it appropriates the idea of that item with which it seems to become assimilated. By ‘idea’ I simply mean a thought of the item that the subject identifies itself with or as.⁶⁹

Given that there is witnessing, identification seems to subtly affect the manner in which things are witnessed, such that one’s thought and perception are imbued with the distinctive flavour of ‘oneself’. David J. Velleman (2002) describes identification along these lines, and

⁶⁹ As discussion in this chapter progresses, we shall see that an appropriated idea or thought cannot overtly appear to the subject as an obvious, separate object of witnessing — although it is still an object.
although he limits his discussion to aspects of the personality, his analysis also works with aspects such as physical attributes (or our idea of them). In his paper “Identification and Identity” (2002) he suggests that when there is a part of one’s personality that “presents a reflexive aspect to [one’s] thought”, that part of the personality is identified with as ‘me’, or is, as he puts it, as part of one’s self-conception at that time (114, 2002). To present a reflexive aspect to one’s thought, according to Velleman, is to view the world through the lens of that aspect, such that its filter seems to become a part of the subject’s first-person perspective. In his own words:

If there is a part of your personality with which you necessarily think about things, then it will be your mental standpoint, always presenting a reflexive aspect to your thought. You will be able to think about this part of your personality as “it,” but only from a perspective in which it continues to function as the thinking “I” – just as you can find a reflection of your visual location “over there” only from a perspective in which it is also “back here.” (114, 2002).

So if a subject (of which we shall call the associated person ‘Arnie’) were to identify itself a physically fit person, then so long as this identity persists, ‘Arnie’ could not help but think about and perceive the world through this self-conception, as a fit person. Arnie’s perspective on the world will be imbued with the idea of ‘fit person’, which may turn shape his pattern of behaviour and desires (for example visiting the gym regularly). It will shape the way he thinks, including his emotions – for example when hearing the statistics of how many unfit adults there are, a surge of pride may accompany the thought ‘I am not one of those people’. And as Velleman rightly points out, so long as an aspect of the personality is being identified with as ‘me’, one will not be able to think about that aspect from a truly detached perspective. Arnie will always think about his fitness through the filter of his conception as a fit person. Identification is essentially reflexive in this manner. In many cases a subject will not be overtly aware that it is identifying as this or that thing. Yet the identification, such as with one’s gender, or as someone with free-will, will be ‘colouring the lens’, modifying the way that the world is perceived.

Identification, as characterised in this analysis, seems to be something that we do with frequency. I propose that these identity-filters (of which there will be many at a time) would, in broad terms, evidence a felt boundary between subject and object in the following way.
When the subject confronts the world, including the world of bodies, thoughts and perceptions, a boundary which is felt to wordlessly differentiate subject (qua witness) from object (qua witnessed), is also, or actually, felt to differentiate the subject-as-appropriated-ideas-qua me, from encountered objects not identified with as ‘me’. The subject who confronts an object as ‘other’, then, does not normally confront the object as merely an impersonal witness from a psycho-physical perspective, but also as features of the body, personality, favourite sports-teams, and most broadly, as ‘me’ (the pre-existing self-entity under which all other features seem subsumed). Dennett has noted that in a certain respect, the boundaries of ‘self’ do not appear static, but as fluid, expanding and contracting with context (1991, 417). In one situation, one’s sense of identity may expand to include a favourite ice-hockey team, feeling that one personally has a stake in winning or losing the match (‘the other’, of course, being the rival team and sometimes, country). In such situations, one’s identification with an item may be described in terms of being ‘part of that item’ with the words ‘we’ and ‘our’ being common, as in “we won the match”. In another situation, one’s sense of identity may shrink away from one’s thoughts and actions to the point of dissociation, as in “That wasn’t the me real talking” (Dennett, 1991, 417). Whether expanded or shrunken, the sense of boundedness seems to turn the subject from a mere point of view into a substantial personal thing.

Now that we have some idea of what is meant by boundedness and how a sense of boundedness is evidenced by identification, we can present further evidence which makes it clearer that we regard ourselves to be bounded ontologically distinct entities. This evidence points to modes or ways in which we, as subjects, commonly identify with items through identifying as (note, not with) that mode itself. These modes of identification will provide solid evidence that we assume there to be a thick boundary between self and other. These modes turn out to include the roles ‘initiator of actions’, ‘thinker of thoughts’ and, most importantly, ‘owner of experiences, actions and thoughts’.

Three Modes of Apparent Self-Identification: This-ness, Autonomy and Consistent Self-Concern

Our ascription of ontological uniqueness (or boundedness) of the self from all other things can be evidenced through the following modes of identification in which a self seems presupposed: (a) the unique importance or value felt in association with its simply being, it
would seem, *this very thing* and (b) felt *causal efficacy* (viz., freely willed initiator of actions, thinker of thoughts) and (c) thoughts, emotions and behaviour which consistently, over time, seem to reflect a concern for one particular entity – the self and (d) the sense of being a personal owner-self replete with requisite attitudes of ‘my-ness’ towards various items. In this section we consider (a)-(c), with a later section being reserved for (d), where we discuss, amongst other things, how (d) is pre-supposed in (a)-(c).

**(A) The Importance of Being ‘This Very Thing’.**

While Galen Strawson (1997, above) lists “ontological independence” as part of the ordinary conception of self, he does not detail how this is to be construed pretheoretically, as a feature most of us take for granted. Both Ryle (1966) and Baron (2000), however, make more explicit reference to variants of (a), which Baron refers to as “this-ness”; a feature to which we shall now turn.

Ryle says that whatever the “I” stands for it is “very important and quite unique” (1966, 31). Baron (2000) also alludes to the self (whose reality he rejects) as something that essentially has a quality of specialness or, as he calls it, “this-ness” (2000, secs 4.2.3, 5.0 & 7.0). “This-ness” is conveyed through the feeling that substituting oneself with another individual who has all the same qualities, leaves out something essential – the value associated with being ‘this very thing’ which is ‘me’. ‘This-ness’ will always involve identification on the part of the subject, sometimes with a specific object, like an author, but as always as accompanied by a more general idea: that of being the one and only me who is the author, etc. This-ness hence reflects a mode of self-identification that can be commonly assumed by a subject. In all such cases, the subject’s identification as the one and only me is suffused with a dispositional sense of value that can be readily felt, should the occasion arise. This value seems to create an almost tangible feeling of boundedness between self and other. An example Baron uses to illustrate such feeling involves that of personal achievement; the desire that something, such as a project, be done by oneself and oneself alone (2000, 7.1.1). For someone writing a novel, for instance, it is generally of no comfort for them to learn that someone else, no matter how similar in attributes, is to take over its writing (unless it is a pre-arranged deal). It matters to him that he, as opposed to someone else, writes the novel. He identifies as its author, and part and parcel of this identification is that he believes and
implicitly desires it to be *he and he alone* who is the author, as opposed to some other subject of experience, no matter how similar. Baron invites us to contrast this with other tasks, such as applying the brakes of a car, by which we feel it does not particularly matter *who* does it (2000, 7.1.1) (although I contend that if there was an accident, it *would* matter who did it, showing the value dispositionally attached to this action).

The sentiment of ‘this-ness’, involving identification as ‘the one and only me’ (and hence a sense of self-identification) is not confined to authoring actions, but also includes the value that we ascribe to being *this very subject* of experience. This dimension is sharply conveyed in a thought experiment by Derek Parfit.\(^{70}\) Imagine that a new form of transportation is invented, by which a machine records every detail of your psycho-physical constitution, replicating your person at the desired destination – before destroying the original. One day, the machine malfunctions, leaving the ‘original you’ (which may itself be a replica) undestroyed. You talk to your doppelganger over a video-link, who informs you that the procedure damaged your heart, leaving you with just a week to live. He tells you not to worry however, since he will be living your life, feeling your feelings, caring for your wife! Any feeling of discomfort you are bound to have is telling of the fact that you value ‘you’ (and for that matter, others) not only for your (or their) qualities, which are present *holus bolus* in the replica, but for simply being, it would seem, *that very thing* (even if *that very thing* is itself a replica from previous travels). One identifies, in other words, as ‘this very subject of experience’, as the *bearer* of one’s psycho-physical qualities. One can feel this sentiment to the extent that the thought of losing one’s ‘this-ness’, such as through replication and destruction, feels akin to the thought of dying. It shows the great deal of importance one places, whether rationally or not, on the idea of numeric distinctness – over and above any qualitative features. Reflecting on the situation through filters heavily associated with ‘being me’, the average subject is thus not thinking: “None of that is who I am, so what does it matter if a replica takes over?” No. He is thinking “I am the subject of these qualities. If a replica takes over, he lives and I die”. One’s ascription of value to being ‘this very thing’ is hence strong evidence that most people deeply identify as being ‘the one and only me’ and hence, as an independent subject of experience. It provides strong evidence

---

\(^{70}\) I have heard this thought experiment, in conversation, being ascribed to Parfit.
for a sense of wordless boundary that is drawn between the subject qua self and all other things, no matter how similar in quality.

**(B) Role: Agent of Actions and Thinker of Thoughts (Autonomy)**

The sense of boundedness is also brought out through considerations pertaining to its causal efficacy. These revolve around the fact that most people take themselves to be autonomous agents, in virtue of their assumed causal powers; thus relating directly to the ‘role-occupiers’ thinker of thoughts, initiator of actions. These roles also point to common modes of assumed self-identification. How, more precisely, do we identify as such thinkers and agents? One way, already mentioned, is through ‘this-ness’; that felt value attached to the idea that I, this particular self, as opposed to some other self, am the agent of certain actions. Another way we construe ourselves to be thinking agents is through the feeling that our deliberate actions are not the result of impersonal factors but rather, of special causal powers pertaining to free will – our free-will. We feel, in other words, that our choices are not blindly determined; but that with any deliberate action, we could have chosen to do otherwise. The feeling that one is able to exert unique causal powers on the world through one’s own free thoughts and actions adds weight to the feeling of being a separate, autonomous entity. Identifying as a (free) thinker and agent would thus plausibly evoke a sense of boundary between our ‘free’ selves and the world with which we interact (including other free agents).

But the feelings of freedom do not seem to stop there. Like ‘this-ness’, the belief in one’s free-will seems to endow those free thoughts and actions with value. One takes particular pride or shame not only in the apparent fact that this self, as opposed to some other self, is the rightful author of an action, but in the ‘fact’ that as an island of special causal power, one is able to author this, as opposed to that kind of action. It is through this feeling of freedom that one feels responsible for their actions. In the extensive literature on ‘free will’, it has been noted that if one truly believed there was no real choice in the matter – that our every action was determined from birth – then one would not fully experience the emotions of pride, shame, guilt, praise or blame, to name but a few. It seems that for these emotions to be properly felt, one must, at some level, buy into the assumption that it is possible to have chosen differently. We do not usually attribute heartfelt praise or blame to behaviours we perceive are mechanistic or random (if we do, then it tends to be through
unconsciously anthropomorphising inanimate objects such as stalling cars and red traffic lights!). The emotional investment in the outcome of one’s actions serves to intensify the sense of boundary between self-as-agent and other (or self-as-thinker and other). The associated roles, ‘thinker of thoughts’ and ‘initiator of actions’ thus depict distinct and repetitive modes in which we, as subjects, identify with things (in the capacity of these roles), thickening the apparent boundary between self and other. And the associated sense of boundedness is best evidenced through the value we attach to being, it would seem, a free author of our actions.

Baron regards the attribution of free-will (viz. the possibility that one could have chosen differently) as integral to the notion of responsibility, and to the notion of self that he explicates. He holds that the kind of causal powers ascribed the agent are moreover not to be found in the impersonally characterised ‘scientific world view’, which he says admits only of explanations involving deterministic or random forces. (2000, 4.2.2). He maintains that due to this, the self cannot be real; while nevertheless conceding that widespread belief in the self, with its defining characteristics of responsibility and ‘this-ness’, is unavoidable. Regardless of whether we agree with Baron’s elevation of the ‘scientific world view’, his reflections on both ‘this-ness’ and ‘autonomy’ have helped to illuminate, from two different angles, the sense of ontological uniqueness that we have. That sense of being a uniquely separate thing, whether as something special, or as something autonomous, is strong evidence for our ascription of boundedness to the self. We can also note its connection with the long-running debate on free-will, and with the fact that many philosophers, such as Kant and Frankfurt, have chosen to identify the most central aspect of our ‘selves’ with ‘the will’.

(C) Consistent Self-Concern.

So far, we have looked at how our reflexive ascription of boundedness to the self is best evidenced through identification. Such evidence includes commonly assumed modes of self-identification, namely, this-ness and autonomy, by which we assume that we are selves qua thinkers of thought or initiators of action. There is another related and powerful source of evidence for the fact that we take ourselves to be a bounded thing. With this-ness and autonomy, our assumption of a boundary between self and other is evidenced in a synchronic or latitudinal fashion. This-ness: at any given time, it seems to matter that it is me rather than
another (including a replica) who lives this very life, completes these very projects, etc. Autonomy: at any given time, I, as opposed to my surroundings, have the power to choose this over that action – and it seems to matter what I, as a free agent, choose. In each of these cases, an assumed boundedness of the self is evidenced latitudinally, through the identification as this rather than that subject, or with being able, it would seem, to choose this over that action, at any given time. But an assumed boundedness of the self can also be evidenced longitudinally, by the very fact that over time, we are most consistently concerned with the welfare (including the goals and aspirations) of what would seem to be our self. The claim is not that we are concerned with the welfare of only our ‘self’ (which would imply that we are selfish) – but rather that, over time, our concerns most consistently (although not always and not exclusively) involve the welfare of what would seem to be our self. This concern is reflected through our emotions, in the way we think and in the courses of action that we choose. The general line of argument is this. If there is evidence that over time, our thoughts, emotions or behaviour most consistently involve the welfare of what would appear to be one’s ‘self’, then this also counts as evidence that we most value our ‘self’. Hence, we ipso facto take ourselves to be that ontologically unique entity which is valued in this way.

Let us consider the emotions; for example, happiness, sadness, fear, surprise, anger, moral indignation, aversion, lust, anxiety, frustration, guilt, pride, vanity, embarrassment, envy, nervousness, compassion, excitement, jealousy, loneliness, disappointment. On the face of it at least, in order to feel most, if not all of these emotions, one has to identify as an ‘I’ who is the subject of the emotions. One cannot, it would seem, feel guilty without buying into an I who feels guilty; or angry, without assuming (if implicitly) that I am angry, and so forth. Most, if not all of these common emotions, seem to include as a part of their ‘aboutness’ not only the object to which the emotion is directed, but an implicit self who is the subject of that emotion. But more than this, they seem to exist because the subject desires what is best for itself, either in terms of avoiding harm or accruing benefits. Imagine that one could not care less what happened to them or what they did. Could one feel guilty or fearful? To feel guilty usually implies that one wishes one hadn’t done a particular deed, which implies that one would have preferred to have acted in a way which avoided causing one’s own (amongst any other) suffering. To feel fear generally implies that one wishes that a particular scenario wouldn’t come to pass, which again betrays the desire for an outcome that
is relatively unharmful to ‘oneself’. Most of the above-listed emotions (except perhaps compassion and pure happiness) involve some investment of one’s happiness in a preferred state of affairs. Whose preferences? George W. Bush’s? Only if it coincides with what one’s ‘self’ would prefer. These emotions are ubiquitous, felt in various grades of intensity many times a day, from desiring to get rid of the smallest itch, to the weighty angst of desiring to get rid of a long thesis (whether one is writing or reading it!). And they almost invariably revolve around what one’s own supposed self would prefer.

This does not make us selfish in the sense that we fail to care also for others. But which other people do we care for most? Strangers or acquaintances? Friends or enemies? It is very unusual to find someone who genuinely cares as much for the welfare of those they don’t know as for those that they do know, or for enemies as much as for friends. The aim here is not to pass moral judgement on whether it is better or worse to care more for those who are the “nearest and dearest”. The aim is rather to observe that this is what most of us do, and that it is part of the evidence that we most consistently care about ‘ourselves’ (if we lose someone we love do we grieve more for the person’s life being cut short, or for our ‘own’ loss?). Nor does it take sophisticated psychology to see that one’s motivations and behaviour tend to be under the sway of the predominant, ‘self’-referencing emotions. In other words, most of us live our lives in such a way that we most consistently act to promote the welfare and happiness of what appear to be our self, as best we can. This simultaneously serves as evidence that we assume ourselves to be a bounded thing, namely, the thing which, over time, we care most consistently about. 71 We might call this longitudinal mode of identification, best evidenced by the emotions, ‘consistent self-concern’.

Evidence for ‘Seeker of Happiness over Suffering’ as an Ascribed Feature of the Self

The evidence for ‘consistent self-concern’ also lends considerable support to the Buddhist contention that the very nature of the self that we identify as being is such that if given the choice, it would seek on the whole to avoid suffering (or, more accurately, dukkha, as

71 Hence William James writes: ‘We know how little it matters to us whether some man, a man taken at large and in the abstract, prove a failure or succeed in life, - he may be hanged for aught we care, - but we know the utter momentousness and terribleness of the alternative when the man is the one whose name we ourselves bear. I must not be a failure, is the very loudest of the voices that clamour in each of our breasts: let fail who may, I at least must succeed. Now the first conclusion which these facts suggest is that each of us is animated by a direct feeling of regard for his own pure principle of individual existence, whatever that may be, taken merely as such’ (1890, 318).
construed in chapter one) and to be as happy as possible, where the notion of happiness can be construed broadly to cover a range of positive emotions characterised by a pleasant affective tone, such as joy, the contentment of goal-achievement, peace and fulfilment. We leave open such questions as to what kind of lifestyle might be most conducive to pleasant emotion, or which pleasant emotions are the ones most worth having. I think it is uncontentious to suppose that most people would on the whole seek to minimise their own dukkha and maximise their own happiness – rather than to maximise their own dukkha or simply not care one way or the other. We are not saying that people are not prepared to undergo privations and suffering to achieve their goals, or that such pain and suffering might not be considered to make the goal even more valuable. But note: the goal, if freely chosen, will rarely be one that we think will make us on the whole less happy and more unhappy. An overall preference to seek happiness and avoid suffering, therefore, seems central to the self that we assume we are. It is worth re-iterating Locke’s passage which picks up on this aspect:

This personality extends itself beyond present existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to itself past actions, just upon the same ground and for the same reason as it does the present. All which is founded in a concern for happiness, the unavoidable concomitant of consciousness; that which is conscious of pleasure and pain desiring that the self that is conscious should be happy. And therefore whatever past actions it cannot reconcile or appropriate to that present self by consciousness, it can be no more concerned in than if they had never been done… (1690, 220). [my italics]

Thus Locke alleges that a concern for happiness is part of the nature of the conscious self and that this fundamental concern is partly what motivates one to identify as the author of one’s actions. Interestingly, he links the ownership of actions to this concern for happiness. It would seem that on Locke’s account, if one didn’t show any concern for an action, then this would be a sign that they did not regard the action as theirs. This accords well with what Buddhism says about ownership, with its link to tanha. To ownership in general, we shall now turn.

Ownership: Perspectival and Personal: Their Relation to Identification

(S.VI) Bhikkhus, there being a self, would there be what belongs to my self?"
"Yes, venerable sir."
"Or there being what belongs to a self, would there be my self?"
"Yes, venerable sir."\(^{72}\)

As I indicated in chapter three, ownership, and hence the role ‘owner of experiences, body etc.’, can be analysed in several ways. There can be ‘possessive ownership’, dictated by social convention as in ‘I own this computer’. Or there can be ‘perspectival ownership’ describing the manner in which objects (such as toothaches, visual qualia and thoughts) can appear to one’s subjective perspective in a way that they appear to no-one else. And there can be ‘personal ownership’ where objects viewed from one’s perspective are felt, implicitly, to belong to a ‘me’, an owner that is neither merely perspectival nor possessive. We saw that Buddhism analyses ‘personal’ ownership in terms of self-identification (see (S.VI) above), such that whenever there is personal ownership or my-ness towards X, there is self-identification (through being the personal owner of X qua me) and, conversely, whenever there is self-identification (through identification with any Y as me or as part of me) certain objects will be personally owned by the self (or mine). By ‘self-identification’, in this context, we mean that the subject not only identifies as, but is identical to the self (or me).

On the Buddhist analysis, the self must always minimally be a personal owner, casting the role ‘personal owner’ as a most ubiquitous mode of self-identification. So, to have a sense of personally owning X, on this analysis, is for a subject to identify as a self by reflexively assuming itself to be the personal owner of X. In this way, the subject harbours a sense of self-identification — whether or not the subject is actually identical to an owner-self. And, conversely, for the subject to identify as a self (by harbouring a sense of self-identification with any Y) is for a subject to assume that some X is personally owned by it — whether or not personal ownership is a real relation. On this analysis, a subject who loses the sense of personal ownership or ‘my-ness’ towards an item (such as a sensation or limb) will no longer identify as the personal owner of that item. And, conversely, a subject who loses its sense of self-identification with respect to an item — no longer assuming it to be ‘me’ or ‘part of me’ — will also lose the sense of ‘my-ness’ towards items that specifically pertained to the ‘old me’. It will no longer harbour a sense of self — of a personal ‘me’ or ‘mine’ — at least with respect to that item.

In what follows, I shall elaborate upon, and to some extent defend this Buddhist analysis, which defines a sense of personal ownership towards some X (as opposed to perspectival or possessive ownership) in terms of a reciprocal relation to a sense of self-identification (through the subject’s identifying itself with some Y). The analysis is one which makes ‘personal owner’ a universal mode of self-identification and hence, central to the assumed self. Given our discussion so far, a sense of personal ownership will thus serve as the most ubiquitous evidence for a sense of the self’s boundedness.

**Where There is ‘Mine’, There is ‘Me’ (viz., the Sense of Being a Personal Owner).**

In this sub-section we suggest that a sense of personal my-ness – the kind that is lost in episodes of anosognosia or depersonalisation – is usefully analysed with reference to a reciprocal sense of self-identification as personal owner or me. On this analysis, a sense of self-identification is not implied simply through the subject’s occupation of the roles ‘perspectival owner’ or ‘possessive owner’. Indeed, a sense of self-identification or me-ness is precisely what seems to distinguish a mere perspectival owner from a personal owner (or more accurately, the sense of being a personal owner). It is easiest to illustrate this with an example. Without resorting to anything as radical as anosognosia, suppose that a self-confessed introvert, Jane, attends a party and has too much to drink. Through the fog of her beer-goggles, she becomes loud and extroverted. Whilst under the effect of alcohol she approaches the world as an extrovert. In the middle of loud argument about the Iraq War she thinks, in an amused and sincere way “Those words are not my own. The beer is doing the talking!” Now we can be immediately sure that Jane remains the perspectival owner of her words and behaviour: the way they sound and the thoughts they convey (about the Iraq War) are unique to her perspective. No one else can learn of her experiences in the same manner that she does. It is, rather, that she does not feel a sense of personal ownership towards her words.

I suggest that we further analyse Jane’s way of thinking by supposing that as a subject, Jane not only feels a lack of personal my-ness towards the thoughts, but she does not identify with the perspectival owner of her extroverted behaviour and hence, not as the personal owner of that behaviour. She does not, in other words, appropriate the idea ‘perspectival owner of the words’ to her current witnessing perspective and so does not
approach the world through this assumed, reflexive 'filter'. As reflected in a passage by
Dennett, it would for example be quite natural for Jane to feel at the time and remark later: "I
didn’t do that! That wasn’t the real me talking. Yes, the words came out of my mouth, but I
refuse to recognise them as my own" (1991, 417) showing that the analysis fits with the way
we speak and behave. Her particular thoughts and perceptions with regard to her words about
the Iraq War appear to her from a unique perspective and yet she does not identify with the
perspectival owner of those thoughts. Jane feels like an owner of her extroverted experiences,
but only in the non-identifying, perspectival sense.

Despite not claiming ownership of her extroverted behaviour, Jane still feels as if the
amused thoughts about her behaviour do belong to her. On the proposed analysis, we can say
that Jane identifies with the perspectival owner of the amused thoughts that pertain to her
state, through appropriating the idea 'perspectival owner of the thoughts' to her current
perspective. Through identifying with the perspectival owner of these thoughts, she identifies
as their personal owner (and so has a sense of self-identification or me-ness with respect to
them). There are more radical examples of where perspectival ownership seems to come
apart from identification with a perspectival owner, such as that of depersonalisation
(mentioned in chapter three). Such a subject, while having first-person access to the feelings
in question, may report feeling that the events "were happening not to me". On our analysis,
while such subjects perspectivally own these experiences, they do not identify with this
perspectival owner as 'me'.

From this analysis, perspectival (or, for that matter, possessive) ownership
corresponds with a notion of perspectival (or possessive) owner that does not imply the
subject's identification of itself with this owner. To see (or hear, or feel...) something, I need
not identify with 'that which sees (or hears or feels...)'. To think something, I need not
identify with 'that which thinks', etc. Of course, most of the time we do identify with that
which sees, hears, thinks and so forth. We identify with them in such a way that we approach
the world as fully-fledged personal owners of our experience. Whenever we harbour a sense
of personal my-ness towards something, such as a perception, action, quale, or person, we are
identifying with its perspectival or possessive owner, by appropriating the idea of this owner

---

Nevertheless, I would argue that subjects of such depersonalisation episodes still appropriate a more general
idea of 'me' to their perspective. It is just that none of their perceptual experiences during the episode seem to
belong to this 'me': more on this in chapter eight.
to our current perspective. The idea of this owner, however subtle, is modifying the way in
which we perceive and think about those things, such that they feel personally owned.

A sense of personal ownership towards X thus always involves that subject’s
identification with the owner of X, whether the identified-with owner is perspectival,
possessive or both\textsuperscript{74}. We must be careful about how we express this. Upon identifying with a
perspectival or possessive owner, we can say, minimally, that a subject approaches the world
as a personal owner of X (with a sense of self-identification). We can also say that the
subject, through feelings of personal ownership towards an item, identifies as, or feels as if
she is, a personal owner of that item. But it is presupposing too much to say that the subject
identifies with a prior personal owner (via a memory of it) since scepticism about the reality
of a personal owner (viz., a self) is at this stage a live option. It is also, for similar reasons,
presupposing too much to say a subject qua personal owner identifies with a prior
perspectival or personal owner (via an appropriated idea of it). Such a suggestion implies that
a personal self is being represented as opposed to presented. One’s overall sense of being a
personal owner may be nothing more than the upshot of identification by a neutral witnessing
subject with a perspectival or possessive owner. On this analysis, we can thus say, without
begging questions, that whenever a subject harbours a sense of personal ownership towards
X, then a subject (minimally a current perspectival owner) identifies itself with (minimally) a
perspectival or possessive owner of X (through appropriating the idea of it to its perspective)
and as a personal owner of X. In other words, where there is a sense of personal my-ness,
there is a sense of being a personal me. To be sure, it would seem that in this process a
personal owner is also being represented – but perhaps it is not actually being represented.

\textbf{Where There is ‘Me’, There is ‘Mine’ (viz, a Sense of Personal Ownership).}

We have suggested an analysis on which feelings of personal ownership towards an item
imply a subject’s identification with a (minimally perspectival or possessive) owner of X and
identification as the personal owner of X. On this analysis, a sense of personal ownership
towards X implies identification (and a sense of self-identification as ‘me’). A subject’s
identification with some Y is what distinguishes, for that subject, a sense of personal from
merely perspectival or possessive ownership of X. This analysis, while no doubt requiring

\textsuperscript{74} What I have to say in this chapter will apply to any other type of non-personal ownership, by which it makes
sense to say ‘that is my x’.
further argument and detail, seems to initially fit with the way we think and speak. We shall now analyse and provide some evidence for the converse relation: that a subject’s identification with any item (with its accompanying sense of self-identification with that item as ‘me’ or ‘part of me’) implies, for that subject, a sense of personal ownership or my-ness towards some other items, X (and hence, given the previous analysis, a sense of being a personal owner).

It is suggested that whenever we identify ourselves with something as ‘me’ or as ‘part of me’ – be this sports-team, body-image or personality – this ‘thickened’ subject or ‘me’ becomes an integrated perspective from which other objects will inevitably feel personally owned (relative to that perspective). And the feelings of personal ownership, pertaining to this identity, will in turn (given our analysis so far) seem to reflect a personal owner qua this identity. The ‘thickened perspective’ will thus itself feel like a subject-owner, a ‘platform me’ from which various objects are considered to be personally ‘mine’. For example, suppose that a person, ‘Sid’, identifies with the role of a punk rocker. In doing so, he appropriates various ideas to his current witnessing perspective such that they colour the very perspective from which he perceives and thinks about the world. The ideas may be of such things as his physical appearance (e.g. blue mohawk), the music he likes (e.g. Sex Pistols) and political ideals (anarchy, Marxism). His identity as punk rocker, moreover, seems to form part of an integrated perspective from which other items, such as his blue hairdye, feel not only possessively, but personally owned by him. His feelings of personal ownership towards the hairdye (and other punk regalia) will in turn reflect his identification as its personal owner qua punk. So long as Sid feels personal ownership towards the hairdye, ‘personal owner-of-the-hairdye’ (and hence the role ‘personal owner’) will seem to be integrated with his overall identity as a punk rocker. Suppose that Sid renounces punk-dom, enrolls in a commerce degree and identifies instead as a businessman. Although remaining a possessive owner of the blue hairdye, he will no longer identify with this owner, and hence, no longer as its personal owner. He is likely to throw the dye away.

I leave open the question as to whether reciprocal identification and personal ownership feelings are to be construed occurrently or dispositionally. For example it might be enough that a subject who identifies with Y be disposed to feel personal ownership feelings in relation to her identification with Y. I suspect that the proposed analysis would be compatible with both readings.
The example of Sid illustrates a case where a specific identification (with the role of ‘punk’) involves a reciprocal sense of personal ownership towards various Xs in relation to this identity (the hairdye), which in turn reflects back to an assumed identity as personal owner (of the hairdye). Indeed, it is hard to see how Sid could identify with the overall role of ‘punk rocker’ without assuming a reciprocal sense of personal ownership towards items relevant to this identity. The boundaries of his assumed identity-role (as punk) will of course shift with context: for example, sometimes Sid’s mohawk will seem to be a part of who he is (perhaps in relation to hairdye which seems personally owned); at other times it will seem personally owned by him (perhaps when he qua punk, shaves part of it off to maintain the look).

Having illustrated this analysis with reference to a specific example, it is now suggested that the analysis applies globally: any identification with Y as ‘me’ or as ‘part of me’ will entail that the subject harbours a sense of personal ownership or my-ness toward certain items that are relevant to this identification with Y. While there is not room to provide exhaustive evidence or arguments for this claim, the aim is to make it sound at least initially convincing. I will first select an array of different everyday examples where it seems to hold true (the aim also being to illustrate the ubiquity of the reciprocal identity-ownership roles). Following this, I will suggest that the three aforementioned modes of identification (this-ness, autonomy, persistent self-concern) all seem to imply that the subject harbours a sense of personal ownership towards items relevant to that mode of identification. If we also accept that a sense of personal ownership implies the sense of being a personal owner (as argued in the previous section) this will show that the role ‘personal owner’ is the most pervasive mode of assumed self-identification and hence, central to the self we assume we are.

Here, then, are some further examples which suggest that when a subject identifies itself with any items Y, there is an accompanying sense of my-ness toward items X that are relevant to the identification:

“I’m hungry! Don’t you touch that food!” Identified with item: hungry body.  
(Assumed to be Personally) Owned: hunger pangs, the food.
(Upon hearing news of how many unfit Americans there are): "I'm not one of those people". Identified-with-item: fit-body. Owned: one's gym-pass, lycra clothes, thoughts and emotions about one's fitness.

“I’m not cornering well on rainy days because my tires are getting bald” (Dennett, 1991, 417). Identified-with item: the motorist. Owned: the tires.

(After giving a terrible conference paper and knowing it) “Well, the audience were not that intelligent anyway.” Identified-with item: a dim-witted, deflated loser. Owned: the extra pangs of misery from trying to suppress this identification; the conference paper, the ideas.

“We did so well to win the footy, mate”. Identified-with item: Aussie football-team. Owned: the feelings of pride at winning the match.

“I’m afraid of losing my mind through this dementia. It has nothing to do with who I am and feels totally alien to me”. Identified-with item: perspectival owner of mind & personality. Owned: the mind & personality, the fear of losing it. NOT PERSONALLY OWNED: the dementia and its effects.

A sense of personal ownership also seems implicated in each of those three major modes of assumed self-identification – this-ness, autonomy and persistent self-concern – modes which most clearly evidence our feelings of the self’s boundedness. For example, a subject’s identification as this very subject, with its dispositional feelings of value, will always involve feelings of personal my-ness towards the relevant situation, actual or imagined. If it matters (to the author) that the novel be written by him and him alone, then there will be feelings of personal my-ness towards the novel as he identifies with its author and possessive owner. If it matters that this life be lived by me and me alone, then it will feel like my personal life that is lived, my life that’s in danger of replication. And if it feels like my life, then I identify as its sole personal owner. Nor is it hard to demonstrate a connection between consistent self-concern and feelings of my-ness. Consistent self-concern requires
that one identify with the person on behalf of whom one is concerned. Part and parcel of this
genral identity is that one feels emotions of concern, emotions that are felt as being
personally mine, reflecting back to an apparent me who is the personal owner and bearer of
those emotions.

With respect to autonomy, the story is similar. So long as I identify as the free thinker
of thoughts or agent of actions, they are always felt to be personally my thoughts or actions.
And if they are felt to be personally my thoughts or actions, then I also identify as their
personal owner. The debate on free-will sometimes focuses solely on the ‘could have acted
otherwise’ aspect, which is fine, so long as it is understood that a feeling of personal
ownership is intrinsic to the feeling of free-will. Unless a subject feels that they personally
own their free thought or action, then the thought or action, even if genuinely free, does not
feel free to them. In the case of tipsy Jane, for instance, her actions qua the extrovert, feel
under the control of beer rather than her own free spirit. She does not identify with the
perspectival owner of her extroverted actions; she does not feel them to be hers. Even if her
actions are free, the absence of personal ownership-feelings toward these actions robs her of
feeling as if she is their separate, autonomous agent. So with respect to these actions, she
does not feel like an island of special causal power amidst a sea of blind, mechanistic causes.
Rather, she feels flooded by this sea, a sea of beer whose effects wash over her character. The
connection between autonomy and identification has been noted by Velleman. He claims that
whenever one acts in a way that can be described as autonomous (with that feeling of free-
will), the action is always rooted in a part of the personality which presents a reflexive aspect
to that person, a part identified with as ‘me’ (115, 2002). In short, no autonomous behaviour
can issue from a part of the person that is not identified with as a personal owner. Insofar as
the feeling of free-will makes a unique contribution to the sense of being a bounded self, it
can do so only if ‘the agent’ also identifies as the personal owner of those actions.

While feelings of agency (qua the feeling of being a freely-willed entity) clearly
require a sense of personal ownership, it is not so clear that a sense of personal ownership
must require feelings of agency. It should be noted that feelings of agency can arise in
connection with the sheer thought of being able to think or act freely upon a situation – one
does not have to physically move about. In this way, autonomy extends to the popular
characterisation of self as ‘thinker of thoughts’. Now it is conceivable that one may be in a
situation where feelings of agency are not activated, while feelings of personal ownership are. Tim Bayne’s example of a jingle running through one’s mind is useful here. He remarks that one need not feel that any authorship or agency with regard to the jingle: indeed, it may persist against one’s wishes. Yet there can plausibly remain a feeling of (personal) my-ness towards the unwanted jingle, at least insofar as it is felt to run through my head (232, 2004). Another example: suppose Ben is lying down, enjoying a pleasant massage. It sounds odd to say that Ben identifies as an ‘agent of the sensations’. It is partly because he is in a receptive mode that the sensations are enjoyable. But he definitely identifies with the perspectival owner of these sensations; he feels them to be his and he wants them to continue. Now it could still turn out that unless one had a disposition to identify as an agent, one could not identify as a personal owner. Perhaps agency is required for a sense of personal ownership in some less obvious manner. Dispositions aside, however, we can gather that the role of ‘personal owner’ is more immediately basic to the self than the role of ‘agent’ or ‘thinker’. It is the sense of personal ownership rather than agency that is easily co-definable with identification and hence, most pertinently linked to the sense of boundedness. It will hence be personal ownership rather than agency that is focused on in the later chapters that concern the ontological status of boundedness.

Integrating Modes of Identification into a Overall Conception of Personality

So far, we have talked about modes of assumed self-identification (this-ness, agency, persistent self-concern and, most broadly, personal ownership) along with specific ideas that, in the capacity of such modes, are appropriated to the current subject’s perspective as ‘me’ or as ‘part of me’. We now say something about how these modes might seem integrated into a general self-conception, a conception that is reactivated and revised (even if just subtly) upon each new interaction between the subject and the world of objects. On this, John Perry offers an insight:

Ordinarily all one’s knowledge about oneself is integrated around a special sort of idea of notion of oneself that we express with “I”. While my perception that the beer is in front of me may not require a representation of myself, the information I acquire

---

76 As I mentioned in a footnote of chapter three, Bayne does not adequately distinguish perspectival from personal ownership. Nor can I find evidence that the distinction is properly articulated elsewhere in the Western literature.
is immediately integrated into self-attached knowledge, that I might express with “I see a beer” or “there is a beer in front of me”. And when I read a piece of e-mail, that says that John Perry’s paper is overdue, I integrate this information into self-attached knowledge, ‘My paper is overdue,” and I realize that it is me that has to get to work. I would think, “There is a beer in front of me, but I have a paper to do. (1998, 9, online)

To express this in the vocabulary of this chapter, “self-attached knowledge” involves both a witnessing subject (which Perry obliquely refers to as an “agent”) and an integrated idea of one’s personality. The most integrated idea of all is the general idea of one’s personality and history as viewed ‘from the inside’ or perspectively. Given that the most general identified-with role (without begging questions) seems to be that of the perspectival owner, it makes sense to suppose that the integrated conception of one’s personality is recalled in the format of either ‘perspectival owner of personality’ or ‘perspectival owner qua personality’. So long as we don’t forget whom we identify as, this general idea (which will vary according to occasion) will be activated every time we perceive and interact with the world. Through identifying with this perspectival owner of or qua the personality, we identify as a personal owner of or qua this personality. Damasio refers to this integrated personality as the “autobiographical self” and he maintains that it is updated upon each new perception (1999, 217-222). So when Sid identifies with the persona of the punk, this persona is not an isolated thing, felt to be had by no one in particular. It becomes embedded into the context of his general self-conception. The owner-of-the-hairdye with whom he (as a witnessing subject) identifies, is not merely the possessive owner-of the-hairdye, nor is it merely the punkish possessive owner of the hairdye. As a witnessing subject, Sid identifies with a perspectival owner qua personality who is punkish possessive owner of the hairdye and hence, as a personal owner in this capacity. And part of his subsequent embracing of the “grey-suit” – a reaction to the “punk” – is again, not a persona to be had in isolation.

Nor are our perceptions, however paltry, felt as isolated occurrences had by no one in particular (when they are, it is often the sign of pathology). Upon each fresh percept, the witness-subject qua perspectival owner of the percept, appropriates to its current perspective not an isolated idea of the role ‘perspectival owner of this percept’ but an integrated idea of ‘perspectival owner qua personality who is the perspectival owner of this percept’. This appropriated idea becomes a filter through which each fresh percept becomes witnessed as personally mine. In the words of John Perry, the integrated notion of one’s personality
becomes attached to the perspective of that which witnesses each new perpect, such that it does not seem merely like “beer in front” but “beer in front of me” (1998, 9, online). When this most general idea of one’s perspectively accessed personality is appropriated to one’s current witnessing perspective, this greatly expands the range of potential objects that can be viewed as personally ‘mine’. Many objects that have felt personally owned in the past, objects as mundane as breathing, are automatically perceived as either personally me or personally mine.

