



UNIVERSITÉ
DE GENÈVE

FACULTÉ DES LETTRES
Département de philosophie



The Unity of Consciousness

A Self-Representational Account

Antoine Rebourg

Supervisor: Professor Fabrice Teroni

Thesis for a Master of Arts in Philosophy with Specialization

June 2017

Contents

1. Introducing phenomenal unity	4
2. Kinds of consciousness, kinds of unity	9
2.1. Creature consciousness vs state consciousness.....	9
2.2. Phenomenal consciousness.....	10
2.2.1. Awareness, attention, reflection.....	11
2.2.2. Focal vs peripheral awareness	12
2.3. Kinds of unity.....	13
3. Mereological accounts	18
3.1. Individuating experiences	18
3.2. Unity relations.....	20
3.2.1. Co-consciousness	20
3.2.2. Subsumption.....	24
3.3. Objections.....	26
3.3.1. From introspection.....	26
3.3.2. Just more content.....	28
3.3.3. Against Block	30
3.4. Conclusions.....	32
4. The One-experience account	35
4.1. Representationalism and transparency	35
4.2. Unity relation: closure.....	37
4.3. Objections.....	39
4.3.1. Against the one-experience view	39
4.3.2. Against transparency.....	42
4.3.3. From the plurality of modes	44
4.4. Conclusions.....	47
5. The Self-Representational account	49
5.1. Self-awareness	49
5.1.1. Higher-order theories.....	52
5.1.2. A same-order theory: self-representationalism.....	55
5.1.3. Naturalizing self-representation	58
5.2. Phenomenal unity	61
5.2.1. Closure revisited.....	61
5.2.2. No more content	63

5.3.	The monitoring role of phenomenal unity.....	65
5.3.1.	Semantic consistency	66
5.3.2.	The normative 'I'	70
5.4.	Conclusions.....	75
6.	Objections and replies.....	77
6.1.	Against the coherence norms	77
6.1.1.	The norms' scope	77
6.1.2.	Non-conceptual creatures and coherence	78
6.2.	Against self-awareness.....	81
6.2.1.	Humean concerns.....	81
6.2.2.	Cartesian implications	82
6.3.	Phenomenal unity and the self	84
6.3.1.	The self and the 'I'	84
6.3.2.	The continuity of the self.....	88
7.	Split-brains	90
7.1.	The data.....	90
7.2.	Is there phenomenal disunity?.....	93
7.2.1.	The automaton interpretation	93
7.2.2.	The phenomenal unity interpretation.....	94
7.2.3.	The phenomenal disunity interpretation	96
7.3.	Is there more than one stream?.....	100
7.3.1.	The two-stream interpretation	100
7.3.2.	The momentary split interpretations	101
7.3.3.	The partially unified stream interpretation.....	105
8.	Conclusions.....	110
	Acknowledgments.....	114
	References	115

1. Introducing phenomenal unity

In the 17th century already, Descartes maintained that the mind was unified. Indeed, he famously premised the thesis of *substance dualism* on this claim:

Je remarque ici, premièrement, qu'il y a une grande différence entre l'esprit et le corps, en ce que le corps, de sa nature, est toujours divisible, et que l'esprit est entièrement indivisible. Car en effet, lorsque je considère mon esprit, c'est-à-dire moi-même en tant que je suis seulement une chose qui pense, je n'y puis distinguer aucunes parties, mais je me conçois comme une chose seule et entière. (Descartes 1647, VI)

The unity of the mind or 'consciousness', in contemporary terms, has been a central issue in the philosophy of mind ever since. Although Kant later refrained from the above Cartesian inference, he too would hold that consciousness exhibits unity, in the sense that a subject of a single act of consciousness represents objects as being tied together into a single complex object. He further took this unity to be a necessary condition to the unity of the self (Brook 2013: §3.4; Hurley 1994: 55). Other influential philosophers such as Hume, Reid, Brentano and James also endorsed the unity of consciousness in one form or the other.¹

Recent discussion of the unity of consciousness is quite vivid and has given rise to a wide spectrum of positions. On one side of this spectrum, some philosophers such as Bayne, Chalmers, Dainton and Tye hold that consciousness is *necessarily* unified; on the other side, philosophers like Dennett claim this unity to be illusory. On intermediate positions, consciousness breaks down on particular occasions (Hurley 1994 and 1998; Lockwood: 1989; Nagel 1971).