Now this is not to stipulate that an overall conception of one’s personality must depend upon access to a rich, personal history through first-person memory. The sense of self does not imply a sense of one’s personal identity over time – at least if that timeframe is longer than 45 seconds. This is perhaps most strikingly illustrated in the case of “David”, a patient of Damasio’s. As the result of a brain trauma, David lost all memory of his rich, autobiographical past and could not recall specific occurrences beyond a 45 second window (Damasio, 1999, 118). He nevertheless identified himself as a personal owner of his actions, for example, being pleased upon winning a game of checkers. While the extent of his identification with things was of course limited, each new perception was still viewed as personally his, and hence, as integrated with his overall, albeit limited, conception of personality (Damasio, 1999, 113-121). Nevertheless, it should be obvious that memories of a rich personal past will in most cases, greatly enhance one’s conception of personality and hence, one’s overall sense of being a unique bounded entity over time, as well as at any one time.

**Identification, Ownership, Boundedness and Tanha**

We have surmised that a sense of self-identification (or me-ness) and the sense of personal ownership (my-ness) are defined in terms of each other, each involving the process of identification. We also have surmised that identification provides the best evidence for a sense of boundedness. Given this, we can surmise that the most broadly assumed mode of self-identification, that of the ‘personal owner’, will provide the most pervasive evidence for the fact that we identify as a bounded, ontologically unique entity. While we reflexively assume that we are a self qua personal bounded owner, I have, in the course of explaining the concept of personal ownership, been careful not to speak in such a way that presupposes the
existence of such an owner. I have said that when we identify as a personal owner of X (hence assuming ourselves to be a personal owner) it may or may not be the case that we also identify with a personal owner of X (if we did, it would imply the reality of the self viz., personal owner). Now, it would seem to us that we do indeed identify with a personal (as well as perspectival) owner of X – but whether we do in fact in fact identify with such an owner remains to be seen. On the Buddhist analysis, which denies the self reality, we only seem to identify with, and hence reflexively represent, a personal owner. On the Buddhist position, the very act of a subject’s identification with the perspectival owner, through appropriation, serves to mentally construct the sense of being a bounded, personal owner. Hence our repeated identification with a perspectival owner of this or that X turns out to not only evidence but to constitute a sense of being a unique bounded personal owner. The Buddhist goal, indeed, can be articulated as losing the sense of being a unique, bounded personal owner. It is the boundedness of the self – as evidenced by a sense of personal me-ness and my-ness – that Buddhism considers primary in contributing to the self’s overall lack of objective reality. In chapter eight, we shall provide independent evidence that the Buddhist analysis is correct; that the sense of boundedness is not in fact accounted for by an actual bounded ontologically unique self, but by those frequent acts of appropriation (hence identification) which appear to evidence, rather than constitute, our sense of being a personal owner.

We might also wonder if there is a further psychological factor that accompanies the reciprocal relation of identification and personal ownership-sense, a factor that can perhaps serve as independent evidence that such identification is taking place. According to Buddhism, there is a further aspect to co-arise with identification and a sense of personal ownership. This aspect is tanha or emotional investment in things being the way we would like them. To harbour a state of tanha is be of such a disposition that one’s happiness or suffering rides upon whether states of affairs are perceived to be in line with one’s desires and preferences (if not in line, then there is suffering or dukkha). As we know from chapters one to three, Buddhism maintains once tanha is eliminated, then so is a sense of personal ownership and self-identification, and hence, the sense of boundedness that is distinctive to the sense of self.
We finish this section by considering some preliminary evidence for the Buddhist claim that a sense of personal ownership and self-identification (hence boundedness) is associated with \textit{tanha}, viz., emotional investment. If we first of all reflect upon cases where the reciprocal senses of self-identification and personal ownership (and hence a sense of boundedness) are most evidently felt, with the strongest senses of ‘me’ and ‘mine’, we will invariably find an ascription of \textit{personal value} to the things identified with and felt as personally owned. Integral to this sense of personal value will be the potential to feel happy or miserable, in proportion to the degree of identification and personal ownership-sense that is borne towards that item. This-ness, for instance, is a major mode of assumed self-identification. Part and parcel of this-ness is the value that we ascribe to our feeling of being \textit{this very subject of experience}. Unless there was emotional investment in being \textit{this very thing}, the thought of a replacement replica would not be upsetting (or exhilarating, if one hated one’s existence). Agency is another area where mattering matters. It seems that the more powerfully we identify ourselves as initiators of certain actions, the more we hope that those actions will turn out as we wish. And the more hope we invest in things turning out well, the more we suffer if they do not. Hence \textit{tanha} seems to accompany a definitive sense of agency.

The longitudinal mode of identification – consistent self-concern – is also value-laden. Consistent self-concern is \textit{defined} in terms of the differential value we ascribe to ‘ourselves’ and to situations – feelings that are integral to \textit{tanha}. It would seem that the more emotion that we invest in ‘ourselves’ or in things that are perspectivally or possessively mine, the stronger our sense of (bounded) self-identification and personal ownership towards those things (and vice versa). This does not mean that one must at all times be explicitly aware of one’s emotional investment in things. As I mentioned earlier, such value exists as a disposition to feel various emotions in relevant circumstances, for example, when the identified with or owned item is in peril. We probably hardly ever think about our eyesight or breathing. It is not until something goes wrong that we suddenly realise the extent to which we identify ourselves with that thing or its perspectival owner. People pay thousands to restore bodily functionings whose smooth running was, before the trouble, taken for granted. And if the operation turns out badly, then provided that one is still alive and thinking, there
will be mental dukkha, to the extent that one identifies themselves with their body (and owner of bank account).

Emotional investment or tanha thus seems implicated in three major modes of assumed self-identification (and hence in personal ownership): this-ness, autonomy and consistent self-concern. Perhaps, as Buddhism would claim, the feeling of differential value towards oneself and various objects is borne out through identification and personal ownership-sense and entirely constitutes our sense of a self-other boundary. As an assumed personal owner, we certainly feel that we are the source of the value-ascriptions, the home of the mattering. As William James puts it: “It [the self] is what welcomes or rejects...It is the home of interest, - not the pleasant or the painful, not even pleasure or pain, as such, but that within us to which pleasure and pain, the pleasant and the painful, speak” (1890, 297-298).

The reciprocal feelings of personal ownership and self-identification seem to tap into that deep vein of self-interest that most of us have, motivating our actions, and providing strong grounds for ascribing a happiness-urge to the self that we assume we are. Locke (1690, 213) may be right to suppose that things get personal when our own happiness is at stake. It would seem that the more something is valued, whether positively or negatively, the stronger our potential feelings of ‘me’ and ‘mine’ with regards to that thing. In chapter eight we argue that Buddhism is correct in supposing that tanha and the co-arising reciprocal senses of self-identification and personal ownership do not only evidence but also help constitute the sense of self-other boundedness. We will suggest that this occurs in such a manner that no actual bounded self need be implicated in explaining the ubiquitous sense of a bounded self.

Attribute: Elusiveness

Roderick Chisholm writes:

The two great traditions of contemporary Western philosophy – ‘phenomenology’ and ‘logical analysis’ – seem to meet, unfortunately, at the extremes. The point of contact is the thesis according to which one is never aware of a subject of experience (1969, 94).

Chisholm is referring to the feature of ‘elusiveness’, which, like ‘boundedness’ and ‘the witness/observer’, is central to the notion of self being explicated. It is alluded to in the selected quotations by Ryle and Damasio. ‘Elusiveness’ pertains to the phenomenon whereby
the thing we take to be ‘the self’ cannot, in its capacity as a *subject*, be directly observed by the subject itself—either through introspection or observation via the five senses. It appears that one always catches an object of observation such as a thought or perception, but never the original subject of observation, in its distinct capacity of being a subject. The subject can never in other words be the focus of attention: the subject is always what focuses the attention. For all that, as Edey (1997) has implied, one has a definite sense of being a subject that is distinct from its objects, which is partly why elusiveness is such a puzzle. The feature of ‘elusiveness’ stands behind most the puzzles on self-knowledge, namely: given that we cannot directly observe the self as a subject, how can we know anything about it? Chisholm singles out Hume as one of the main thinkers who made ‘elusiveness’ famous, helping to spark the above-mentioned traditions of debate. Hume wrote:

> For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never catch *myself* at any time without a perception, and never can observe anything but the perception.’(1739, 162).

It should be noted that Hume did not explicitly point to ‘elusiveness’ as a feature integral to the assumed self—his exercise was not one of explicating the assumed idea of self, as we are trying to do now. Rather, his analysis presupposed this assumed idea of self—which featured qualities of ‘endurance’ and ‘invariability’ (still to be discussed). Upon failing to introspectively locate anything bearing these features, he sought to explain, in a manner consistent with his empiricism, the origin of such ‘mistaken ideas’ that we have about the self. We can thus infer that for Hume, ‘elusiveness’ was not so much a feature he grappled with in and of itself, as it was a feature that beset the very nature of his inquiry. It created for him the puzzle of where one would obtain the ‘ideas’ of uninterruptedness and invariability, given the elusiveness of these qualities (or rather, a *thing* with these qualities) to introspection. If the origin of all ideas is sensory, and all things sensory are observed as interrupted and variable, then from whence arise the contrary ideas? Any answers to this question would automatically bear upon the elusiveness issue, since the latter is woven into the problem. Common interpretation has it that Hume was ultimately dissatisfied with his empirically based ‘solution’.
‘Elusiveness’ and ‘the possibility of self-knowledge’ remain major puzzles to both analytic and continental philosophy. Many explanations have been offered – some of which deny that the ‘self’ is problematically elusive\(^{77}\) – but none of which are, to my knowledge, universally agreed upon. ‘Elusiveness’ will be returned to later in the thesis, when I attempt to account for this phenomenon in terms of a notion of consciousness derived from the concept of witnessing. At this stage of the inquiry, it is sufficient to point out that ‘elusiveness’, evidenced by the existence of these puzzles, is central to the self we assume we are.

**Attribute: Unity (Singularity)**

...although the things which I perceive or imagine are perhaps nothing at all apart from me [and in themselves], I am nevertheless assured that those modes of consciousness which I call perceptions and imaginations, in as far only as they are modes of consciousness, exist in me. (Descartes, 1641, 42).

I take ‘singularity’ or ‘unity’ to refer, firstly, to that aspect by which one unquestioningly regards all the different ‘roles’ (observer, owner, agent, thinker) to be occupied by the very same thing; viz the ‘self’ I take to be I or me. For instance, we do not reflexively assume the owner (whether personally or perspectivally construed) to be a different self from the ‘agent’, ‘thinker’ or ‘observer’. All are felt to be the very same subject of experience – the very same me – whether the different roles seem to manifest at one particular time (synchronously) or over time (diachronically). This becomes apparent through reflecting on the fact that the roles are not usually identified as in isolation, but are often fluid and overlapping, with different, changing weightings. For despite this change, there remains the sense of a single underlying ‘self’ that is not altered by the changing roles. Emotions are useful indicators of this point. Each emotion may involve the activation of several ‘roles’ at once, and there may be more than one emotion at a time. The embarrassed person may for instance identify simultaneously as the personal owner and past initiator of the shameful actions, as well as the thinker of the current thoughts. At any moment, while seeming to occupy a variety of such roles, feeling a range of shifting emotions and perceptions, one has the background impression that in spite of this affective and perceptual variation, there is only one underlying self that is the subject

of them all. Another way of putting it is that there appears to be only one ‘point of view’ or perspective. The feeling is not of a tangle of different underlying selves and points of view, waxing and waning, or even of compartments of an underlying self, each dedicated to its own particular role. In fact it is very hard, perhaps even impossible, to imagine what this would feel like. The closest real-life examples are of those suffering with ‘multiple personality syndrome’ but even then, there is a unity of roles within each ‘persona’ who happens to be ‘in control’ at the time.

Another way to approach ‘unity’ is by looking to contemporary philosophical puzzles that relate to what is known as the ‘unity of consciousness’, or what Hume called “simplicity” (1739, 163). This debate appears to focus more on objects of consciousness than on the different assumed roles of a ‘self’, but the underlying principle is the same. A popular formulation of the unity phenomenon, relevant to this context, is as follows. One may experience multiple sensations within different sense modalities (such as the taste and texture of popcorn along with the sound and sight of a movie) and yet one automatically feels, from the first-person perspective, that they belong to the very same self, (in its capacity of being a subject) rather than to different subject-selves. Another way of putting the point is to say that at a given time, whenever a variation of perspectively owned objects are assumed to be personally mine, they are also assumed to belong to a single, unified me. The self, in this manner, would seem to be a unified rather than differentiated entity.

Unity must not be confused with boundedness. It is tempting to suppose that because ‘unity’ seems to unite or bind disparate objects to a single point of view, then that affirmation of singularity must simultaneously affirm that subjective perspective’s nature as a bounded entity, separate from other things. Unity and boundedness, are, however polar opposites: in a slogan, unity unites, boundedness divides. Boundedness is defined with reference to the self-other boundary: it is evidenced by identification with some, but not other items as ‘me’. Unity, on the other hand, brings together objects that seem separate and divided, subsuming them under one conscious canopy. Now it may well turn out that unity and boundedness are importantly related; for example, that without unity there cannot be a sense of boundedness or identification. Boundedness thus seems, for instance, to involve a separation between a

---

78 A recent formulation of the unity puzzle relevant to this thesis (and to be discussed later) is found in Bayne and Chalmers (2003).
unified self and the rest of the world. However, this does not mean that unity and boundedness amount to the same thing. On the face of it, at least, unity and boundedness are two different aspects to the self that we assume are.

The philosophical question behind the puzzle of unity concerns what it is that is responsible for this apparent unity, whether at any one time, or over time. For example, could 'synchronic unity' reduce to the mere fact that the objects occur at the very same time or is there more to such unity than synchronicity? Could the unity be in fact an illusion? We will be returning to the 'unity of consciousness' later in the project. At this stage, we can note that as with 'elusiveness', reflection on experience, together with the existence of an associated philosophical puzzle is evidence that the phenomenon of unity is a feature of the self we assume we are.

**Attributes: Endurance and Invariability.**

The features of 'endurance' and 'invariability', the former known in Hume's writings as 'uninterruptedness', are aspects of what Hume called 'identity'. In the debate on the unity of consciousness, identity is implicated in the phenomenon and puzzle of diachronic unity. The debate assumes an impression or intuition of the very same, unchanging consciousness or self that unifies the various, multifarious conscious states, not only at any one time (synchronously), but also over time (diachronically), whether from one moment to the next, or for the span of a lifetime. The puzzle is where we get these intuitions of endurance (note, not perdurance) and invariability, given that, as Hume has noted, we can only detect a flux of rapidly changing perceptions when we introspectively "enter most intimately into what I call myself" (1739, 162). This of course ties in with the puzzle of elusiveness, since these endurance and invariability, should we intuit their reality, would seem quite elusive to the usual methods of observation such as introspection and perception. After trying to account for our apparent intuitions of both synchronic unity ("simplicity") and diachronic unity (involving identity) with reference such factors as the imagination, a dissatisfied Hume concludes in a postscript:

But all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought and consciousness. I cannot discover any theory, which gives me satisfaction on this head...(1740, 175).
In this section, it is asked whether Hume was right in supposing that identity is indeed a central aspect to the sense of self being elucidated. Do we really have intuitions of our selves as entities that are enduring and invariable? We begin by looking to the aspect of endurance.

Endurance.

Do we in any way take ourselves to have enduring, uninterrupted existence? I take my cue from Galen Strawson (1997) in addressing this question on two levels; namely, whether we attribute endurance to ourselves in (1) over large tracts of time, and (2) on an experiential, moment-to-moment basis. Turning first to (1), we have to first counter an immediate objection. In an earlier section we noted that a sense of self does not require a sense of existing over tracts of time that are longer than say, 45 seconds. What, then, are we to make of the suggestion that longer-term identity is attributed to the self we assume we are? I believe the right thing to say is that while a sense of longer-term identity is not necessary to a sense of self, it could still be (and usually is) sufficient for harbouring a sense of self and so a highly useful part of the definition. Long-term endurance, then, is still something that could qualify as important to the self we assume we are, so long as it is understood in its capacity as a sufficient rather than necessary part of that assumption.

With this proviso in place, we need not look too far for evidence of apparent longer-term numerical identity. First of all, there is the famous philosophical puzzle of personal identity, revolving around such questions as: what makes this person at t the same as that person at t-2, 25 years ago? Most formulations of the puzzle simply take it for granted that we have an intuition or sense of the self’s long-term numerical sameness: their job is to explain this intuition. Further evidence points to a close connection with the aspect of unity mentioned earlier. It was noted that we unquestioningly attribute all our different ‘roles’ (eg ‘actor’, ‘owner’, ‘thinker’, ‘knower’) to the very same subject of experience that is ‘me’, whether at any one time, or over long tracts of time. We do not take the underlying sense of self to numerically differ with the varying roles and experiences; with this self being the ‘owner’, that self being the ‘actor’; new selves being created every moment at the speed of thought, or upon awakening from deep sleep. The experience rather seems to be of
fluctuating experiences, through different roles, from the one unified perspective; a perspective befitting to a single selfy subject of experience.

A further source of evidence for assumed numerical longer-term identity appeals to the nature of the felt emotions. In most occasions when such emotions as guilt, fear, anxiety, disappointment, nervousness, joyful anticipation arise, their occurrence presupposes a feeling of numerical identity between the current ‘self’ who is feeling that emotion, and a past or future ‘self’ – a point noted by Canfield (1990, 19-20). Suppose that one feels guilty over an action committed at an earlier time. It is part of the nature of that guilt to assume that the current self, feeling guilty, is identical to a past self who initiated the act in question. Unless this identity was assumed, it would be difficult, if not impossible, to feel or explain the guilt. Suppose that one is nervous of an impending exam. It is part of the nature of that nervousness to believe that some subject, namely oneself, is going to sit a future exam and perhaps not do as well as one would hope. Many occurrences of emotion presuppose an identity between what would seem to be one’s current self and a past or future self. Later in the project I will be suggesting that the very having of these emotions is not merely evidential but is also partly constitutive of the sense of numerically identical self over time.

Let us now turn to (2) and consider some evidence for conscious moment-to-moment endurance of the self over time. This is the phenomenon whereby the self is not merely assumed or obliquely felt to have longer-term numerical identity, but is somehow actively experienced as enduring and unceasing, from moment to conscious moment. I say ‘somehow’ because as Hume discovered, such endurance is not obvious. The endurance seems elusive in that it cannot be directly perceived or introspected as we perceive and introspect the usual objects of consciousness. I say ‘conscious’ moment because there are episodes, such as deep dreamless sleep, where endurance of self, even if present, is not apparent. For our purposes it is enough to argue that between such apparent lapses in consciousness, there are lengthy episodes where one is conscious and that for the duration of these episodes, the ‘self’, through its conscious or witnessing aspect, has the character of moment-to-moment endurance as opposed to perdurance through rapidly pulsing blips. In favour of the ‘perdurance’ view, Galen Strawson holds that while we might superficially attribute endurance of consciousness to our experience, a little reflection will reveal, from the
first-person perspective, that our moment-to-moment experience is not one of enduring conscious existence. Rather, he says:

My claim is not just that there can be radical disjunction at the level of subject matter [presumably, types of experience]. Switches of subject matter could be absolute, and still be seamless in the sense that they involved no sensed temporal gap or felt interruption of consciousness. It seems to me, however, that such experience of temporal seamlessness is relatively rare. When I am alone and thinking I find that my fundamental experience of consciousness is one of \textit{repeated returns into consciousness from a state of complete, if momentary, unconsciousness}...the situation is best described, it seems to me, by saying that consciousness is continually \textit{restarting}. There isn’t a basic substrate (as it were) of continuous consciousness interrupted by various lapses and doglegs. (1997, 422).

Strawson’s claim (which, ironically, he aligns with Buddhism in a footnote) seems unconvincing for the following reasons. First, I hold that it seems logically impossible to \textit{directly} experience what Strawson calls ‘repeated returns into consciousness from a state of complete, if momentary, unconsciousness’. Direct experience implies some concurrent conscious awareness either \textit{of} or \textit{with} that mode of experience – unless we are distorting the word ‘experience’ to allow for unconscious episodes (where there is nothing it is like to experience it). On the conventional usage, then, to directly experience unconsciousness is a contradiction in terms, since in order to directly experience X one must be conscious. One cannot, therefore, directly experience episodes of unconsciousness. It is like trying to search for darkness with a flashlight.

Perhaps Galen Strawson is alluding to \textit{indirect} experience, such as that inferred through memory. One may for instance emerge from a general anaesthetic with the queer feeling that one’s consciousness has, as Strawson would put it, “bang[ed] out of nothingness” (1997, 422). Perhaps this queeriness is due to remembering how things were just before losing consciousness (such the hands of the clockface) and then noting, seemingly a moment later, that things are alarmingly different – four hours have elapsed. Does ordinary waking experience seem, on reflection, to be like a series of mini general anaesthetics? I doubt it. It is true that one can have moments of distraction and reverie, sometimes being unable to actively recall a few moments of conscious activity (famously when driving a car). But this seems to confuse poor attentiveness and memory (during which there was experience, although forgotten) with nothing being experienced at all. So long as one is
awake, it is highly unlikely that there is any moment when there is no object of consciousness, including those that are inattentively apprehended. And so long as there is some object of consciousness, whether direct or peripheral, we can be sure that there is consciousness or “inner life” (whether we identify consciousness with ‘witnessing’ or with the stream of objects that are witnessed). It is very unusual for a person’s mind, whilst awake, to be altogether free from thoughts or background emotions. Even if there was a temporary lull in thought processes, then at least one of the five senses will be bombarded with sensory stimuli – sounds, tastes, sights, smells, tactile experiences, not to mention an almost continuous flux of proprioceptive sensations. It is true that the sensations may rapidly arise and pass away. But importantly, there will hardly, if ever, be a moment when no sensory stimuli are present, or at least a moment when consciousness ceases for long enough for us to infer, as with a general anaesthetic, ‘nothing was happening then’.

Perhaps Strawson means to assert that each object stimulates its own unique moment of consciousness, so that there seems, on reflection, to be many numerically different but overlapping moments of consciousness banging out of nothingness. While this may turn out to be true in theory, it still jars with ordinary, everyday experience. For it seems, to me at least, as if the objects are unequivocally presented to one and the very same consciousness – to me, or to my consciousness. That, after all, is what spawns the puzzle of unity; the puzzle of how disparate objects seem to belong to the numerically same consciousness, whether at any one time, or over time. If this didn’t seem to be the case, then there wouldn’t be such a puzzle. It would also seem to suggest that consciousness is more fittingly identified with the witnessing aspect of ‘self’, than with the disparate witnessed objects, which come and go (more on this later). I thus conclude that Strawson is quite mistaken to suppose that consciousness, from a first-person perspective, seems gappy in the way that he suggests. For the reasons mentioned above, it seems to endure rather than perdure. It would seem, in particular, to be the witnessing aspect of self that seems to endure on a moment-to-moment basis.

Nor is it a merely impersonal witnessing that seems to endure. We have already surmised that a sense of personal ownership towards objects of witnessing – and hence an identity as their personal owner – is the usual case with regard to a person’s conscious life. To lose the sense of personal ownership or my-ness towards such objects as ordinary
perceptions, no matter how paltry, is often symptomatic of pathology (for example, depersonalisation or anosognosia). It seems reasonable to suppose, therefore, that it is a witnessing subject qua personal owner – that entity to whom the perceptions and other objects of consciousness seem to belong – who is assumed to endure, whether from moment-to-moment, or over a longer span of time.

Invariability

Invariability is that aspect of the self that bespeaks of qualitative identity over time (whether longer-term or moment-to-moment) despite any changes to personality and perceptions. Immediate evidence for our assumption of the self as invariable may be pointed out in the way we think and speak. When say such things as ‘I have changed’ it seems that we imply some underlying ‘I’ who is unchanging bearer of the change. Nor does this seem like a mere convention: in saying this, it does not seem as if the essential ‘me’ at an earlier time t-1 is literally different in its essence to the current ‘me’ at time t. It feels, rather, as if ‘I’ points to an essential, unique but subtle me-making quality that belies the change and bestows numerical identity, such that ‘there could only be one of me’. As Damasio puts it:

In all the kinds of self we can consider one notion always commands centre stage: the notion of a bounded, single individual that changes ever so gently across time, but, somehow, seems to stay the same...(1999, 134).

Given that the objects of consciousness, including those identified with as ‘me’, qualitatively change over time, it seems appropriate to look to witnessing feature to pinpoint what seems invariable about the self. Because it will likely be elusive, and perhaps always co-occurring with attention-absorbing, altering objects of consciousness, the qualitatively neutral witnessing aspect may be understandably difficult to pinpoint, if it is indeed a component of experience. What further evidence could there be for our assumption of the self as invariable?

Further evidence could lie in the fact that the subject-like, witnessing aspect of self, which appears to stand apart from its witnessed objects, seems to also carry its own non-intentional phenomenal feel as a self-subject (such that there seems, even if subtlety, something it is like to be a self). This phenomenal feel that the self seems bring to ordinary
conscious states does not seem reducible to qualia pertaining to objects of consciousness. If the phenomenal feel of 'me' cannot be reduced to object-qualia, then it would seem that there could not be any discoverable qualitative parameter along which it could vary over time. This would lend to the self an impression of invariability.

**Attribute: Unconstructedness**

'Unconstructedness' describes that aspect by which the self stands apart from, and is not constructed by, the kind of objects that can be perspectively owned – thoughts, ideas, emotions, perceptions, experiences and so forth. This aspect flows naturally from everything that has been said about the self in this chapter: the fact that the self appears as a bounded, unified, enduring, invariable, elusive subject of experience. In ordinary conscious states, the elusive subject always appears separate to any objects – as that which observes, perceives, witnesses, thinks – never as that which is observed, perceived, witnessed or thought. In appearing to stand apart from its objects in this way, the self, with its subject-like character, purports to be of an essential nature that is not ontologically contributed to by such objects at any given time. Insofar as first-person experience leads us to forge a subject/object distinction – with the subject qualified by its selfy roles and attributes – it would seem as if there is something it is like to be this elusive, bounded self. The intrinsic phenomenal, subject-like character of this self would seem to be unborrowed by any observable objects of experience. To be sure, the self may appear to originate or 'house' such objects, as thoughts, but always in a context where the objects appear as non-identical to the self.

This remains the case, even while the subject is identifying with various objects as 'me' through appropriating ideas of them to its perspective. Contrary to what one might initially think, appropriation does not overtly interfere with the impression of the self's unconstructedness. While appropriated ideas are still, ontologically speaking, objects of consciousness – of the sort that can be perspectively owned – their status as such objects becomes effectively invisible to the identifying subject. If they did not become invisible, but still appeared as overt objects to the subject, then they would not have been appropriated to the perspective of the subject (they could be separate thoughts about the appropriated idea, for instance). When 'the self' confronts the world as 'other', any appropriated idea will always be qua unified elusive subject, never qua separate observed object. This curious
shrouding effect that appropriation has upon the selected object-ideas thus does nothing to impede (and perhaps even enhances) the subject’s impression that as a bounded self, its ontology is free from any potential contribution by those appropriated object-ideas. The subject hence retains the impression that the self, the thing they assume they are, is unconstructed.

The subject’s appropriation of ideas to its perspective does not imply that the self is actually something that is constructed, in any essential way, by those appropriated ideas. While appropriation serves as the best evidence for the fact that we assume our selves to be bounded – evidence that does not seem to interfere with the self’s unconstructedness – it does not prove that the boundedness is in fact constituted by such appropriations. From the first-person perspective, indeed, it would seem as if a ready-bound self is identifying with various items, rather than as if the identifications are constituting the very boundary between self and the world. Buddhism holds that appropriation does constitute the sense of boundedness – but this is moving away from our concern of characterising the self we assume we are, to a different concern of how that assumption is generated.

There are further matters to be clear about. By ascribing unconstructedness to the self that we purport to be, we are not saying that the self purports to be the type of thing that could, at a given time, exist in isolation of all objects of consciousness. Nothing about the assumed self seems to preclude the possibility of the self being something that at any given time will necessarily co-exist with its overtly observed objects. Perhaps consciousness with a sense of self is inherently dualistic, such that whenever there appears to be a self, there must also appear to be objects that are observed by the self. What is being claimed, rather, pertains to that very sense of dualism: the unbridgeable gap that is felt between self (with its intrinsic non-intentional phenomenology) and object in a given state of consciousness, at a given moment.

Nor are we saying that unconstructedness per se amounts to boundedness (although unconstructedness of the self implies boundedness). Unconstructedness simply pertains to the fact that the self qua witnessing subject does not seem constructed by the kind of mental objects that are witnessed, just as diffuse light does not seem reducible to the objects it illuminates. Boundedness (e.g., as evidenced through identification) seems to add a definitive, although elusive border of ‘me’ around this witnessing, as a cloud of dust might seem to add
shape to the diffuse light. Just as the collective light-and-dust can seem to offer a unified source of illumination for other objects, the subject-and-identified-with-item can seem, from the first-person perspective, to be a unified source of observing, perceiving or thinking about other objects. Unconstructedness pertains strictly to that ‘illuminating’ or witnessing aspect by which the subject – bounded or not – does not seem to be constituted by any observable mental objects. Unconstructedness would hence also characterise the feeling of perspectival without personal ownership.

**Conclusion**

It has been argued in this chapter that most people do indeed assume themselves to be self-entities of the kind alluded to by both Buddhism and Western philosophers. That is, most people assume themselves to be a witnessing or conscious subject, which occupies the roles of observer/witness, owner (perspectival and personal), freely-willed agent of actions and thinker of thoughts, and happiness-seeker/avoider of suffering. People also assume this same self-entity to bear the attributes of ontological uniqueness/boundedness, elusiveness, unity, endurance, invariability and unconstructedness. We already know that Buddhism denies objective reality to this self as a whole. As it happens, many Western philosophers have also denied objective reality to this self and in the next chapter we look to some such evidence. But before doing this, we investigate what it means to deny reality to such a self and whether such a denial could actually mean the same thing in Buddhism and the West, given the different metaphysical assumptions.
Chapter Five: How Do We Construe “The Self Lacks Reality”?

Introduction

We have argued that most people assume themselves to be a witnessing or conscious subject which (1) occupies the roles of observer, owner (perspectival and personal), freely-willed agent of actions and thinker of thoughts, and seeker of happiness over suffering and which (2) bears the attributes of boundedness, elusiveness, unity, endurance, invariability and unconstructuredness. The aim of this chapter is to (a) gain a clearer understanding of what it would mean to deny reality to this self, and to (b) provide evidence that the self’s reality has, in this particular sense, been denied by several leading Western thinkers (as well as by Buddhism). In context of the wider project, this is an important step, since we cannot argue that the self lacks reality – and hence defend the Buddhist position – unless we know what it means for the self to lack reality. It will emerge that the self would lack reality by being constructed and illusory (concepts that we spend some time explaining). By referring to Western thinkers who deny the self reality by arguing it is constructed and illusory, we demonstrate the relevance of this project to Western concerns: deeming the self constructed and illusory is not peculiar to Buddhism. Within the context of this framework, however, there are major differences to emerge between Buddhist and Western ontologies of self. We have noted in chapter three that while Buddhism regards boundedness as the main reason for the self’s overall illusory status, it does not deny reality to a number of further ascribed features such as unity, endurance and invariability. We shall see, by contrast, that the Western counterparts do effectively deny reality to these and other features considered real in Buddhism, by regarding them as constructed. This will point to a significant divergence in the extent to which the reality of self is denied by Buddhism and the West (a divergence not

---

79 No philosophers – or Buddhists for that matter – have explicitly described a notion of self with all the features depicted in the above formulation. However, the fact that they refer to something along the lines of a unified, unchanging, personalised (hence bounded) subject of experience is clear evidence they are concerned with a reflexively assumed self. From the arguments of chapter four, they are also, ipso facto, concerned with the self characterised above.
emphasised in the comparative literature). It is the Buddhist rather than 'Western' account of no-self that this project will, in later chapters, seek to defend.

**What does it Mean to Deny Reality to the Self?**

If 'reality' pertains simply to the metaphysical standard by which things are granted or denied existence, then we may have a difficult time finding an agreed-upon standard by which the self could be said to lack reality. We have already seen that the metaphysics of Buddhism, in relation to which the self is deemed to lack existence, involves the unconditioned nibbana. In Western philosophy, whose metaphysical standards would generally preclude the reality of nibbana, denying existence to the self may mean something very different. For example, when Richard Baron (2000) says that the self is unreal, the standards by which he construes the self to not exist are those informed by what is sometimes known as the 'scientific image'. On the scientific image, the sum total of things regarded as real are things that can be revealed through careful scientific method. Such methods, says Baron (who seems to adopt the additional stance of eliminative materialist) reveal no posits of 'folk' psychology, including thoughts, emotions, and an observing, thinking self-entity with this-ness and autonomy (2000, sec 2-3). Western philosophers who reject Baron’s ‘elevation’ of the scientific worldview – allowing thoughts and feelings into their ontology – may deny the self reality because (being elusive) the self is not introspectively (or externally) observable. Such a line seemed to be taken by Hume (1739).

Given the divergent metaphysical systems that can inform someone’s favoured standard of reality, can there be any convergence in the way that different thinkers, regardless of background metaphysics, deny the self reality? I believe that there can be such a convergence. The self has sometimes been claimed, in both Buddhism and the West, to be illusory. Once we articulate how the self is deemed illusory (first by describing what is meant, most broadly, by ‘illusion’) it will become clear how the self could be denied reality in a way that stays neutral on the background metaphysics.

**What is an Illusion?**

For our purposes, we shall work with a notion of ‘illusion’ that is broad. In essence,

---

80 I thank Jane McKessar for first drawing my attention to this point.
Where X purports to exist in manner F, to subject S, X-as-F is illusory iff X does not really exist in manner F\textsuperscript{81}.

This formulation will cover not only standard perceptual illusions (such as the Mueller-Lyer illusion) but also delusions (such as that of the man who believes he is being monitored by aliens) and hallucinations, such as that of the pink elephant squatting on the Sydney Harbour Bridge (did you see it too?). We choose this broad formulation because the manner in which the self can be illusory, while nicely captured by this formulation, does not fit into any of the conventional sub-categories (as will shortly be demonstrated). Let us now say more about this formulation.

All illusions (including delusions and hallucinations) will involve appearances to the perspective of a conscious subject who has particular perceptual or cognitive faculties. Indeed, unless we radically changed the meaning of the word ‘illusion’, it is not at all clear how an X-as-F could count as illusory and yet X not appear (or be disposed to appear) in a manner F to a subject – a manner that will conflict with X’s actual ontological status. By the formulation ‘Where X purports to exist in manner F, X-as-F is illusory iff X does not really exist in manner F’ we mean the following: X appears to a subject in such a way that X seems to ‘wear on its sleeve’ a particular manner of existing, F; a manner which turns out to conflict with X’s actual manner of existing, as not-F. In this way the appearances of X-as-F are cognitive, openly conveying a message about X’s manner of existence – a message that is false. For example, suppose that X is ‘two lines of unequal length’ and F is ‘existing independently of the subject’. In the case of the Mueller-Lyer illusion, a message of X-as-F is conveyed to the subject: there is the appearance of there being two lines of unequal length. What the appearance conveys – X-as-F – will be illusory because the actual external situation does not reflect X existing in manner F (as would seem to be the case) but reflects X existing in a manner that is not-F (namely as the mere content of an appearance).

A person harbouring an illusion need not be fooled by it in such a way that they wholly believe X to exist in the purported manner, F. Those familiar with the Mueller-Lyer illusion will not actually believe the lines to be of objectively uneven lengths. Yet the lines will still visually appear to them as objectively uneven in length. It is in this capacity that the

\textsuperscript{81} I am grateful to John Baker for helping me to formulate this definition.
phenomenon is classed as a standard optical illusion. Someone on hallucinogenic drugs who ‘hears’ a chorus of angels as purporting to exist from heaven above (as opposed to from their head) will not, should they know that this to be the effect of drugs, believe that they are really hearing a heavenly chorus. The heavenly chorus (the content of a hallucination) will nevertheless be illusory – purporting to exist (as straight from heaven – hence independent of perception) in a manner in which it does not actually exist (it exists only in their minds). Sometimes, however, the illusion cuts deeper into one’s cognitive set so that the very appearance of X involves central beliefs about X’s mode of existence. The schizophrenic who hears voices in his head and is ignorant of his condition is likely to completely buy into the objective reality of these voices, believing, for example, that the voice is of God telling him to go on a mission. When beliefs as major as this are drawn into the illusion, the phenomenon tends to be classed as a delusion. The voices from God nevertheless qualify as illusory on our definition, since they purport to exist in a manner (namely, as really from God and hence independent of perception) in which they do not actually exist. It is just that the mechanism of ‘purporting’ draws upon a deeper set of beliefs in the subject’s cognitive set.

Sometimes an illusion qua delusion will be almost entirely doxastic or convictional, involving no specific sensory-perceptual dimension. “Jim” may for instance be gripped by the conviction that aliens are watching him, even though he has not been presented with any ‘perceptual cues’ that bespeak of this impression. He nevertheless has a distinct sense of being watched by aliens, where ‘sense’ is construed in the manner outlined in chapter one. We may recall that ‘sense of X’ in this context involves a general conscious impression such that there is something it is like for a subject to have an impression of X. However, ‘sense’ is not a success-term – the X sensed need not have objective reality, but may be illusory. In the case of Jim who has a sense of being watched by aliens, the watching aliens need not – and do not – have the ontological status they purport to have, namely that of existing outside of Jim’s mind. The fact that they purport to have such status, from the perspective of poor Jim, renders the watching aliens to be illusory.

In its normal usage, the term ‘delusion’ has connotations of involving a subject’s deeper beliefs, together with a cognitive malfunction or abnormality – such as in the case of Jim. The term ‘illusion’, at least in its more restricted usage, has connotations of being perceptually oriented and not involving malfunction or abnormality. (Indeed, if one failed to be optically
appreciate the Mueller-Lyer illusion, then that would be a sign of abnormality or malfunction). When it comes to articulating how the self could possibly be illusory, however, it is apparent that while involving elements from both camps, it will not seem to fit exclusively into one or the other. The self, if lacking reality, would seem like a delusion in the following respect: its appearance or ‘sense’, as we determined in chapter one, would involve a reflexively assumed belief in the self’s existence – a belief which turns out to be false. And yet unlike a delusion, this reflexively assumed false belief will not be considered, at least in the West, to involve a cognitive malfunction or abnormality. Indeed it is statistically very normal. But nor would the sense of self obviously class as a standard illusion since, like the sense of being watched by aliens, it is doxastically rather than perceptually oriented. And it would not, of course, class as a hallucination. For this reason, I have formulated the broader definition to capture the way in which the self might, plausibly, be classed as illusory. We can hence surmise, on our broad definition of illusion, that should the self that is sensed not actually exist in its purported manner, then the self will count as illusory. The self will be purporting to exist in a manner in which it does not actually exist. (We shall later specify just what this purported manner would amount to).

Before going further, we need to avert a possible confusion by pointing to an ambiguity in the way that the term ‘illusion’ can plausibly be deployed. We need to distinguish illusion qua the appearance-of-(X-as-F)-to-S (the vehicle) from illusion qua X-as-F (the content, viz., thing conveyed through the appearance-of-(X-as-F)-to-S). While our stipulation defines the term ‘illusion’ in the latter sense (through the word ‘illusory’), the former sense is also common. With the Mueller-Lyer illusion, for instance, the very appearance of two objectively uneven lines to a subject is commonly termed ‘an illusion’. But the illusion qua appearance is not the thing whose existence is in question. Rather, the thing whose existence in question is the content of that appearance, namely, the illusion qua two objectively uneven lines. Indeed, unless the illusion qua appearance-vehicle (to S) existed, there could not be the illusion qua content-of-appearance, viz., two objectively uneven lines. A similar story applies to the other kinds of illusion. The term ‘illusion’ can refer to the existing appearance of purportedly external voices, such that the subject can be said to be harbouring such an appearance, or be in the grip of an illusion. In this manner of speaking, the illusion exists as a real something that is ‘had’ by the subject. But we can also say, as we have predominantly been doing in this
discussion, that the voices, since they purport to exist externally, are an illusion (or illusory). 