What is meant, exactly by the thesis that our consciousness is unified? My inquiry, like that of the theorists just mentioned, will be concerned with *phenomenal* consciousness. Thus my first order of business will obviously be to define phenomenal consciousness; one cannot hope to properly account for phenomenal unity without primarily defining what it is for a mental state to be phenomenally conscious.

While this has proved to be one of thorniest challenge of the philosophy of mind, and will indeed be the object of extended discussion throughout my work, I join Searle in thinking that a definition in common-sense terms might be sufficient – at least for now – to identify the target of our inquiry. Phenomenal consciousness bears the following features:

Qualitativeness: every conscious state that you undergo has a qualitative feel to it. There is something it is like to undergo each of them.

¹ For a thorough historical survey of these views, see Brook & Raymond 2006: 1056 and 2010: §1).

Subjectivity: every conscious state exists only when experienced by a certain subject, like an animal or a human being. In other words, consciousness has a first-person mode of existence. (Searle 2000: 559 ff.)²

Of course, this off-the-bat definition consists at best in necessary features of consciousness and will have to be further specified. I will tentatively complete and refine it as I treat the existing accounts of the unity of phenomenal consciousness. Thus my work will treat these two issues in parallel.

The unity of consciousness elicits two distinct questions. The first is whether a subject's consciousness is unified and continuous *over time* – whether it is *diachronically* unified. The second is whether a subject's consciousness is unified *at a given time* – whether it is *synchronically* unified. Although answers to either one certainly influence answers to the other, I will focus exclusively on the problem of synchronic unity. We will quickly see that it is plenty enough to have on our plate.

With these clarifications in hand, we can now adumbrate what is meant by the thesis that our consciousness is unified at a time. In our everyday conscious life, we entertain a great variety of conscious mental states, or *experiences*.³ Imagine what it is like to attend the concert of your all-time favorite singer Céline Dion. Among the crowd, you experience your environment through all *senses*. You contemplate the gorgeous silhouette of the artist on stage, and most particularly her glowing hair and broad smile; you hear her smooth voice interpret *Pour que tu m'aimes encore*; you feel the contact of the fellow fans that surround you; you smell the subtle mix of sweat, perfume and excitement that emanates from this crowd; you relish the lingering taste of the beer that you chugged down just the instant before. But your conscious experiences extend well beyond these perceptions. For instance, you undergo an *emotion* of sheer joy as you finally participate in this long-awaited event. Accordingly, in addition to understanding the lyrics of the song, you entertain the *thought* that you are incredibly lucky to be there. Finally, you subtly feel your own body being carried by your legs and being mildly tired from all the dancing and wriggling. All these experiences are *conscious*, in the sense that there is something it is like to undergo them; they each have a certain feel to them.

The thesis of phenomenal unity states that there is something it is like for you to enjoy *all* your experiences *together* at any instant, like the ones I have just described. Behind this is the rough idea that you do not live them in a piecemeal manner, but that they collectively form one global conscious experience.

² By contrast, eggplants and molecules have a *third-person* mode of existence: they can exist in the absence of any living creatures.

³ I will use the expressions 'conscious mental state', 'conscious state' or more simply 'experience' indifferently throughout my work.

Despite its *prima facie* intuitive appeal, the thesis of phenomenal unity appears rather mysterious and calls for clarifications at least on the following questions:

- 1) Under what conditions can a mental state be said to be (phenomenally) *conscious*, as opposed to ‘non-conscious’?
- 2) What is it that individuates a conscious state from another?
- 3) What is it exactly for a subject’s conscious states to be *unified* with each other? What is the *nature* of this relationship? What does it *achieve*?

Before cutting into these much-debated issues, a few remarks are in order. First, before evaluating the idea that consciousness is unified, should we ask ourselves whether consciousness *exists*? As Brook and Raymont (2010: §3) highlight, this question is independent from that of phenomenal unity. We may well indeed restrain ourselves from any commitment to the existence of consciousness, just as we may give an account of what Santa Claus is like without affirming his (or her) existence. But in order to forestall the objection that the unity debate is an idle one, I will endorse the widely accepted view that there is indeed something which we appropriately call ‘consciousness’ and about which we can make true statements. I am in other words what Dainton (2000: 21) calls a ‘robust realist’ about consciousness.