*Qua* content-of-appearance, the voices do not exist in the externally existing manner that they, through their appearance, purport to exist. And regarding the illusion of self (should the self be an illusion), we already know that the actual sense of self, which would serve as the vehicle/appearance for the illusory content, would be real enough. Indeed, unless there really *was* a sense of self, then there could not be a conveyed *self* that may turn out to not exist in its purported manner. Unless I stipulate otherwise, I shall continue to use the term ‘illusion’ or ‘illusory’, to talk about the X-as-F (whose existence is in question) rather than the sense or appearance of X-as-F (whose existence is not in question).

**The Self as Construct and Illusion**

When considering how the self’s purported manner of existence may conflict with its actual ontological status, it is useful to first analyse in more detail how this scenario may apply to phenomena that do not involve the self, and then see if this analysis applies to the self. In the other examples of illusion we have looked at, there has been an implicit standard of reality by which (a) the item in question has, from the perspective of the subject, purported to exist and (b) the item is deemed, from a perspective outside of the subject, to not exist. This standard of reality is invariably a reality that is *external to the perspective of the observing subject*—regardless of how the details of this reality are cashed out. With the Mueller-Lyer illusion, for instance, the standard of reality by which the content *<two objectively uneven lines>* seems optically real to the observer, is precisely a reality that seems, from the observer’s perspective, to exist *beyond* her perspective as an observer. *<Two objectively uneven lines>* wears on its sleeve the message, “measure me with a ruler and I’ll be uneven!” It purports to have a reality that will meet any measuring criteria that would prove it an objective, perspective-independent, fact. It is by reference to this perspective-independent standard of reality that *<two objectively uneven lines>* proves to be an illusion, not existing in the perspective-independent manner that it purports, from the subject’s perspective, to exist. Likewise, the person who ‘hears’ voices from God presumes (with greater doxastic depth) that the voices-of-God exist beyond the mere confines of his perspective. This purported perspective-independent reality serves as the

---

82 I adopt the convention of referring to the *<content of appearance>* in these special brackets, when the context renders it unsuitable to commit to one ontological status or another.
standard by which the voices-of-God get judged (by outsiders) as illusory; as not existing in the manner that they purport, from the subject’s perspective, to exist.

As well as this, the same system of perspective-independent reality (regardless of its details) can offer an alternative explanation for the very appearance of two objectively uneven lines or the very appearance of God’s voice, an explanation that does not appeal anywhere to <two uneven lines> or to <voices from God>. The explanation may include such factors as neurochemical imbalance, but importantly, it will always appeal to cognitive and conscious components that directly contribute to the mental or subjective impression of the entity in question. Such an explanation will hence appeal to such factors as memory, imagination, qualia, thoughts and perceptions, with perspectively owned objects of consciousness (whether attended to or not) figuring in the explanation. In this way, we can say that the content <two objectively uneven lines> or <voices from God> is nothing other than a mental construct, not existing beyond its subjective, perspectively owned impression. The illusory status comes about from the fact that they purport to have an unconstructed reality, that is, to exist beyond their subjective, perspectively owned impression.

Let us define more precisely the notion of ‘mental construct’ (and its contrast-class of unconstructed X) and then the notion of ‘illusion’ in relation to this:

**D1:** Constructed X: X is a construct iff X does not exist beyond the appearance or presentation, of which it is the content.

**D1:** Unconstructed X: X is unconstructed iff X does exist beyond the appearance or presentation, of which it is the content.

**D1:** Illusory X: X is illusory iff X is a construct and yet purports to be unconstructed. That is, X does not exist beyond the appearance or presentation, of which it is the content, and yet X purports to exist beyond the appearance or presentation, of which it is the content.

Suppose I ‘hear’ a voice. The voice is the content of a sensory-like presentation. I appear to hear a voice and I assume that the voice ‘heard’ has an existence independent of its manner of appearing to me. Suppose that the voice does in fact exist in its purported manner and is independent of me, represented by its auditory appearance. On our analysis, the voice turns out
to be unconstructed and non-illusory, as we should expect. Suppose that I now ‘hear’ a voice that sounds just like the first and which also purports to exist independently of its auditory appearance. The voice, however, is a figment of my imagination, and is hence a construct that does not exist beyond the auditory appearance, of which it is the content. Its existence actually depends upon input from cognitive factors contributing to the hallucination – factors which do not include an external voice in their ontology. On our analysis, the voice will be illusory (as we should expect) since it purports to exist in a manner (namely as an external, unconstructed voice) in which it does not actually exist (it is a cognitive construct). The analysis also allow for constructs that are not illusory. Much of the material one thinks about does not exist beyond those thought-vehicles of which it is the content – and nor does it purport to exist beyond those thought-vehicles. For example I plan to visit the philosophy department at UWA<sup>83</sup> before I leave Perth. I imagine walking around the department, saying goodbye, but not for a moment am I under the impression that the scenario <saying goodbye to people at UWA> exists beyond my imagination of which it is the content.

Does the proposed analysis offer a way in which the self can be considered to not exist in a manner in which it purports to exist, by virtue of actually being constructed? The self purports, from the first-person perspective, to be a unified, elusive, enduring, bounded, subject-entity that thinks the thoughts, owns the perceptions, initiates the actions – always appearing in opposition to its objects – never as constructed by them. Yet it could quite conceivably turn out that the unconstructed entity we have an impression of being is in fact an entity which is the content of an appearance whose existence is constructed by those very factors that it purports to stand apart from, namely, thoughts, emotions and perceptions etc. Going by the above definitions of ‘construct’ and ‘illusion’, we should be able to say from this that the self, while purporting to be unconstructed – existing beyond the appearance, of which it is the content – is in fact constructed, viz., not existing beyond the appearance, of which it is the content.

But this does not, upon reflection, seem like the right thing to say. The self is a peculiar kind of entity, in that it seems to represent its own objective existence both transparently and reflexively by simply being the subjective entity that it is. The self does not purport to exist beyond its own appearance. So while on one hand, the sense of self does

<sup>83</sup> The University of Western Australia.
convey the message <conscious, witnessing, bounded, elusive, unified subject (etc.) who is owner, actor, thinker (etc.)>, on the other hand, there does not seem, phenomenologically, to be a gap between the message conveyed (<the self>) and that which conveys the message (the sense of self). The self seems to be its own ontological ambassador, with the sense of self appearing to be none other than the self. Just as there is no felt gap between the subject *per se* and the self (as discussed in chapter one) there seems to be no felt gap between the self one purports to be, and the sense of being that self (we will shortly see why there cannot seem to be this gap). What it is like to be a conscious self who thinks thoughts etc., seems no different from what it is like to have a sense of being a self who thinks thoughts etc. The self we assume we are simply *is* a conscious entity with certain roles and attributes, its lived, internal what-it-is-like-to-be-a-self-ness (in relation to its observed objects) packed into its ontology, so to speak. To have a sense of being this entity does not seem to require an additional meta-sense, such that there seems something it is like to be an entity, of which there is something it is like to be that entity. In short, the self purports to not exist beyond the very appearance (the sense of self) of which it is the content, since the sense of self seems to reflexively represent the self’s existence as if it *were* that very self.  

The upshot is that the self comes out, on the above definitions, as purporting to be constructed – not unconstructed, as we should expect. Before we modify the definitions to accommodate the way that the self could truly be constructed, unconstructed or illusory, it is worth reiterating that in its present form, the definitions seem to work well for Xs that are not the self. Such Xs differ from the self in that their manner of being constructed, unconstructed or illusory will always *presuppose* a felt or sensed distinction between the X that is the content, and the appearance in which X is the content. For example, when I mistake the voice that I ‘hear’, to be something outside of the appearance of which it is the content, the mistake is not that I fail to differentiate appearance from its content: the mistake is that I suppose the voice to have an existence beyond of the appearance of which it is the content. That is, I suppose the voice to be unconstructed, not in a fashion that presupposes its auditory appearance to be none other than the voice that is heard, but in a fashion that presupposes *this very distinction* between voice and its auditory presentation. Suppose that I now imagine (in

---

84 This would also explain why, after a paper presentation at UWA, I got virtually laughed off the stage for suggesting there was a difference between the self and the sense of self. Phenomenologically, there does not seem to be a difference.
a non-illusory way) someone singing, such that <a singing voice> is the content of a presentation whose medium is imagination. The singing voice is a construct. Once again, I am never under the impression that the singing voice is at one with its imaginary presentation: the medium through which I imagine the voice seems like one thing, the voice imagined, another.

We might wonder if there is something that unites those Xs to which the above definitions of constructed, unconstructed and illusory successfully apply – something that may shed further light on why the definitions fail to apply to the self. Indeed there is. It would seem that the Xs to which the above definitions of ‘constructed’, ‘unconstructed’ and ‘illusory’ successfully apply are all objects of which a subject can be conscious. <Voices from God>, <philosophy department scenarios>, <objectively uneven lines> and <pink elephants> are all obviously objects of consciousness. Perhaps, then, the above definitions fail to apply to the self precisely because the self does not purport to be an object, but a subject of consciousness (a subject that is a bounded, unified entity etc.). This purporting to be the subject of consciousness, by the very definition of ‘subject’, means purporting to stand apart from perspectively owned objects that appear to it, objects such as thoughts, perceptions, and imaginations. In purporting to stand apart from such objects that appear to it, there can be no felt gap between the self that purports to be unified, bounded subject of experience and the sense of being that entity. If there were a felt gap, such that the self appeared as an attentively observable object to an observing subject that sensed it – then it would no longer be the self. Part and parcel of selfhood is that the self presents itself as being at one with the observing subject – of which there is something it is like to be that observing subject – never as an observable, separate object. It is in this capacity that the self is presented as being an unconstructed, subjective entity – unconstructed by any observable object that could be perspectively owned by it – objects that include thoughts, feelings, imaginations and so forth.

We have discerned, hence, that the self, purporting to be unified with the observing subject, will purport to appear in such a manner that its very presentation (viz., sense of self) is none other than the self. The sense of self, seeming identical to the ontology of self, will hence purport to be stand apart from its observable objects and be unconstructed by them in
this fashion. In view of this, we are now ready propose the following additions to our
definitions such that they can accommodate the subject-like self as well as objects:

**D2: Constructed X:** X is constructed iff X is the content of an appearance or
presentation that is contributed to by objects of consciousness such as
thoughts, feelings, perceptions and emotions AND X does not exist beyond
the appearance or presentation, of which it is the content.

**D2: Unconstructed X:** X is unconstructed iff EITHER X is the content of an
appearance or presentation that is NOT contributed to by objects of consciousness
such as thoughts, feelings, perceptions and emotions, OR X does exist beyond the
appearance or presentation, of which it is the content.

**D2: Illusory X:** X is illusory iff X is constructed and purports to be unconstructed.
That is, X is illusory iff EITHER X is the content of an appearance or presentation
that is contributed to by objects of consciousness and yet, X purports to be the content
of an appearance or presentation which is not contributed to by objects of
consciousness, OR X does not exist beyond the appearance or presentation, of which it is the content.

These definitions allow us to express the way in which the self, as well as objects, can be
constructed, unconstructed, or illusory. If the self is a construct, then it will exist as the
content of a presentation that is contributed to by objects of consciousness such as thoughts,
and will not exist beyond *this* presentation, of which it is the content. But because the self
purports to be *unconstructed*, that is, to be the content of a presentation *not* contributed to by
objects of consciousness (such as thoughts), then the self, if actually constructed, will be
illusory. The self-as-unconstructed, viz., the self, will be purporting to exist in a manner in
which it does not actually exist — as something unconstructed when it is in fact a construct,
viz., the mere content of a presentation (the sense of self) contributed to by objects of
consciousness. It could hence conceivably turn out that the unconstructed entity we deeply

---

85 Do after-images count as illusory or constructed on these definitions? I would say that so long as the subject
is clear upon their status as afterimages, they would not count as constructs (hence not as illusions) since they
seem to exist as the kind of raw sensory ‘stuff’ of which mental constructs are built. Should the subject,
however, mistake an afterimage for (say) a green blob on the wall, then there would be an illusion at hand, since
the experience of the afterimage is now carrying the content <green blob on wall> and hence purporting to exist
beyond the appearance of which it is the content.
assume we are is in fact an entity whose existence is constructed by those very factors that it
purports to stand apart from, namely, thoughts, emotions and perceptions etc. The self that
we have a sense of being would, on such a scenario, be no more than the content of an
appearance that is contributed to by these factors. The sense of self would not, as it would
appear from the first-person perspective, be accounted for by the unconstructed self (namely
by being the unconstructed self) – it would instead be accounted for by factors that do not
appeal to ‘the unconstructed self’ in their ontology. Rather than the self thinking the
thoughts, the thoughts would be helping to think the self.

An Intersection of Agreement for Those who Deny Reality to the Self

The definitions and analysis provide a clear intersection of agreement for thinkers who deny
reality to the self, despite different informing metaphysics. They can all agree that the self
fundamentally lacks reality because it is an illusion, purporting to exist in a manner in which
it turns out to not exist. They can all agree that the self purports to be exist as something that
is unconstructed by thoughts, perceptions and so forth (unconstructedness, after all, is
essential to the self’s formulation). They can all agree that the self’s purported manner of
existence, as unconstructed, conflicts with its actual manner of existence, as constructed.
They can all agree that the appearance-with-content (namely, the sense of self) will be
accounted for not by an unconstructed self (which would seem to account for it), but by
cognitive components such as thoughts, ideas, perceptions, emotions and witnessing. In
short, there can be enough agreement upon the manner in which the self fails to exist –
through its being deemed as constructed and illusory – to make universal sense of the claim
that the self lacks reality.

This convergent way of denying reality to the self still manages to accommodate the
different reasons for which the self can be said to lack reality. The differences will lie in the
accounts of how and why the self is constructed (and hence illusory). For example, someone
taking a Humean or Jamesian line may claim that the self comes to be constructed through the
fact that its identity (endurance and invariability) and unity has no reality beyond a bundle of
thoughts and perceptions (etc.), which contribute to the appearance of unity and identity.
Should such an account be correct, then this, together with the fact that the unity and identity
are purported to be unconstructed features of the self (not contributed to by thoughts etc.),
would render the self, as whole, to be a construct and hence illusion. Someone taking a
Buddhist line, on the other hand, may say that the self comes to be constructed through the fact
that its boundedness, through the process of identification, is a product of nibbanic witnessing
appropriating various ideas to its perspective. They will claim that through the mechanism of
appropriation, the impersonal, nibbanic witnessing comes to be felt as a personal bounded self,
rendering the boundedness (because it depends upon contribution from thoughts etc.) to be a
construct. Should such an account be correct, then this, together with the fact the boundedness is *purported* to be an unconstructed feature of the self, would render the self, as a whole, to be a
construct and hence illusion.

The following schema provides three-fold way of thinking about the scenario of “self
as constructed and illusory”. While broadly equivalent, each way provides a different angle
or approach to construing the self as constructed and illusory. The schema will be useful to
reflect upon when looking at evidence from Western philosophers who deny reality to the
self.

1) Through its various roles (e.g., as thinker of thoughts, owner of experiences, initiator of
actions), we take the self to be a conscious subject-entity, which is ontologically separate
from, hence unconstructed by, thoughts, perception and other experiences. In reality, the
self is not ontologically separate from the thoughts etc, but is the content of an idea that is
created, at least in part, by the thoughts, perceptions and so forth. Thus the self does not
precede or create the thoughts etc; rather the thoughts etc., go towards creating the idea of
it. In this way, the fundamental duality between a thinking, perceiving, bounded subject
on one side and its apprehended objects, thoughts and ideas on the other, is constructed
and hence illusory. This has bearing on the various roles and aspects as they are applied
to self.

2) We take the thoughts etc to be owned – perspectivally and personally – by a self, when in
reality they are not owned by a self. The idea that we, the self, own our thoughts and
perceptions etc., is caused at least in part by the edifice of thought and perception (etc.)
that comprises the sense of self, rather than by a thought-independent owner, the self.
Similar considerations apply for other such roles as actor and thinker. We take the self to
be the initiator of (free) actions and thinker of thoughts. In reality, there is no self playing these roles; 'actor' and 'thinker' are, at least in part, a fictitious invention of the thoughts (etc). Neither is there a bounded witness-self who knows the experiences (although Buddhist accounts construe witnessing to be real enough).

3) We believe that there is a bounded, enduring and invariable, unconstructed entity which, through its various roles, unifies the thoughts and experiences as belonging to, and issuing from, it. In reality, there is no such thing, but only a flux of thought and perception along with mental faculties such as memory and imagination. The Buddhist account also includes witnessing, which is construed as enduring and invariable, a source of the apparent unity. But importantly, there is no room in this picture, whether painted by East or West, for an entity described as 'the self' that serves to unify the thoughts. If there is a genuine principle of unity, then this principle is not grounded in the self-entity.

**Western Thinkers Who Deny Reality to the Self**

From chapters three and four (and the brief recapitulation in this chapter) we should have gained a clear enough idea of how Buddhism regards the self to be constructed and illusory, with boundedness and personal ownership (through identification) being considered the primary reason for the self's constructed and illusory status. The Buddhist account is notable in that most of the features ascribed to the self such as witnessing, perspectival ownership, unity, elusiveness, endurance, invariability and unconstructedness, are not, in and of themselves, considered to be thoroughly constructed or illusory. Insofar as they are undistorted by the impression of boundedness, these features, in and of themselves, are considered to be the content of an appearance that is not contributed to by objects of consciousness. The role of these features in maintaining the overall illusion of self is therefore only a secondary one, such that they form the ontological background against which a sense of boundedness and personal ownership can take hold. We shall see that several leading Western thinkers also regard the self to be constructed and illusory. From their quotations to follow, it will be evident that while agreeing with Buddhism that the self as a whole is constructed (and hence illusory), these Westerners provide a very different story as to how the self is constructed. The features of unity, endurance and invariability are
often enlisted as the *primary* reason why the self, as a whole, is to be regarded as constructed and therefore illusory. These features, in and of themselves, are considered to be entirely constructed.

We begin with David Hume (1739). Let us first recall the evidence that Hume was indeed concerned with the notion of self we have been discussing all along. When he writes: “But the self or person is not any one impression, but that to which our several impressions and ideas are supposed to have a reference” (1739, 161-162) it seems he is indeed capturing the relevant notion of self, viz., a entity which *has* thoughts, ideas and perceptions, but is not identical to them. Hume also ascribes “simplicity” to the self (1739, 163), a principle of unity by which a subject’s diverse thoughts and perceptions are owned by a single entity. Most famously, Hume’s discussion supposes a common ascription of diachronic unity or *identity* to the self, by which we are compelled to “suppose ourselves possessed of an invariable and uninterrupted existence through the whole course of our lives” (1739, 163) (notions discussed under ‘invariability’ and ‘endurance’ in chapter four). Now the actual state of affairs, according to Hume, is quite contrary to how things appear. Instead of there being a constant, underlying entity which unites the varying perceptions and accounts for their identity, there is merely:

> a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement (1739, 162)....There is properly no simplicity in it [the mind] at any one time, nor *identity* in different, whatever natural propension we may have to imagine that simplicity and identity (1739, 163).

He goes on to argue that:

> The identity which we ascribe to the mind of man is only a fictitious [viz., constructed] one, and of a like kind with that which we ascribe to vegetables and animal bodies. It cannot therefore have a different origin, but must proceed from a like operation in the imagination upon like objects. (1739, 168)...identity [and simplicity] is nothing really belonging to these different perceptions, and uniting them together, but is merely a quality which we attribute to them, because of the union of their ideas in the imagination when we reflect upon them (1739, 169).
Hume clearly regards the appearance of the self's unity and identity to be accounted for not by factors that include actual unconstructed unity and identity, but by mental factors such as "the union of their ideas in the imagination when we reflect upon them". The self, on Hume's theory, is a mental construct, in virtue of its constructed unity and identity. The illusion is hence that of an unconstructed entity, viz., the self, that is not diverse and in flux (as a construct should seem to be) but is simple, constant and invariable, the source of the apparent synchronic and diachronic unity (identity). While not all sceptics of the self's reality single out synchronic and diachronic unity (identity) as a primary source of the self-illusion, it would seem that these features are, nevertheless, singled out most often. While details of their accounts vary, the following philosophers, like Hume, regard unity and identity (amongst other features) to be wholly constructed.

Moving on now to William James, recall that in *The Principles of Psychology*, James writes:

If the stream [of consciousness] as a whole is identified with the Self far more than any outward thing, a certain portion of the stream abstracted from the rest is so identified in an altogether peculiar degree, and is felt by all men as a sort of innermost centre within the circle, of sanctuary within the citadel, constituted by the subjective life as a whole. Compared with this element of the stream, the other parts, even of the subjective life, seem transient external possessions, of which each in turn can be disowned, whilst that which disowns them remains. Now, what is this self of all the other selves?

Probably all men would describe it in much the same way up to a certain point. They would call it the active element in all consciousness; saying that whatever qualities a man's feelings may possess, or whatever content his thought may include, there is a spiritual something in him which seems to go out to meet these qualities and contents, whilst they seem to come in to be received by it. It is what welcomes or rejects. It presides over the perception of sensations, and by giving or withholding its [p. 298] assent it influences the movements they tend to arouse. It is the home of interest, - not the pleasant or the painful, not even pleasure or pain, as such, but that within us to which pleasure and pain, the pleasant and the painful, speak. It is the source of effort and attention, and the place from which appear to emanate the fiats of the will...

Being more incessantly there than any other single element of the mental life, the other elements end by seeming to accrete round it and to belong to it. It becomes opposed to them as the permanent is opposed to the changing and inconstant (1890, 297-298).
This "spiritual self" says James, is empirically known to us through bodily feelings – especially around the head and throat (1890, 301) – although we may not be explicitly aware of this fact. But he concedes that such feelings are not enough to account for what it is that unites (I would say, 'seems to unite'), both synchronically and diachronically, the diversity of experience we call our 'own'. In the following passage, James identifies the ‘roles’ of ‘(perspectival and personal) owner’, ‘knower (witness)’, ‘thinker’, ‘initiator of actions’, and the source of synchronic and diachronic unity (hence identity) with a section of the ‘stream of consciousness’, namely, whatever thought one may have at the present moment:

...the real, present onlooking, remembering, 'judging thought' or identifying 'section' of the stream. This is what collects, - 'owns' some of the past facts which it surveys, and disowns the rest, - and so makes a unity that is actualized and anchored and does not merely float in the blue air of possibility (1890, 338) ...It is the Thought to whom the various 'constituents' are known. That Thought is a vehicle of choice as well as of cognition; and among the choices it makes are these appropriations, or repudiations, of its 'own'. (1890, 340)...The I which knows [knows the empirical aggregates] cannot itself be an [p. 401] aggregate, neither for psychological purposes need it be considered to be an unchanging metaphysical entity like the Soul, or a principle like the pure Ego, viewed as 'out of time.' It is a Thought, at each moment different from that of the last moment, but appropriate of the latter, together with all that the latter called its own. (1890, 400-1).

Like Hume, James attempts to fully account for “that appearance of never-lapsing ownership for which common-sense contends” (1890, 339) – without positing an enduring, invariable metaphysical ‘Owner’ that would do the work of the self (1890, 338-39). His solution (as with Hume) is to propose that such a self amounts only to content that is presented by mental objects (the “judging thought”), as opposed to being something substantial that underlies and

\[86\] One gets the distinct impression that the “judging thought” involves witnessing, although James never makes explicit the notion of witnessing, in and of itself. If James’s “judging thought” did involve witnessing, and witnessing was acknowledged, as Buddhism claims, to be the source of unconstructed unity and endurance, then James’s account (although more detailed) would be very similar to what is being advanced in this project. As it is, the account offered by this project, which will explicitly acknowledge the influence of Damasio, can be considered to in some respect build upon those ideas by James that have strongly influenced Damasio. These will include, especially, the role of emotion in constructing the sense of boundedness ascribed to self.

\[87\] James makes it clear that the unity and endurance we ascribe to the self is something unconstructed:

"...common-sense insists that the unity of all the selves is not a mere appearance of similarity or continuity, ascertained after the fact. She is sure that it involves a real belonging to a real Owner, to a pure spiritual entity of some kind. Relation to this entity is what makes the self's constituents stick together as they do for thought" (1890, 337).
precedes them. When he rejects the idea that the self offers a grounding for the impression of unity, endurance and invariability (something Buddhism would agree with) James also implicitly rejects the idea that those impressions could have any grounding whatsoever in unconstructed reality (setting his account apart from that of Buddhism). He attributes their impression entirely to the “judging thought” whose nature it is to change from one moment to the next. Through this move, James’s account renders these and other features of the self to be the content of an appearance contributed to by thought, rather than the content of an appearance that has no such contribution. Through being a product rather than a precedent of thought, the nature of self is thus rendered constructed. Going by this account, the self will be illusory insofar as it appears to underlie and precede thought, rather than to exist, as it actually does, as the content of thought.

James’s account has significantly influenced that of three contemporary thinkers: Owen Flanagan (1992), Antonio Damasio (1999) and Daniel Dennett. Flanagan believes that it is a mistake to suppose that there is:

...an “I” that stands behind all conscious experience [and]...constitutes the core of the self, our conscious control centre, the source of all action plans, and the agent for whom all experiences accrue before being filed for future reference or discarded...the mind’s “I” is an illusion (1992, 177)...
The illusion is that there are two things: on one side, a self, an ego, an “I,” that organizes experience, originates action, and accounts for our unchanging identity as persons and, on the other side, the stream of experience. If this view is misleading, what is the better view? The better view is that what there is, and all there is, is the stream of experience. “Preposterous! What then does the thinking?” comes the response. The answer is that “the thoughts themselves are the thinkers” (James, 1892, 83)...

...We are egoless. This, of course, sounds crazy, so I have a fair amount of explaining and comforting to do. (1992, 178).

Flanagan builds his explanation of our sense of self on “James’s idea that the self is...an after-the-fact construction, not a before-the-fact condition for the possibility of experience” (1992, 177, 178). On Flanagan’s account, the self is constructed and hence illusory in the same overall manner as that proposed by James and Hume (although details of their accounts

---

88 This is not to deny that James’s position differs from Hume’s on the details of how the stream of thought preserves the impression of unity and identity. According to James, Hume did not properly account for the sense of sameness that one has over time (1890, 352).

differ). There is however one point, in connection with Flanagan, which should be clarified. When he mentions the illusory status of “the mind’s ‘I’” – equivalent to the self – he sometimes gives the impression that this could be merely a philosopher’s illusion, “a tempting idea” which, with the help of precise thinking, can be eradicated. (1992, 177, 182-83). While there have indeed been many philosophers whose theories have given the self a serious unconstructed status that matches the content of our common conviction (often with extra bells and whistles, such as immortality) it should never be forgotten that this common deep-seated conviction is what spurred such philosophers on in the first place. Disabusing oneself of such a philosopher’s illusion – if the self is indeed an illusion – would therefore not entail disabusing oneself (qua subject) of the illusion conveyed in harbouring the sense of the self, namely, the conviction that one really is an ontologically unique, unconstructed thinker of thoughts etc. It would rather entail replacing any incorrect philosophical account, one that accords with our common conviction, with a correct philosophical account of how it is that we – including the well-informed philosopher – come to have this deep-seated but mistaken conviction.

Antonio Damasio (1999), who has also acknowledged the influence of James, sees the problem of consciousness as being divided into two. The first problem, he says, concerns how we get a “movie-in-the-brain” with its multi-modal sensory input culminating in familiar mental images that convey aspects of physical objects as well as the thoughts we have about them (1999, 9). The components of such images are known as ‘qualia’ – raw sensory qualities such as the whiteness or fragrance of a rose. Explaining how the physical brain could give rise to such qualities occupies a large field in the philosophy of consciousness. The second problem of consciousness, according to Damasio, concerns how the brain “engenders a sense of self in the act of knowing [about the world through such images]” (1999, 9).

The solution for this second problem... requires the understanding of how the images of an object and of the complex matrix of relations, reactions, and plans related to it are sensed as the unmistakable mental property of an automatic owner who, for all intents and purposes, is an observer, a perceiver, a knower, a thinker, and a potential actor (1999, 10-11).
Like Hume, James and Flanagan, Damasio does not think that this sense of being the observer and owner of the mental images is explained by there really being such a self that plays the role of owner and observer:

This second problem is all the more intriguing since we can be certain that the solution traditionally proposed for it – a homunculus creature who is in charge of the knowing – is patently incorrect. There is no homunculus, either metaphysical or in the brain, sitting in the Cartesian theatre as an audience of one and waiting for objects to step into the light... In effect, the second problem is that of generating the appearance of an owner and observer for the movie within the movie... (1999, 11).

In essence, Damasio holds that mental images from the somata-sensory modalities, hence, objects of consciousness, “constitute knowing and sense of self” (1999, 128) and that “the essence of core consciousness is the very thought of you – the very feeling of you – as an individual being involved in the process of knowing of your own existence and of the existence of others” (1999, 127). The very thought and feeling of oneself is conveyed in “storytelling”, “from non-verbal imagetic to verbal literary” by which the brain, for adaptational reasons, “map[s] what happens over time inside our organism, around our organism, to and with our organism, one thing followed by another thing, causing another thing, endlessly” (1999, 189). Our most basic sense of self, according to Damasio, is thus not produced by an actual, pre-existing the self, but by images generated by the brain, which tell wordless, primordial stories about what happens to the organism as it goes about engaging with the world:

Whether we like it or not, the human mind is constantly being split, like a house divided, between the part that stands for the known and the part that stands for the knower... The story contained in the images of core consciousness is not told by some clever homunculus. Nor is the story really told by you as a self because the core you is only born as the story is told, within the story itself. You exist as a mental being when primordial stories are being told, and only then; as long as primordial stories are being told, and only then. You are the music while the music lasts (1999, 191).

Damasio regards that virtually all features of the features of the self – knowing (witnessing), elusiveness, unity, identity (over time), agency, (perspectival and personal) ownership and boundedness – to be entire mental constructs, in and of themselves. It is hence in quite an
extensive manner that Damasio means “the core you is only born as the story is told”. In the context of modern neuro-psychology (Damasio is first and foremost a neurologist), he elaborates upon certain features of James’s account, in particular, how boundedness of self is constructed primarily by the emotions pertaining to individual concern (corresponding closely with the Buddhist notion of tanha). In a later chapter, I describe in more detail this aspect of Damasio’s theory, since I believe that it provides independent support for the Buddhist contention of how boundedness lends the self its overall constructed status.

Let us finally consider Dennett. In his chapter “The Reality of Selves” (in Consciousness Explained) Dennett wastes no time in dismissing the reality of self:

Are there entities, either in our brains, or over and above our brains, that control our bodies, think our thoughts, make our decisions? Of course not! Such an idea is either empirical idiocy (James’s “pontifical neuron”) or metaphysical claptrap (Ryle’s “ghost in the machine”) (1991, 413).

Damasio has noted that Dennett’s positive account of consciousness, against which the self is deemed an illusion, is one which depicts consciousness and the self as a “post-language phenomenon” (Damasio, 1999, 188). Dennett thus conceives of the self in its broader capacity of being an “owner of personality”, a personal owner of a rich and detailed history that belongs exclusively to it. Like James, he believes that the sense of such a self is grounded not in an actual self, but in a mental construct:

But the strangest and most wonderful constructions in the whole animal world are the amazing, intricate constructions made by the primate, Homo sapiens. Each normal individual of this species makes a self. Out of its brain it spins a web of words and deeds and, like other creatures, it doesn’t have to know what its doing; it just does it...Our tales are spun, but for the most part we don’t spin them; they spin us. Our human consciousness, and our narrative selfhood, is their product, not their source (1999, 416)...These strings or streams of narrative issue forth as if from a single source...their effect on any audience is to encourage them to (try to) posit a unified agent whose words they are, about whom they are: in short, to posit a centre of narrative gravity (1999, 418).

It should be clear enough from this that on Dennett’s account, as with the preceding ones, a self, qua underlying entity that unifies our experience and creates our thoughts, is both constructed and illusory.
Buddhist and the Western Accounts of ‘No-self': Summarising the Similarity and Differences

We have thus outlined a number of well-known Western thinkers with accounts that converge with that of Buddhism insofar as the self is regarded, as a whole, to be a construct and hence illusion. We are also able to see how the accounts diverge, namely, through the manner in which the self is considered to be constructed. The Western accounts diverge from the Buddhist account insofar as they regard many ascribed features – viewed as inherently unconstructed by Buddhism – to be thorough constructs. Unity, endurance and invariability are most commonly alluded to in this capacity. Buddhism regards these features to be constructed only to the extent that their impression is distorted through the mistaken assumption, on behalf of witnessing, that it is a bounded, personal owner. To the extent that they are not distorted by the impression of boundedness, however, these features (as we saw in chapter three) are what unconstructed nibbanic consciousness, via the witnessing, brings to the sense of self. Our analysis should therefore enable us to see more clearly the danger of hasty alignments between Buddhist and Humean denials of the self. Despite an overall similarity between the Humean and Buddhist accounts of self, there are also very important differences.

In light of this analysis, it will be useful to re-iterate our analogy that was used in the Introduction to help cognize the difference between the Buddhist and the typically Western account of no-self. Suppose that two people dream about a shrill voice. The shrill voice is in both cases a construct, the content of an appearance contributed to by thoughts, images and so forth. We can suppose that the first dream is woven around the sound of an alarm clock, which lends <the shrill voice>, as the content of the appearance, its piercing quality. To the extent that the shrillness is contributed to by the dream-independent alarm-sound, the shrillness is not, in itself, a mental construct. It is only a construct to the extent that it is ascribed, in the dream, to the shrill voice and is distorted through this assumption. We can suppose that in the second dream, the sound of the shrill voice is not woven around any alarm but is dreamt up entirely. The shrillness ascribed to the voice, as well as the shrill voice, is a mental construct. <The shrill voice>, contributed to by alarm-sound, is analogous to <the self> as construed on the Buddhist account, with many of its contributing features inherently
unconstructed$^{90}$. <The shrill voice> that is entirely dreamt up is more analogous to <the self> as construed on the Western accounts (especially Damasio’s), with many, if not all of the key features entirely constructed.

**Elimination, Reduction and Causal Roles**

We finish this chapter by clarifying a couple of matters that, in relation to the possible status of self-as-construct-and-illusion, could potentially cause confusion. The first matter concerns a potential confusion between elimination and reduction. If the self were to be an illusion, would it be eliminated or reduced, given the proposed account? When the reality of an entity is denied, it is generally said in philosophical parlance to be *eliminated* from the ontology defined by those standards of reality. When the ontology of an entity is fully explained in terms of a more basic ontology, it is said to be *reduced*, but not necessarily eliminated. In context of the above account, we can say that the self, as an illusion, has in one way been eliminated and in another way reduced. To see how this is so, we need to be aware of the distinction between the *vehicle* of thought and perception (etc.) and the *content* of thought and perception (etc.) – a distinction already referred in passing. The vehicle refers to the actual content-bearing thoughts and perceptions, as present-time occurrences, while the content refers to what the thoughts and perceptions are *of* or *about* (just as a book with its inky letters are vehicles for the content of what the book is about).$^{91}$ So we can say of a thought *qua* vehicle “his thought of a unicorn occurred at 10-20 am”. But we cannot always say the same thing about its content, for example, that “the unicorn which he thought about occurred at 10-20 am”. As William James has noted (1890, 297), sometimes the thoughts will be about ‘things’ that exist *only* in imagination, such as unicorns, demonstrating that the thoughts can partake in a different order of existence from that which they are about.

---

$^{90}$ Although unconstructed in a different manner to that of the shrill sound.

$^{91}$ On this distinction, William James writes: “…in everyone, at an early age, the distinction between thought as such, [p. 297] and what it is ‘of’ or ‘about,’ has become familiar to the mind. The deeper grounds for this discrimination may possibly be hard to find; but superficial grounds are plenty and near at hand. Almost anyone will tell us that thought is a different sort of existence from things, because many sorts of thought are of no things - e.g., pleasures, pains, and emotions; others are of non-existent things - errors and fictions; others again of existent things, but in a form that is symbolic and does not resemble them - abstract ideas and concepts; whilst in the thoughts that do resemble the things they are ‘of’ (percepts, sensations), we can feel, alongside of the thing known, the thought of it going on as an altogether separate act and operation in the mind.”(1890, 297-98).
Suppose Jim believes that he is being watched by aliens. <Being watched by aliens> reduces only to the content or aboutness of Jim's belief and to nothing more. In this respect, <being watched by aliens> is reduced and not eliminated from the relevant ontology, which concerns only Jim's thought-content. However, <being watched by aliens> does not reduce to anything beyond Jim's thought-content, not to any situation involving aliens, and not to Jim's thoughts qua vehicles, just as <Sherlock Holmes is a detective> does not reduce to any situation in the non-fictional world, including inky shapes on a book page. With reference to these standards of reality, <being watched by aliens> is eliminated from the ontology. What we can say, however, is that Jim's belief that he is watched by aliens reduces to the patterns of thought and perception which serve as vehicles for this content, and that these may in turn have a neuro-biological explanation.

Similarly, we can say on the proposed analysis that the self, if illusory, will be reduced to the content of an appearance contributed to by thought and perception. It is only with reference to this standard of reality that the self has not been eliminated. This is to say no more, and no less, than the self is a construct. Given this, it is incorrect to say that the self is reduced to anything that is unconstructed, including the vehicles of thought and perception that help convey this content. In this way, I think it is a category mistake for James to say, in literal terms, that "the thoughts themselves are the thinkers" if by "thinker" he means something with the attributes of an unconstructed self. An unconstructed thinker has been eliminated from that very reality (a reality that includes thought-vehicles) whose standards have been employed to deem the self a construct and hence illusion. What is reduced to the content-bearing vehicles of thought, perception, and perhaps witnessing, is not the self but rather our sense of the self. Dennett (1991) and Damasio (1999) are reasonably careful to distinguish in this way between vehicle and content. Damasio consistently refers to the sense of the self as being reduced to mental images (thoughts) produced by the brain. He does not intend his technical terms "core self" and "autobiographical self" to refer to the self, but to those mental and neurobiological patterns which ground the sense of the self. And Dennett, while not as consistent as Damasio, writes: "Thus we do build up a defining story about ourselves, organised around a sort of basic blip of self-representation... The blip isn't a self,

of course; it’s a representation of a self” (1991, 429). It would be more accurate, however, for Dennett to say that the blip is a presentation of a self, given that ‘representation’ has connotations of depicting an independently existing entity.

Another potential (and related) source of confusion concerns the possible causal role that the self plays in our manner of thought and behaviour. Sometimes a strong link is posited between the self and human survival, such that stripped of this fiction (viz., construct) “an individual human being is as incomplete as a bird without its feathers, a turtle without its shell” (Dennett, 1991, 416). Perhaps this means that the self has an indispensable causal role to play in our actual thoughts and behaviour. Or if, like the Buddhists, we think it not indispensable, then does the self not at least have a strong influence on how we think and behave? We need to be clear, however, on what is actually playing the causal role, should the self be illusory. Consider Jim with his belief that aliens are watching him. What is it that makes Jim tremble with fear and cower under his bed? Is it the aliens who are watching him — or simply his belief that he is being watched by aliens? Surely, the latter. Similarly, if we accept that the self is as illusory as the aliens, then we must also accept that the self cannot be influencing our thoughts and behaviour. What would play this causal role is rather our conviction or sense that we are such a self.

**Conclusion**

In this chapter we articulated a way in which parties from different metaphysical backgrounds can uniformly consider the self to lack reality. The self, if it lacked reality, would lack reality by being illusory, that is, by purporting to exist in a manner in which it does not actually exist. The self purports to exist in a manner that is inherently unconstructed by any thoughts, perceptions or observable objects of consciousness that can be perspectivally owned by it. That is, it purports to be that thing which precedes and thinks the thoughts, always standing apart from them in this fashion. If illusory, then the self would exist in a manner that is constructed. It would amount to nothing but the content of an appearance that is contributed to by those very objects that it purports to stand apart from. The thoughts will be helping to think the self rather than the self thinking the thoughts.

---

93 Damasio also regards the sense of self to be ‘indispensable’ for human survival (1999, 304-305).
Within this converging framework of self-as-construct-and-illusion, we looked at differences in the way that the self can be deemed a construct (and hence illusion). The Western accounts considered in this chapter – those of Hume, James, Dennett, Flanagan and Damasio – all showed evidence of regarding the self as more thoroughly constructed than in Buddhism. In particular, the Western thinkers regarded the individual features of unity, endurance and invariability to be thorough constructs. The Buddhist account of self, as we know, considers boundedness (as brought about through a sense of personal ownership, agency and self-identification) to be the primary source of the self’s illusory status. To the extent that their impression is not distorted by that of boundedness, the (mis-ascribed) features of witnessing, unity, elusiveness, endurance, invariability and unconstructedness are considered, in and of themselves, to be unconstructed.

It is a primary goal of this project to demonstrate that the Buddhist account of self is correct, with a metaphysically neutral ‘witnessing’ referred to in lieu of nibbanic consciousness. In the following chapter I begin to advance this position by arguing that the notion of witnessing must be analysed in such a way that implicates those features considered unconstructed in Buddhism. This will make it more likely that witnessing, as advanced on the Buddhist account, will ‘carry’ these features – as unconstructed – into an overall construct of the self.
Chapter Six: Linking Problems of Consciousness with Awareness

Introduction

Let us briefly recapitulate how Buddhism construes the ontological status of self along with each of its ‘contributing’ features. While the self as a whole is regarded in Buddhism to be a construct and hence illusion – in agreement with several leading Western thinkers – many features attributed to the sense of self, unlike in the West, are not in themselves considered constructed or illusory (to the extent that their impression remains undistorted by boundedness). The attributed non-illusory features include witnessing (or witness-consciousness), unity, endurance, invariability, unconstructedness and, I shall argue, elusiveness. The feature primarily responsible for the self’s constructed and illusory status is boundedness as exemplified through the role of personal owner.

On the Buddhist account, the unconstructed features of unity, endurance and so forth are ‘imported’ into the sense of self through witness-consciousness. The witness-consciousness, characterised by these features, is what we have termed in Buddhism ‘pre-nibbanic consciousness’ – nibbanic consciousness as partially ‘covered’ by the illusion of self; or nibbanic consciousness as ‘duped’ into reflexively believing it is a bounded personal owner-self. While intrinsically identical to nibbanic consciousness, pre-nibbanic consciousness appears, through factors extrinsic to its nature, as if it were just ordinary witness-consciousness – a component of the self. If this Buddhist account were correct, then upon analysing the relation between features ascribed to the self (suspending judgments about their deeper ontology) we should expect to find, at the very least, a link existing between ordinary witness-consciousness and its alleged ‘characterising features’ of unity and

---

94 I am grateful to Tim Bayne for his feedback on the material in this chapter.
95 As it was mentioned in chapters two and three, Buddhism would also regard the sought-after happiness, that is felt on occasion, to be grounded in the pure unconstructed nature of the mind, viz., in nibbanic consciousness. Due to constraints of space we will not be addressing this claim in the project. The urge to seek happiness over suffering, however, would be regarded in Buddhism as integral to the illusion of self so long as it caught up with the belief that dukkha-free happiness can be obtained through an ‘ideal configuration’ of objects. This belief is what drives and indeed helps constitute tanha – and hence, on the Buddhist position, helps contribute to the sense of a bounded self, rendering the self a construct. The role of the happiness-seeking urge in its relation to tanha and the sense of self will be returned to in chapter eight.
so forth. This link should not extend to boundedness, personal ownership or agency (recalling from chapter four that agency depends upon personal ownership of action).

While it is not the ambition of this project to argue directly for the metaphysically robust reality of (unconditioned) nibbanic consciousness, I hope to establish that a scenario that would be implied by this reality will hold. In this chapter I will argue, hence, that Buddhism is right in supposing that there is indeed a special link existing between ordinary witness-consciousness and the aforementioned features, when witness-consciousness is construed as having a non-intentional phenomenology (something the self purports to have). I will argue that it is part and parcel of the very concept of such ‘felt’ witness-consciousness — or awareness, as I shall call it — that it be construed with reference to the concepts of unity, elusiveness, endurance, invariability and unconstructedness. The concepts of these features will in turn be analysed with reference to the concept of awareness. I suggest that because it singularly captures so many important features — features that have generated longstanding puzzles of subjectivity — the concept of awareness, characterised in this fashion, ought to be a central concept of consciousness in the philosophy of mind. In the West, the bulk of philosophical attention tends to be focused upon concepts of consciousness, which, I shall argue, fail to properly capture these puzzling features.

I will then argue that the link between the concept of awareness and those of its features does not extend to the concept of boundedness (central to the notion of self), which depicts that feature of ontological uniqueness as borne out through the role of personal owner. While boundedness must be analysed with reference to awareness (simply because the self, as a bounded personal owner, will be also a witnessing subject that is unified and so forth), the reverse implication will not hold. Awareness will not implicate boundedness, demonstrating that a conceptual link will hold only between those concepts of the self’s attributes that Buddhism ascribes to (pre-nibbanic) witnessing. Such a scenario will also increase the likelihood that boundedness will underlie the self’s constructed status — if the self is indeed constructed — a Buddhist contention to be properly argued for in chapters eight and nine.

We must note that this chapter’s task is not to demonstrate that awareness (viz., felt witness-consciousness) or its specifying features of unity and so forth have the unconstructed reality that they purport to have, and would have to have if the Buddhist account was correct.
The task of demonstrating the independent unconstructed reality of awareness and so forth is postponed until chapter seven. Since this current chapter's discussion will remain neutral on the ontology of awareness and associated features, the reader should imagine the clause ‘if it exists’ occurring after each instance of the term ‘awareness’, ‘unity’, ‘elusiveness’ and so forth.

1. Awareness as a Concept of Consciousness

Before we can define awareness through the features of unity, elusiveness and so forth – and these features in terms of awareness – the core concept of awareness must itself be clarified. I will introduce the core concept of awareness by first making some general remarks about it and then by providing examples of how a related, but more metaphysically loaded version of that concept is used in Advaita Vedanta. I will then define the core concept of awareness more precisely.

The term ‘awareness’ will be used in this project to denote a kind of witness-consciousness that has an intrinsic phenomenal character. This immediately contrasts with how the term ‘awareness’ is often used in Western philosophy. The term ‘awareness’ is often employed to refer to specific functions that consciousness may have (such as its disposition to cause patterns of behaviour) – with any intrinsically phenomenal dimension being either ignored or denied. In the context of this thesis, however, ‘awareness’ will be employed as a term-of-art to refer to witnessing with intrinsic phenomenal character. The concept is similar to, although not as metaphysically loaded as, a central concept in the Eastern tradition of Advaita Vedanta. The concept goes by several names in this tradition: as well as the word translated as ‘awareness’, there are also the translations ‘witnessing’, ‘conscious presence’, ‘Self’ and ‘I’. Advaita Vedanta emphasises this concept in a way that Buddhism does not, which is why the following quotations are drawn from this tradition (although we have already demonstrated awareness to be every bit as central to the philosophy of Buddhism). The following quotation, then, is from the words of a reportedly ‘enlightened sage’ (‘Arahant’ in Buddhist terms), Nisargadatta Maharaj. In the tradition of Advaita Vedanta,

---

96 For example, see David Chalmers (1996, 28-31) and Barry Dainton (2002). This functional notion of awareness is basically witnessing sans non-intentional phenomenology and is addressed later in this chapter under ‘higher-order theories’ of consciousness.
this passage illustrates the more metaphysically loaded version of ‘awareness’ in use (viz., ‘witness’, ‘conscious presence’):

...There is nothing that what you call the “witness” has to do; witnessing happens purely by itself. ...whatever experiences you have had, whatever you have seen of the world, has all been witnessed. But that one to whom the witnessing takes place is entirely separate from that which is witnessed...(1994, 67).

...And ultimately what happens? The conscious presence alone remains. That is, there is just conscious presence, not “I” or “you,” or anything. I repeat: it is total presence; that is, total manifestation – not I, you or any individual...This consciousness, which is within the body and therefore has mistakenly assumed that it is the body, gradually realises its true nature, namely that it is only conscious presence without any inherent individual aspects. Finally, it considers itself the conscious presence of the total manifestation, and all individuality is lost. ...When you ask questions, ask them on the basis that you are not the body-mind, but that you are the conscious presence (1994, 87)

Ramana Maharshi, from the tradition of Advaita Vedanta, is also believed by many to have been an enlightened Arahant. What he terms the ‘Self’ or ‘I’ pertains not to what we have been calling ‘the self’ but to awareness in roughly the sense that I am using the term:

The nature of the Self or ‘I’ must be illumination. You perceive all modifications [of objects in consciousness] and their absence [of objects in consciousness]. How? To say that you get the illumination from another would raise the question how he got it and there would be no end to the chain of reasoning. So you yourself are the illumination. The usual illustration of this is the following. You make all kinds of sweets of various ingredients and in various shapes and they all taste sweet because there is sugar in all of them and sweetness is the nature of sugar. And in the same way all experiences and the absence of them contain the illumination which is the nature of the Self. Without the Self they cannot be experienced, just as without sugar no one of the articles you make can taste sweet....First one sees the Self as objects, then one sees the Self as void, then one sees the Self as the Self; only in this last case there is no seeing because seeing is being (1971, 133).

While a handful of current Western thinkers (often familiar with Eastern traditions) give attention to a usually less metaphysically loaded version of this concept: for example Deikman (1996), Forman (1998), Shear (1996), (1999), Woodruff Smith (1986) and
Wilber’s ‘awareness’, as portrayed here, is not a mainstream notion in contemporary Western philosophy of mind (although historically, such thinkers as Moore and Hamilton may have endorsed versions of it). Some thinkers, such as Caston (2002) and Kriegel (2004), have followed in the tradition of Brentano and/or Aristotle by alluding to a notion of consciousness that includes but is not limited to awareness. This Brentanian/Aristotelian notion builds object-directedness into the concept of consciousness, such that a person’s witness-consciousness is inevitably focused upon other objects, whilst having an elusive, sidelined sense of its own presence. What I am calling ‘awareness’ here may hence be similar to what these thinkers regard as the ‘sideshow’ (although see possible differences in footnote 99 below). For worse and for better, the object-directed or intentional dimension is excluded from my definition. ‘For worse’ because the relation between awareness and its objects is left virtually unexplored in this chapter (a task I unabashedly leave to the Brentanians and Aristotelians). ‘For better’ because the Brentanians/Aristotelians do not, I believe, do a clear job in articulating the character of this ‘sideshow’: their depictions tend to be vague. I hope to provide clarity in precisely this area, and in a way that links the

---

97 I am aware (from conversation with Mark Edwards) that Ken Wilber talks about the notion of awareness or the ‘Witness’, as he calls it, in many of his written works. I have not had time to give Wilber’s pioneering work the attention it seems to deserve and hope to rectify this in the future.

98 Sir William Hamilton, quoted in C.O. Evans (1970, 39), writes: “Consciousness may be compared to an internal light, by means of which, and which alone, what passes in the mind is rendered visible. Consciousness is simple, - is not composed of parts, either similar or dissimilar. It always resembles itself, differing only in the degrees of its intensity; thus, there are not various kinds of consciousness, although there are various kinds of mental modes, or states, of which we are conscious.” G.E. Moore, also quoted in Evans (1970, 39), writes: “We have then in every sensation two distinct elements, one which I call consciousness, and another which I call the object of consciousness. This must be so if the sensation of blue and the sensation of green, though different in one respect, are alike in another: blue is one object of sensation and green is another, and consciousness, which both sensations have in common, is different from either...The moment we try to fix our attention upon consciousness and to see what, distinctly, it is, it seems to vanish: it seems as if we had before us a mere emptiness. When we try to introspect the sensation of blue, all we can see is the blue: the other element is as if it were diaphanous.”

99 I do not buy into the analysis (a la Brentano) that witnessing gains its intrinsic phenomenal character by taking itself as its own secondary intentional object (the primary object being the object of perception) – if by ‘object’ we mean something whose qualities are able, in principle, to be attended to via the five sensory modalities and/or conceptual thought (including emotion, imagination and memory). On the level of how things seem (the level that concerns us here) witnessing can never observe itself as an object in this sense – as will emerge in the section on ‘elusiveness’.

100 For example Victor Caston, when discussing Aristotle’s theory of consciousness, maintains that Aristotle “parts company” with theorists who suppose the phenomenal character of an experience to be “exhausted” by its intentional properties (2002, 790). Of this elusive and yet phenomenal “on the side” dimension, Caston writes: “Significantly, Aristotle can only tag this phenomenal property by reference to the perceptible [viz., intentional] quality it is about... in the end, he may not be able to say much more than this. He is referring here to what perceiving such qualities is like, and he finds that we can only express the character of such experience ‘on this side’, as it were, of reports of our primary intentional experiences of objects. That is simply part of the
concept of awareness to several major puzzles in the philosophy of mind. Moreover, if Buddhism and Advaita Vedanta are correct, then awareness need not depend upon the presence of objects that one is aware of: an object-oriented status will not be essential to awareness. While the definition I offer remains neutral on this point – awareness may or may not require objects to be aware of – it notably does not rule out the possibility of objectless awareness through prematurely stipulating awareness to be co-present with objects.

The core concept of awareness, to be developed further in this chapter, is as follows:

awareness in this sense

1. pertains to ‘subject’ side of apparent subject/object dichotomy
2. involves witnessing (the modus operandi of a subject).
3. has own intrinsic phenomenal character (present in all conscious states) that is peculiar to its subjective status and not reducible to intentional content (pertaining to objects).

We have already spoken at some length about 1 and 2 (in earlier chapters) and so will not be adding much more except to clarify a further distinction in connection with 2. The main focus of discussion will be on 3. We should be reminded that when talking about awareness in this chapter, we are remaining neutral on its ontological status, in particular, on whether awareness is real or illusory.

The Distinction between Awareness and Attention

I would like to further clarify the distinction between attention and awareness since this distinction will be alluded to later in the chapter. Consider what happens when one attends to an object. The object is the focus of attention, so what we call ‘witnessing’ is directed mostly towards the object. Nevertheless we can argue that there is inattentive witnessing of such objects as background noise and the feeling of one’s feet on the floor. Evidence of inattentive witnessing is one’s ability to recollect being dimly aware of these items all along. The fact that it makes sense to speak of inattentive awareness or to say “there was witnessing of the elusiveness of awareness” (2002, 791). In a later section of this chapter (on ‘elusiveness’) I footnote an example where Kriegel’s (2004) analysis of the ‘sideshow’ lacks clarity of exposition (by seeming to confuse subject with object).
noise though I did not attend to it” suggests that witnessing is more basic than attention.¹⁰¹ A useful analogy for the relation between awareness *qua* witnessing and attention is that of a flashlight, whose beam can be concentrated and focused on various objects, while nevertheless casting a dimmer pool of light on objects not focused on. Witnessing can be likened to the light in general, while attention can be compared to the focusing or directing of the beam on this object and that. Even with the beam focused, there is a surrounding, dimmer circle of light, illuminating objects not focused on, analogous to the inattentive witnessing. The analogy of witnessing or awareness as light has been employed many times, in both East and West, to illustrate different aspects of awareness (for example see Ramana Maharshi, above). It is in many ways a useful analogy.

The Intrinsic Phenomenal Character of Awareness

Awareness is intrinsically experiential, imparting, to use the metaphor used by Ramana Maharshi (1971, 133), a common “flavour” to all conscious states. Barry Dainton (2002, 32) refers to it as “tangible” (as opposed to “pure” which lacks any phenomenal character)¹⁰². The important matter is that its phenomenal character is non-intentional. This means that the phenomenology is not reducible to the spectrum of experiential qualities (whether attentively or inattentively noticed) that characterise the differentiated objects of experience and that can be attended to in *principle*. Hence it excludes the qualia specifically associated with deliverances of the five senses, emotions, conceptual thought, memory or imagination. There is, nevertheless, *something it is like* to be aware. The phenomenal character is that of pertaining to a subject *qua* witnessing – a quality which seems, as we have already discussed, to be part and parcel of the self in its capacity as a witnessing subject. An integral feature of the witnessing subject, mentioned in earlier chapters, is that in viewing objects (internal or external) as distinct from itself, the subject seems to be an altogether different type of thing – a thing whose *raison d’être* is to observe and never to be observed *qua* subject as an item of attention. The self, purporting to be a bounded witnessing subject (etc.), seems

¹⁰¹ Both C.O. Evans (1970) and Kriegel (2004) make this distinction and take it further, analysing the phenomenal character of ‘on the side’ witnessing (what I shortly term ‘subjective sense of presence’) in terms of the content of inattentive purview. In the section on ‘elusiveness’ I argue that this kind of approach does not work.

¹⁰² After making this distinction between “tangible” and “pure” awareness, Dainton (2002) goes on to discuss, and dismiss, the “pure” variety (equivalent to a higher-order theory of consciousness). All the problems he points out seem due, precisely, to its non-tangible character. He does not discuss the “tangible” variety.
take on this unobservability and hence appears intrinsically un-objectlike (since to be an object is to be observable as an object of attention). Yet there seems, despite this, to be something it is like to be a observing subject-self – evidenced by the very fact that we pre-theoretically forge a subject/object distinction – and suggesting an immediate sense of one’s existence qua observer that does not overtly reduce to observable objects and their qualities.

Through the arguments of this chapter, it should emerge that the essence of this non-intentional phenomenology (or what seems to be such an essence) can be characterised through witnessing and its co-defining features without any need to refer to what boundedness, through the role ‘personal owner’, would bring to the sense of self. In later chapters it will be suggested that those factors that primarily create the impression of boundedness seem to impart a ‘flavour’ of ‘me’ to the generic phenomenal character of witnessing, such that it seems as if there could only be one unique self. This unique ‘flavour’ of oneself, however, will be something that is constructed, like a spectral hue that is added to pure ‘luminous’ witnessing. It will not be intrinsic to the phenomenal character of witnessing. Our current strategy is to show that such witness-consciousness with its generic, impersonal non-intentional phenomenology (that we call ‘awareness’) is something whose concept demands that it be unified, elusive, enduring, invariable and unconstructed – while these features must in turn, through their very concepts (as featured in traditional puzzles of subjectivity), implicate the generic non-personal awareness. This conceptually constrained ‘co-specification’ of awareness and its features will, I shall argue, involve no reference to boundedness, personal ownership or the self.

When initially characterising this non-intentional phenomenal dimension to awareness (prior to defining it in terms of unity etc.) there are immediate difficulties of description, since the language of phenomenology is devoted to describing objects of experience. Nevertheless, it is worth trying. Maharaj (1994, 87) refers to it as “conscious presence”, Arthur Deikman (1996, 350) describes awareness as “the subjective sense of our existence” and C.O Evans (1970, 150) talks about a “lively sense of presence” that is integral to our sense of self. I think these phrases can be combined to produce one that may accord quite closely with intuition (although intuitions are not being assumed here to necessarily reflect an objective reality to this phenomenal dimension). I propose that we call it ‘the subjective sense of presence’. I prefer ‘presence’ to ‘existence’ because ‘presence’ conveys
the double meaning of being present with respect to space (as in the perpetual here) and present with respect to time (as in the perpetual now) – and we arguably experience a subjective sense of presence through both the here and the now. With regard to time, especially, our sense of the present moment seems to emanate either from or through the subject, rather than being derivable from objects. That is to say, the present moment, of which we have a definite sense, does not seem to be found amongst the objects of consciousness; rather, objects (such as perceptions) seem to occur to a subject in the present moment. The sense of present-momentness is perhaps the easiest way to discern the phenomenology of awareness, and the most obvious clue that there could be more to conscious experience than object-qualia. Awareness thus conveys a type of experience – a subjective sense of presence – that is described independently of, and is unmediated by, any specific quality pertaining to objects, outer or inner.

Having articulated the core concept of awareness, a descendent of the Advaita Vedantin notion, it should be clear that by ‘awareness’ we are not meaning something that could be non-phenomenal. Sometimes we say “I am aware that P”, meaning, “I know that P”. Or we say “I am aware that O” meaning “O is available to me for verbal report and control of my behaviour”. While awareness as specified in this chapter will no doubt be related to these functional uses of the term, ‘awareness’ is to be interpreted here as non-intentional phenomenal witnessing, felt as a subjective sense of presence. I shall sometimes refer to awareness in this capacity as ‘witnessing presence’. It is also worth being reminded that in keeping with concerns set out in chapters one, two and three awareness will not be assumed to necessarily occupy such a category as ‘activity’ or ‘state’: this possibility will remain an open one.

2: Linking Problems of Consciousness with ‘Awareness’.

In this section it will be argued that the features of elusiveness, synchronic unity, endurance, invariability and unconstructedness, in virtue of the problems they generate for attempts to develop an account of consciousness, are to be specified with reference to awareness, viz., witnessing presence. Awareness must, in turn, be specified with reference to these aspects. Through their associated puzzles, these apparent aspects or features serve to capture much of what seems puzzling about consciousness as a quintessentially subjective
phenomenon. If these features are to be specified with reference to awareness, then the concept of awareness, if also characterised with reference to these features, will encapsulate much of what seems quintessentially subjective about consciousness.

The strategy of this section will be to first outline each of the features as they occur in their context as traditional puzzles for consciousness and subjectivity in the philosophy of mind. In doing this, we will be staying neutral on whether subjectivity, consciousness or the notion of a subject is to be defined with reference to awareness or witnessing, unless the puzzle specifically alludes to consciousness in that manner. The next stage will be to demonstrate that awareness – rather than some rival notion of consciousness – does indeed best fit the notion of consciousness being alluded to in the puzzles. We should be absolutely clear that at this stage the aim is merely to show that the concept of awareness is most fittingly implicated in articulating the relevant puzzles, and that these features, in turn, qualify the manner in which awareness purports to exist. The aim is not at this stage to show that elusiveness, unity, awareness and so forth have any independent reality.

Elusiveness

In chapter four I quoted a section from the following passage by Roderick Chisholm (1969) where he draws attention to the centrality of elusiveness as a problem of consciousness and frames the problem as one about subjectivity in particular. Here is the passage in full:

To recapitulate: elusiveness pertains to the fact that one is never aware of a subject of experience, or more accurately, to the fact that the apparent subject cannot observe itself as an observing subject that is simultaneously its own object of experience. The apparent

---

103 Chisholm (1969, 95) thinks that “both groups have lost their way” in supposing there to be a genuine problem here.
subject is systematically elusive to its own observational purview. On the face of it, we might wonder why there is a puzzle at all. Elusiveness is a ubiquitous phenomenon, not confined to subjects of experience, and most cases of it are innocuous. For example, no metaphysical inquiry has been launched into why an eye cannot directly see itself, or why one cannot jump on the head of their own shadow. Such facts are a matter of logic or epistemology, and philosophers such as Ryle (1966) have argued that people are confused if they think that there is more to 'the elusive subject' than this.

Even if Ryle is ultimately correct, the fact of the matter is that philosophers from “the two great traditions” do think that there is more to ‘the elusive subject’ than a mere logical puzzle. Our question is: why do they think that? Amongst the panoply of elusive phenomena, why is the subject singled out as the puzzling case? The answer is that while on one hand, we as apparent subjects systematically elude ourselves from our introspective or perceptual observation, on the other hand we have some intuition - a conscious sense or feeling - that we do, simultaneously and immediately, experience ourselves as subjects. What is puzzling, then, is that we have an immediate sense of our own subjective existence but that this subjective existence can never be directly observed. The elusiveness of the subject thus seems to involve a phenomenological contradiction. Here is a formulation of the elusiveness phenomenon, in light of the puzzling feature that we have just outlined:

**Elusiveness of the subject of consciousness:** On the one hand, the apparent subject is elusive to itself. That is, it cannot directly observe itself, *qua* subject. On the other hand, we have an intuition that the subject *qua* subject is immediately experienced. Hence the puzzle. **The subject** seems to be that which is both elusive to observation and yet experienced.

Now we have, during the course of this project, defined the subject as a specific perspectival locus for witnessing, in relation to which objects are witnessed. For the current purposes, however, it would be presupposing too much to define the subject in terms of witnessing. We want, after all, to *demonstrate* that a kind of witnessing, in particular, a non-intentional phenomenal witnessing (viz., awareness), is indeed the *modus operandi* of subjecthood.

---

104 I am grateful to Hartley Slater for making this Rylean point to me so persuasively that it forced me to sharpen my formulation of the puzzle.
referred to in the puzzle of elusiveness. Now, some may at this point persist in their Rylean objection that we do not need to refer to a particular modus operandi of subjecthood in order to determine that the subject of the above puzzle is elusive. Just as a video camera cannot, as a matter of logic, directly observe (that is, video) itself, the subject, as the locus of an observing perspective, cannot directly observe itself. The deeper modus operandi, such as whether the observer (or observing) is in fact conscious, is irrelevant to the observer’s ‘elusive’ status. While this is in one sense undeniable, it does not, once again, capture why elusiveness of the observing subject should seem any more puzzling than that of the video camera – and yet it does seem to be more puzzling. It seems more puzzling with the subject precisely because, unlike with the video camera, there is the intuition of an experiential dimension to the elusive observing subject that is suggestive of a phenomenological contradiction. It would seem, then, that the apparent phenomenological contradiction forces us to formulate the puzzle in a more searching way, namely: what is it about the subject – in virtue of its modus operandi – that seems both elusive to observation and yet experienced? It is hence the modus operandi of the subject, rather than the subject per se, that will concern our discussion.

Our task, then, is that of answering the question: What concept of consciousness best captures that consciousness that would serve as the subject’s modus operandi in the above puzzle-formulation, by seeming to be both elusive to observation and yet experienced? The strategy in both this section and the next (on synchronic unity) will be to first see if popular rival concepts of consciousness fit the bill of the subject’s modus operandi (or whatever notion of consciousness is underlined) in the requisite puzzle. One concept will be popular although unlikely, while the other concept(s) may initially seem to fit the bill and be a close rival to the concept of awareness. We will then argue that the concept of awareness is the only concept of consciousness that truly fits, and hence, best captures the feature that distinguishes the relevant puzzle.

The first concept of consciousness we will consider, for the elusiveness puzzle, is a popular one in the philosophy of mind, which will be termed ‘phenomenal object consciousness’.105 This notion equates consciousness with its phenomenal objects;

---

105 Any theory of consciousness, which takes seriously (a) the apparent reality of phenomenal object properties (or qualia) and (b) analyses consciousness in terms of these properties, will satisfy this definition. It is usually
specifically, qualia that pertain to the five senses and cognition (including specific feelings pertaining to thoughts and emotions). On this notion, a moment of consciousness will be exhausted by whatever collection of phenomenal objects one is conscious of at a given time: for example, redness, dizziness, sadness, the feel of a chair.

Phenomenal object consciousness can be easily dismissed as a candidate for the subject's modus operandi in the elusiveness puzzle. While such qualia are puzzling in their own right, the puzzle is not that they seem to systematically elude observation. Indeed, it is because such qualia are so typically unelusive and observationally 'in our face' that theorists are able to ponder over such questions as why red and green seem complementary colours, or how the physical brain can give rise to such a variety of novel qualities. Elusiveness pertains rather to the subject of these phenomenal objects, a subject whose modus operandi seems to escape the purview of normal, in-the-face observation. 'Phenomenal object consciousness' is hence a non-starter, pertaining to the wrong side of the subject-object dichotomy – to the objects apprehended rather than to the apprehending subject. It cannot therefore substitute the modus operandi of 'subject' in our articulation of the elusiveness puzzle.

Higher-order theories of consciousness are also quite common in contemporary Western philosophy. On these theories (or at least on typical renditions of them) a mental state is conscious by virtue of the fact that it is the content of a higher-order mental state, a state to which the content is presented. This higher-order state, be it a thought or perception, confers consciousness upon its directed content without actually being intrinsically conscious itself (Caston, 2002, 753-755). The higher-order state would satisfy the elusiveness requirement, since its inevitable directedness at other mental states would mean that it could not observe itself (if conscious, it is always in virtue of being the target of another higher-

---

this kind of consciousness that theoreticians have in mind when they address what Chalmers (1996a) calls "the hard problem of consciousness" (namely, how complex physical systems such as the brain can be associated with qualia). Chalmers himself introduces the concept of consciousness by listing a wide catalogue of sensory and perceptual experience, concluding the section by saying "this brief look at the rich varieties of conscious experience should help focus attention on just what it is that is under discussion [in the hard problem]" (1996a, 11). It should be noted, however, that he also includes in his list the "deep and intangible" "phenomenology of self", which he likens to a "background hum" (1996a, 10) so it is not always clear whether he identifies the character of conscious experience exclusively with object-qualia. Were Chalmers to analyse consciousness in terms of any qualia, whether pertaining to object or subject, then the definition would become too broad to specifically substitute 'consciousness' in the elusiveness, unity and other puzzles to be discussed in this chapter. According to Caston (2002, 754) – whose synopsis of higher-order theories I have briefly summarised above – versions of the higher-order theory of consciousness have been defended by Armstrong, Lycan and Rosenthal. Dainton (2002, 32-33) also discusses this theory (viz., awareness, of the "pure" variety) adding Churchland, Carruthers and Locke to the list of its advocates.
However, the higher-order state clearly does not satisfy the requirements for the elusiveness *puzzle*, since there seems, on versions of the theory I have examined, to be nothing it is like to be in such a state *per se*. Since it *ex hypothesi* seems to lack any intrinsic phenomenology, a higher-order mental state could thus add none of its *own* intrinsic phenomenology to any given conscious experience. The phenomenal dimension to a given conscious state will be exhausted by those observable objects to which the higher-order state is directed. Nor would the phenomenal objects, for reasons recently articulated, satisfy the requirements of the elusiveness puzzle: while experienced, they would not be elusive to observation. Hence ‘higher-order consciousness’ – whether construed with emphasis on the ‘content-made-conscious’ or on the higher-order ‘consciousness-conferring mental state’ – cannot be that aspect of the subject that seems, simultaneously, to be both unobservable and yet experienced.

Let us now consider a more serious contender; that of “unprojected (or peripheral) consciousness”, proposed as a candidate for ‘the subject’ (its designated *modus operandi*) by C.O. Evans (1970). Evans writes: “I give the name ‘unprojected consciousness’ to those elements of consciousness that together make up the background of consciousness when attention is paid to an object” (1970, 104). Basically, Evans (1970, 104-107) holds that our field of consciousness is polarised into two areas; that of the objects being attended to – “the foreground” – and that to which we are not attending but are nevertheless consciously aware – “the background”, or, as he calls it, “unprojected consciousness”. Evans regards this two-fold structure of attention to be a defining feature of consciousness: there is, he thinks, no such thing as a conscious state that lacks this structure. Even the haziest of daydreams involves attending to some retinue of mental objects, however unconnected they may seem. Objects of attention occur, and must occur, he says, against a background of unprojected consciousness that is composed of such elements as thoughts and feelings. Any mental contents which happen to constitute a moment of unprojected consciousness need not be there the next moment: they are constantly shifting, some becoming objects of attention, others ceasing to be elements of consciousness. What is important is that although the contents are mobile, the two-fold structure of consciousness into foreground and background remains fixed. Our awareness of elements within unprojected consciousness is of course dim. At the time of experience, we do not individually differentiate the elements of unprojected
consciousness otherwise we would, by definition, be attending to them. Rather, he claims, its experience is of an undifferentiated whole which lends phenomenal character to our consciousness, but can be discerned as separate elements only in retrospect, for example:

...when eventually the pain in the blistered heel is noticed its presence need not come as a complete surprise. The person may recollect that he had after all been dimly aware of it all along. It is this form of awareness to which I am referring when I claim that the elements of unprojected consciousness are elements of awareness. If they were not, there would be no justification in including them in consciousness-unprojected or otherwise. (1970, 107).

Now it makes no sense, says Evans, to attend to elements in unprojected consciousness. For the moment we shift our attention to elements in the background of awareness, they become, by definition, objects of attention in the foreground, while a new lot of elements form the background. For this reason, he says that unprojected consciousness, while experiential, is systematically elusive - to attention (1970, 148-149). Since he regards a defining feature of 'the subject' (which he refers to as "the self") to be its elusive yet experiential nature - and so in accordance with the above specification of elusiveness puzzle - he sees it as a natural step to identify that aspect of subject-self with unprojected consciousness:

There is a parallel with the 'I' which, as Ryle has argued, is systematically elusive too in the sense that the 'I' cannot be objectified by attention (1970, 149)... I am suggesting that it is indeed true that the subject-self can never itself become an object of experience, and I am maintaining that this has nothing to do with the nature of the self - transcendental or empirical. It is put down to nothing other than the way attention operates. This enables it to be asserted in all consistency both that the self as subject is experiential, and that it is never presented as an object of experience. Furthermore, it obviates the necessity of treating the self as something unknown in itself. The theory overcomes the paradox that the self, although discoverable in experience, is never an object of experience, and in the process removes the main prop holding up The Pure Ego Theory of the Self. The essence of the matter is, on my view, that the self is experiential (i.e. is composed of elements of consciousness), but is never known as an object of experience. This is one of the factors that accounts for the view that the self lies behind its experiences. It also explains why we have such a lively sense of the presence of the self, and why we are so nonplussed by denials of the self's existence... (1970, 150).

At first glance, the elements that at any one time form an unprojected consciousness nicely fit the specification for modus operandi of 'subject' in the elusiveness puzzle. The elements are elusive, in that attention can never be focused upon them without them ceasing to be
elements of unprojected consciousness. At the same time they lend to our conscious life a rich background phenomenology, fitting the bill of being experiential as well as elusive. On closer inspection, however, unprojected consciousness is an unsuitable substitute for the modus operandi of subject in the elusiveness puzzle – at least on the level of how things appear, the level that concerns us here. The modus operandi of an apparent subject is that of apprehending, observing, noticing, attending, perceiving or introspecting (no, we have not yet mentioned awareness!). That is, the apparent subject must carry a disposition for apprehending objects, whatever the mode of apprehension. This is not stipulating that there must always be objects of consciousness; but that if there are such objects, then the apparent subject must always appear to apprehend them in some manner or other. The problem with identifying the subject’s modus operandi with unprojected consciousness is that the latter does not fit this bill, since its elements do not seem, under any stretch of the imagination, to apprehend objects. It does not remotely seem, for instance, as if a peripherally registered pain in the heel, along with the rest of unattended-to elements, is involved in attending to the objects of attention, such as a speech. And yet that is how it should seem, if unprojected consciousness is to substitute the modus operandi of ‘subject’ in the elusiveness puzzle. Rather, it seems as if a subject – one and the same subject – is, via its modus operandi, inattentively noticing the pain in the heel etc., while attending to a speech. This is the subject whose modus operandi seems problematically elusive.

On the level of appearances then, the elements of unprojected consciousness are far more suitably identified with observable objects of consciousness than with the elusive modus operandi of an observing subject. Indeed, any item which may at time t1 form an element in unprojected consciousness, such as a pain in the heel, can be attended to the very next second – which clearly renders the item to be an object by this project’s definition. And even while concurrently elusive to attention, the elements of unprojected consciousness are not elusive to broader modes of apprehension such as inattentive noticing – modes of apprehension that seem, in and of themselves, to be elusive to their own observation. So while Evan’s concept of unprojected consciousness does capture something elusive and
experiential, it is ill suited to convey the *modus operandi* of the subject at the heart of the elusiveness puzzle. We must look elsewhere for a suitable concept of consciousness.107

Let us now consider whether our concept of awareness — viz., witnessing presence — can convey the kind of consciousness that is implicated in the subject’s puzzlingly elusive yet experiential *modus operandi*. We can immediately note that awareness, through its component of witnessing, involves the broadest mode of conscious apprehension, whether attentive or inattentive. So unlike unprojected consciousness, awareness is in the right ballpark to fit what is required by the *modus operandi* of the subject. Awareness also satisfies the central requirements of the elusiveness puzzle: its very nature must imply a phenomenological elusiveness. Phenomenologically, awareness is (by its definition) felt as a subjective sense of presence, and so is experiential in its intrinsic character. This sense of presence, moreover, can never be occurrently witnessed, attentively or inattentively, as an object of consciousness. For its non-intentional phenomenology guarantees that it not be borrowed from any object-qualia — and hence from any qualia that can be attended to in principle.108 The phenomenology of witnessing is hence elusive to own observation in the manner that is required.

Awareness is thus a concept of consciousness that portrays something both intrinsically elusive to observation and yet experiential. Unlike unprojected consciousness, which is also elusive and experiential, the concept of awareness is naturally implicated in the elusiveness puzzle, capturing what is distinctive about the subject’s *modus operandi*. The concept of awareness therefore successfully captures the kind of consciousness that is alluded to in the puzzle of elusiveness. Given that this elusive yet phenomenal feature seems inherent to all conscious states, generating this puzzle, then so too, must awareness.

---

107 A similar problem may arise in connection with Uriah Kriegel’s (2004) account when he treats what he terms “intransitive self consciousness” (viz., the phenomenology of self *qua* subject), as just another element of peripheral awareness. Kriegel’s account seems to initially improve on Evans’s account insofar as Kriegel identifies only a select part of unprojected consciousness (rather than the whole of it) with the phenomenology of awareness; namely, with that aspect pertaining to one’s peripheral *knowledge that* one is perceiving, thinking etc. However, Kriegel does not think that this intransitive ‘awareness of oneself’, as he calls it, is fundamentally different (in kind) to peripheral awareness of other things (2004, 195), holding, even, that attentive awareness of ourselves as *subjects* is possible, although relatively rare (2004, 194). Kriegel’s approach fails to capture the elusive *‘modus operandi’* of *‘subject’* since it blatantly casts the so-called *‘subject’* as an *object*: that is, as something that can, in principle, be attended to.

108 I thank Graeme Priest for his objection at a conference, which led me to clarify this point. (Before I had merely said that the witnessing and not the sense of presence was elusive to itself, which led to confusion).
The Synchronic Unity of Consciousness.

We turn now to the 'synchronic unity of consciousness'. Tim Bayne and David Chalmers have recently articulated this phenomenon in the following way:

At any given time, a subject has a multiplicity of conscious experiences. A subject might simultaneously have visual experiences of a red book and a green tree, auditory experiences of birds singing, bodily sensations of a faint hunger and a sharp pain in the shoulder, the emotional experience of a certain melancholy, while having a stream of conscious thoughts about the nature of reality. These experiences are distinct from each other: a subject could experience the red book without the singing birds, and the could experience the singing birds without the red book. But at the same time, the experiences seem to be tied together in a deep way. They seem to be unified, by being aspects of a single encompassing state of consciousness (2003, 23).

From this we can formulate the phenomenon that underlies the puzzle of synchronic unity:

There appears to be a synchronic unity of consciousness when, at any one time, a multiplicity of objects in consciousness seems unified to a subject by being aspects of a single encompassing state of consciousness.

The puzzle is how normal experience, so obviously characterised by a multiplicity of objects, can at the same time seem deeply unified to a subject. As with elusiveness, the unity phenomenon is puzzling because it seems to point to a phenomenological contradiction. What concept of consciousness best captures the underlined “single encompassing state of consciousness” and hence, best conveys the source of apparent conscious unity that belies the obvious complexity? As with elusiveness, we shall countenance a few possibilities, beginning with the least plausible: ‘phenomenal object consciousness’.