Secondly, in talking about consciousness in general I will assume that our conscious lives are mostly alike; I will therefore indulge, as most theorists do, in inductive inferences from introspection and others’ introspective reports to human consciousness as a type. This is of course not to deny that there are intersubjective differences in experience. But it is incumbent on the skeptics to show that anyone but me is a zombie, or that consciousness, and more particularly its structure and phenomenology, significantly differs across conscious subjects.

Thirdly, granted consciousness exists and is broadly similar across subjects, accounts about its unity need not make ‘[any] assumption about the ‘deep’ nature of consciousness, such as the relationship between consciousness and the physical world’ (Bayne 2010: 4). Here too I will however assume a neutral or ‘moderate’ naturalism (Dainton 2000: 10) according to which conscious experience is tightly linked to the physical world and the brain in particular (Searle 2000: 566-8).

Finally, I will follow Bayne (2010: 12) in laying out the constraints that a satisfying theory of unity ought to meet. First, it must be *substantive*, i.e. non-trivial with regard to one’s account of consciousness; secondly, it must be *plausible* as it should go without saying; thirdly, it must be *informative*, meaning that it should enrich our understanding of consciousness. I wish to make two quick additions (which are more or less tacitly assumed by Bayne) to these criteria. We should expect a satisfying theory of phenomenal unity to have some sort of *explanatory power* regarding mental phenomena. Moreover, its scope should be *comprehensive*, in the sense that it should account for the

unity of *all*, and not just a portion of, the ‘components’ – whatever they are – of a subject’s consciousness at a given time.

What are the reasons for investigating the problem of the unity of consciousness? Before commencing my inquiry, I can hazard at least three. First, as Descartes suggests, the unified structure of a subject’s consciousness may have implications with respect to material substance. Secondly, the unity of our consciousness may be the touchstone of other forms of unities that conscious creatures display. Indeed, it may explain why subjects act and reflect as one with more or less efficiency, and above all regard themselves and other subjects as one. More importantly, the unity of a creature’s consciousness is commonly taken, in one form or the other, to be inextricably linked to the questions of *what* and *how many* this creature *is*. In other words, the unity of consciousness may play a central role in this creature’s identity, or ‘self’, as was already defended in Kant’s work. Thirdly, assessing the conditions under which our consciousness is or is not unified may shed light on particular neuro-pathological cases, such as so-called ‘split-brain’ patients. Indeed, the conscious life that these subjects exhibit seems to differ from that of cognitively intact subjects of experience. Thus, even if our endeavor turns out to yield little motivation for the view that consciousness is unified, one can reasonably expect it to account for the difference between ‘normal’ consciousness and those particular pathological cases.

My discussion is organized in four steps. In chapter 2, I start off by laying out important distinctions intended to clarify the myriad ways in which different things are said to be conscious. I then quickly survey the variety of senses in which consciousness may be said to be unified, in order to put the thesis of phenomenal unity into perspective.

In chapters 3 and 4 I present two dominant accounts of phenomenal unity. The first, endorsed by Bayne, Chalmers and Dainton, rests on a mereological approach to consciousness while the second, Tye’s, rebukes this approach. After presenting each account, I highlight a few problems that it incurs; some bear on the theory of phenomenal consciousness that is assumed, others more specifically on the conception of phenomenal unity which is built thereon. While subjecting these accounts to critical scrutiny, I retain certain claims that sound to me promising and pivotal in the debate of phenomenal unity. Among other things, I endorse a broad version of representationalism on which an experience is best construed as a *representation*, consisting in a content represented in a psychological mode.

In chapter 5, I conciliate my representationalist inclinations with the long-standing (and still popular, as Searle shows) idea that consciousness essentially bears a *subjective character*. This allows me to propose an account of phenomenal unity on which the subject’s experiences are *representationally* unified by being the conjoint objects of a single act of *self-awareness*. This proposal is then enhanced by the claim that this unity is normatively constrained. In chapter 6, I address some

of the objections that could be made to my account on the following fronts: the normativity of our experiences, self-awareness and the relation of phenomenal unity to the self.

Finally, in chapter 7, as is customary among phenomenal unity theorists, I derive from my account an interpretation of split-brain experiments. This last step will provide one last occasion to stress the shortcomings of the other existing theories of phenomenal unity.

2. Kinds of consciousness, kinds of unity

Consciousness is, to borrow Block's (1995) famous phrase, a 'mongrel concept'. It is indeed endowed with a rich variety of meanings which the following cases should illustrate:

- (1) Yesterday after work, Philippe went out on a pub crawl and ended up seriously inebriated. On his way home, he passed out and lied unconscious for several hours on the sidewalk.
- (2) After undergoing an important surgical procedure for which it was sedated, my mom's cat regained consciousness.
- (3) As Mowgli fights the tiger, he is conscious of its slick moves.
- (4) As Maria chitchats with Sam, she becomes conscious of her attraction to him.
- (5) As Maria chitchats with Sam, her attraction to him becomes conscious.
- (6) Riding my bike to the university, my mind wanders and I am unconscious of what is going on around me. Nevertheless I seamlessly make it to my destination, which suggests that I am conscious of what is going on around me.⁴

These cases allow us to make the following observations. First, the property of being conscious is ascribed to different sorts of entities: in (1)–(4) it is ascribed to *creatures*, whereas in (5)–(6) it is ascribed to *mental states*. Second, there are two syntactic uses of the predicate 'is conscious': in (1)–(2) it is *intransitive* – the subject S is (un)conscious *simpliciter* – whereas in (3)–(4) it is *transitive* – S is conscious *of* something. Third, and foremost, the apparent plausibility and coherence of (6) suggests that it puts to use two distinct concepts of consciousness, which obtain under different conditions. In what follows I sketch a rough taxonomy that should help us distinguish between these various kinds of consciousness, and determine which is relevant to the thesis of the unity of consciousness.

2.1. Creature consciousness vs state consciousness

A first kind of consciousness is ascribed to organisms of all sorts, be it my friend Philippe or my mom's cat or a baby; it is therefore called 'creature consciousness'. A conscious creature is *sentient* and *awake* (Rosenthal 1997). But sentience, understood as its mere *capacity* or disposition to sense and respond to the world, may be insufficient to grant a creature the property of being conscious: consider the case (1) of Philippe, who is arguably sentient but unconscious nonetheless. Thus in (1)

⁴ Partly inspired from Tye 2003: 2-5.

and (2) consciousness seems better defined as wakefulness, construed as the exercising of this capacity such that the creature is *normally alert* (Van Gulick 2014: §2.1.), as opposed to asleep or comatose (Carruthers 2016: §1.). One can also say that a creature is conscious when it is *responsive*, that is, when it is ‘processing information about the external world and responding to it’ (Tye 2003: 6). Such definitions of this intransitive use of creature consciousness are fairly uncontroversial.

By contrast, cases (3)–(6) outline a transitive sense of creature consciousness. Mowgli is conscious *of* the tiger’s moves, Maria is conscious *of* her attraction for Sam, and so on. The fact that statement (5) seems a legitimate inference from (4) indicates that a creature’s transitive consciousness can be couched in terms of *mental states* consciousness. Consequently, one can say of Mowgli’s perceptual experience and of Maria’s attraction that they are conscious. This is what the question of the unity of consciousness is about: we aim to account for how the various conscious states of a given subject such as Mowgli, Maria or yourself are unified together at a time. But before engaging in this inquiry, it must be specified just what it means for a mental state to be conscious. And this, contrary to creature consciousness, is the subject of heated debate.

2.2. Phenomenal consciousness

A mental state is phenomenally conscious (simply ‘conscious’ henceforth) if there is *something that it is like* to be in it, i.e. if it has a certain qualitative ‘feel’ (Nagel 1974: 436; Block 1995: 228-30). As a terminological clarification, one should add that a mental state’s being conscious ‘is a matter of the state’s having phenomenal character, ... where this in turn is a matter of its having what are variously called experiential properties, phenomenal properties, sensory properties, or ‘qualia’ (Shoemaker 2003: 59).