For similar reasons as before, talk of ‘phenomenal object consciousness’ is poorly suited to substitute the phrase ‘a single encompassing state of consciousness’. Basically, the conveyed qualia lie on the wrong side of the subject/object divide. Object-qualia appear as the complexities that are unified to a subject in an encompassing state of consciousness, not as the single state of consciousness in which the complexities seem unified. Indeed, the encompassing state of consciousness does not appear as an object of consciousness at all. If it did appear that way, then it is likely that Hume, desperately seeking the source of unity that he termed “simplicity”, would have located it amongst his disparate perceptions. We can
therefore easily dismiss 'phenomenal object consciousness' as a candidate to articulate the unity phenomenon that underpins the puzzle.

Talk of 'higher-order consciousness' is also unsuited to capture what is meant by the phrase 'single encompassing state of consciousness'. The unity of consciousness, like elusiveness, is puzzling because it presents as something whose presence is intuited subjectively and pre-theoretically, from the first-person perspective. Hence the apparent phenomenological contradiction: a unity of consciousness that subjectively seems to belie the complexity of its content (the philosophical challenge being to explain this puzzling appearance). One does not have to be a theoretician or philosopher to discern, in a given conscious state, a sense of unity underlying the multiplicity: unity does not seem like merely an abstract inference a la Kant. If it did seem like a mere abstract inference, then unity would never have been puzzling in the way that it is. Now, a higher-order mental state may well be unifying, but because there is by definition no intrinsic phenomenal feel to this unity – the phenomenal dimension to consciousness being exhausted by the multiplicity of objects to which the state is directed – the notion 'higher-order consciousness' (in virtue of the higher-order mental state) cannot substitute that notion of consciousness implicated in the unity puzzle.

Another possibility lies in talking in terms of 'access consciousness', as the notion is defined by Ned Block (1995) and explored by Bayne and Chalmers (2003). Bayne and Chalmers explore the notion of access consciousness for reasons that are similar to those outlined in this section, namely, to see whether the notion captures what seems unified about consciousness. Before even looking at their account, we can dismiss talk of 'access consciousness' as a suitable for unpacking what is meant by consciousness in the unity puzzle. Like higher-order consciousness, access consciousness is not defined with reference to any intrinsic phenomenal state. Rather, it is defined with reference to causal and functional roles in the cognitive economy. For a subject to be access-conscious of an object is for that object, viz., the content of a mental state, to be available for verbal report, reasoning and voluntary behaviour. For example, if I am access-conscious of a book, then that entails I can say things like "there's a book" and do such things as pick it up and return it to the library.

---

Since there need not be anything it is like for me to execute this exercise, then any unity bestowed by access consciousness will not, by simple virtue of its being bestowed by *access consciousness*, be phenomenally felt (see Bayne and Chalmers, 2003, 28).

Having dismissed the terminology of ‘access consciousness’ as a suitable candidate for unpacking the notion of consciousness in the above-formulated unity puzzle, it is nevertheless an instructive exercise to see how Bayne and Chalmers (2003) go about dismissing it. Their analysis, we shall see, will shed further light upon both the character of unity puzzle, and their motivation for adopting the notion of consciousness that they do think is suitable. Now, we may ask just what it is for a state of consciousness to be unified in accordance with access consciousness. How could access consciousness serve as a single encompassing state of consciousness that unifies a multiplicity of objects? Bayne and Chalmers suggest the following formulation. For a set (or pair) of objects to be access-unified, that is, unified in a single state of access consciousness, they must be available for verbal report, reasoning and voluntary behaviour, not only individually, but conjointly (2003, 31). Suppose that elements in their example above – the singing birds, the pain in the shoulder and the red book – were access-unified to a subject. Their conjoint access would be graphically illustrated if the subject were to deliberately use their unpainful shoulder to pick up the book and fling it at the singing birds while yelling “damned Australian ravens!” In such a case, one would clearly be acting upon cognitive content pertaining to each of these objects in an integrated, rather than merely individual manner.

Notwithstanding our previous considerations, the term ‘access consciousness’ may seem like a suitable substitute for the term ‘consciousness’ in the unity puzzle, since when we are disposed to *behave* as if our contents of consciousness are unified, we usually *feel* that they are unified. However, as Bayne and Chalmers point out, the problem with access unity is that it can appear to break down – and it seems counterintuitive to suppose that in such scenarios, consciousness will feel or appear *diss*unified. In other words, it seems that access unity can come apart from what we intuitively call ‘the unity of consciousness’. They cite the following experiment by George Sperling (in 1960) as a clear example of where access-unity

---

110 I have noticed with interest that ravens (although many call them “crows”) in Australia vocalise their cries with a loud, drawn-out nasal twang that seems to capture what is quintessentially *Australian* about the local accent. When congregating in groups, the volume of their ‘conversations’ can reach such a pitch that it has been known to irritate members of even the ‘calmest’ sector of society, the Buddhist sangha!
may crumble. In this experiment, a subject is shown, for a brief duration of 250ms, a grid containing three rows with four letters per row. A tone then sounds and the subject is asked to recite the letters in either any given row or in all three rows. When asked to recite any given row, for example the second row, the subject will correctly report on average 3.3 out of the four letters. When asked to recite all three rows, however, the subject will recall, on average, only 4.5 out the 12 letters. A plausible interpretation of these findings is that with regard to all three rows, access unity has broken down. Access unity breaks down when elements of consciousness are individually but not conjointly available for report. The fact that the subject is able to recall (with reasonable accuracy) any given row, suggests that each of the rows are, at a given time, individually available for verbal report. The subject is thus access-conscious of each of rows 1-3. The additional fact that the subject cannot accurately recall all three rows suggests that while each row is individually available for report, they are not jointly available. The three rows, of which the subject is access-conscious, are not access-unified to them. This would then seem to be a case where access unity has broken down (2003, 35).

Now, if the unity of consciousness were defined with reference to access-consciousness, viz, ‘access-unity’, then we would have to say, with regard to the above case, that the unity of consciousness will have appeared to have broken down. But this just doesn’t seem plausible. It seems highly implausible that from the phenomenal perspective of the subject undergoing the experiment, their conscious contents will have at any stage appeared to have become disunified. As Bayne and Chalmers point out, perhaps the whole grid will appear fuzzy to them, or not present itself as an entire grid, but whatever the experience, it is very unlikely that their field of consciousness will seem disunified (2003, 36). This crucially ties into a point that the authors make in another part of their paper, namely, that even if the unity of consciousness could actually break down, it certainly doesn’t subjectively seem like something that could break down (2003, 37). Unity seems to be a deep and necessary feature of any conscious state, its breakdown impossible to subjectively imagine. The problem with access unity is that its breakdown is all too easy to imagine – and it probably actually happens. As we are concerned here with the level of appearances, namely, with how the unity phenomenon seems to display itself, such considerations serve to strike out ‘access unity’ as
a plausible term to substitute ‘single encompassing state of consciousness’ in the puzzle of unity.

We will now consider a concept of consciousness that Bayne and Chalmers do think accurately conveys what is behind ‘the unity of consciousness’. Unlike the concepts expressed by the terms ‘access consciousness’ or ‘higher-order consciousness’, the proposed concept is of something inherently phenomenal, involving what Ned Block (1995) refers to as “phenomenal consciousness”. The authors write that a mental state is phenomenally conscious “when there is something it is like to be in that state...being in that state involves some sort of subjective experience” (2003, 28). Importantly, it would seem that this “subjective experience” is not necessarily restricted to qualia pertaining to objects of consciousness – it involves any dimension that is experiential. If there turns out to be phenomenology that pertains to the apparent subject of experience (through its modus operandi), then the concept of phenomenal consciousness will also allude to that. This immediately puts the concept of phenomenal consciousness at an advantage over that of phenomenal object consciousness in capturing what is meant by the term ‘single unifying state of consciousness’. Now, Bayne and Chalmers continue their analysis of unity by referring to what they call “subsumption” (2003, 26-27). A subsumptive state of consciousness – whether access or phenomenal – is a state which, at any given moment, subsumes or encompasses all of a subject’s experiences: auditory, visual, emotional, proprioceptive etc. Subsumption is non-trivial because the subsuming state is not merely the conjunction of all these experiences, but is a single conscious state in its own right (2003, 27). It may, they say, be thought of as the “subject’s conscious field” which serves to unify the disparate elements, or as “the singularity behind the multiplicity” (2003, 27). Now given that access consciousness fails to properly capture the single unifying state, a natural step is to analyse the phrase ‘subsumptive consciousness’ in terms of the phrase ‘phenomenal consciousness’. A subsumptive, phenomenally conscious state will thus be a state of consciousness in which there is something it is like for a subject to be in a single encompassing state that unifies the disparate elements. In the words of the authors:

A set of conscious states is phenomenally unified if there is something it is like for a subject to have all the members of the set at once, and if this phenomenology subsumes the phenomenology of the individual states. (2003, 32). A phenomenal state A subsumes phenomenal state B when what it is like to have A and B simultaneously is the same as what it is like to have A. (2002, 41).

To summarise this in my own (albeit similar) words, the notion phenomenal unity might be defined in terms of phenomenal subsumptive consciousness thus:

A set of conscious states will be phenomenally unified (at a time) when there is something it is like to be in a single subsuming state of consciousness that encompasses all the individual states.

On this analysis, the 'what-it-is-likeness' of the subsumptive state captures the apparent singularity behind the multiplicity in any given conscious state. We can immediately gather that subsumptive phenomenal unity must be something that has an intrinsic subjective character (pertaining to the what-it-is-likeness of its singularity) that is not overtly reducible to the multifarious character of its objects. The unity will hence, qua that unity, seem to bestow its own phenomenal dimension of singularity to any conscious unified state. So far so good. The notion of phenomenal subsumptive consciousness also carries an advantage over that of access consciousness in that it builds in the stipulation that it can never seem like a consciousness that is disunified. For if a state of consciousness were to seem disunified (from the perspective of the subject) then that would imply that there is no longer something it is like for that subject to be in a single subsuming state of consciousness. The state would no longer seem single but fractured, which would imply that it is no longer a state of phenomenal subsumptive consciousness. Hence, the term 'phenomenal subsumptive consciousness' nicely captures our intuition that the unity of consciousness is not something that can ever seem to break down. And yet the analysis does not prematurely guarantee the independent reality of this unity principle. It allows that the principle of unity – viz., phenomenal subsumptive consciousness – could turn out to be illusory.

Can we be more specific about the phenomenology of the subsumptive state? We know that part and parcel of the notion of subsumption is that the character of this unifying state is not overtly reducible to, nor derived from, merely the conjuncts of the state, but is a state in and of itself. Given the analysis so far, this must involve a phenomenal state in and of
itself. And given this, we might ask: from where could the phenomenal character of this state appear to be derived, if not merely from characteristics of objects that comprise the conjuncts? The answer must be: from that pertaining to the apparent subject. And the phenomenal character pertaining to the subject is strongly indicative of its *modus operandi*: awareness. Hence, we have good initial reason to suppose that awareness is the concept of consciousness being implicitly appealed to when analysing what is meant by talk of a 'single encompassing state of consciousness'.

However we have yet to establish a direct link between awareness and the synchronic unity of consciousness. Does awareness intrinsically yield a felt singularity that would appear to belie any observed multiplicity? The question is best explored by countenancing the possibility that awareness, as a witnessing sense of presence, could seem disunified at a time. On reflection it could not. Suppose awareness were to appear, from the first person perspective, to be disunified to consciousness at time, t. On such a scenario, there would seem to be disunified phenomenal witnessing at time t. But to what perspective could this *seem* to be so – if not from the built-in first-person perspective of witnessing? We would have it that phenomenal witnessing appears as disunified to ... unified witnessing. Why *unified* witnessing? For the following reason: if witness-consciousness is that which apprehends any disunified X as a disunified X, in a *single conscious moment*, then that witness-consciousness must itself, at that moment, not seem disunified in relation to the disunified X that appears to it. In short, witnessing can never seem disunified at a time, since witnessing would have to seem unified in order to apprehend its own disunity, which is impossible. Awareness also satisfies the criterion of being a subjectively felt unity – a singularity behind any multiplicity – by virtue of its subjective sense of presence. The fact that witnessing can never appear as disunified at a time is not, hence, to be interpreted as evidence of a mere inferred unifying vacuum or nothingness (which would also never appear as disunified at a time since there would be no appearance *per se*). It is a *positive presence* that is felt to be necessarily unified at a time. Awareness, in virtue of its felt witnessing, is therefore a suitable candidate to capture the apparent singularity behind the multiplicity, in a given conscious state.

I hold there is good reason to prefer the concept of awareness to that of phenomenal subsumptive consciousness in capturing what is meant by the phrase 'single encompassing
state of consciousness’ in the unity puzzle. The concept of awareness is simply more informative. The concept of phenomenal subsumptive consciousness is on the right track for it serves to tell us that there is something it is like to be in a single conscious state that encompasses the multiplicity of subsumed objects. The concept of awareness tells us in addition that this ‘something it is like’ pertains to the modus operandi of a subject (rather than any object), is of the nature of witnessing, and is felt as a subjective sense of presence. That sense of presence, if awareness exists, would involve an immediate sense of the perpetual here and now. In the previous section, we learnt that despite being experiential, it appears elusive to its own observation. And we now learn that awareness involves a sense of being that which, at any given time, will unify disparate objects in single field of consciousness. In the next section, which articulates the puzzle of what we have called ‘diachronic unity’, we will argue that the concept of awareness is implicated in describing the associated features of endurance and invariability. If successful, these arguments will render the concept of awareness to be yet more informative – and most significantly, a common concept that is appealed to when formulating all these central puzzles of consciousness. For the time being, we can note that awareness is the concept of consciousness that, by virtue of being more informative than (and indeed subsuming) the concept of subsumptive phenomenal consciousness, is most fittingly implicated in the puzzle of synchronic unity. Given that unity seems to feature in all conscious states, then so too, does awareness.

**Diachronic Unity of Consciousness: Endurance and Invariability**

Let us consider whether the notion of awareness is to be appealed to in articulating the puzzle of diachronic unity, through which arise the features of endurance and invariability. When unity of consciousness is conceived of as persisting over time, hence diachronically, then this gives rise to a commonly held intuition, namely, that from one conscious moment to the next it is the very same unchanging consciousness that unifies the various conscious states.

*Whenever conscious states are apprehended as changing from one moment to the next, it seems that we can intuit an elusive yet unifying, enduring, invariable consciousness that observes the change.*
What concept of consciousness is implicated here?

Explaining what accounted for the intuition behind the puzzle was a problem that Hume famously grappled with. After trying to explain this intuition by appealing to such factors as memory and imagination, a dissatisfied Hume concluded in a footnote:

But all my hopes vanish, when I come to explain the principles, that unite our successive perceptions in our thought and consciousness. I cannot discover any theory, which gives me satisfaction on this head...(1740, 175).

Hume could not abandon his intuition that, although contents of consciousness are changing, there at the same time seems to be an underlying unifying principle that persists, consciously registering the change, not itself changing. We can surmise that Hume was seeking a conscious state or mode that involved the element of synchronic unity – ‘simplicity’ as he called it – unifying the otherwise disparate elements of consciousness at a given time. We can also surmise that the state he sought was elusive to his introspective gaze – otherwise Hume would have located it amongst his flux of rapidly altering percepts.

We have already argued that the concept of awareness is implicated in articulating the puzzles of both elusiveness and synchronic unity. The elusiveness dimension would accord with Hume’s failure to locate unity amongst the discontinuous, changing objects of consciousness. The phenomenal dimension, viz., a subjective sense of unifying presence, would accord with his persisting intuition that even though he cannot locate the source of unity, any account that ignores it will be incomplete. Given that the synchronic unity of consciousness is to be described with reference to the concept of awareness, there is a strong initial reason to suppose that the diachronic unity of consciousness should also involve reference to awareness. But the dimensions of numerical identity (endurance) and qualitative identity (invariability), with which Hume was also concerned, must still be independently specified. Because these features appear to qualify the principle of unity (hence the term ‘diachronic unity’) and because ‘unity’ is articulated with reference to awareness, it will be most parsimonious – and in keeping with how things seem – to also specify the terms ‘endurance’ and ‘invariability’ with reference to awareness. Hence the strategy here will differ from that adopted when articulating synchronic unity and elusiveness. Instead of first seeing whether rival concepts of consciousness can be appealed to in describing what is
meant by the terms ‘endurance’ and ‘invariability’, we will jump straight to the exercise of seeing whether awareness implies the features that are specified by these terms. If it does imply them, then we will have prior parsimonious reason to select the concept of awareness as ideally depicting the notion of consciousness alluded to in the puzzle of diachronic unity.

**Endurance:** In a given state of apparently enduring consciousness which registers change in other conscious elements can awareness ever seem unenduring, viz., perduring or gappy? If it could seem perduring, then it would not seem suited to capture the notion of consciousness above. But we can quickly determine that awareness can never seem numerically unenduring in this manner. As the most general mode of conscious apprehending, witnessing will be ‘set up’ to notice comings and goings of conscious objects – but never its own coming and going. This is not to deny that from an ‘objective’ or third-person perspective, witnessing could actually be inferred to arise and pass away. However to say that witnessing directly notices its own absence, in a given state of enduring consciousness, is to commit to a contradiction. For what could possibly notice the absence of witnessing, from the first-person perspective, except witnessing? Witnessing would have to be present to notice its absence! Therefore, awareness can never be something that seems to come and go in any given state of enduring consciousness. Awareness will rather seem to be that enduring consciousness. Nor does the witnessing equate to a mere absence, which would also not seem to perdure. The subjective sense of presence bestows the enduring witnessing with a positive ‘lived’ dimension, which would create the puzzle of numerical continuity that had the empirically minded Hume so stumped. From this (and previous considerations) we can conclude that the concept of awareness is well suited to capture what is meant by ‘endurance of consciousness’ in the puzzle of diachronic unity – which is a puzzle about how consciousness seems.

**Invariability:** In a given state of apparently enduring consciousness that notices variation in other conscious elements, can awareness ever seem, essentially, to vary in quality? If it could, then its concept would be ill suited to capture the notion of consciousness above. On the face of it, awareness might seem to vary. After all, there is bright attentive witnessing and dim inattentive witnessing, perceptual witnessing (through five different senses), cognitive witnessing (to do with thinking, remembering, imagining, having emotions) and introspective witnessing (to do with noticing one’s own thought-processes
etc). But to deduce from this the fact that witnessing seems qualitatively variable in its essence is to overlook the possibility of a factor common to this set of apprehendings. They all seem, after all, to be types of witnessing – which further articulates the basic intuition of something invariable underlying (and yet discernible within) each conscious state. And we are in effect asking whether the concept of awareness yields a common qualitative element in virtue of which all these modes of apprehension would seem to be types of witnessing. Is there a first-personally discernible (hence phenomenal) feature about awareness that could characterise witnessing qua witnessing, as opposed to merely witnessing qua attending, perceiving, introspecting, and so forth?

On our definition of awareness, there is indeed a common phenomenal element that characterises witnessing qua witnessing. The subjective sense of presence, which is unifying, elusive and present-centred, must appear as the same neutral quality in any mode of conscious apprehension. Hence it cannot seem to alter from one moment to the next, despite the alterations in modes of perception etc. Moreover, given that awareness does not appear to derive its phenomenal character from the changing objects of consciousness, there would seem to be no parameters along which awareness could seem to alter in phenomenal quality from one moment to the next. It is useful to compare awareness once again to the beam of a flashlight. Although the beam may vary in intensity, either at one time or over time, the essential quality of luminosity stays the same. Luminosity is the deeper quality that unites the focused beam and its dimmer circle of surrounding light, just as the subjective sense of presence (unified and elusive) is a deeper quality that unites each mode and moment of witnessing – should awareness be real. And while the flashlight-beam with its surrounding circle of dimmer light is a suitable analogy for attentive versus inattentive witnessing, luminous spectral colours may offer a suitable analogy for the different perceptual, cognitive and introspective modes of awareness. For despite their diversity, the spectral colours are also united by the deeper ontological fact of their luminosity.

There is a further type of reflection that seems to reveal awareness to be an invariable element to all conscious states. The very fact that we are able to immediately know, from the first-person perspective, that we are perceiving (e.g., seeing or hearing or feeling) or introspecting thoughts, or attending to something sharply, would suggest that this mode of knowing is not in itself limited to one of these categories of apprehension. Rather it is neutral
between them, which suggests generic witnessing, viz., awareness. Caston (2002) has recently interpreted Aristotle’s theory of consciousness in a way that supports this idea. On this reading, Aristotle argues for a perceptual awareness that has its own intrinsic, felt character (2002, 759) that is immediate (not mediated by any representation (2002, 778)) and yet indirect (primarily directed at objects other than itself, thus elusive (2002, 785-787)). Through such generic awareness, integral to any act of perception or introspection, we immediately and reflexively know which mode of perception or introspection is being employed. As I mentioned near the beginning of this chapter, this Aristotelian or Brentanian notion of consciousness is reasonably similar to the notion of ‘awareness’ being developed in this chapter. The main difference is that the concept of awareness developed in this chapter (a) does not build any object-directedness into its definition (including taking itself as a secondary object ‘on the side’) and (b) is, I hope, clearer in articulating the subjective character of the phenomenal witnessing, through its linkage to the puzzles of consciousness.

From our considerations in this section, we can thus conclude that reference to awareness, in virtue of its elusive, unifying, witnessing sense of presence, seems well suited to capture that elusively felt dimension of invariability: the concept of awareness certainly implies such invariability. Given that awareness is implicated in the apparent features which endurance and invariability seem to qualify – elusiveness and synchronic unity – then it seems most economical to suppose that awareness is the concept of consciousness which best captures what is meant by the phrase ‘unifying, enduring, unchanging consciousness that observes the change’ in the puzzle of diachronic unity. Since diachronic unity seems central to all conscious states, then so too, does awareness.

Unconstructedness

While no particular philosophical puzzles are associated with the feature that I am referring to as ‘unconstructedness’, this feature is implicit in each of those we have considered. For example, just as elusiveness of the subject would not seem particularly puzzling if the subject’s modus operandi lacked an experiential dimension (as with a video camera), it would equally not seem puzzling if the modus operandi appeared like just another observable object, for example, as the mere content of a thought or perception. In other words, if the

---

112 Kriegel (2004), (forthcoming), also develops his account along these Aristotelian/Brentanian lines.
modus operandi of subject appeared as overtly constructed – with its phenomenal dimension appearing contributed to by various observable thoughts or perceptions – then there would not be an elusiveness puzzle. The modus operandi, seeming observable, would no longer seem elusive to observation. It is by virtue of the very fact that the modus operandi of ‘subject’ seems unconstructed by any observable objects – and yet experienced alongside them – that there is a puzzle of elusiveness.

With synchronic unity, a similar story applies. If the ‘singular encompassing state of consciousness’ appeared as an overt construct – its subjective sense of presence appearing contributed to by various observable thoughts and perceptions – then there would be no puzzle of synchronic unity. It is only because that singular unifying state does not present as an observable object of consciousness – but yet seems experienced – that synchronic unity becomes a deeply puzzling phenomenon. The ‘unity of consciousness’, at least as formulated through the puzzle articulated in this chapter, thus involves the impression of being unconstructed by any observable mental contents. The endurance and invariability of consciousness, besides seeming elusive and unifying (and hence unconstructed from this angle) are also puzzling in virtue of the fact that the consciousness they appear to qualify seems unconstructed. It does not introspectively seem as if any overtly constructed content of consciousness – such as that pertaining to thoughts and perceptions – is either enduring or invariable. Any careful observation (such as Hume’s famous introspection) will reveal such mental content to alter from one moment to the next. Hence the consciousness that seems enduring and invariable does not seem constructed, but unconstructed. In virtue of its non-intentional subjective sense of presence – implicated in each of the aforementioned puzzles – awareness, if it exists, will therefore be unconstructed.

3. The Concept of Awareness as a Central Concept of Consciousness

The concept of awareness is implicitly appealed to when describing what seems quintessentially subjective and in many ways puzzling about consciousness. Bayne and Chalmers (56-57, 2003) have already suggested that synchronic unity should perhaps be built into the basic concept of consciousness, such that any explanation of consciousness must involve an explanation of the synchronic unity of consciousness. My suggestion is simply to take this idea further and also build diachronic unity (with endurance and invariability),
elusiveness and unconstructedness into a basic concept of consciousness. Since the concept of awareness both implies and is implied by these puzzling subjective features, it should be a primary concept that is addressed in any theory purporting to concern itself with the subjective character of consciousness. The “hard problem of consciousness” will indeed be best addressed if one works with a concept of consciousness that singularly captures its hardest problems.

4. Why Boundedness is Not Implied By Awareness

Having established that a conceptual link exists between the concept of awareness and those delineating the features Buddhism regards as essential to awareness, the next step is to show that the link does not extend to the concept of boundedness as played out through the role ‘personal owner’. If the concept of awareness did imply the concept of boundedness in the same way that it does the concepts of elusiveness etc., then the Buddhist account of consciousness and no-self could not possibly be correct. On the Buddhist position, the self is deemed constructed in terms of its boundedness through the role ‘personal owner’, with unconstructed input from those features of consciousness that we have co-specified with awareness. Should the concept of boundedness also be implied by the concept of awareness, then any ontological status attributed to awareness and its co-specifying features – whether as unconstructed, constructed or illusory – would also have to be attributed to boundedness. With boundedness (through the role personal owner) on the same ontological decking as awareness et al, there could be no hope of defending an account of no-self where the contributing consciousness (viz., co-specified awareness) is unconstructed, whilst boundedness and personal ownership are constructed. In this final section of the chapter, I set out to demonstrate that the concept of boundedness does not follow from the concept of awareness.

We have already argued (in chapter four) that our impression of being a bounded, separate, ontologically unique entity is best evidenced through an assumed ‘self-identification’ whose broadest mode is played out through the role of personal owner. Through identifying as a personal owner (and hence minimally with a perspectival owner), the subject assumes itself to be a ‘me’ with reciprocal feelings of ‘mine’ towards various other objects that relate to the identity-role. The feeling of being a unique and personal owner
qua 'me' will, should the self be real, be most suitably reflective of the self’s ontologically unique, bounded status.

What we have to argue, therefore, is that the concept of awareness does not imply the concept of boundedness as it is borne out through this role of 'personal owner'. Boundedness will of course implicate awareness. If it did not, then there could not be any sense of a bounded self. By its very definition, the self is a personalised, bounded subject that has unity, elusiveness and so forth. We want to know whether the (co-specified) awareness and modus operandi of subject would, in and of itself, necessitate boundedness. If we can conceive of awareness without a sense of boundedness, then this will be enough to demonstrate that the concept of awareness does not imply the concept of boundedness.

We can show quite easily that the concept of awareness does not imply that of boundedness. While we have not yet investigated the psychological possibility of awareness sans a sense of personal boundedness, we can certainly conceive of this possibility, which is all we need for our purposes. In chapter four we considered various pathologies (anosognosia and depersonalisation) involving subjects to which a portion of perspectivally owned objects did not, at a time, seem personally owned. All we have to conceive of is a 'global' case where, at a time, no perspectivally (or possessively) owned objects seem personally owned by its subject. From our analysis in chapter four we can infer that with no sense of personal ownership towards anything, there would be no reciprocal sense of being a personal owner, no sense of (personal) 'me' and 'mine', and hence no sense of boundedness. Newborn infants and primitive organisms may conceivably harbour awareness sans any sense of boundedness. Now we have not yet discussed whether, as a matter of psychological fact, such cases exist - or whether a loss in the sense of boundedness would entail a loss of actual boundedness. But the very fact that we can conceive of and indeed investigate the psychological possibility of awareness without a sense of boundedness, shows that the concept of awareness does not intrinsically imply the concept of boundedness as it does the concepts of elusiveness, unity and so forth.

**Conclusion**

In this chapter we have aimed to establish, in accordance with Buddhism, that the concept of awareness, viz., witnessing with an intrinsic subjective sense of presence, is to be co-
specified in terms of the concepts of elusiveness, unity, endurance, invariability and unconstructedness. Through capturing these puzzling subjective features of consciousness, awareness serves to capture what seems quintessentially subjective about consciousness itself. For this reason, I believe that awareness (as opposed to popular rival concepts) ought to be a central concept of consciousness in its own right. I also aimed to demonstrate, in accordance with Buddhism, that the concept of awareness does not imply the concept of boundedness that is associated with a self's central role of personal owner.

We have not as yet established (in accordance with Buddhism) the independent reality of awareness or its co-specifying features. For this to be established, awareness must be shown to exist in the manner that it purports to exist – as non-illusory and unconstructed. The next chapter will thus involve an argument that the co-specified awareness, integral to our sense of self, is unconstructed because it has non-illusory status.
Introduction

In this chapter we aim to establish the non-illusory status of awareness and its intrinsic features (whose concepts are implied by the concept of awareness). For this non-illusory status to be established (which is needed if awareness is to have the independent reality ascribed to it by Buddhism) awareness must be shown to exist in the manner it purports to exist, namely, in the manner specified in the previous chapter. Awareness is specified as a witnessing subjective presence whose non-intentional phenomenology is felt as present-centred, unifying (both at any one time, or from moment to moment), yet elusive to observation. As something whose phenomenology is unborrowed from objects of consciousness, awareness, if it exists, must exist as completely unconstructed by thought or perceptual content. If apparent awareness, perhaps in virtue of one or more of its defining features (that form part of its content or 'aboutness') turned out to owe its existence to thought-content rather than to (unconstructed) awareness itself, then that would render awareness constructed and illusory and hence lacking in independent reality. For example, if the analyses of 'no-self by Western thinkers mentioned in chapter five (such as Hume (1739) and James (1890)) were correct, with the features of unity, endurance and invariability being constructed, then awareness, implying as it does these features, would similarly have to be viewed as being constructed and hence illusory. This chapter will first argue that the co-specified awareness, central to our sense of self, has unconstructed reality. We then address a possible objection from those who may still deny the independent reality of awareness by advocating eliminative materialism. Finally, we consider the influence of what I call the ‘object-knowledge thesis’, which has, in my opinion, thwarted the popularity of awareness as a concept in the philosophy of mind.

The Central Argument

The central argument is as follows. We have established (above) that if awareness is a construct, then it must be an illusion. If it can be shown that awareness is not (and perhaps cannot possibly) be an illusion then it will follow, by modus tollens, that awareness is not (or cannot possibly be) a construct. To see that awareness is not an illusion we need to first
recapitulate what, ontologically speaking, is involved in the generation of an illusion. In chapter five, we distinguished between two common usages of the term 'illusion', namely, (a) illusion \textit{qua} vehicle, viz., appearance-of-content-to-subject-S and (b) illusion \textit{qua} content-of-appearance-to-S. Now from the outset we defined 'illusory X-as-F' as an X that, from the perspective of S, purports to exist in a manner F when it does not actually exist in manner F. It should hence be clear that 'illusory X-as-F' pertains to (b), namely, to illusion-\textit{qua}-content rather than to (a), namely, illusion-\textit{qua}-vehicle. Indeed, unless we accepted the non-illusory reality of the illusion-vehicle – the \textit{appearance-of-(X-as-F)-to-S} – there could be no such phenomenon as \textit{content-of-appearance}, X-as-F, which turns out to fall short of its purported mode of existence. In other words, there has to be an actual, non-illusory appearance as an event occurring to subject S at time t, in order for there to \textit{be} any content of that appearance conveyed to S which could, on further examination, turn out to be illusory. For example, in the case of the Mueller-Lyer illusion, the appearance of two objectively uneven lines to a subject is not illusory: what is illusory is the conveyed content \textit{<two objectively uneven lines>}. And with the self, what various thinkers deem illusory is not the feeling of being a self (viz., the appearance or sense of self), but the self (the content) which that feeling conveys. The \textit{appearance-of-(X-as-F)-to-S-at-t} is hence not illusory in our defined sense: it exists as the vehicle \textit{for} the illusory content.

Given this, what are we to say about awareness, with the features in terms of which it is specified? Like the self, awareness is peculiar in that its appearance-vehicle and content seem, from the perspective of the subject, to coincide: there does not, phenomenologically speaking, seem to be a \textit{gap} between the content which awareness intrinsically conveys (viz., a witnessing unified, elusive, enduring, invariable sense of presence with a status that is unconstructed) and the vehicle of subjective appearance which conveys that content. From this, we may be tempted to conclude that since the \textit{appearance} of awareness must be non-illusory, then for that very reason the content \textit{<awareness>} (with all its co-specifying features), which seems to transparently characterise the nature of that appearance itself, must likewise be non-illusory. This move, however, is fallacious. Appearances can be deceptive: while certain content (such as unconstructedness) may phenomenologically \textit{seem} to intrinsically characterise an appearance, the appearance may conceivably turn out, intrinsically, to lack that feature (which would render the feature/content illusory). After all,
if we were to conclude that awareness is non-illusory (and hence unconstructed, as it seems to be) on the simple basis of how it phenomenologically appears to the average subject (as united with its content) then we should also have to conclude that the self, whose very presentation seems to suffuse awareness with the content <boundedness>, is non-illusory. In the following chapter I hope to demonstrate how boundedness does not intrinsically characterise the ontology or appearance of awareness – even though it commonly seems, from the first-person perspective, to characterise awareness in this fashion. (We have already shown that the concept of boundedness is not in fact implied by the concept of awareness).

Nevertheless we may still ask whether there is some particular feature to awareness whose very existence is implicated in the ontology of any appearance such that without this feature, there could not be an appearance – including the appearance-vehicle for an illusion. When we put the question this way, it becomes clear that there is at least one such a feature to awareness, namely, the central witnessing component implicated in the ‘appearing-to-subject-S-at-t’ part of the illusion-vehicle. In order for any illusion to take hold, there must be a subject, and the subject necessarily involves a witnessing (first-person) perspective to which potentially illusory content can appear. The witnessing cannot itself exist as illusory content-of-appearance, since witnessing (with its subjective sense of presence) is built into the underlying vehicle of appearance that would carry any illusory content. Since the very concept of witnessing (with its subjective sense of presence) demands that the denoted witnessing presence qua witnessing presence can never manifest as other than elusive, unified, enduring, invariable and unconstructed, then these features will inherit the non-illusory status of the witnessing presence to which they are intrinsic. All the aspects that awareness intrinsically brings to the sense of self will hence be non-illusory. This implication notably does not extend to boundedness which, as we argued in chapter six, is not implied by the concept of awareness (although it would usually seem, phenomenologically, to be intrinsic to the awareness we experience). If awareness is non-illusory, then it cannot be constructed, since we have already determined that if awareness were constructed then it would have to be illusory. The non-illusory status of awareness will guarantee its unconstructed status and hence, that awareness exists, intrinsically speaking, in the manner that it purports to exist; namely as the content of an appearance that is not contributed to by any objects of consciousness such as thoughts, feelings, emotions or perceptions. Awareness
underlies its perspectively owned objects – as it seems to do – the objects do not construct awareness. The claims made by Hume (1739), James (1890), Dennett (1991), Flanagan (1992) and Damasio (1999), to the effect that unity, endurance and invariability are constructed by such objects, are therefore false.

**The Spectre of Eliminative Materialism**

We have argued that awareness can be neither illusory nor constructed. Some philosophers may nevertheless claim that no subjective or mental phenomena can have any reality, because awareness, implicated in any subjective phenomena, is not real. It is hard to see how such a view could be developed unless it was, in the end, some version of eliminative materialism. Defenders of this position may claim that subjective witnessing, along with all mental phenomena (including illusion-vehicles), is a byproduct of the brain, lacking in any reality. However, there are serious problems with this position, a position that also happens nowadays to be rather unpopular in the philosophy of mind. It is unclear, without resorting to dogma, how one could construe the manner in which awareness and its subjective (perspectively owned) contents could lack reality. We have demonstrated that it is not an option for anyone, including the eliminativist, to declare awareness to be either illusory or constructed – two major ways in which something can, by parties of differing metaphysical commitments, be plausibly deemed to lack reality. This means that if awareness were to lack reality, it could not do so in such a way that its actual ontological status would conflict with its purported mode of existence.

To this, the eliminativist might respond that awareness *is* illusory since its very appearance, as purporting to exist simpliciter, conflicts with the eliminativist standard of reality, which denies existence to appearances. This response, however, would fail to peg awareness as an illusion – even if the eliminativist standard of reality turned out to be correct. It would first of all unacceptably alter the meaning of the term ‘illusion’ such that the structure of an illusion would no longer be confined to a non-illusory appearance-vehicle that carries illusory content. Even if we were to accept this distortion to the term ‘illusion’, awareness, if illusory on this count, would have to *purport to exist* with reference to some standard of reality relevant to that of eliminative materialism. The standard of reality relevant to the eliminativist materialist is, minimally, one whose ontology excludes mental
phenomena such as appearances. But the phenomenon of awareness is silent with reference to this standard of reality – it does not wear on its sleeve ‘purports to have independent reality in a system of metaphysics which excludes the existence of mental phenomena’. The manner in which awareness purports to exist is simply as the (non-intentional) content of an appearance (viz., that of unified, elusive witnessing, etc) not contributed to by perspectively owned objects such as thoughts and emotions. Awareness cannot be deemed illusory by purporting to exist in a manner that relates to the agenda of the eliminativist materialist.

A possible step towards denying reality to awareness, given this restriction, may be to argue that awareness (along with all other mental phenomena) is epiphenomenal, lacking in any genuine causal power (all the causal work being done by underlying neurology). If awareness were epiphenomenal, then this would not render it illusory, since awareness (a) cannot be illusory for reasons already mentioned, and, more specifically, (b) does not ‘wear on its sleeve’ any story about its metaphysical underpinning, including the ultimate status of its causal efficacy in relation to the brain. The concept of awareness would suffer no internal contradiction from the possible scenario of awareness supervening upon neurological properties, which do any causally relevant work. Suppose that awareness were epiphenomenal in this way. This would not be enough to guarantee the non-reality of awareness, since awareness could still plausibly be construed to exist as a real enough effect of the neurological states to which it is reduced.113 There would need to be an additional claim, which declares any causally impotent X to lack reality. One may for instance stipulate, as Richard Baron (2000) may do, that only causally active, objectively measurable elements have a place in the elite ‘scientific world view’ and that anything which is not a part of this scientific framework lacks reality. The problem with this approach is that it seems to violate the sensible philosophical dictum: preserve appearances unless there is over-riding reason not to. To my knowledge eliminative materialism does not provide over-riding reason to renounce what seems so obvious; namely that subjective phenomena, epiphenomenal or not, are as much a part of the world as ‘scientifically respectable’ properties. It is probably for this reason that eliminative materialism never gained popularity in the philosophy of mind.

113 Imagine two worlds, W and W*. W has conscious phenomena epiphenomenally caused by neurological events; W* does not: everyone there is a zombie. The only way to articulate the difference between W and W* is by appealing to the conscious epiphenomena, which would suggest that conscious phenomena have prima facie reality, even if they are causally impotent.
Philosophers of mind nowadays are generally more interested in explaining rather than denying the reality of mental phenomena. The history of philosophy has shown that repeated attempts to banish subjectivity to the 'attic of the mind' because it is methodologically difficult to deal with, has not made subjectivity – including the witnessing perspective from which we approach the world and its scientifically respectable properties – disappear. The mind-body problem is still at large, and facing up to the reality of awareness – whether epiphenomenal or not – seems a necessary step in resolving the mind-body problem.