We generally endow a conscious state such as Mowgli’s in (3) with at least the following properties: it is *intentional*, i.e. *about* something (here, about the tiger’s moves); it guides the subject’s reasoning and behavior (here, it is what makes Mowgli cautious and move fast); it has a physical or natural description (it ‘corresponds to’ certain oscillations in some part of the brain).

Yet the phenomenal property of Mowgli’s state – that is, what makes it *conscious* in the sense that we intend in everyday contexts – seems to have no logical connections to statements couched in intentional, functional or natural terms. This phenomenal property is thus what generates the famous ‘hard problem’ (Chalmers 1995) or ‘explanatory gap’ (Levine 1983) of consciousness.

2.2.1. Awareness, attention, reflection

Theorists frequently appeal to the notion of *awareness* in order to explain that of phenomenal consciousness: they say that a mental state is conscious when the subject is aware of it (Dainton 2000: 29). However, one must distinguish between awareness of a mental state and awareness of its *content*.

Thus (3) amounts to stating that Mowgli is aware of the content of his visual state, namely the tiger's slick moves. Here, the property of being aware of is to be understood in an 'attentive and recognitional sense' (Dainton 2000: 30); in other words, Mowgli's awareness of the tiger's moves just is Mowgli's *paying attention to* the tiger's moves. But it is not obvious that Mowgli's visual state is conscious in virtue of the fact that Mowgli is aware, in the same attentive sense, of this visual state *qua* mental state. Indeed, contrast (3) with (4). Contrary to Mowgli, Maria is aware of her attraction for Sam in the sense that she pays attention to her *mental state*. Since this attention is directed 'inwards', so to say, into her own mental life, it is generally referred to as an act of *reflective* awareness. So while Mowgli may be aware of his visual experience, he seems not to be reflectively aware of it.

Although reflective awareness is a particular kind of attention, it differs from it in two respects. First, while we constantly pay attention to the content of our experiences, we are only punctually reflectively aware of our experience. To put it differently, we normally do not focus on our experience, but rather on the properties and facts of the world *therethrough*, as shows Mowgli's case. A second difference between attention and reflective awareness is that the latter generally (though not always) non-inferentially gives rise to a judgment (Dainton 2000: 34). So while Mowgli's consciousness of the tiger's moves does not bring about any judgment with regard to this experience, Maria's reflective awareness of her attraction yields a judgment that she is attracted to Sam.

Still, there is a tight connection between attention and reflective awareness. When a subject is attentively aware of some content, it is normally possible for this experience to be an object of reflective awareness (Rosenthal: 1997:730). So, although Mowgli's awareness is not, during his fight, the object of reflection – he does not think that he is having a visual experience of the tiger's moves –, it *could* be.

Most theorists assimilate reflective awareness to *introspection*, or 'introspective awareness' (Tye 2003: 6). Introspection is 'attentive, deliberately focused consciousness of one's mental states' (Rosenthal *Ibid.*); it involves some sort of decision on the subject's part. However, it seems that you could spontaneously come to be aware of your experiences under certain circumstances. Maria's awareness of her attraction to Sam could result from her intention to probe her emotions; she would then be introspecting. But the awareness of her attraction could as well spontaneously 'pop up' in her experience; and this reflective awareness would not qualify as introspection.

2.2.2. Focal vs peripheral awareness

Case (6), finally, indicates that the scope of the subject's attention does not exhaust that of her awareness (and thus that of her consciousness). (6) can be rephrased as the following incoherent-sounding statement: while absent-mindedly riding my bike to work, I am both not aware and aware of my immediate surroundings (the cars, the pedestrians, etc.).

Transpiring from this observation is the idea that awareness comes in degrees. We can schematically distinguish between *focal* and *peripheral* awareness (Kriegel 2004: 190). This distinction is most vividly illustrated by visual experience: I am currently focally aware of the computer screen, and peripherally aware of the books on its left side. This contrast is due to the structure of the visual apparatus: in the eye, the fovea determines what we are focally aware of. But it applies just as well to other modalities of experiences, and *across* modalities. When absent-mindedly riding my bike, I am focally aware of what my mind wanders about but I am not *wholly unaware* of my surroundings; rather, I am *peripherally* aware of them, which is overwhelmingly suggested by the fact that my eyes are open and that I manage not to collide with any other vehicles. This remark foreshadows the idea that what one is aware of, even peripherally, has a certain impact on one's behavior (more on this in chapter 3).

A couple of metaphors should help the focal/peripheral distinction sink in. Focal awareness, which we normally call attention, is *serial*: it is directed at the content of only two or three modalities at most at a time (Bayne 2010: 79). For instance, as I type these words I am currently focally aware of the thought that I am currently typing, the computer screen that I am seeing and the sound of the keys. These experiences have each a certain phenomenal property: there something it is like to undergo them. However, there is far more to my phenomenal experience than these phenomenal properties: attention 'moves like a searchlight through the phenomenal field, but it doesn't define that field – there is plenty of phenomenology that falls outside of its beam' (O'Brien & Opie 1998: 383; see also Lockwood 1989: 162-3 for a similar metaphor). The phenomenology that falls outside this beam may be seen as the 'experiential penumbra' (Bayne: *Ibid*) or 'phenomenal background'. It is so called because it is more subtly experienced by the subject. In my case, it is composed *inter alia* of a certain proprioceptive feedback of my body,⁵ as well as a mild pain in the back and a vague feeling of relief at the prospect of wrapping up this section. There is a faint phenomenal feel to each of these mental states, which is why they are said to be conscious.

⁵ Conclusive proof of the constancy of such feedback in the subject phenomenal background can be found in Sacks 1987: 564 (quoted in O'Brien & Opie 1998: 384).

Notice that my attention constantly shifts from contents onto others, and that these variations involve variations in the phenomenal character of my experiences (Hill 1991: 118-26; Tye 2003: 10). For instance, the very act of mentioning the mild pain in the back that was in my phenomenal background an instant ago brought this pain to the foreground and made it more 'present', more phenomenally salient than it was before I paid attention to it.

Finally, if reflective awareness is regarded as a special kind of attention, it follows that we can just as well distinguish between focal and peripheral reflective awareness. In fact, I will later (in chapter 5) explore and defend the idea that while we are certainly not focally aware of our experiences at all times, we may be constantly peripherally aware of them.

For now, these numerous distinctions are intended to stress that our conscious experiences, at any given time, capture a great variety of features of the world through many modalities: perceptions, thoughts, bodily feelings and emotions. With this variety of experiences comes a very rich phenomenology that is by far not exhausted by the scope of our focal awareness.

This first characterization of phenomenal consciousness has involved the notions of awareness and attention, but admittedly these notions in turn cry out for explanations: what is it exactly for a subject to be aware of an experience, or of a given content? One will imminently consider different analyses of phenomenal consciousness, notably in terms of representations and functions. However, this broad characterization should suffice to provide us with an intuitive grasp on the notion.

Let us now get back to our inquiry. The thesis that a subject's consciousness is unified at a time is concerned with state consciousness, and more precisely with *phenomenal consciousness*. It states that all of a subject's conscious mental states are phenomenally unified. In other words, it is the thesis that there is something it is like for this subject to be in all these states *together*, in a perfectly unfragmented fashion.

2.3. Kinds of unity

Before elaborating on this precise thesis, one ought to consider the various senses of 'unity' in which a subject's experiences, or at least some of them, may be said to be unified. It may turn out that other theses of unity are substantive, plausible and informative, and thus fit our criteria for a satisfying theory of consciousness unity. If not, a survey of potential contenders is still likely to provide useful insights into what the phenomenal unity of consciousness is *not* about, as well as on the relations between the phenomenal unity and other unities.

Subject unity: a set of mental states are 'subject-unified' when they are entertained by the same subject at the same time. In other words, all my experiences are subject-unified at t just in virtue of the fact that they are *mine*, and not another subjects'. This is what Dainton (2000: 25) and Bayne (2004: 220) respectively dub 'consubjectivity' and 'co-ownership'. But consider the generic thesis of phenomenal unity: it holds that for any subject S , at any time t , S 's experiences are unified. If 'unified' is understood as 'subject-unified', then the unity thesis is 'not very informative' (Dainton *ibid*), not to say trivially true, since the unity of the subject S is stipulated in the very definition of the thesis. A further problem is that a subject is generally individuated by means of his experiences (Brook and Raymont 2010: §6.4). This apparent circularity will shortly dissipate with the distinction between a subject, understood as a physical basis, and a *self*, defined by a set of experiences (see sections 3.1. and 6.3.). Anyhow, there are intimate connections between subject unity and phenomenal unity. Bayne and Chalmers (2003: 26) hold that subject unity entails phenomenal unity. Indeed, as we shall see shortly, Bayne (2010) postulates *prima facie* that phenomenal unity is a *necessary* feature of S 's experiences.