**The Object-Knowledge Thesis**

While there have been some who have advocated the importance of the concept of awareness (or something similar – see beginning of chapter six), the Western philosophy of mind has not been that receptive to the concept of awareness as defined in this project. The reluctance to give a central place to some version of the concept in the philosophy of mind has partly been because awareness has been viewed as ephemeral and difficult to define and – perhaps the other side of the coin – at odds with a tacit but widespread assumption which I shall call the object-knowledge thesis. The object-knowledge thesis states that all knowledge and experience must be derived from the object side of the apparent subject-object dichotomy (remembering that objects are, in principle, able to be attended to). It rules out the possibility of unmediated knowledge or experience that pertains specifically to the subject-side of the apparent subject-object dichotomy – in particular, to its modus operandi of awareness. With an epistemic tradition based on a bifurcation of rationalism and empiricism, it is unsurprising that the object-knowledge thesis often goes unchallenged in Western philosophy. Despite their differences, both empiricism and rationalism rest upon the object-knowledge thesis. Empiricism maintains that all things we can know must ultimately derive from deliverances of the five senses; rationalism holds that not all things we can know must derive from the five senses but can be contributed to by the intellect. Both empiricism and rationalism limit the source-pool of knowledge and experience to deliverances from either the senses or the intellect, hence ruling out the possibility of unmediated experiential contribution from that pertaining to the subject's modus operandi. (It should be evident that the concept of 'knowledge' is to be understood synecdochically, that is, as representing a broad group of related epistemic notions, such as apprehending, experiencing, intuiting and so forth).
I think it is not an exaggeration to say that part of what has been so puzzling about problems of consciousness (relating to its subjectivity) has been an unquestioning allegiance to the object-knowledge thesis. In other words, the adoption of this thesis has made the problems harder than they have to be. For instance, Hume (1739), locked into his empiricism, could not countenance the possibility that his intuitions of unity (simplicity) and identity (continuity and invariability) could be sourced in anything other than the products and mechanism of his empirical perception. Had he admitted to the reality of awareness as their source, he could have moved forward in his explanation of these intuitions, rather than admitting defeat. And Kant (1787), while allowing that synchronic and diachronic unity did have a source in the noumenal (non-empirical) subject, did not bestow the subject with any intrinsic phenomenal character that could be brought to conscious experience. Kant’s analysis did not therefore account for the persisting intuition that there is something it is like to be in a unified state of consciousness whose phenomenal character is not entirely reducible to features of the objects unified. Unity is not, a la Kant, merely knowable to the minds of philosophers as the conclusion of some transcendental inference that must hold if conscious experience is to be possible. Unity is apparent to first-person experience – everyone’s experience – and it is because it is apparent that it has continued to puzzle philosophers. Especially philosophers who cling to the object-knowledge thesis, which rules out the possibility of experience sans senses and reason (or intuitions and concepts, in Kant’s terminology).

Several concepts of consciousness, which appear initially similar to the concept of awareness, turn out not to have implied a rejection of the object-knowledge thesis – at least on my understanding of these concepts. I am thinking in particular of the Brentanian and/or Aristotelian concepts of consciousness that have been recently developed by thinkers such as Kriegel (2004), (forthcoming), and Caston (2002). The similarity that these concepts bear to the concept of awareness is that the phenomenal character of the referred-to consciousness, on these theories, is taken seriously and is recognised as not being exhausted by the character of those objects to which the consciousness is attentively and primarily directed at a given moment. The phenomenal character of consciousness is nevertheless, on these positions, accounted for by consciousness inattentively taking itself as its own reflexive ‘on the side’ object – and hence as an observable, peripheral object. Consciousness is not, like awareness,
to be credited with its own intrinsic experiential ‘luminosity’, but must take even itself as an (inattentively observed) object, with the consciousness-object not differing, in token, from the subject-consciousness that observes it.

I have already argued that such positions are at odds with how witness-consciousness presents itself. We have established that witness-consciousness, to which such theories clearly allude, must present as intrinsically elusive to its own observation, whether that observation happens to be attentive or inattentive. As I argued in the section on ‘elusiveness’ (in chapter six), such consciousness presents as that which observes objects, attentively or inattentively – it never presents as an object that is observed by consciousness, whether attentively or inattentively. This is another way of saying that witness-consciousness pertains essentially to the subject side of the apparent subject/object dichotomy, not to the object side – a claim reinforced by each of the puzzles considered in chapter six. Allowing witness-consciousness to be or seem like its own observable object fails to account for our intuition about its intrinsic phenomenal character (as these theories hope to do) because the claim seems, in essence, to make no sense. I do not know how to meaningfully construe the claim that something could present as both subject-like (to do with the modus operandi of witnessing) and object-like (as a witnessed object that can be attended to in principle) unless the terms ‘subject’ and ‘object’ are altered to no longer reflect the intuitively fundamental subject/object distinction. Altering the terms, however, would seem to quite literally change the subject and hence sidestep the issue, which is not desirable, given that advocates of such theories seek to frame and explain the way that consciousness appears (and consciousness appears as intrinsically subject-like rather than object-like). It is far more intuitive to renounce the object-knowledge thesis by simply allowing witness-consciousness (viz., awareness) to have an intrinsic phenomenal character that is underived from any observable objects, including itself.

A recent philosophical position that could serve to challenge the object-knowledge thesis is that developed through the notion of “subsumptive phenomenal consciousness”, proposed by Bayne and Chalmers (2003). Their position seems at odds with the object-knowledge thesis. For when unpacked, it states that there is something it is like, at a time, to be in a single unified state of consciousness, and that this something it is like is not reducible to ‘atomistic’ phenomenal states associated with the (sensory or mental) qualia of objects. It
therefore seems to admit to the reality of subject-knowledge. I have argued that the notion of awareness developed in this thesis encompasses and is more informative than 'subsumptive phenomenal consciousness' since it also builds in elusiveness and (moment-to-moment) diachronic unity. I would suggest, furthermore, that 'awareness' is deserving of a central place in the philosophy of mind, not only because it captures a notion of consciousness common to these puzzles about subjectivity. Accepting the concept of awareness as central also goes some way towards solving the puzzles. We have seen that a persistent obstacle to the solution of the 'elusiveness' and 'unity' puzzles has been the assumed object-knowledge thesis, which does not take seriously the possibility of unmediated subject-knowledge, and hence awareness. Accepting the reality of awareness involves a rejection of this common assumption and hence the removal of a common obstacle, permitting one to freely acknowledge what is behind these puzzles in the first place. It also has a significant bearing on what is known as the "hard problem of consciousness", whose solution aims to explain the place of consciousness in the world. Tackling the hard problem must first involve defining what consciousness is and I hope to have argued convincingly that as a quintessentially subjective phenomenon consciousness is awareness. (The subjectivity of consciousness is what allegedly makes it a hard problem). I believe that once the object-knowledge thesis is rejected, and the reality of awareness accepted, better progress will be made in dealing with the hard problem of consciousness. For any explanation of consciousness qua awareness will simultaneously impact upon the phenomena of elusiveness and unity (both synchronic and diachronic), thus accounting for several puzzling features at once.
Chapter Eight: How the Self could be a Construct.

Introduction

It has so far been argued that a subset of features attributed to the self – the self that most of us pretheoretically assume ourselves to be – is not constructed or illusory. I have thus argued (in chapter six) for a concept of awareness which is to be specified in terms of this set of features and I have argued that the features in question – (each purporting to be unconstructed) namely; unity, elusiveness, endurance and invariability – are to be specified in terms of awareness, viz., a witnessing, subjective sense of presence (which also purports to be unconstructed). Awareness and its features are thus co-specified in terms of each other. This co-specification notably does not extend to the concept of boundedness. I have furthermore argued (in chapter seven) that the co-specified awareness does not merely purport to be unconstructed (such that if it exists, it must be unconstructed in nature) but that it really is unconstructed in its nature, as it seems to be. In virtue of its unconstructed status, the co-specified awareness can be deemed a non-illusory facet of the mind, contributing its non-illusory features to the sense of self.

Now it has been established in chapter five that the more complex entity, the self, is held by Buddhism and some leading Western thinkers, to be constructed and illusory. Given that the self cannot be constructed or illusory simply in virtue of the co-specified awareness, then the self, if constructed, probably owes its constructed (and hence illusory) status to the feature of boundedness, as Buddhism would proclaim\(^\text{114}\). Boundedness is that aspect by which a subject’s witnessing presence – or we can now say, awareness – is identical to an ontologically unique subject with personalised boundaries that delineate a me; viz., a self, from the rest of the world. It was argued in chapter four that our sense of being a bounded entity, viz., the subject’s sense of being identical to awareness-as-bounded, is most widely borne out through the subject’s identifying as a ‘personal owner’ – a role which must be

\(^{114}\) The urge to seek happiness and avoid suffering – evidenced in the section on ‘consistent self-concern’ in chapter four – is one amongst several sources of evidence for the sense of self-other boundedness (and hence a sense of self). Given this, the ontological status of self (whether constructed or unconstructed) will be more immediately connected to that of self-other boundedness than to factors which provide evidence for the sense of such boundedness. Later in this chapter, it will be suggested, in fact, that consistent self-concern helps to actively construct the sense of self-other boundedness (without any need for positing a self) – hence this aspect commonly attributed to the self will not be ignored in our discussion.
assumed before one can identify as a ‘thinker of thoughts’ or ‘agent of actions’. The role ‘personal owner’ was demonstrated to be the broadest mode of self-identification, such that there is a self if and only if there is personal owner. Should awareness-as-bounded turn out to be a genuine, non-illusory entity, the role personal owner – and hence the self, would also be non-illusory. Awareness-as-bounded, equivalent to the bounded self that is a personal owner, would have the unconstructed reality that it purports to have. Should awareness-as-bounded turn out to be illusory, however, then the bounded self that is a personal owner would not have the unconstructed reality that it purports to have: it would be constructed. Our sense of being a bounded, personal owner, viz., our sense of self, would have to be explained by phenomena that do not appeal to the self in their ontology.

We have already demonstrated that the concept of boundedness does not, unlike the concepts of unity, elusiveness and so forth, follow from the concept of awareness, since it is possible to conceive of awareness without boundedness. This suffices to show that, on a prima facie level, awareness is not as tightly associated with boundedness as it is with elusiveness, unity and so forth. That fact that we can conceive of awareness without boundedness, however, does not show that awareness is not, as a matter of psychological fact, intrinsically bounded through the role of ‘personal owner’. Conceivability does not imply actuality. Perhaps awareness must in fact occur as part and parcel of an unconstructed, bounded, personal owner, viz., as a self – a fact that is not obvious through merely analysing the concept of awareness. Perhaps, in other words, the bounded self (equivalent to awareness-as-bounded) is unconstructed as it purports to be, and hence not illusory. In this chapter we investigate evidence to the contrary, which suggests, convincingly, that the bounded-self-as-unconstructed, in virtue of its boundedness, is in fact not unconstructed – as it purports to be – but constructed. That is to say (in view of our analysis in chapter five) we shall be suggesting that the self, which purports to be the content of an appearance or presentation that is not contributed to by objects of consciousness, is in reality, the content of an appearance or presentation that is contributed to by objects of consciousness. In other words we shall be arguing that the self (or awareness-as-bounded), primarily in virtue of its boundedness, is constructed and hence, illusory.

In more detail, this chapter is structured as follows. We begin by re-iterating how each (established) intrinsic feature of awareness partakes in the witnessing’s ‘dressing up’ as
a specific identity and hence, a self. This will not only serve as a useful reminder of the arguments in chapter four – outlining the kind of entity we take ourselves to be – but will also serve as a description of how these features of awareness, discussed in isolation in chapters six and seven, seem personally bound up in the overall impression of a self.

We then turn to our argument that awareness-as-bounded viz., the self or personal owner, is constructed – an argument that will proceed in several stages. The first stage aims to rule out a scenario which would serve as strong evidence that awareness-as-bounded, viz., a self, is unconstructed. The scenario we aim to rule out is one where awareness always co-occurs with a sense of self-identification and personal ownership; a scenario that would strongly suggest that self-identification (hence the role ‘personal owner’) is psychologically intrinsic to awareness. In this section, then, we consider a case from neuropathological literature which I believe suggests that a subject’s awareness may globally occur without an accompanying sense of personal ownership (and the discussion is preceded by considering cases that are sometimes misdescribed as cases where someone lacks a sense of self).

Having ruled out the scenario which might be taken as strong evidence that awareness-as-bounded is unconstructed, we then consider a possible picture of how a sense of boundedness could psychologically arise in such a manner that is consistent with that boundedness being constructed. In doing this, we re-introduce the Buddhist notion of tanha viz., emotional investment, in an attempt to show how the psychological presence of tanha and its related thoughts could plausibly help construct our sense of self/other boundedness. Initial support for this picture can be found in the fact, reported by Damasio (1999), that all signs of emotion (implicated in what he calls “individual concern” (1999, 304)) appear to be lacking in those patients whose sense of self is suspended whilst still being aware. The way Damasio talks about individual concern suggests strongly that what he is referring to with this phrase is what Buddhism refers to as tanha.

The co-occurrence of tanha and the sense of personal boundedness, however, is not by itself enough to demonstrate that the bounded self, viz., personal owner, is contributed to by thoughts and feelings pertaining to tanha. Just as the self does not purport to be conscious the whole time (for example during deep sleep), it could likewise be the case that, under various conditions, the self’s capacities for its reflexive representation as a personal owner become suspended. Nothing about the concept of self rules out the possibility that emotional
investment or *tanha* may simply amount to this potentially fluctuating capacity. The next stage of the argument for the self’s constructed status, then, aims to show that this picture of an unconstructed self, as the source of our self-sense, is a relatively implausible picture. The argument essentially appeals to the principle of parsimony: it seems that the relevant data of felt boundedness can be explained in a manner that does not appeal to the more contentious metaphysical entity of *self*; rather it appeals to an account (by Damasio, 1999) that utilises scientifically credible ideas and evidence. On Damasio’s account, the sense of personal boundedness historically arises from adaptive pressure on evolving organisms to maintain their homeostatic boundaries. He contends that the sense of self is a conscious analogue of the organism’s homeostatic boundaries whose current status is mapped in specific areas of the brain. The conscious analogue of self exists as the content of (neurobiologically correlated) thoughts and emotions (and their feelings) whose steady proliferation is psychologically driven by boundary-preserving emotional investment (what Damasio calls “individual concern” (1999, 304)). It will be clear that Damasio’s account of the apparent boundedness of self, which makes no appeal to an unconstructed self, gives independent credibility to the Buddhist account. If the arguments so far have worked, then we will have provided good reason to suppose that the boundedness of self is constructed and hence illusory.

The principle of parsimony, of course, does not dictate that we give preference to a favoured scientific theory if that theory fails to explain all the relevant facts. The shortfall in Damasio’s account lies in its failure to adequately explain such facts as the appearance of unity, invariability, endurance, witnessing and elusiveness. As I argued in chapter seven, these features must be unconstructed through the fact that they are integral, by conceptual necessity, to the demonstrably unconstructed awareness (viz., witnessing presence). Damasio’s attempt to explain them as constructs will thus fail. As a prelude to the account of self-as-construct that will be offered in chapter nine, I show how the notion of identification, reintroduced from chapter four, dovetails nicely into Damasio’s account of personal boundedness as constructed. The process of identification essentially allows the natural features of awareness – unity, elusiveness and so forth – to become apparently integrated into the thought and emotional content whose ‘storyline’ (through appropriated ideas) seems to
presuppose a bounded self. This provides the basis from which we will present our model of the 'two-tiered illusion of self' in chapter nine.

**Recapping Evidence that Awareness Purports to be a Bounded Self**

Boundedness pertains to that factor by virtue of which a subject's witnessing presence comes 'dressed up to the party of life' as a specific identity with boundaries – as this *very* self who is ontologically unique, separate from the rest of the world. We have established that the most pervasively *assumed* identity-role to evidence a sense of boundedness is the role 'personal owner' (with its reciprocal feelings of 'me' and 'mine'). Let us then review some evidence that this suggestion is indeed accurate.

Each co-defining feature of the witnessing presence seems to partake of this 'dressing-up', such that the features seem to qualify not merely the bare, impersonal witnessing presence, but rather the bounded personal self as a whole. In light of chapters six and seven, which provided some detailed discussion of the features of awareness, it is worth at this stage being reminded of how it seems as if it is a bounded self-entity that displays these features – thus reiterating what was argued for in chapter four. It is not, hence, merely bare perspectival awareness (viz., a perpectival owner), but a personalised bounded self as a whole that feels like a witnessing presence with a unique first-person perspective. The perspective of the witnessing presence thus appears to be the *self*'s perspective. Moreover, the personalised self, not just bare awareness, seems to *be* that unified entity to which at a given conscious moment one's thoughts and perceptions seem unified. Even when thinking about aspects of my personality, it is always, as Velleman (2002, 114) puts it, *with* those parts of the personality being currently identified with as 'me'. Those very parts of the personality that I'm currently thinking with never seem to appear as separate, observable objects of consciousness, but always appear *at one* with the witnessing presence who is observing or thinking. The observing *self* would thus appear to also inherit the observational elusiveness of witnessing presence, as well as its unity. It hence seems as if a personal, bounded, unified, elusive witnessing *self* – not a mere impersonal, unified, elusive, witnessing presence – is confronting the world with its objects as 'other'.

Nor, I suggested, does this elusive, unified witnessing self seem (from the viewpoint of the subject) to pop in and out of existence over a given stretch of conscious time. From
moment to moment, it reflexively seems to the subject as if it is the \textit{self}, not merely impersonal awareness, who is continuous rather than gappy, enduring in the manner that witnessing endures. It also seems to the subject as if it is a \textit{self} – not merely an impersonal witnessing presence – who partakes in longer-term numerical identity such as when waking up from deep sleep, or when remembering an earlier stage of one’s life. And despite the obvious qualitative changes in personality over time, it still appears to the subject as if that qualitatively invariable thread running through one’s life pertains not merely to an impersonal witnessing presence, but to a personalised ‘me’ (of which there could only be one). There hence reflexively seems, to the subject, to be an elusive, hard-to-define yet invariable quality of \textit{me-ness} that unifies all its thoughts and perceptions as belonging to ‘me’, not only at one time, but over a lifetime. It would thus seem as if a bounded self takes on the invariability that is native to the witnessing presence. Finally, nothing about the bounded self, as reflected upon from the first-person perspective, would seem to be overtly constructed by thoughts, emotions or perceptions that are harboured at a given time. This apparent unconstructedness of the self is a direct upshot of its apparent elusiveness and unity. Thoughts (etc.) \textit{qua} thoughts will always appear to be owned or initiated by the independently existing unified self, and hence always present to the first-person perspective as separate \textit{from} the self – as objects rather than a subject of awareness.

\textbf{Is the Bounded Self a Construct?}

\textbf{Can Awareness Exist without Sense of Bounded Self?}

If awareness must, as a matter of psychological necessity, co-occur with a \textit{sense} of bounded self (viz, a sense of self-identification as personal owner) – a suggestion which \textit{seems} to be implied by much of the last section – then this will provide powerful reason to suppose that awareness must co-occur with an actual self. Awareness, unavoidably presented as a bounded self, will effectively \textit{be} the bounded self, with the unconstructed status of awareness being conferred to the self as a whole. In this section we aim to rule out the possibility that awareness must always occur with a sense of self, by seeking an actual example, from the neuropsychological literature, of awareness \textit{sans} sense of self. We will see why typical cases of depersonalisation, sometimes offered as examples of where the sense of self is missing, do not serve as good examples – but why epileptic automatism, in contrast, \textit{is} a good example.
While these findings will not by themselves establish that awareness actually does lack the boundedness of a self (which would, remember, imply the self is a construct) they will be an important first step. If awareness need not co-occur with a sense of self, there is initial reason to suppose that it need not occur with a self, either.

The examples of two different psychopathologies, anosognosia and depersonalisation, were used in chapter four to illustrate the difference between perspectival and personal ownership. While anosognosia involves the subject’s ignorance of a specific disorder such as paralysis (such that one may deny feeling personal ownership towards one’s paralysed arm) depersonalisation seems to affect the subject in a more global way. According to the DSM-IV Manual (2004, online)\textsuperscript{115} depersonalisation involves the “persistent or recurrent experiences of feeling detached from, and as if one is an outside observer of, one's mental processes or body”. Another psychiatric manual (2004, online) adds, “This disorder is characterised by feelings of unreality, that your body does not belong to you, or that you are constantly in a dreamlike state”\textsuperscript{116}. From this, it may be tempting to conclude that during typical episodes of depersonalisation, one’s sense of self is entirely absent, since it would appear from the above definitions that feelings of personal ownership and self-identification towards one’s body or mental processes will be suspended or absent. However, if the individual cases and further descriptions that I have read about typify depersonalisation, then while not denying that the sense of self may be somewhat compromised, there is reason to question whether a sense of self will in fact be missing during such episodes. For a sense of self to be present at time t, at least in the manner that the notion has been developed in this project, it is enough that there be some item – internal or external – towards which the subject harbours the reciprocal senses of ‘me’ (through identification as its personal owner) and ‘mine’ (through a sense of personal ownership towards the item). So long as there is any sense of personal ‘me’ or ‘mine’ with regard to any item, then the subject will harbour a sense of self per se, even if the subject should lack a sense of self-identification and/or personal ownership towards all other items that are, at that time t, within their conscious purview.

\textsuperscript{115} "BehaveNet - Clinical Capsule: DSM IV: Depersonalization Disorder" (2004)
http://www.behavenet.com/capsules/disorders/depersdis.htm

\textsuperscript{116} "AllPsych Online: The Virtual Psychology Classroom (Depersonalization Disorder)" (2004)
http://allpsych.com/disorders/dissociative/depersonalization.html
My suspicion, with regards to depersonalisation, is therefore as follows. The cases that I have read about all suggest that the person suffering an episode of depersonalisation, at a given time, feels a lack of personal ownership (my-ness) towards some and perhaps even most of their perspectively owned experiences. But there is also evidence that there are at least some such items towards which they, as subjects, harbour a sense of self-identification and/or personal ownership – which will guarantee, on the account developed in this project, that they also harbour a sense of self. The items in question invariably involve a pod of recurring negative emotions – even if blunted – that pertain to their realisation that they are in a depersonalised state: they realise there is something wrong and they wish the state and its attendant sensations would go away. This ‘wishing the state would go away’ is a clear instance of emotional investment or what Buddhism would refer to as tanha. The negative emotions arise because the person is in a situation that they wish were otherwise. In chapter four, emotional investment was enlisted amongst the strongest evidence that a subject identifies itself as a bounded personal owner (in this case, by identifying with the perspectival owner of the depersonalised state and the thoughts and emotions pertaining to it). Consider the following extract from a chronic sufferer of depersonalisation:

I feel as though my Self has "pulled back into my body," that my Self does not "fill me out to my skin." I literally feel that when touching something, I can feel it, yet feel it from a distance as I am removed from myself, or stuck too far inside myself.

Please note, at no time do I actually believe these feelings are real. I know I shouldn't feel the way I do; I want these horrible sensations to go away. In psychiatric

---

117 This is not to deny that those suffering from depersonalization may not on the whole experience a dulling of emotions. According to the Depersonalization Research Unit team, depersonalization may in fact often arise as a coping mechanism to avoid painful negative emotions arising from a variety of traumatic or aversive situations in childhood or adulthood. The resultant avoidance of emotional arousal has led to a globalised blunting of emotional response. Many sufferers of depersonalization state that a major goal of treatment for them is to regain the ability to experience emotions, although they may also express concerns about their ability to cope and fear being ‘overwhelmed.’ (2001, 132). What this suggests, then, is that although such patients experience emotions less often and with less intensity than usual, they are still concerned about recovering from their condition – and recovering in a such way that avoids trauma. Support for this is found in people’s own accounts of depersonalization on the website: http://www.depersonalization.info/overview.html. Such people commonly report feeling a lack of emotion in response to triggers normally expected to cause emotion (for example, a friend’s death), and yet they invariably speak as if they long for their overall state to be overcome, sometimes even attempting suicide to escape their predicament. As one person put it: “There is a lot of "pain" in being able to feel nothing. Most people will never understand that, but they don't know what it's like to not exist. They haven't died and been forced to live through years of nightmare” (from the above website: “personal stories”).
ter-minology my "reality testing is intact," I have "insight" into what's happening to me and I know it shouldn't be happening. I am not delusional or psychotic in any way. (2004)\textsuperscript{118}

From this passage, I would surmise that the author has suffered a shrinking but not dissolution to the felt boundaries of self. She does not fully identify with the normal bodily sensations that are typically but tacitly assumed to be integrated with the self. Yet it would seem that she still identifies clearly with the perspectival owner of her "horrible sensations" since she wishes the feelings to "go away". The fact that she feels upset suggests that she harbours a sense of my-ness towards the sensations, which in turn indicates that she identifies herself (qua subject) as a 'me', viz., as the bounded personal owner of the sensations. I would thus claim that whenever emotions such as chronic anxiety, depression or fear accompany an episode of depersonalisation – a fact that is often reported in online discussions of the syndrome – we can reasonably infer that the depersonalised subject is identifying itself with the perspectival owner of the emotions and hence as the personal owner of those negative emotions and hence, as a self. The author of the passage above also appears to identify with – and value – the perspective who has the insight into what is happening: it is this which prevents her from being delusional about her condition. In fact, according to Los Angeles psychiatrist Oscar Janigar, "reality testing remaining intact" is partially constitutive of the disorder's DSM-IV definition (2004).\textsuperscript{119} The sense of terrible isolation often reported with depersonalisation is further testimony to the fact that a sense of bounded self is still there – in fact with boundaries tightened to such an abnormal degree that the subject feels as if there is an impenetrable barrier between their shrunken self and the world. As another patient puts it: "It’s like I fall deep within myself. I look at my mind from within and feel both trapped and puzzled about the strangeness of my existence." (op cit).

There are more extreme forms of depersonalisation disorder – the most extreme being Cotard’s Syndrome.\textsuperscript{120} This complex condition cannot be summarised in a paragraph, but the primary symptom is the delusion that one is dead, implying an absence of 'intact reality testing'. While Tim Bayne (2004, 232-233) holds that cases of Cotard’s Syndrome are

\textsuperscript{118} From A.E. Gale: “Dreamchild: Living with Chronic Depersonalization” at http://www.dreamchild.net/
\textsuperscript{119} “Strangers to our Selves" (2004) http://www.depersonalization.info/overview.html
\textsuperscript{120} For a website that has links to reputable articles on Cotard's syndrome, see (2004) http://groups.msn.com/TheAutismHomePage/cotardlinks.msnw
evidence that a sense of personal ownership or ‘my-ness’ towards one’s thoughts is missing. I am sceptical that this is the case. What makes me suspect that there is not a lack of identification as the personal owner of some X, is, once again, the seeming frequency with which such cases co-occur with negative emotions or delusions such as those of immortality. According to the autism homepage website:

Berrios & Luque (March 1995) performed a statistical analysis on 100 cases of Cotard’s syndrome. Cotard Syndrome was found to affect men and women in equal numbers and severity. The elderly were more likely to develop the syndrome. Depression was reported in 89% of subjects, anxiety in 65% of subjects, and guilt in 63% of the subjects. Delusions were categorized thusly: nihilistic delusions concerning the body (86%), nonexistence (69%), hypochondriacal delusions (58%), and delusions of immortality (55%).

If this quotation is anything to go by, I would suggest that the typical Cotard’s patient does not cease to identify with things, but has highly abnormal patterns of identification. That is, he does not identify with a living body, as most of us do, but identifies with a dead body, appropriating an idea ‘owner of the dead body’ to his current witnessing perspective. Thus he believes that he (qua self) is literally dead, sometimes to the point where he can ‘feel’ the worms of decomposition. The negative emotions are simply in keeping with the assumed identity: believing that one is dead and rotting is not likely to be a pleasant experience (and usually prolonged negative emotions precede Cotard’s Syndrome). The high co-occurrence of negative emotions (even if blunted) is, I suggest, evidence that the patient identifies as their personal owner, showing little doubt that he believes it is he who is dead. On the basis of this, I would surmise that as with the less extreme forms of depersonalisation, an intact sense of self is typically present in those who suffer from Cotard’s Syndrome.

There are, however, pathologies that do seem to provide far clearer evidence that the sense of bounded self (and hence personal ownership) can be properly lost, if just temporarily. During such episodes, there is no evidence of ‘intact reality testing’ and no evidence that the subject is emotionally upset about what has befallen him. One classic example of such an episode is known as an epileptic automatism, which can be brought on by a brain seizure. Antonio Damasio (1999) has observed and described in some detail the

behaviour of patients who have undergone such episodes. The patient may be having a perfectly normal conversation and then suddenly without warning,

freeze whatever other movement he was performing, and stare blankly, his eyes focused on nothing, his face devoid of any expression – a meaningless mask. The patient would remain awake...This state of suspended animation might last for as little as three seconds. ...The longer it lasts, the more likely it...will be followed by absence automatism, which, once again, can take a few seconds or many. ....As the patient unfreezes he looks about...his face remains a blank...he drinks from the glass on the table, smacks his lips, fumbles with his clothes, gets up, turns around, moves towards the door, opens it...then walks down the hallway. ...the automatism episode would come to an end and the patient would look bewildered, wherever he would be at that moment. ...would have no recollection whatsoever of the intervening time [during the episode]. (199, 96-97).

According to Damasio, the patient during the episode,

would have remained awake and attentive enough to process the object that came next into his perceptual purview, but inasmuch as we can deduce from the situation, that is all that would go on in the mind. There would have been no plan, no forethought, no sense of an individual organism wishing, wanting, considering, believing. There would have been no sense of self, no identifiable person with a past and an anticipated future – specifically, no core self and no autobiographical self. (1999, 98).

We can note that Damasio’s notion of “core self” corresponds closely with what I have been calling the ‘sense of self’, viz., the basic sense of being an elusive, bounded, individual owner, thinker and actor. The ownership that concerns Damasio is clearly personal, involving a distinct sense of ‘me’ and ‘mine’.122 The “autobiographical self” is simply an extension of the core self, involving a sense of one’s long-term extended personal history with the ability to imagine one’s existence into the anticipated future. We may recall from chapter four that another of Damasio’s patients, ‘David’, with his 45-second memory, had an impaired autobiographical self but intact core self (1999, 117-121). The patient undergoing an episode of epileptic automatism, however, has not even this much window for reflection – he merely

122 Other evidence for this is that early in his book, Damasio writes: “In parallel with representing the printed words and displaying the conceptual knowledge required to understand what I wrote, your mind also displays something else, something sufficient to indicate, moment by moment, that you rather than anyone else are doing the reading and understanding of the text. ...[This presence signifies] you, as observer of the things imaged, owner of the things imaged, potential actor on the things imaged” (1999, 10). On page 127, there is also clear indication that Damasio is concerned with personal ownership.
acts “within the microcontext of the moment” (1999, 98). Nothing would seem to be reflected upon or committed to memory – or at least memory that he can normally access. There is hence no evidence that he identifies with any aspect of his experience as ‘me’, and so no evidence that he harbours a sense of personal ownership or ‘my-ness’ towards his experience. In addition to this, all signs of emotion, Damasio notes, are entirely absent – and he holds that emotions are reliable indicators for consciousness with a sense of self (1999, 99-100). I think it is reasonable to infer, as Damasio does, that during such an episode, the sense of a bounded individual self is absent. I think that is also reasonable to infer, from the fact that the patient is awake and minimally responsive to his environment, that awareness is present.

There are other pathologies mentioned by Damasio, which also point to an absence of he terms a “core self” (equivalent to a sense of self) with the presence of awareness: akinetic mutism and advanced cases of Alzheimer’s disease (1999, 101-106). To describe these in detail will not add much to the discussion, but to mention them in passing is useful, since it adds weight to the empirical evidence that a person’s awareness can psychologically exist without a reflexive sense of a bounded self. Awareness sans sense of self is not, it would seem, a mere conceptual possibility. Such cases make it more plausible to suppose that newborns and an array of primitive organisms will also harbour awareness without a sense of bounded self.

Damasio’s findings would appear to indicate, then, that a person can harbour awareness without reflexively sensing this awareness as being identical to a bounded entity, viz., a self. This does not imply that awareness is not in fact an intrinsically bounded self, with its innate quality of boundedness (enabled through identification) simply obscured or disabled on occasion by certain cognitive malfunction (just as, on the Buddhist account, the intrinsic happiness of nibbanic consciousness is generally obscured by tanha and the sense of self). However, this discussion does seem to rule out what would be strong evidence for an unconstructed-self scenario, namely, that of awareness always transparently exhibiting a sense of boundedness. In the next section we consider the further details of how the sense of bounded self might plausibly be contributed to by cognitive factors that would render it a mental construct.
The Next Steps to Arguing that the Self is a Construct

We have established a first vital step in arguing that the bounded self is constructed: we argued this by ruling out the scenario that whenever there is awareness there must be a sense of self (a scenario that would powerfully suggest the bounded self inherits the unconstructed status of awareness). Cases of epileptic automatism seemed to best exemplify what would most plausibly be described as awareness without a sense of self. This scenario did not however make clear enough how cognitive input into the self - should the self be constructed - would pan out on a psychological level. How could the thoughts, emotions and so forth help create a sense of bounded self in lieu of an actual self? In the next two sections we advance the case for the self’s possible constructedness by (a) describing the Buddhist account of how emotions and attendant thoughts - subtending *tanha* - could help create a sense of bounded self and (b) outlining a theory of self by Damasio which both endorses the central role of felt emotion (in relation to *tanha*) and bypasses the need to posit an independent self in explaining how such a sense of bounded self comes about. In the section following this, we see how Damasio’s theory might be incorporated into the theory of identification, advanced from chapter four, to result in a basis for the two-tiered illusion of self (proposed in chapter nine).

How Emotions and *Tanha* might Plausibly Contribute to the Impression of a Bounded Self.

Any convincing account of self-as-construct must provide a plausible picture of how the sense of bounded self could arise through contribution from thought-processes, emotions and so forth, instead of from an actual unconstructed self – the entity from which it appears to arise. While providing such a picture will not alone prove the non-existence of an underlying self, it will weaken the need to posit a self in order to explain the sense of self. Providing such a picture should also clarify, on a psychological level, the potential role of various inputs into the sense of self, such as thought-patterns, emotions and related dispositions.

Now from our discussion so far, we have had some reason to suppose that emotion will play a significant role in the presentation of any sense of self and hence in any potential account of the self-as-construct. For example, in chapter four it was argued that the presence of most emotions (where there is apparent concern for a *self*) is enough to indicate that a
sense of bounded self exists (more on this soon). This argument was used to justify the current chapter’s suggestion that cases of depersonalisation are not good examples of awareness sans sense of self. And we reviewed evidence, in this current chapter, that a lack of emotion correlates with an absent sense of self, as exemplified in cases of epileptic automatism. The picture about to be presented will be one in which emotions, cognitively construed, are plausibly seen as both sufficient indicators and necessary contributors to the sense of a self-other boundary, such that an actual bounded self need not be brought in to explain the sense of bounded self. The picture is based upon my analysis of Buddhism.

In doing this, it will first be useful to review from chapter four our account of how the presence of most kinds of emotion in a given psychological state is sufficient to indicate a sense of self-other boundedness and hence, a sense of self. It is not hard to see how these emotions, through their cognitive content, could actively contribute to the sense of self-other boundedness. Having done this, we will then need to address the issue of how a sense of self will often seem present in the absence of obvious emotion. The picture must accordingly render the emotions subtle or ubiquitous enough to allow for that ongoing sense of bounded self that we experience. The Buddhist notion of tanha – what we have called ‘emotional investment’ – is relevant to this story.

In chapter four (in the section on ‘consistent self-concern’) we noted that the presence of most kinds of emotion indicate that a subject harbours a sense of self-other boundedness. It is important to understand that emotion is not in this context being treated as merely a bunch of affects in isolation, but as a cognitive state that incorporates those affects. As cognitive states, most emotions involve, as part of their ontology, a train of thoughts and perceptions with content usually depicting what is the emotion is directed at, lending emotions their characteristic ‘aboutness’ or intentionality (although some moods might lack this). In chapter four it was suggested that the content also reflexively depicts – or seems to depict – a self who is having the emotion (this is also generally true of moods). We noted that most emotions not only reflect a background desire to be happy or avoid suffering (a further feature of the self) but a background desire for one’s self to be happy or to avoid suffering. To feel guilty for example usually implies that one wishes that one had performed an alternative action – one with a happier guilt-free outcome for oneself. To feel nervous at the thought of one’s impending speech implies that one would like one’s own speech to be well
received and is concerned it may turn out badly for oneself, as opposed to any old person. One could not feel nervous, moreover, unless one’s current witnessing perspective identified with the future owner of the speech as a self, believing it is the very same self as this current one who will bear the consequences. Importantly, feeling nervous implies a felt separation between one’s identified-as self and the envisaged possible outcome. How could one feel nervous without at least implicitly cognising the possibility of the ‘self’s’ alienation from success? Or guilty without believing the ‘self’ is regrettably not the owner of a more desired action? Or proud without cognising the lesser entity the ‘self’ might have been? All such emotion thus seems to point to, as a part of their content, an unspoken, felt boundary between the identified-as self on one hand and the desired or undesired scenario on the other, as it is perceived or imagined by the witnessing subject. From this picture it is not hard to see, objectively at least, how such emotions might help construct the bounded self, such that the entity we assume we are exists as part of the content of the ‘self-concerned emotions.

However, while such obviously felt emotions as nervousness, guilt and pride clearly exhibit and possibly contribute to the sense of a self-other boundary, it is also true that throughout a given day, the attention of an average person is not likely to be grabbed all the time, or even most of the time, by the presence of such emotions. Many portions of waking life tend to be rather emotionally neutral or fleeting, with a background mood whose tone might vary from motivated to bored. During such relatively neutral episodes, we can grant that the sense of ourselves in relation to perceived objects will not be as obviously felt in its boundedness, as it will be in the presence of strong emotions. Nevertheless the sense of self does not seem to disappear during such episodes of relative neutrality. If it did disappear, then we would expect the experience of perception-whilst-feeling-emotionally-neutral to significantly differ from that of perception-while-feeling-emotionally-charged – perhaps we should feel regularly depersonalised! – yet the situation does not present as being this way. Objects of perception seem presented to one and the same underlying self, emotionally charged or not. They all seem to be equally my perceptions. From my first-person perspective it seems to be exactly the same ‘me’ looking neutrally now at a blade of grass, as the ‘me’ who felt anxious earlier today. Now if emotion is supposed to be necessary for maintaining a sense of self, how then do we explain such tracts of time where there seems to be an ongoing sense of self without obviously felt emotion?
It is here that the Buddhist notion of *tanha* becomes relevant. I have been talking about most emotions as exemplifying a sense of boundedness and we may be wondering what common factor unites the 'bounded-self emotions'. We have noted that all such emotions involve the implicit depiction of a bounded self as their bearer, together with a background desire to avoid suffering or be happy. The general psychological drive that is implied by this combination – the background desire for *one’s alleged self* to be happy or avoid suffering – is what the Buddhism calls ‘*tanha*’, which I earlier suggested is best translated as ‘emotional investment’. To repeat the definition from chapter one: *tanha* conveys an emotional attachment to objects being a particular way, such that (a) the assumed self’s happiness (emotional well-being) or absence of mental *dukkha* (mental dissatisfaction or suffering) seems to depend upon a desired configuration of objects (whether imagined or actual) such that (b) if they do not configure the way one would like, there is *mental dukkha* of a particular emotional nuance (for example, that of anger, sadness, anxiety, disappointment or guilt). *Tanha* is hence a mental disposition to want the world to conform to one’s desires – what one takes to be one’s own desires – whether the desires are bodily or mental in their origin. It is a disposition to emotionally invest in the satisfaction of a desire; such that one’s emotional state is affected by whether the desire is fulfilled. When the desire is not fulfilled, there is mental suffering or dissatisfaction of a particular emotional nuance, such as mild annoyance, disappointment or anger.\(^\text{123}\) Conversely, when the desire is fulfilled, there can be such positive emotions as achievement, relief and joy.

The majority of human emotions will involve some dependency of mental well-being or dissatisfaction upon circumstance (i.e., whether the circumstance conforms to one’s desire), which implies a mindset with *tanha*. But a mindset with *tanha* does not imply the obvious exhibition of emotion. If at any waking moment one’s emotional state, be this positive, negative or neutral, is dependent upon changeable circumstance, then whether one knows this or not, there is *tanha* present, according to Buddhism. We can for instance know whether there is *tanha* in our mindset if we sincerely answer “yes” to the question: “at a given waking moment, would any imaginable circumstance elicit happiness or mental suffering (given that I remain alert to what is happening)?” My suspicion is that most people who can understand this question will answer “yes”.

\(^\text{123}\) See chapter one for a discussion of *dukkha*. 
We might now wonder how tanha could actively maintain the sense of a self-other boundary in the absence of obviously felt positive or negative emotion. What causal effect could the very presence of tanha have upon our psyche and behavior such that it propels an urge to seek one’s ‘own’ happiness or avoid one’s ‘own’ suffering? How could this sense of personal ‘own’ arise when strong emotion is absent? Buddhism holds that tanha serves to exert an ongoing influence on one’s thoughts and behavior in such a way that the presence of emotion need not always be obvious. Tanha has sometimes been likened to a psychological itch that is never relieved for long. Perhaps a way to become more explicitly aware of this ‘itch’ is suddenly stop in one’s tracks, physically ‘freeze’ and do nothing but observe one’s mental states. It could be that within seconds, the itch of tanha will make itself known, demanding to be ‘scratched’. There may for example be the urge to finish one’s task, with the thoughts “I won’t be happy till this task is done!” and “I don’t like being still, I’m getting uncomfortable, I want to change position!” If stopped for long enough, the mental suffering, with stronger emotions (and sense of an ‘I’) may really start to mount. It is also quite possible, on this picture, that ‘scratching the itch’ by acting on tanha as we go about our business may be done before ‘the itch’ reaches our full attention with full-blown emotion. It is quite possible that ‘the itch’ and its ‘scratching’ will often not reach our full attention at all, with mere thought of such emotions – involving in their content what may happen ‘to me’ if ‘the itch’ is not ‘scratched’ – guiding behavior on an inattentive level. This whole process may, in other words, often occur on the cognitive level of what Evans has called “unprojected consciousness”. It is in this way that the Buddhist notion of tanha – the urge for ‘oneself’ to be happy and avoid suffering – can conceivably drive our usual thoughts and behavior in a subtler manner.

In influencing our behavior on this subtler level (in the background of unattended-to awareness), we can see how tanha, operating with the mere threat of unpleasant emotion or anticipated reward of pleasant emotion could endorse the sense of self-other boundedness. There will be no need to actually experience the full-blown emotions; the very thought of having them will be enough to effect action. And the very thought of having them – whether this thought is attended to or not – will activate the sense of self-other boundedness since it is after all ‘my’ itch that demands to be scratched. In this way, tanha, the ongoing desire for ‘oneself’ to be happy or avoid suffering, will automatically involve a sense of boundedness;
the identified-as self on the one hand, and the perceived or imagined desired or undesired scenario, on the other. Take the changing of physical position; something one generally does with minimal cognizance. This action may well be motivated at least partly by tanha, if the unfulfilment of the physical desire to change position were to result in mental suffering (as it usually would). The mental suffering, were it to manifest itself, would involve cognitive content that implies a desire for the ‘self’s’ separation from the physical pain (e.g., the thought “I wish it were otherwise”), implying a sense of self-other boundedness. Given this, it may well be the case that with most people, the urge to change position will not only be motivated by the instinctive desire to avoid physical pain, but by the more cognitively loaded desire to avoid discomfort to oneself. The disposition to feel mental suffering at not changing position will be integral to this desire, and thus be causally implicated in the actions taken to change position. Part and parcel of this cognitively loaded desire will be a sense of self-other boundedness, given that it is ‘my’ suffering that is sought to be avoided (where there’s a sense of ‘mine’ there must be a sense of ‘me’). The mental discomfort almost unconsciously avoided in something as simple as changing position, may thus involve a presentation of self-other boundedness that is likely to completely escape our notice.

Should this Buddhist analysis be correct – and it seems on the face of it plausible enough – then it seems reasonable to suppose that during our waking lives, a sense of self-other boundary will be almost continually being activated as we seek – attentively and inattentively – to bridge a gap between our current situation of relative or potential discomfort, and the one we imagine will make us happier. As we go about our life, there will be an ongoing sense of ourselves as separate, bounded entities, in perceived or imagined relation to the desired or undesired situation. The cognitive medium for such perceivings and imaginings will be the emotions that imply tanha, including those that are more subtle or imagined (which is most of the emotions). According to Buddhism, this ongoing tanha, viz., emotional investment, will not only indicate but also help create and drive the sense of self-other boundedness. The sense of self-other boundedness, in turn, will fuel tanha. Unless I identify as a bounded ‘me’, then how could I care most about whether ‘my’ desires are fulfilled? Tanha and the sense of self-other boundedness are hence plausibly co-dependent, on the Buddhist position.
Now we might ask how this account would handle one's interaction with parts of the environment that do not elicit any emotion – perceived or imagined. For example I could not care less about whether this blade of grass were to exist or not; yet I still identify myself as something decisively separate from the blade of grass. It is reasonable to suppose that if the Buddhist theory is correct, then the ongoing mindset of tanha, involving a constant (although not exclusive) look-out for 'number one', will be enough to create a general perceptual tropism, such that we involuntarily view even emotionally neutral items in our perceptual purview as decisively ‘other’ from the ‘self’ who perceives them. Tanha (viz., emotional investment) will thus, on the Buddhist position, fuel the subject’s overall disposition to identify itself with the perspectival owner of its perceptions, such that the perceptions are felt as personally ‘belonging’ to the subject. In this way, then, we can see how the sense of self-other boundedness could be maintained without any need to posit an independent self.

On the Buddhist account, then, emotions, whether actual or imagined, attended to or not, are both sufficient and necessary indicators of the sense of a self-other boundary. They serve, in addition, to actively create that sense of boundary (although they cannot do this in the absence of awareness). On this account, we would infer that cases of epileptic automatism lack not only evidence of obvious emotion (for example anxiety), but also evidence of emotional investment viz., tanha. So what we previously took to be separate evidence for an absent sense of self – for example a lack of future-oriented behavior – would, on the Buddhist account, double as evidence for a lack of self-concern; self-concern being the hallmark of tanha. There would in other words, be no evidence in such cases of a background desire that one’s ‘self’ be happy or avoid suffering, and hence no evidence that such persons identify as the personal owner of their past, present or future experience. This would amount to positive evidence that a sense of bounded self is, in such cases, absent.

\[124\] Perhaps in a similar fashion, a person with a phobia can develop a perceptual tropism, such that all items are perceived with an implicit 'filter' as to whether it is the feared object or not.

\[125\] The ultimate goal of Buddhist practice is to eliminate tanha and with it, the sense of bounded self that harbours even this much propensity to have thoughts of ‘me’ and ‘mine’.
Damasio’s Neurobiologically Informed, Parsimonious Theory of Boundedness as Constructed from Emotion, Feelings and Thoughts

We now consider some independent support for the Buddhist analysis presented above. First of all, if the Buddhist analysis were correct in supposing emotional investment to be necessary for a sense of self, then we would expect to find no cases where tanha is absent, but a sense of self present. No cases, that is, of a healthy sense of ‘me’ and ‘mine’ coupled with a lack of disposition to feel happiness or suffering in relation to any perceived or imagined states of affairs. As far as I know, no such case has been discovered. In support of this contention, Damasio has concluded from his findings on epileptic automatism and other such pathologies that “emotions [and feelings of them] and core consciousness tend to go together, in the literal sense, by being present together and absent together” (1999, 100). “Core consciousness” on Damasio’s theory basically involves awareness coupled with a basic sense of self (1999, 16) – although he never alludes to the ‘awareness’ component as a separate factor. However, it is still possible that tanha, viz., emotional investment, is necessary for the (non-object-like) reflexive representation of an independent, bounded self, such that a lack of tanha will simply mean that the self, while still existing, is not being reflexively represented. On this scenario, which may be favoured by those who believe in the existence of eternal souls, we would expect to find tanha co-existing with a sense of self. Cases such as epileptic automatism would simply mean that the usual faculties for letting the self be reflexively known to itself are suspended – just as they are suspended during deep sleep. In such instances the self would not cease to exist just because the sense of self is interrupted through a lack of wakefulness or emotional investment.

The time has finally come to dethrone this spectre of a realist theory of self. Damasio’s constructivist theory, as I mentioned in the introduction, is well-suited to the task. As with the Buddhist account, it proceeds by giving a convincing story of the apparent boundedness of self in a way that requires absolutely no reference to an unconstructed, bounded self, which Damasio sometimes terms a ‘homunculus’ (1999, 11). As with the Buddhist account, a mindset of tanha – which Damasio refers to as “an individual concern which permeates all aspects of thought processing, focuses all problem-solving activities, and inspires the ensuing solutions” (1999, 304) – is considered central to actively generating and maintaining the sense of self-other boundedness. In contrast to the Buddhist account,
however, Damasio buttresses the theory's psychological dimension with neurobiological support, usually from his own laboratory (he is a neurologist). From this perspective, he offers a persuasive theory of why a sense of self might have arisen in the organism's evolutionary history. Later we shall see how advantages of the Buddhist (and Advaita Vedantin) account in explaining other alleged attributes of the self (unity, elusiveness etc.,) can be used to strengthen what turn out to be weaknesses in Damasio's overall theory. But first, I shall outline that aspect of Damasio's theory which I think has great merit; specifically his account of how apparent self-other boundedness comes about. I shall do this in a manner that is detailed enough to convey the main points while leaving out most of the technical aspects.

The sense of self, on Damasio's theory, is basically a mental analogue of the organism's physical boundaries, which grows out of a biological need to maintain them. (This biological need, we shall soon see, is played out in emotions). Hence:

I have come to conclude that the organism, as represented inside its own brain, is a likely biological forerunner for what eventually becomes the elusive sense of self. The deep roots for the [sense of] self, including the elaborate [sense of] self which encompasses identity and personhood, are to be found in the ensemble of brain devices which continuously and nonconsciously maintain the body state within the narrow range and relative stability required for survival. (1999, 22).

This ensemble of brain devices, which Damasio calls the "proto-self", is nonconscious. It is the neural co-ordination centre for the organism's homeostatic regulation. The sense of bounded self arises when, for reasons of biological advantage, the proto-self is represented, which involves the organism becoming conscious not only of the world around it, but of itself as an elusive entity interacting with the world. The sense of the self's individual boundedness and stability has its origins in the genuine but relative stability of the represented proto-self as it maintains homeostatic regulation within the organism's physical boundaries (1999, 168-194). However, I would explicitly add that the conscious representation of the proto-self as a self is not exactly veridical, but involves a cognitive exaggeration of its boundedness and stability, as well as an additional sense of its being a conscious entity who is an owner, observer and agent. If it were not for this exaggeration and embellishment, then the self would effectively be the biological proto-self that is represented
rather than the mental construct Damasio claims it is. This point, which Damasio himself has not made clear enough, is important to bear in mind – it will be returned to.

The role of the emotions, on Damasio’s theory, becomes relevant when we consider biological advantages that may be procured through representation of the proto-self as a self. While Damasio refers to emotions by familiar names and categories (e.g., the six “primary” emotions, including happiness, anger, sadness; the social “secondary” emotions such as jealousy, embarrassment, pride; and the “background” emotions such as excitement, calmness, depression (1999, 51)) his use of the term ‘emotion’ is unconventional. For Damasio, ‘emotion’ does not denote the familiar first-personal feeling of these emotion-categories: for this he reserves the term ‘feelings’\(^{126}\). ‘Emotion’ denotes instead the “complicated collections of chemical and neural responses” that have a regulatory role to play in helping the organism to maintain its boundaries (1999, 51). They are part and parcel of homeostatic regulation and can be thought of as ‘sandwiched’ between the most basic regulatory mechanisms, such as reflexes, and the “devices of high reason” (1999, 54). The neurology of emotions will hence overlap with the neurology of the proto-self – a point that Damasio confirms (1999, 100)\(^{127}\). The most basic function of emotions is to assist the organism in maintaining its boundaries by avoiding sources of danger as well as pursuing sources of energy, sex or shelter (1999, 54). Now the effectiveness of emotions will be obviously improved, he suggests, if at least some of them become felt (from the first-person perspective) in a way that (a) connects them, through conditioning, to the fundamental motivating drives of pleasure and pain and (b) associates these feelings, through conditioning, with objects in the environment (1999, 54-55). In fact, I think that Damasio would hold that all emotions must at least have the capacity to be felt. Hence he writes:

As a result of powerful learning mechanisms such as conditioning, emotions of all shades help connect homeostatic regulation and survival “values” to numerous events and objects .... Emotions are inseparable from the idea of reward or punishment, of

---

\(^{126}\) “It is through feelings, which are inwardly directed and private, that emotions, which are outwardly directed and public, begin their impact on the mind...” (1999, 36). See also page 42.

\(^{127}\) This part of Damasio’s theory (as with many other parts) is not merely speculative but has empirical support. It explains why, when the sense of self is impaired, the felt and observable emotions are also impaired: the same neurological machinery underpins them (1999, 100).
pleasure or pain, of approach or withdrawal, of personal advantage and disadvantage...(1999, 54-55).

As the organism evolves to become still more complex, its physical boundaries will become increasingly vulnerable, requiring devices more ingenious than blind conditioning to protect them. This is where foresight – requiring consciousness with a sense of self – comes in. The capacity to plan ahead or represent past scenarios means the organism need not wait until danger is immanent or the energy source proximate, in order for it to act accordingly.

Foresight, reason and extended memory enable the organism to mentally review its options for action and response – and then choose what is most beneficial to it. The capacity of the organism to represent itself (viz., organism) as a past, present or future self, is, on Damasio’s theory, an adaptation that grows out of the organism’s need to foresee and plan its future in an optimal way. How could reviewing one’s future be viscerally motivating unless one thought it was one’s own future at stake? Now this process of mental reviewing and choosing, according to Damasio, is not an exercise of pure reason, but is very much grounded in the emotions which have, through conditioning and memory, become associated with certain options. I am for example thinking right now of whether to take a nice walk on the beach – or do my thesis. Fierce feelings of guilt (associated with pain) have immediately arisen alongside the thought “Do your thesis or chapter seven [now eight] will never be done by the end of this month!” Two predominant emotions felt as guilt – and the desire to get chapter seven finished – are what keep me glued to this seat right now (hence avoiding the pain and pursuing longer-term pleasure). The emotions that arise in connection with my imagined future are of course not felt to be anyone’s old emotion – they are felt to be mine. I after all seem to be imagining myself walking on the beach or finishing the chapter and bearing the consequences of each. So as well as allowing the organism to represent itself as a past or future entity, consciousness with a sense of self, Damasio holds, allows the organism’s key guiding emotions, in the form of feelings, to be immediately known to it (1999, 81). This enables the organism to feel more of an incentive to act on them – as it will need to do when imagining past or future scenarios. And the organism’s emotions (as feelings) cannot be known to it – at least in this extra-motivating manner – unless it represents itself as the elusive and personally concerned bearer of the emotional feelings;

\[128\] See for example, Damasio, 1999, 139-140, 284-285, 303-304.
hence, as a bounded self (199, 285, 302-304) So whenever emotions, or for that matter any other mental or perceptual states are felt to be ‘mine’, the proto-self (sharing neurology with emotions) is being represented as a ‘me’, viz., as a self who has – and can act in accordance with – these emotions and perceptions etc.

Now we saw that what compels the organism to be initially motivated by the feelings of emotion (and hence to protect its boundaries) are the conditioned associations of the emotions with pleasure and pain. This initial motivation (and effectiveness at boundary-preservation) is greatly enhanced by a feeling of overriding and asymmetrical concern that is bestowed by the sense of self (which Buddhism calls ‘tanha’). The ‘self’ feels, in the words of William James “the home of interest, - not the pleasant or the painful, not even pleasure or pain, as such, but that within us to which pleasure and pain, the pleasant and the painful, speak” (297, 1890). This overriding concern will ensure that the organism-subject (identifying as a self) is primarily motivated by its pleasure and pain; looking out for number one, so as to protect its boundaries – both psychological and biological. Hence Damasio writes:

I would say that consciousness, as currently designed, constrains the world of imagination to be first and foremost about the individual, about an individual organism, about the self in the broad sense of the term…Perhaps the secret behind the efficacy of consciousness is selfness. In short, the power of consciousness comes from the effective connection it establishes between the biological machinery of individual life-regulation [the proto-self] and the biological machinery of thought. That connection is the basis for the creation of an individual concern which permeates all aspects of thought processing, focuses all problem-solving activities, and inspires the ensuing solutions. Consciousness is valuable because it centres knowledge on the life of an individual organism. (1999, 304).

To summarise: on Damasio’s theory, the sense of self, integral to “core consciousness”, grows out of a biologically driven need for feelings of emotions to be known to the organism having them. Knowing emotions – which requires thought and feeling – enables the organism to plan effectively for its future, by envisaging itself as the past and future subject of positive or negative scenarios. The ongoing motivation to avoid the harmful and pursue the beneficial is made possible by an on-going individual concern that permeates the organism’s psyche, a concern which is felt through the psychological medium of these thoughts and feelings. Biologically, this concern is grounded in the innate drive for homeostatic boundary-
preservation; psychologically, this concern manifests as a sense of wanting the best for one’s self. On a psychological level, the concern effectively creates the sense of self-other boundedness and the sense of boundedness in turn, perpetuates the concern. As with the Buddhist account, *tanha* and a sense of the self’s boundaries are co-dependent.

Now at any waking moment, one’s impression of the self as owner, thinker and actor will, on Damasio’s theory, be generated by a stream of perceptions, thoughts and attendant emotions which represent not only the objects attended to, but the organism who is attending to them. It represents the organism via this impression of a self who engages with the objects. And just as Buddhism contends, the cognitively construed emotions (with thoughts and feelings) that subtend *tanha* and allow for a continuing sense of self are abundant, with many of them relatively subtle, imagined and unattended to. Damasio writes:

In effect, normal human behaviour exhibits a continuity of emotions induced by a continuity of thoughts. The contents of those thoughts, and they are usually parallel and simultaneous contents, include objects with which the organism is actually engaged or objects recalled from memory as well as feelings of the emotions that have just occurred. In turn, many of these “streams” of thought – of actual objects, of recalled objects, and of feelings – can induce emotions, from background to secondary, with or without our cognizance. The continuous exhibition of emotion derives from this overabundance of inducers, known and not known, simple and not so simple. (1999, 93).

It is as if, Damasio contends, the thoughts and felt emotions are both verbally and non-verbally telling one that they pertain to me, a self, and that I had better heed them to maximise my welfare. Damasio uses the term ‘images’ to convey the cognitive medium by which the thoughts and emotions are consciously known to one. (A more familiar word might be *qualia*). Hence the thoughts and emotions, through mental imagery, tell the organism a story of itself, except that in this story, the organism does not feature as merely a biological entity, but as a self engaging with the world (1999, 188-192). While the self is depicted in these images as the story’s protagonist, there is no real self (or homunculus, as Damasio puts it) who is telling the story. The self, as such, is an illusion. Hence:

You know you exist because the narrative exhibits you a protagonist in the act of knowing. (1999, 172) … The story contained in the images of core consciousness is not
[however] told by some clever homunculus. Nor is the story really told by you as a self because the core you is only born as the story is told, within the story itself. You exist as a mental being when primordial stories are being told, and only then... You are the music while the music lasts (1999, 191).

The constructed, hence illusory status of self is thus made plain on Damasio’s account although, perhaps because he has not defined the notion of an illusion, Damasio never says that the self is an illusion.

I hope I have provided enough detail to convey at least some of what is appealing about Damasio’s position (naturally I will not have done it full justice in these few paragraphs). In my opinion, his theory offers powerful reason to suppose that positing a thought-independent bounded self to explain our impression of the self’s boundedness is superfluous – both diachronically and synchronically. From an evolutionary (diachronic) perspective, Damasio provides a convincing (although in places speculative) account of how the sense of personal ownership and hence boundedness might arise from adaptational pressure on the organism to plan for its future with the aid of emotional feedback. He also provides convincing reason to suppose that the bounded self synchronically arises through the content of felt emotions and thoughts which ‘tell the story’ of such a self who is having them. (Evidence: knock out emotion and the sense of self disappears). The self-referencing thoughts and emotions, driven by concern for one’s ‘own’ welfare, are plausibly ubiquitous enough to allow for such a sense of bounded self to be sustained. Due to conditioning which associates emotional reaction with numerous objects, internal or external, the relevant triggers for emotion and thought will be profuse, including the self-referencing thoughts and emotions which will themselves become triggers.

In summary, Damasio’s theory gives independent credibility to the Buddhist proposal that the self is a construct whose existence and continuation as a personal, bounded entity psychologically depends upon tanha, that pervasive sense of concern for one’s ‘own’ welfare qua self. This tanha, both driving and borne out by a retinue of (often subtle, unattended to) felt emotion and thought, serves also to protect the organism’s physical boundaries (for example by making decisions to avoid immanent danger). In virtue of its boundedness, the self will therefore be a construct; the content of its appearance contributed to by objects of
consciousness, namely, the content of those thoughts and emotions that subtend the *tanha*. Now if the self is a construct then it must also be illusory. For the self, as we know, purports to be an unconstructed entity who *has* the thoughts, emotions and perceptions rather than something that is any way created by them.

**The Shortfalls of Damasio's Theory**

While Damasio has, I believe, provided a plausible theory of how the self comes to exhibit boundedness, there are gaps in his account of the self as a whole. Indeed, Damasio does not focus enough on how the representation of the proto-self as a self ends up exaggerated or embellished with *other features* attributed to the self. So while stable, relatively speaking, the proto-self is not enduring or invariable in the uncompromising manner that the self purports to be. While its neurons work together to serve the individual organism, the proto-self is not unified in the indestructible way that the self purports to be. Nor are we conscious of the proto-self — a point that Damasio freely admits (1999, 174-175). And if we are not conscious of the proto-self then the proto-self can hardly be elusive to its own observation (coupled with a sense of its own existence). Hence as a purportedly conscious entity, the self displays features that are either exaggerated or not belonging at all to the represented proto-self.

While Damasio does attempt to deal with elusiveness (1999, 172) and to some extent endurance (1999, 176) his account of how these features arise as integrated with the bounded self is unconvincing. If my arguments so far are correct, then his account of these features must fall short, since its very framework ignores explicit reference to awareness, the source of these aforementioned features. Being intrinsic to awareness, these features cannot be entirely constructed, as Damasio proclaims. When he talks about "the knower" or "consciousness", moreover, it is always in the context of awareness with a sense of self. His theory does not cater for awareness or witnessing simpliciter.\(^\text{129}\)

Given that there is good reason to suppose that Damasio's theory of the self's boundedness is basically correct, I shall aim to propose an account that will incorporate this part of Damasio's theory. The purpose is not so much to challenge as to extend an aspect of

---

\(^{129}\) I inferred that epileptic automatisms involve awareness *sans* sense of self, from the fact that Damasio reported wakefulness *sans* sense of self in such patients — and wakefulness implies awareness. Awareness and wakefulness are not synonymous, however, since there is awareness during dream-sleep (and possibly deep sleep).
Damasio’s theory in a way that I think will accommodate awareness and hence properly account for other features (mis)attributed to the self. Now it must straightaway be noted that this proposed account will not be at the neurobiologically informed level of Damasio’s; it will be proposed at a “higher” psychological level. There is a reason for being optimistic that the grafting will work: what will be proposed is based on my interpretation of what Buddhism but especially Advaita Vedanta has to offer to the story of self. Buddhist theory of apparent self-other boundedness, as we saw, found much support in Damasio.

The part of Damasio’s theory that I propose to extend, then, is the synchronic aspect, where the sense of bounded self arises from the ubiquitous stream of thoughts and emotions which seem, first-personally, to imply the existence of an individual, bounded self as their thought-independent bearer. I propose to extend the account by reintroducing the notion of identification that was proposed in chapter four and reviewed in the introduction. I will first recap and elaborate upon the notion of identification in light of our current conclusion that the self is a construct and illusion. I will then attempt to integrate this analysis with the relevant aspects of Damasio’s account.

Identification Revisited and Revised in Light of Self as Illusion

Identification, as we recall, separates ‘perspectival’ from alleged ‘personal’ ownership. For something, X, to be perspectivally owned by subject S is for X to appear to X from a first-person perspective (qua witnessing presence) that cannot be had, either contingently or necessarily, by another subject. ‘Subject S’ is effectively the perspectival owner of X; awareness as it appears from a body-bound sensory-mental, spatio-temporal perspective.

Note that there is nothing potentially misleading about the perspectival owner in itself; what is potentially misleading (from a Buddhist perspective) is the common assumption that awareness must, as a matter psychological necessity, occur from this perspective (an assumption notably imported, through identification, into our sense of being a personal owner). Hence the notion of perspectival ownership effectively builds in the contingent and undisputable fact of awareness normally appearing from the sensory-mental perspective of a body-mind.

For something perspectivally owned by S to also seem personally owned by S is, minimally, for S to identify with the perspectival owner of that thing, X. For S to identify
with the perspectival owner of X is for S to appropriate an idea (often a memory) of that perspectival owner of X to its current witnessing perspective, in a way such that the current witnessing perspective seems at one with this idea, taking itself to be that perspectival owner (for example an owner of one's thoughts). Whenever S's current witnessing perspective takes itself to be the perspectival owner of X, feelings of personal ownership will, on this theory, automatically arise towards X. The familiar feelings of personal ownership or my-ness will seem in turn, to wordlessly indicate not just a perspectival but a personal owner who has X—a self, in other words. In this way we can see how the minimally contingent bodily perspective of awareness—the perspectival owner—becomes folded into the sense of a personal self. Now in chapter four, we left open the possibility that those feelings of personal ownership actually do indicate a personal owner or self who has them. We hence left open the possibility that S's current witnessing perspective might actually be part and parcel of a self (that awareness might be bounded, in other words). On such a possibility, S's identification with a perspectival owner of X would amount to identification with a personal owner of X: identification would hence double as the veridical and reflexive (non-object-like) representation of a self. This would then imply that when we seem to recall ourselves (viz a self) doing Y (involving apparent identification with personal owner of action Y) then we are, quite literally, recalling our self doing Y. Or when we imagine what we assume is our future self doing Z then we are literally imagining our self doing Z.

Now if the self is a construct, as we've argued that it is, then a subject's identification with a perspectival owner (whether past or future) will not be coupled with the representation of a self. Identification, which elicits feelings of personal ownership, will only seem to represent a personal owner, viz., a self. So when we seem to recall ourselves doing Y, what will actually occur, on this analysis, is the following. The current witnessing perspective (a perspectival owner) will be identifying with the perspectival owner of action Y by appropriating an idea of this owner (viz., the first-person memory of S who did Y) to its current perspective. This appropriation will elicit characteristic feelings of personal ownership towards action Y. These feelings of personal ownership or my-ness will seem to point to a personal owner, me, who both performed action Y and is now recalling it with sentiments of my-ness. But there will be no such personal owner who performed or is recalling action Y. Subject S will hence identify as—but not with—the personal owner of Y.
along with the feelings, thoughts, emotions and perceptions that attend it. The assumption of a self as ‘personal owner and agent of Y’ will thus be created from feelings of personal ownership; and the feelings of personal ownership, on this analysis, will be an upshot of the current witnessing presence appropriating to its perspective the idea ‘perspectival owner of Y’. The theory, then, is that identification through appropriation somehow serves to generate feelings of personal ownership that bamboozle the current witnessing perspective into assuming that the appropriated idea, along with its own witnessing nature, is part and parcel of a personal self. This is the illusion of self.

It will help to illustrate this with an example. Suppose Ben wins a race, and then feels proud because of it. On our analysis, Ben’s feeling of pride and attendant thoughts “I won the race, yippee!” indicate that he feels personal ownership towards winning the race and his proud feelings. The feelings of pride are a give-away sign that as a current witnessing perspective, Ben identifies with the perspectival owner and agent who won the race (appropriating those memories to his current perspective) and hence as the personal owner and agent who won. He would not feel proud unless he believed it was he who won the race. But is this belief a correct one? Not literally. He believes that he qua personal self won the race – that is how the event seems remembered from his perspective. His believing this much, however, has already bought into the feelings of personal ownership that tell him “I, a personal owner and agent called ‘Ben’, won the race, and I, the same personal self who won the race am feeling mighty proud because of it”. But we have argued that such feelings mislead. There is in fact no such self-entity who could have won the race or now believes that he won it. There is only, at any one time, a perspectival witnessing presence with a bundle of thoughts and emotions that collectively comprise the sense of a self who has them. There is no actual personal self who has these emotions: that is what it means to say the self is an illusion.

How Theory of Identification will be Integrated Into Damasio’s Analysis of Self as Construct

We should now be able to see how the proposed theory of identification dovetails into Damasio’s (and the Buddhist) analysis of the self-impression co-arising with thoughts and emotions which both comprise tanha and tell the story of a ‘me’ who has them. In a given
moment of awareness with a sense of self (what Damasio calls “core consciousness”), the feelings of personal ownership which depict a self will, on our proposed theory, consist partly – not entirely – of those discrete tanha-driven thoughts and emotions (presented in characteristic imagery) that both Damasio and Buddhism claim to be ubiquitous. It is this component of the self-depicting personal-ownership feelings – the emotions and their attendant thoughts – that actively contributes to the sense of a personal boundary between oneself and the rest of the world. It is this component that Damasio has done such a good job of explaining. But – and here is where the proposed account expands on Damasio – the feelings of personal ownership which tell the story of a self, will also involve input from the current awareness which identifies as the self who is personal owner of those feelings. After all, the thoughts and emotions that help comprise tanha are objects of awareness, and there can be no objects of awareness without awareness. The input from awareness will account for the origin of many features attributed to the self as a whole – those features which qualify awareness simpliciter. They will be the features which Damasio’s account does not properly explain, namely: unity, elusiveness, continuity, invariability, unconstructedness, and, of course, witnessing presence.

The feelings of personal ownership, on our account, arise from awareness’s identification with a past or future perspectival owner, via an appropriation of an idea (such as a memory) of that owner, to the current perspective of awareness. We should note that the appropriated idea of the perspectival owner, whether pertaining to the past or future, will always be something that itself occurs in the present, along with the awareness that appropriates it. In terms of Damasio’s theory, it makes sense to suppose that this appropriation of a perspectival-owner idea as a current self will simultaneously involve a representation of the updated non-conscious proto-self as a current self. Perhaps what I have been calling ‘appropriation’, then, doubles as a form of representation, possibly with the

---

130 Although the word ‘identification’ was avoided so as to explain Damasio’s theory in his own terms, identification is essentially what would, on his theory, bestow advantage to the organism in the transition between emotions merely being felt, and their being felt to belong to oneself. It would mark the transition between perspectival ownership of feelings (reacting instinctively to physical pleasure or pain) to a sense of personally owning them, such that they feel properly ‘mine’. There would hence be identification with this perspectival owner of these feelings. While identification would herald the onset of that sense of personal boundedness that typifies the self, it would build on the existing tropism of a perspectival owner. As Damasio’s theory makes plain, without this tropism, there could be nothing for the sense of self or identity to build upon. The perspectival owner is first building block in the sense of self.
proto-self being represented and the ideas representing the perspectival owner, physical and mental sides of the same coin.

While not being so bold as to suggest that a representation of the perspectival owner (viz. the appropriated idea of it) would reduce in any way to the non-conscious proto-self that is represented as a self, the possibility is an interesting one. Damasio’s account is peppered with evidence that the neurobiological correlates to the sense of bounded self – psychologically experienced via the retinue of self-referencing thoughts and emotions (in awareness, we add) – will involve the proto-self (for example, 1999, 100). And we may recall that Damasio provides evidence to suppose that the proto-self shares its neural substate with the (non-conscious) emotions whose job it is to maintain homeostatic regulation (1999, 100). It is important to realise, however, that integrating the account of identification with these aspects of Damasio’s account, does not force a physicalist hand. What it does do is to place the ontology of the sense of self in a neurobiological context, which takes seriously the emotions of tanha on both a conscious and non-conscious level – regardless of whether one opts for a reduction of one level to the other. We shall see that this has favourable implications with regard to objectively testing for the presence or absence of tanha – and hence for a sense of self. At the same time, the integration acknowledges the crucial input of awareness with the intrinsic features it has to offer to the self as a whole.

There is a distinct advantage to arise from integrating the Damasio/Buddhist theory of a tanha-driven self with the notion of identification. It provides an acid-test for not only determining whether a person has a sense of self, but a test for determining the particular objects towards which one has a sense of personal ownership and hence reciprocal self-identification. They will be those objects which are (a) perspectivally or possessively owned and (b) whose perceived status affects the actual or potential happiness or mental suffering of that person. In chapter four, this potential link between identification and value was mentioned when we noted that the best evidence for objects seeming personally owned (reflecting a sense of personal boundedness) is the emotional value invested in them. Recall that three modes of identification which strongly evidenced a sense of personal boundedness – ‘this-ness’, ‘autonomy’ and ‘persistent self concern’ – involved the witnessing subject investing emotion in various objects. It was also noted that the patients with anosognosia, as
referred to in the paper by Nikalinakos (2004), lacked any sign of emotional investment in
the limb that was felt by them to be personally unowned.

In light of the arguments of this chapter, we can now confidently assert that emotional
investment towards a perspectivally or possessively owned X is a both necessary and
sufficient sign that one identifies as X’s personal owner. The presence of tanha in relation to
some such X guarantees identification as X’s personal owner. The absence of tanha towards
X guarantees a lack of identification as X’s personal owner. The presence of tanha towards
any X at all guarantees identification as a personal owner and hence that one has a sense of
self. The absence of tanha towards any X at all will guarantee a lack of identification as a
personal owner and hence indicate that one lacks a sense of self. With backing from
Damasio’s theory, the presence of tanha may well have (or potentially have) neuro-
physiological correlates, which would make the presence or absence of tanha (hence the
sense of self) empirically testable from a third-person perspective. This will have important
implications for testing the credibility of the Buddhist claim that practices such as meditation
gradually eliminate the sense of a bounded self (while retaining awareness).\(^{131}\)

Before moving on to describe the ontology of self-as-construct-and-illusion in more
detail (in the next chapter) a disclaimer is in order. Having asserted how this current account
might improve on an aspect of Damasio’s, I will not pretend that in its current formulation, it
will offer a full explanation of how the self as whole comes to exhibit the features that it
does. In short, I propose to skip over the precise mechanism of how the current witnessing
perspective actually does ‘appropriate’ (or represent) the recalled or imagined idea of a
perspectival owner to its perspective with the resultant feelings of personal ownership. (I
confess to finding it too difficult at this current stage of my investigations). The precise
mechanism of appropriation will be largely treated, for purposes of this discussion, like a
black box. Somehow, black-box appropriation performs an ‘alchemy’ which seems to fuse
(or should I say ‘confuse’) the current awareness with a bundle of thoughts and emotions
(including the appropriated idea) resulting in the impression of a unified, elusive, bounded,
enduring, invariable, unconstructed self. The feelings of personal ownership somehow
manage to blind the current awareness from the fact that it is nothing more than awareness
along with a bundle of thoughts, memories and emotions etc. They fool the awareness into

\(^{131}\) I say more about this in the concluding section entitled “Glimpses Beyond”.
reflexively assuming that it is a single, personal, bounded owner, viz., an unconstructed self that owns and generates observed thoughts and emotions. And the personal bounded self seems, as part of this illusion, to take on the features that rightly belong to only the unconstructed awareness — the unity, elusiveness and so forth. In the following and final chapter, I outline a more specific way to describe the illusion of self as being contributed to by two ‘feeders’ — the unconstructed awareness — and the elements (thoughts and emotions) that create the reflexively assumed impression that this awareness is a bounded self.
Chapter Nine: The Two-Tiered Illusion of Self

So far we have been describing the illusion of self in broad terms that are compatible with different ways in which the self could turn out to be illusory – for example in virtue of its unity, elusiveness, boundedness etc. Thus we have been saying that the self-as-unconstructed is illusory since, while it purports to be unconstructed, it is, as a matter of fact, constructed (a formulation which allows for different stories about its constructed status). Given our more detailed analysis in chapter eight which pins the source of the self’s constructed status on boundedness, we might wonder if there a more specific way in which to describe the illusion of self. There is indeed. What emerges from our analysis is what I shall call a ‘two-tiered illusion of self’. Now it is important to understand that the ‘two tiers’ to the self-illusion do not correspond to two ontologically separate processes that occur during identification. The two tiers are rather to be understood as describing the same illusory phenomenon – the self – from two different angles. The two angles correspond to the two main ‘feeders’ into the illusion of self: the ‘boundedness feeder’ on one hand and the ‘awareness feeder’ on the other.

The first tier of the illusion consists of input from the ‘boundedness feeder’ and can be described as a person’s (co-specified) awareness, actually non-bounded in its nature, purporting to be a bounded self. Since awareness is not inherently bounded in this way then it cannot be the bounded self that it purports to be: hence the illusion of self, viz., awareness-as-bounded. The second tier of the illusion consists of input from the ‘awareness feeder’. It can be described as the bounded self, actually a construct, purporting to harbour those (co-specified) features that rightly belong only to the unconstructed awareness. Since the self is not inherently unconstructed in the manner it purports to be, then it cannot be an unconstructed self: hence the illusion of self, viz., self-as-unconstructed.

The idea behind this ‘two-tiered’ illusion of self is not original to this project, for it is explicitly pre-figured in Advaita Vedanta (and, I have argued, implicitly in Buddhism). In the following passage, the leading figure of this tradition, Sankara, (788-820 CE ), speaks of awareness (akin to undifferentiated fire) taking itself to be a bounded entity (akin to heated
iron), by adopting features associated with this bounded entity (for example agency, personal ownership):

Allied to the intellect, just a part of itself, although the true self of everything, and beyond the limitations of such an existence, [awareness] identifies itself with this illusory self - as if clay were to identify itself with earthen jars. (1968, 190)

In conjunction with such additional qualities, the supreme self [awareness] seems to manifest the same characteristics, just as the undifferentiated fire seems to take on the qualities of the iron it heats. (1968, 191)

And the bounded self through which awareness (akin to the sun in the next quotation) is reflected, appears, through identification as the whole (sun reflected in jar-water), to take on what is inherent to awareness:

The ignorant see the reflection of the sun in the water of a jar and think it is the sun itself. In the same way [he] sees the reflection of consciousness [viz., awareness] in its associated qualities [to do with boundedness] and mistakenly identifies himself with it. (1968, 218)

The natural state of awareness is, however, unfettered by illusions of boundedness. To realise its true nature is to see through this illusion:

The wise man ignores jar, water and the sun's reflection in it, and sees the self-illuminating sun itself which gives light to all three but is independent of them. (1968, 219).

I shall now elaborate in more detail upon how to construe the two-tiered illusion of self such that we see how it is borne out through (a) co-specified awareness in general and (b) each feature intrinsic to awareness – witnessing presence, unconstructedness, elusiveness, unity, endurance and invariability. During this exercise, some repetition and overlap will be natural: for example the explanation of how each feature becomes ‘bounded’ (in tier one) will involve the process of identification which casts ‘the self’ as protagonist of a wordless story (to borrow Damasio’s turn of phrase). It should be understood that the process of identification will automatically enfold the (minimally contingent) psycho-physical perspective of awareness, viz., the subject, into the illusion of self, so that the impression will be of a self who is a perspectival as well as personal owner. This account should serve to draw many threads of argument from previous chapters into the beginning of a unified,
patterned weave. I say ‘beginning’ because it will be obvious that the proposed account of the self’s ontology is sketchy, particularly in its ‘black-box’ explanation of how appropriation works. Nevertheless, one must be reminded that other areas which might seem flimsy in this presentation are in fact recapitulations of earlier argument – for example the evidence that we identify as a self with each of the enlisted features, and the justification for supposing that the self qua that feature is illusory.

Let us now outline how the model of the two-tiered self-illusion applies to the co-specified awareness in general, which will serve as a template for examining in more detail how the model applies to each specific feature of awareness.