Gestalt unity: Gestalt unity is the phenomenon whereby the experience of two different objects creates a third experience whose object differs greatly from the mere association of the first two. A famous example of this phenomenon is Rubin's vase, which is 'made up' out of two faces in profile. Thus, were one of the two faces removed, the experience of the vase would be replaced by the radically different experience of a face. To take another example the other way around, there may be cases where feeling a pain in one's shoulder while also experiencing a stomachache would engender a unique sort of pain that could not be experienced without the stomachache. In such cases, one could say that the pain in one's shoulder is 'gestalt-unified' with the stomachache (Bayne & Chalmers 2003: 43). Gestalt unity, however, is quite rare and does not extend beyond perception; one cannot expect it to shed light on phenomenal unity.

Objectual unity: a set of conscious states are *objectually unified* when they are directed at the same object. The most obvious examples of such unity are found in our perceptual experience. When walking down the street you simultaneously see the red light ahead and hear a dog barking behind you, your visual experience of the red light is not objectually unified with your auditory experience of the bark. But when you look at a particular blue car driving by, your experiences of the car's different features, namely its shape, color, motion and sound, are all objectually unified by virtue of being directed at one and the same object. This well-known phenomenon, called (*feature-*)*binding*, has yet to be properly elucidated (Brook & Raymont 2010: §2.2): how is it that the perception of color, shape, contours and so on, which are detected in different parts of the visual cortex, are somehow tied

together in such manner that the subject experiences a single three-dimensional object? Moreover, your visual experience of the car is one that combines the car's various parts together: 'the visual experience of a spatially extended thing [...] is a synthesis of visual experiences of parts of that thing' (Shoemaker 1996: 177). Your experience, finally, achieves 'multisensory integration': the car's visual *and* auditory features are bound together (Stein et al. 2002: 227).⁶

How does the binding phenomenon relate to our inquiry? Some regard binding as a motivation for endorsing the thesis that consciousness is unified:

The interest of the binding problem is that it looks like this problem might give us in microcosm a way of studying the nature of consciousness because just as the visual system binds all of the different stimulus inputs into a single unified visual percept, so the entire brain somehow unites all of the variety of our different stimulus inputs into a single unified conscious experience. (Searle 2000: 562)

However, objectual unity is not as such what the thesis of phenomenal unity is about; for while we undoubtedly have objectually unified experiences such as those of the blue car driving by, most of our experiences at a certain time are directed at different objects and are thus not unified in this way. For instance, your desire for an ice cream, or your thinking about the immortality of the soul, could not be objectually unified in any plausible sense with your visual experience of the car. This kind of unity is at best limited to perceptual experiences. Hence it does not have the comprehensive scope that the thesis of phenomenal unity is supposed to have.⁷

Spatial unity: a set of conscious states are spatially unified if '(i) each has a spatial representational content, and (ii) the representational content of each is *comparable*, in the sense that the objects represented are represented as being in spatial relations to each other' (Bayne & Chalmers 2003: 4, emphasis in original). For instance, my auditory experience of the keyboard is not directed at the same object as my visual experience of a computer screen, but both experiences are spatially unified in that they represent objects as being part of the same space: I can locate the keyboard and the screen as being adjacent to each other. Importantly, representational contents need not be part of the same *visual field* for experiences to be spatially unified. The mnemonic experience of my bike locked in the street is spatially unified with my visual experience of the computer in front of me insofar as I can locate both experienced items in the same space and thereby compare them in size, shape and so on. This is why Brook and Raymond (2006: 1057) name this kind of unity 'unity of consciousness of objects in the world'.

⁶ Note that this kind of unity is not restricted to the content of factive perceptions. If I hallucinated a red book, my experience of redness and that of rectangularity would still be unified (Bayne & Chalmers 2003: 3).