The first tier of the self-illusion consists of input from the ‘boundedness feeder’ where the self is construed as ‘awareness-as-bounded’. Evidence that we do indeed identify ourselves as awareness-as-bounded has been extensively provided in chapter four this project (and recapitulated at the beginning of chapter eight). This ‘boundedness tier’ to the self-illusion focuses on the fact that awareness purports to exhibit a feature that is befitting to the self, namely, that of being a bounded personalised entity. Our explanation of this impression is one that appeals to contribution to the subject’s awareness from objects of consciousness (i.e., thoughts and emotions subtending tanha) such that the impression of a bounded self is created. In this ‘boundedness’ tier, awareness-as-bounded comes out as illusory because awareness purports to be bounded when awareness is not in fact bounded in its nature. Awareness is not bounded because we have shown that: (a) awareness in itself is unconstructed (from the arguments of chapters six and seven) while (b) awareness-as-bounded is constructed, viz., the content of an appearance contributed to by objects of consciousness – in particular, thoughts and emotions subtending tanha (chapter eight). Since awareness is unconstructed, it cannot be bounded (since awareness-as-bounded is a construct) which renders the self, viz., awareness-as-bounded, to be an illusion.

The second tier of the self-illusion consists of input from the ‘awareness feeder’ where the self (equivalent to awareness-as-bounded) is construed as the self-as-unconstructed. Evidence that we identify as an unconstructed entity has been presented in chapter four (and recapitulated in chapter eight). This ‘unconstructedness’ tier to the self-illusion focuses on the fact that the self, viz., bounded awareness, purports to exhibit the unconstructed status that is befitting to awareness itself. Our explanation of this impression
is one that appeals to genuine unconstructed input from the awareness feeder (chapters six and seven) such that the impression of an unconstructed self will be generated. In this tier, self-as-unconstructed comes out as illusory because the self purports to be unconstructed when as a matter of fact the self is a construct. The self is not actually unconstructed (as it purports to be) because, as we demonstrated in chapter eight, the boundedness aspect of self-as-unconstructed is the content of an appearance that is contributed to by objects of consciousness, rather than existing independently of these objects as it purports to do. Hence the self-as-unconstructed is an illusion.

We shall now apply this general schema of the two-tiered self-illusion to awareness in light of each of its specific features: witnessing presence, unconstructedness, elusiveness, unity, endurance and invariability.

**Witnessing Presence**

With regards to witnessing presence, the first tier of the self-illusion consists of input from the ‘boundedness feeder’ where the self is construed as ‘witnessing-presence-awareness-as-bounded’. Evidence that we do indeed reflexively construe witnessing presence to be bounded resides in the fact that feelings of this-ness and dispositional value seem to point to ‘this very presence who witnesses’. In other words, the witnessing-presence aspect of awareness does not normally present as something that is non-bounded and impersonal. Our explanation for this apparently personal boundedness of witnessing presence is that through the process of identification (driven by tanha), the ‘storyline’ told by current emotions (obvious or subtle, attentive or inattentive) with attendant thoughts and perceptions ‘casts’ the natural input of witnessing presence as a bounded self, perhaps in similar way that dreams can ‘cast’ real alarm sounds into their storyline as shrill voices. In this boundedness tier, the self, viz., witnessing-presence-awareness-as-bounded comes out as illusory because witnessing-presence-awareness purports to be bounded when witnessing-presence-awareness is not in fact bounded in its nature. Witnessing-presence-awareness is not bounded in its nature because we have shown that: (a) witnessing presence in itself is unconstructed (from the arguments of chapters six and seven) while (b) witnessing-presence-awareness-as-bounded is constructed, viz., the content of an appearance contributed to by objects of consciousness – in particular, thoughts and emotions subtending tanha (see explanation
above). Since witnessing presence-awareness is unconstructed, it cannot be bounded (since witnessing-presence-awareness-as-bounded is a construct) which renders the self, viz., witnessing-presence-awareness-as-bounded to be an illusion.

With regards to witnessing presence, the second tier of the self-illusion consists of input from the ‘awareness feeder’ where the self, viz., <bounded-witnessing-presence-awareness>, is presented as being as unconstructed in the manner of pure awareness. The evidence that we construe the self, viz., <bounded-witnessing-presence-awareness>, to be unconstructed is that from the viewpoint of ‘myself’, it would seem that whatever it is that makes my witnessing-presence seem uniquely and personally mine will be unobservable as an object – (it cannot be attended to) – and so it does not seem as if any objects are helping to construct the boundaries of this witnessing perspective. Our explanation of this impression is one that appeals to genuine unconstructed input from the awareness feeder such that an impression of an unconstructed bounded witnessing presence is generated. Whenever a current witnessing presence appropriates an idea to its perspective, the native contribution from awareness to resultant feelings of personal ownership will naturally infuse ‘the personal owner’ with that subjective sense of presence. The witnessing presence is what lends the apparent self its first-person, subjective perspective, imbuing it with the puzzles of consciousness. Because a ‘personal witnessing presence’ is all that most of us have experienced (at least to the best of our memory) we normally cannot imaginatively separate pure first-person feelings from personal first-person feelings – just as a salt-water fish cannot imaginatively separate the sensations of pure water from the sensations of the salt-water to which it is accustomed. To the usual conscious mind, witness-presence awareness and ‘me’ thus present as a basic unit.

In this ‘awareness’ tier, the self, viz., bounded-witnessing-presence-awareness-as-unconstructed comes out as illusory because the bounded-witnessing-presence-awareness purports to be unconstructed when as a matter of fact, as demonstrated in chapter eight, it is a construct. The boundedness aspect of witnessing-presence-as-unconstructed is the content of an appearance that is contributed to by objects of consciousness that subtend tanha, in particular, thoughts and emotions. Evidence for this constructedness is that a suspension of emotion and its attendant thoughts results in the suspension of the impression that the
witnessing perspective is bounded. Hence the self, viz., bounded-witnessing-presence-awareness-as-unconstructed, is an illusion.

**Unconstructedness**

With regards to unconstructedness, the **first tier** of the self-illusion consists of input from the ‘**boundedness feeder**’ where the self is construed as ‘unconstructed-awareness-as-bounded’. Evidence that we do indeed reflexively construe unconstructed awareness to be bounded resides in the fact that feelings of personal ownership such as this-ness convey a personal, as opposed to a merely perspectival subject who stands behind and owns these feelings (is not constructed by them). Our explanation for this apparent boundedness of unconstructed awareness is that through the process of identification (driven by tanha) the ‘storyline’ told by current emotions (obvious or subtle, attentive or inattentive) with attendant thoughts and perceptions ‘casts’ the natural input of unconstructed awareness as a bounded self, perhaps in similar way that dreams can ‘cast’ real alarm sounds into their storyline as shrill voices. In this boundedness tier, unconstructed-awareness-as-bounded comes out as illusory because unconstructed awareness purports to be bounded, when in fact unconstructed awareness is not bounded in its nature. Unconstructed awareness is not bounded because (a) unconstructed awareness in itself is unconstructed (from the arguments of chapters six and seven) and (b) unconstructed-awareness-as-bounded is constructed, viz., the content of an appearance contributed to by objects of consciousness – in particular, thoughts and emotions subtending tanha (see explanation above). Since unconstructed-awareness is unconstructed in actuality, it cannot be bounded (since unconstructed-awareness-as-bounded is a construct) which renders unconstructed-awareness-as-bounded to be an illusion.

With regards to unconstructed awareness, the **second tier** of the self-illusion consists of input from the ‘**awareness feeder**’ where the self, viz., <bounded unconstructed awareness>, is presented as being unconstructed, in the manner of pure awareness. The evidence that we construe the self, viz. <bounded unconstructed awareness> to be unconstructed is that from the viewpoint of ‘myself’, it never feels as if the self as a whole is something that is being constructed by current emotions and their attendant thoughts and perceptions. The self is always depicted as their personal, ontologically unique, owner, standing apart from them as a separate entity. Our explanation of this impression is that in
the process of identification, the real unconstructedness of native awareness filters into the net impression of the bounded self-entity. In this tier, the self, viz., bounded-unconstructed-awareness-as-unconstructed comes out as illusory because the bounded-unconstructed-awareness purports to be unconstructed when as a matter of fact, as demonstrated in chapter eight, it is a construct. The boundedness aspect of awareness-as-unconstructed is the content of an appearance that is contributed to by objects of consciousness that subtend tanha, in particular, thoughts and emotions. Evidence for this constructedness is that a suspension of emotion and its attendant thoughts results in the suspension of the impression that the unconstructed awareness is bounded. Hence the self, viz., bounded-unconstructed-awareness-as-unconstructed, is an illusion.

Unity

With regards to unity, the first tier of the self-illusion consists of input from the 'boundedness feeder' where the self is construed as 'unified-awareness-as-bounded'. Evidence that we do indeed reflexively construe unified awareness to be bounded resides in the fact that any thoughts and impressions at a given waking moment will seem to present themselves not merely to a unified awareness but to a unified personal awareness. The thoughts will hence not merely seem to be perspectivally mine, but personally mine, indicating that I identify as a personal unified owner who has them. The self as a bounded entity is felt to be that unified entity which unites such items as its personal own. Our explanation for this apparent boundedness of unified awareness is that through the process of identification (driven by tanha) the 'storyline' told by current emotions (obvious or subtle, attentive or inattentive) with attendant thoughts and perceptions 'casts' the natural input of unified awareness as a bounded self, perhaps in similar way that dreams can 'cast' real alarm sounds into their storyline as shrill voices. In this boundedness tier, unified-awareness-as-bounded comes out as illusory because unified awareness purports to be bounded, when in fact unified awareness is not bounded in its nature. Unified awareness is not bounded because (a) unified awareness in itself is unconstructed (from the arguments of chapters six and seven) while (b) unified-awareness-as-bounded is constructed, viz., the content of an appearance contributed to by objects of consciousness – in particular, thoughts and emotions subtending tanha (see explanation above). Since unified-awareness is not constructed in
actuality, it cannot be bounded (since unified-awareness-as-bounded is a construct) which renders the self, viz., unified-awareness-as-bounded, to be an illusion.

With regards to unified awareness, the second tier of the self-illusion consists of input from the ‘awareness feeder’ where the self, viz., <bounded unified awareness>, is presented as being unconstructed in the manner of pure awareness. The evidence that we construe the self, viz., <bounded unified awareness> to be unconstructed, is that it would seem to ‘me’ that ‘I’, a unified self to whom impressions belong, am necessarily unified, with no impression of any potential cracks in this unity (an impression which would be present if unity seemed constructed). As a result, it would seem to me that I couldn’t really imagine what it would be like for my self to become disunified. Our explanation of this impression will be that during the process of identification, somehow the mechanism of appropriation employs the native unity of awareness to ‘fuse itself together’ with any appropriated ideas, such that the feelings of personal ownership seem to unquestioningly emanate from a personal owner who is this necessarily unified entity. In this tier, the self, viz., bounded-unified-awareness-as-unconstructed comes out as illusory because the bounded-unified-awareness purports to be unconstructed when as a matter of fact it is, as demonstrated in chapter eight, a construct. The boundedness aspect of unified-awareness-as-unconstructed is the content of an appearance that is contributed to by objects of consciousness that subtend tanha, in particular, thoughts and emotions. This is evidenced in the fact that the apparent unity of self can be objectively proved to lack the necessity that it subjectively seems to have. Cases of epileptic automatism (with their suspension of thought and emotion) together with our background conclusion that the bounded self is a construct, are testimony to the fact that the ‘unity of self’ can come apart. Hence the self, viz., bounded-unified-awareness-as-unconstructed, is an illusion.

**Elusiveness**

With regards to elusiveness, the first tier of the self-illusion consists of input from the ‘boundedness feeder’ where the self is construed as ‘elusive-awareness-as-bounded’. Evidence that we do indeed reflexively construe elusive awareness to be bounded resides in the fact that it is not only awareness that seems elusive to its own observation but personal, bounded awareness. Hence boundaries seem part and parcel of what is elusive about
awareness. While having a sense of its ‘own’ boundedness, it is a seemingly personal awareness that can never manage to observe its boundaries as an object of thought or perception. The boundaries appear to qualify an elusive, observing, personal subject – as opposed to a non-elusive, observable object. Our explanation for this apparent boundedness of elusive awareness lies in the stream of thoughts and emotions that (through *tanha*) contribute a sense of personal ownership or my-ness towards any object of consciousness. While the lion’s share of the ‘aboutness’ of these thoughts and emotions is the main target of the storyline, (e.g., our impression of the objects to which the thoughts and emotions seem directed), a portion of this ‘aboutness’ is reserved for the implicit self who is having them, viz., the personal owner of the thoughts and emotions. This sidelined section of the storyline’s ‘aboutness’ always portrays the self as an individual protagonist who is definitively *separate* from the target object, and this helps to create a distinct impression of self-other boundedness.

In this ‘boundedness’ tier, the self, viz., *elusive-awareness-as-bounded* comes out as illusory because elusive awareness purports to be bounded, when in fact elusive awareness is not bounded in its nature. Elusive awareness is not bounded because (a) elusive awareness in itself is unconstructed (from the arguments of chapters six and seven) while (b) elusive-awareness-as-bounded is constructed, viz., the content of an appearance contributed to by objects of consciousness – in particular, thoughts and emotions subtending *tanha* (see explanation above). Since elusive awareness is not constructed in actuality, it cannot be bounded (since elusive-awareness-as-bounded is a construct) which renders the self, viz., elusive-awareness-as-bounded, to be an illusion.

With regards to elusive awareness, the second tier of the self-illusion consists of input from the ‘awareness feeder’ where the self, viz., *<bounded elusive awareness>*< is presented as being *unconstructed* in the manner of pure awareness. The evidence that we construe the self, viz., *<bounded elusive awareness>* to be *unconstructed* lies in the fact that it does not appear as if the ‘bounded awareness’ can ever observe its ‘own’ boundaries as a non-elusive object – which it would need to be able to do if the bounded elusive awareness were to *seem* constructed by objects of consciousness. Our explanation of this impression is that the elusiveness, which seems to make the bounded self-entity sidelined and unobservable, will in fact emanate from the elusive awareness that has contributed to the
storyline of its 'protagonist' as a genuine alarm-sound might contribute to the shrill voice in a dream. Through the mechanism of appropriation, awareness will be somehow blinded to the fact that 'its elusive boundaries' are nothing more than the projected and imagined content of thoughts and emotions, together with its own native input of elusiveness.

In this tier, the self, viz., *bounded-elusive-awareness-as-unconstructed* comes out as *illusory* because the *bounded-elusive-awareness* purports to be unconstructed when as a matter of fact it is, as demonstrated in chapter eight, a construct. The boundedness aspect of elusive-awareness-as-unconstructed is the content of an appearance that is contributed to by objects of consciousness that subtend *tanha*, in particular, thoughts and emotions. There are hence no elusive boundaries of self, since there is no self that is bounded. There is only elusive awareness together with a bundle of non-elusive thoughts, emotions and ideas that, with the help of elusive witnessing, tell the story of an elusive self. Hence the self, viz., *bounded-elusive-awareness-as-unconstructed*, is an illusion.

**Endurance**

With regards to endurance, the first tier of the self-illusion consists of input from the 'boundedness feeder' where the self is construed as 'enduring-awareness-as-bounded' both from moment-to-moment (as non-perduring consciousness) and during longer time-spans (as assumed in personal identity). Evidence that we do indeed reflexively construe non-perduring awareness to be *bounded* lies in the fact that the element of our consciousness which seems both lived (not a mere absence) and cannot be observed as perduring, seems identical to that element of our consciousness which is unified, elusive and personal. On a longer-term basis, it seems to be numerically one and the same bounded personal self that is living this life with a history and anticipated future.

Our explanation for this apparent *boundedness* of enduring awareness is as follows. Regarding the appearance of *bounded, non-perduring awareness*, the constant stream of ideas appropriated to the perspective of awareness through the process of identification will, through their content, provide a 'thickness' or impression of *me-ness* to the sense of self from moment to conscious moment. The stream of emotions and attendant thoughts also do much to explain the sense of a bounded self with *longer-term endurance* or identity. Their content tells the story of a continuing self who existed in the past and is planning for its future. That
is because the current awareness identifies not merely with a perspectival owner of the latest percep or action, but with a perspectival owner of an integrated historical personality *qua* the owner of that latest percep or action. Most emotions, moreover, have built into them the idea of a numerically continuous self who has them. Having these emotions requires that one identify not only as their current personal owner, but also as the owner of a past or future scenario that was either perceived or is anticipated to trigger the emotion. For example one could not feel guilty unless one believed that an earlier self—numerically identical to the current self—performed a regrettable action. One could not feel nervous at an impending exam unless one implicitly believed that a future self who will sit the exam is identical to the current self who fears sitting it. The more subtle presence of *tanha*, implicated in such emotions, also has the subject reflexively pre-supposing that it is an enduring bounded self. It requires that one identify as the personal owner of past or future scenarios that are perceived, attentively or inattentively, to fulfill or frustrate the desires upon which actual or imagined emotions depend.

In this boundedness tier, the self, viz., *enduring-awareness-as-bounded* comes out as illusory because enduring awareness purports to be bounded, when in fact enduring awareness is not bounded in its nature. Enduring awareness is not bounded because (a) *enduring awareness* in itself (at least the non-perduring version) is unconstructed (from the arguments of chapters six and seven) while (b) *enduring-awareness-as-bounded* is constructed, viz., the content of an appearance contributed to by objects of consciousness—in particular, thoughts and emotions subtending *tanha* (see explanation above). Since *enduring*-(non-perduring)-awareness is not constructed in actuality, it cannot be bounded (since *enduring-awareness-as-bounded* is a construct) which renders the self, viz., *enduring-awareness-as-bounded* to be an illusion. There is no *bounded*, enduring awareness, only lived awareness with no positive evidence of non-endurance whilst awake.

With regards to enduring awareness, the second tier of the self-illusion consists of input from the ‘*awareness feeder*’ where the self, viz., *<bounded enduring awareness>*<>, is presented as being *unconstructed*, in the manner of pure awareness. The evidence that we construe the self, viz., *<bounded enduring awareness>* to be *unconstructed*, lies in the fact that the ideas that are appropriated to the perspective of current awareness—ideas that actually serve to construct the boundaries of self—never seems to exhibit any
actual discontinuity – i.e. they are never directly observed to arise and pass away from consciousness. If the self’s endurance did seem overtly constructed – dependent on impermanent objects of consciousness – then the self’s ‘endurance’ would seem more like *perdurance* (akin to Galen Strawson’s “pearl view” of the self, 1997, 424-425) with it a decisively ‘grainy’ feel to its existence over time (perhaps, as mentioned in chapter four, like a string of mini general anesthetics).

Our explanation for the impression of the non-perduring and hence unconstructed-like endurance of a bounded self is that the process of identification will render any appropriated idea to seem integral to awareness (hence unified and elusive), which has a masking effect on its actual coming and going from awareness over time. If the appropriated idea could be directly observed to come and go – hence observed as discontinuous – then it would not seem integral to awareness in the first place. Also, whenever current awareness identifies itself with a past or imagined perspectival owner, the native endurance of current unconstructed awareness will make it seem as if this identified-as self – whether of distant past, imagined future or immediately preceding moment – is itself uninterruptedly existing at every new conscious moment. This native endurance will be committed to memory, so it seems as if the recalled or imagined self is an enduring rather than static or perduring entity.

In this tier, the self, viz., *bounded-enduring-awareness-as-unconstructed* comes out as illusory because the *bounded-enduring-awareness* purports to be unconstructed when as a matter of fact it is, as demonstrated in chapter eight, a construct. The boundedness aspect of enduring-awareness-as-unconstructed is the content of an appearance that is contributed to by objects of consciousness that subtend *tanha*, in particular, thoughts and emotions. When we reflect in further detail, it is especially clear that the longer-term numerically enduring bounded self (the kind often alluded to in debates on personal identity) is, due to its constructed status, not as it seems. It must have gaps in its existence since the identified-with ideas, even if continuously overlapping during waking hours, do not occur 24 hours a day. While we cannot rule out the possibility of witnessing during deep sleep, there is solid evidence that there are no thoughts during that period. No thoughts imply no identification, and hence on our analysis, no constructed self. The ‘self’ that awakens from deep sleep cannot, therefore, be numerically identical to the self who fell asleep – contrary to how it
would seem. Hence the self, viz., bounded-enduring-awareness-as-unconstructed, is an illusion.

**Invariability**

With regards to invariability, the first tier of the self-illusion consists of input from the 'boundedness feeder' where the self is construed as 'invariable-awareness-as-bounded'. Evidence that we do indeed reflexively construe invariable awareness to be bounded lies in the fact that the quality of same-ness that pervades conscious experience is not merely impersonal but seems imbued with the essence of me, of who I am. This quality of 'me-ness' seems to persist, despite changes to my personality over time. No one else could possibly have this quality—it seems uniquely integral to my identity.

Our explanation for this apparent boundedness of invariable awareness is as follows. Feelings of personal ownership (explicit or implicit) towards various Xs at a time trigger a reciprocal sense of being a personal owner such that the process imparts, as it were, a 'tone' or 'hue' to the (otherwise colourless) subjective sense of presence. To use a crude metaphor, it is as if each instance of identification (of which there are many) spills a pod of dye into the colourless substrate of awareness contributing to the resultant impression of a 'unique me'. The dye seems to 'colour up' the awareness with the shade of the current (and often recurrent) appropriated ideas such that the subject feels integrated and defined in relation to that 'colour', some times more noticeably than others (think of punkish persona versus owner of a breath). And like genuine dye, the mixing of many colours will result in the impression of a uniform colour. New colourful additions will soon swirl into this old colour—or near enough so that the change goes unnoticed. The uniform colour, then, is 'me', viz., the sense of a personal owner of or qua one's general personality as it stands in relation to some aspect of the world, X. This 'personal owner' is the upshot of the awareness identifying with a perspectival owner of (or qua) personality qua owner of some X (e.g. a percept). The identification will have triggered feelings of personal ownership towards X, implying that sense of being X's integrated personal owner. The resultant sense of a self is thus reactivated and revised upon every new identification (on the part of awareness) with a perspectival owner of (or qua) personality qua owner of X, radical or mundane. Because of the genuine overall resemblance between one memory of personality and the next (along with the fact
that both have awareness as identical invariable ‘substrate’) the subtle changes are likely to be glossed over in the manner that Hume describes (1739, 169-170). Even a sudden overlay of reds and purples (for example becoming a born-again Christian) will not fully disguise the underlying ‘colour of me’ (hence the sense that I have become a born-again Christian) and will sooner or later blend into the general ‘dye-pool’.

The emotions and their attendant thoughts markedly contribute to the impression of it being the invariably same bounded me existing over time since they elicit storylines of other times that we felt that way. For example, when feeling depressed, we are likely to remember previous times we felt depressed, which will seem to create arcs of identity to the previous owners of the depression. Immediately there will seem to be a solid, numerically continuous and invariable me who is defined and hence bounded by the implicit quality of feeling depressed. (Which may explain why when someone jovially says ‘that mood won’t last’ it is hard to believe!).

In this boundedness tier, the self, viz., invariable-awareness-as-bounded comes out as illusory because invariable awareness purports to be bounded, when in fact invariable awareness is not bounded in its nature. Invariable awareness is not bounded because (a) invariable awareness in itself is unconstructed (from the arguments of chapters six and seven) while (b) invariable-awareness-as-bounded is constructed, viz., the content of an appearance contributed to by objects of consciousness – in particular, thoughts and emotions subtending tanha (see explanation above). Since invariable awareness is not constructed in actuality, it cannot be bounded (since invariable-awareness-as-bounded is a construct) which renders the self, viz., invariable-awareness-as-bounded to be an illusion.

With regards to invariable awareness, the second tier of the self-illusion consists of input from the ‘awareness feeder’ where the self, viz., <bounded invariable awareness>, is presented as being unconstructed, in the manner of pure awareness. The evidence that we construe the self, viz., <bounded invariable awareness> to be unconstructed, lies in the fact that the underlying quality of me-ness seems to remain qualitatively the same throughout my life, despite changes to personality. If the ‘invariability’ seemed overtly constructed by objects, then the ‘I’ would seem to change its essence to reflect the different identified-with objects at different times. This would effectively ‘water down’ any sense of invariability in a similar manner to how perdurance would ‘water down’ any sense of the self’s continuity.
The self does not, however, present itself (through its endurance and invariability) as being 'watered down' in this fashion. Our explanation for the impression of this unconstructed invariable awareness of a bounded self is that the native unconstructed awareness serves as the invariable and colourless 'substrate' into which the dyes of identity are spilt. Without this 'substrate' of awareness there could be no such sense of continuity or invariability – short term or long. Just as spectral colours split by a prism share the deeper quality of luminosity, each moment of 'dyed awareness', although different in some sense, will through a lifetime share the deeper invariable quality of awareness. The mechanism of identification – with the help of factors alluded to in tier one – somehow manages to juxtapose the genuine invariability of awareness onto ‘dye-level’, such that it feels as if the personal ‘dye of self’ is of an invariable hue. This impression may be enhanced by the natural ‘change-blinding’ effect that identification has upon the actual coming and going of identified-with ideas from a subject’s awareness. So when we identify with a perspectival owner of X, there may be the tacit impression that we have always been the personal owner of X – because the explicit introduction of that object into our sense of identity will have gone unnoticed.

In this ‘unconstructedness’ tier, the self, viz., bounded-invariable-awareness-as-unconstructed comes out as illusory because the bounded-invariable-awareness purports to be unconstructed when as a matter of fact it is, as demonstrated in chapter eight, a construct. The boundedness aspect of invariable-awareness-as-unconstructed is the content of an appearance that is contributed to by objects of consciousness that subtend tanha, in particular, thoughts and emotions. The ‘dye of self’ is not invariable in the unconstructed manner of pure awareness. During waking hours, its hue is constantly and subtly changing. Hence-invariable-awareness-as-unconstructed is an illusion.

Summary

The two-tiered illusion of self that I have proposed is not of course complete in its detail, although it does appeal to the arguments of previous chapters in its support. It offers an outline for a theory of self, which, following Buddhism and Advaita Vedanta, takes seriously the input of co-specified awareness as not constructed or illusory. By alluding to the process of identification (through which boundedness comes about), the account also builds a perspectival owner into the overall construct of self.
Conclusion

In this chapter, we have drawn upon other chapters of the project to provide an overarching account of how the ontology of self as construct and illusion – with both constructed and unconstructed contributors – could come together in what I have termed the 'two-tiered illusion of self'. The constructed contributor, boundedness, is kept alive by tanha or emotional investment. In chapter eight I found strong support for this part of the theory from both East (Buddhism) and West (Damasio, 1999). The multifaceted role given to unconstructed awareness in the ontology of ‘no-self’, however, is what sets this proposed theory apart from many Western counterparts – theories proposed by Hume, James, Flanagan, Dennett and Damasio – that also render the self to be an overall construct and illusion. The contribution to the self-illusion from unconstructed awareness is notably what aligns this project’s theory of the self with those accounts that I have extracted from Buddhism and Advaita Vedanta. It is these Eastern traditions to which I owe all the fundamental ideas of this project, culminating in ‘the two-tiered illusion of self’; ideas that I have tried to defend within a framework of Western analytic philosophy.
Glimpses Beyond

In this project we argued for a model of self-as-illusion, a model that draws upon mental input that is constructed (viz., *tanha*, leading to boundedness) and unconstructed (underlying awareness co-specified with elusiveness, unity, endurance and so forth). By not treating unity and endurance (in particular) as the primary reason for the self’s constructed and illusory status, the model of self-as-illusion proposed here – prefigured in the East – significantly departs from the way the existence of self has been typically denied in the West. The departure can been seen as significant not only because the model is itself unconventional, but because the model sets the stage for further inquiry in a way that other models of the self cannot accommodate. I will conclude this project by saying a few words about this line of inquiry which I think holds a lot of promise.

The two-tiered model of self-illusion can be used as a foundation to begin investigating a question whose eventual answer could have great significance for the Western philosophy of mind. The question is whether it is *psychologically* possible that a person’s awareness, with all its co-specifying features, could somehow divest itself of the self-illusion by directly ‘waking up’ to its real nature, just as one might wake up and realise that the ringing sound comes from the alarm clock and not the dreamt-of shrill voice. It has already been suggested, through such cases as epileptic automatism, that awareness can come apart from the sense of self. Damasio’s patients, who seemed to temporarily lose the sense of themselves as bounded entities, did not cease witnessing. They remained awake and minimally responsive to their immediate surroundings. Yet there was not a shred of evidence that such patients – at least the ones Damasio described – had *any clue* as to what was happening. Stuck in the micro-context of the moment, there seemed to be no ‘mental room’ for them to harbour the realisation that an illusion of self was overthrown (and no ‘mental room’ for the usual identification). Many subjects of depersonalisation, on the other hand, seemed to retain their mental acuity (recall part of the DSM-IV definition: “reality testing remains intact”) and while evidently not identifying as the personal owner of various (perspectivally owned) thoughts and perceptions, still harboured negative emotions about their general predicament. In this project, we argued that such emotions evidence a sense of
personal ownership – and hence a sense of self. Once again, these cases came across as being pathological, as impairing the well-being and normal functioning of the recipient.

What we are curious about, then, is whether awareness might ‘wake up’ from the illusion of self such that contra Damasio’s patients, the separation of awareness from self-sense happens in a manner that is non-pathological and preserving of mental acuity and, contra typical depersonalisation, in a way that, of course, preserves no residual sense of self. Such that a person (qua their awareness) will actively and suddenly know, with full mental acuity, that the illusion of self has been overthrown. This ‘awakening’ will involve the direct, clear and immediate realisation that awareness is not – and never was – intrinsically bounded in the way that the self is. It makes sense to suppose that on a psychological level, such ‘awakening’ would be somewhat akin to a global depersonalisation – truly global – where none of one’s perspectivally owned experiences, while clearly apprehended, are felt as being personally ‘mine’. Unaccompanied by negative emotions, or by such impressions as being dead and “falling deep within oneself”; lacking any sense of ‘me’ or ‘mine’ whatsoever, this mode of awareness would genuinely lack an accompanying sense of self.

The psychological possibility of awareness ‘awakening’ from the illusion of bounded self might be dismissed as fanciful speculation were it not for the fact that it is deemed integral to the goal of practice in Buddhism and Advaita Vedanta; that of reflexively ‘realising the true nature’ of so-called nibbanic consciousness. Nibbanic consciousness, as we may recall from chapter two, is characterised in the Buddhist suttas as involving a mode of consciousness (or awareness, we can now say) where there is no tanha or emotional investment in any item and hence, on our theory, no sense of self. It is notable that the occasional person who has reportedly ‘realised’ nibbanic consciousness (an ‘Arahant’) is often described as having “woken up to their real nature”. This mode of knowing is reported to involve immense peace, happiness and mental clarity with complete freedom from the possibility of mental suffering or dukkha. It is clearly depicted as being non-pathological.

In chapter one we outlined the Four Noble Truths in an attempt to relate Buddhist philosophy to the concerns of this project. The Fourth Noble Truth sets out a path of practice known as “The Noble Eightfold Path”, involving methods called ‘insight-wisdom’ (panna), ‘meditation’ (samadhi) and ‘virtue’ (sila). In light of our model of the two-tiered self-illusion – a model implied by the Buddhist account of consciousness and no-self – this
system of training can be understood as one whose goal is to (a) dismantle the co-dependent sense of self and *tanha* while (b) simultaneously uncovering the real character of the native awareness such that it comes to reflexively know itself as it is itself (as opposed to it mistakenly assuming it is a self).

In view of this, I see the natural line of inquiry as being one that addresses the following questions. Is it possible for a person’s awareness to be divested of the self-illusion in such a manner that the person will suffer no pathological impairment, *contra* to what might be inferred from Damasio’s theory? Could the methods of the Noble Eightfold Path, which purport to effect this non-pathological transformation, be amenable to philosophical scrutiny such that their prima-facie plausibility might at least be established? I am fully convinced that these are questions on which significant philosophical headway can be made without undue departure from the traditions of analytic philosophy. My optimism stems partly from an edited-out chapter that addressed exactly these issues. My particular focus was on how the method of meditation could help undermine the sense of self in such a way that gives reason to suppose that pathology would be averted. The following is a brief synopsis of my thinking on this subject so far.

The general idea is that meditation would work, at least in part, by ‘reprogramming’ our usual patterns of attention so that the attention would no longer be compulsively ‘captured and lost’ in the content of those ‘storylines’ needed to preserve the sense of a bounded self. I argued, in line with Buddhism (and Damasio)\(^\text{132}\), that our ordinary patterns of attention are, to borrow a phrase from Hume, a “slave to the passions”, the passions being emotions and thoughts that help constitute *tanha* and the sense of a bounded self. While

---

\(^\text{132}\) Damasio writes: “To some degree, the message implied in the conscious state is: “Focused attention must be paid to X.” Consciousness [with sense of self] results in enhanced wakefulness and focused attention, both of which improve image processing for certain contents and can thus help optimize immediate and planned responses. (1999, 182) ...Emotion is critical for the appropriate direction of attention since it provides an automated signal about the organism’s past experience with given objects and thus provides a basis for assigning or withholding attention relative to a given object. (1999, 273) ...the consequences of having emotion and attention are entirely related to the fundamental business of managing life within the organism...” (1999, 274). On Damasio’s model, attention will quite naturally be directed by the emotions whose job it is to make sure – via the sense of self – that the bulk of focused-on objects are in some way relevant to the organism’s welfare. For this reason, Damasio conjectures that the neurological mechanisms governing attention and those processing emotion and homeostatic balance will be in the same vicinity (1999, 274). What emerges, then, is that the propensity for a person’s awareness to selectively attend to things of perceived interest to the apparent self – integral to *tanha* viz., emotional investment – is necessary for maintaining the sense of self-other boundedness.
enslaved, the attention is repeatedly drawn into storylines that implicitly depict the self as protagonist of recalled or imagined scenarios in the past and future. In the broadest terms, I proposed that Buddhist meditation would work via a system of training that gradually alters the patterns of attention so that they are no longer a ‘slave to the passions’. If such methods enable the attention to become knowingly freed from the grip of such deep-rooted conditioning, then it would make sense to suppose that the quality of attention will, as a result, become very powerful and present-centred. If attention is longer drawn compulsively and selectively to objects of perceived relevance to ‘self’, with its imagined timezones, then the tanha (emotional investment) which relies upon this ‘enslaved’ pattern of attention will be destroyed. When tanha is destroyed, then so too, on the two-tiered model of self-illusion, will the sense of self-other boundedness. The non-bounded and unconstructed awareness, which is focused during attention – and which serves as the vehicle for meditation – will remain intact.

I suggested that a differential quality in attentive capacity could partially explain why a loss to the sense of self through methods of meditation would be non-pathological, while a loss to the sense of self through neurological trauma would be pathological. In all the pathological cases (such as epileptic automatism, Alzheimer’s disease or akinetic mutism) Damasio has explicitly noted that the quality of attention is abnormally low. He has also noted that higher-quality attention is a reliable indicator of mental acuity (1999, 182-183). Because of this, the present-centredness of awareness through meditation would be just a distant cousin of that present-centredness forced upon patients undergoing an epileptic automatism. High quality attention is not (to my knowledge) found in any recognised pathologies that fully suspend the sense of a self. While this factor will not be enough to explain how a person could autonomously survive in the absence of what Damasio calls “over-riding individual concern”, the calibre of attention purportedly attained by those advanced in their meditation practice could well provide an important clue in understanding how pathology might be averted.

Within the framework of this project, a philosophical inquiry into how meditation might work (to undermine the sense of bounded self) could, and would probably need to, be supplemented by empirical research. Longitudinal studies may for instance be conducted on serious meditation practitioners to see whether their mode of practice does in fact undermine
their sense of self. In chapter eight we argued, with support from Damasio, that there is a sense of self if and only if there is *tanha* – the investment of emotion in various desired situations such that there is mental *dukkha* if the desires are frustrated. Damasio’s account suggested that such ‘boundary-preserving’ emotions are neurophysiological, with distinct regions of the brain and body involved in their activation. Given this supposition, it does not seem unreasonable to suppose that the presence or absence of *tanha* should in practice be empirically testable on several levels, from a subtle neurophysiological level (for example those concerning what Damasio calls the “proto-self”) to a more overtly physical level (for example galvanic skin-conductance tests in response to actual and imagined scenarios).

If it turned out that the sense of self could indeed be non-pathologically lost through meditation practice, the implications for the philosophy of mind would be far-reaching – regardless of whether an unconditioned *nibbanic consciousness* would truly be uncovered (an investigation that would in itself be entirely worth pursuing). For one thing, it would challenge the persuasive idea, forwarded by Damasio, that a sense of self is necessary for well-functioning consciousness. Damasio writes:

> When the mental aspect of self is suspended, the advantages of consciousness soon disappear. Individual life regulation is no longer possible in a complex environment. In the full personal social sense, individuals remain capable of basic and immediate bodily maintenance. But their connection to the environment on which they depend is broken down, and, because of the breakdown, they cannot sustain such bodily maintenance. In fact, left to their own devices, death would ensue in a matter of hours because bodily maintenance would collapse. This, and comparable examples, would suggest that a state of consciousness which encompasses a sense of self as conceptualized in this book is indispensable for survival (1999, 203-204).

Damasio’s conjecture is based upon his observation of such cases as epileptic automatism, Alzheimer’s Disease and akinetic mutism – and as far as these cases go, it is a reasonable conjecture to make. He is not entirely unaware, however, that an unexplored avenue exists, since at one point he mentions, somewhat hesitantly, that “impaired extended consciousness possibly contributes to the dissolution of self associated with states of depersonalization and with states of mystical [sic] selflessness” (216, 1999). Damasio shows no indication of

---

133 An experiment utilizing procedures for measuring emotion (involving galvanic skin conductance tests) is described by Damasio (1999) on pages 49-50. Similar procedures could be employed for measuring the presence or absence of *tanha*. 
actually having *researched* those cases that could provide the best counterexamples to his idea that a sense of self is *always* needed for unimpaired consciousness.\(^{134}\) If there *could* be a different and non-pathological way for the sense of self to be lost, but with the consciousness remaining unimpaired, then this would compel the philosopher of mind to think about consciousness and its capacities in a new and different light. For instance, one would have to explain how a person is motivated to act when there is *no* emotional investment in the results of any action. On this matter, the Eastern accounts of how ‘Arahants’ act in the world are telling. Not only are such persons reportedly able to ‘get by’: their ordinary actions – those normally associated with autonomy and survival – are depicted, without exception, as proceeding more effortlessly and efficiently than the comparable actions from persons *with* a sense of self. From an ethical perspective, their conduct is invariably described as exemplary in its virtue, wisdom and compassion – perhaps as close to Aristotle’s *phronimos* as one is likely to get. Should these reports be correct, they would raise a plethora of questions, including the question of how it would be possible to feel compassion – and *dukkha*-free happiness for that matter – in the absence of any ‘self’-propelled emotions. To turn Damasio’s conjecture on its head: could there be something about the ubiquitous sense of self and *tanha* that actually *prevents* consciousness from seeming to function at its best?

With these thoughts we leave the discussion.

\(^{134}\) I must confess to cringing each time I see that meme-encrusted word “mystical” being used to describe those ‘selfless’ modes of consciousness depicted in this project. It does *nothing* to dispel connotations of such consciousness being so mysterious as to be entirely out of philosophy of mind’s orbit and entirely unrelated to the consciousness of ‘Joe Bloggs’!
Bibliography


Bodhi, Bhikkhu. (1994) "Dhamma & Non-Duality - 1." Buddhist Publication Society Newsletter. 27


http://plato.stanford.edu/entries/emotion/.

http://www.chass.utoronto.ca/~sousa/subjectivity.html


http://ase.tufts.edu/cogstud/papers/chalmersdeb3df.htm

http://ase.tufts.edu/cogstud/papers/chalmers.htm. 4-6.

http://www.depersonalization.info/overview.html


Online version (slightly different) at http://www.imprint.co.uk/online/edey.html

http://www.mentalstates.net/SOC.html


http://www.zynet.co.uk/imprint/Forman.html. 185-201.


http://www.dreamchild.net/


http://www.enabling.org/ia/vipassana/Archive/G/GunaratanaMindfulnessIPE/