⁷ A related version of object unity is *conceptual unity*, which obtains 'when various objects or events can in principle be recognized to share some one general attribute: to fall under some one concept, to be diverse instances of some common property, tokens of some common type' (Hurley 1998: 90). However, such unity is equally too local to be of relevance to a theory of consciousness unity.

However, one quickly realizes that the scope of spatial unity is barely wider than that of objectual unity. The view that *perceptual* experiences are necessarily spatially unified may bear some plausibility (Tye 2003: 12), but it seems unable to unify other modalities of experience. It is not clear, for instance, how depression, or a belief about evidential justification, could have a spatially located content that would allow for meaningful comparisons with perceptual states. If this is true, then these sorts of experience are simply not spatially unified with the other perceptual experiences that are simultaneously entertained by the subject.

Neurophysiological unity: a set of conscious states may be said to be neurophysiologically unified if and only if the realization of these states ‘involves a single area or mechanism in the brain’ (Bayne & Chalmers 2003: 27; Tye 2003: 12).

There is no need to assess the truth of neurophysiological unity in order to determine the kind of relation that links it to the phenomenal unity of consciousness. The former pertains to a natural, sub-personal level whereas the latter concerns the subjective, personal level. As many philosophers have observed, it would be fallacious to project the structure and properties of one level on to the other (Hurley 2003: §4): it is obvious that when you see a blue car your brain need not be blue in order to represent the car’s blueness. As Connolly et al. (2011: 6) note, this point is vividly illustrated by the phenomenon of feature-binding considered above: some perceptual experiences are deemed unified because their *contents*, such as the car’s various features, are unified; yet one knows these experiences to be realized by *multiple* neural structures. We can therefore apply this observation to phenomenal consciousness: from the multiplicity, and perhaps disunity, of the neural correlates of a set of conscious states, we should not infer that these states are phenomenally disunified.

What comes out of these various kinds of unity at work in consciousness is that they cannot unify the entirety of the subject’s experiences at a given time. By contrast, the thesis of the phenomenal unity of consciousness is all-inclusive. Recall the initial example of your Céline Dion concert. The thesis of phenomenal unity states that at every instant, all your conscious experiences of all modalities – including (at least) your perceptions, bodily feelings, emotions and thoughts – are somehow tied together in such a way that you enjoy them together, as if they formed *one* total conscious experience. Given its exhaustive scope, one should not hope to analyze phenomenal unity with other kinds of unity.⁸

⁸ That is, unless the unity of consciousness is analyzed in terms of the *disjunction* of all the other kinds of unity, as some propose (see Hill 1991 and 2014: 501-2; also Masrour 2014 for the ‘Connectivity Thesis’). However, the very idea of a disjunctive analysis strips the unity off its genuineness and appeal. Furthermore, there is good chance that some of these unity relations have little to do with *phenomenal* consciousness.

Assessing the plausibility of the thesis of phenomenal unity will require that the following questions be answered:

- 1) What *is* an experience? What is it made of? How is it individuated from other experiences?
- 2) What is the *nature* of the relation that holds a subject's experiences together?
- 3) What are the main motivations for, and implications of, the thesis of phenomenal unity?

Replies to questions 1) cleave into two opposite approaches to consciousness. Bayne's, Chalmers and Dainton's is *mereological*, in the sense that they take the subject's 'global' state of consciousness at a time to be composed of smaller experiences – what Brook and Raymond (2010) call 'experiential parts' –, the individuation of which will be addressed first-hand. On this approach, you can distinguish between your visual experience of Céline Dion on stage, your auditory experience of her smooth voice, your emotion of delight and so on, because each is an experience in its own right (Bayne 2010: 5-6).

The opposite approach admits of no experiential parts. It finds its origin in James, who regarded the 'mind-stuff' theory, according to which 'our mental states are composite in structure, made up of smaller states conjoined', as 'one of the obscurest of assumptions' (James 1890: 145-6). Most recently championed by Tye (2003), it is the view that a subject enjoys just one complex conscious experience at a time. It can already be anticipated that the two approaches yield very different answers for questions 2) and 3), but I will not spoil these for you. Let us now take a critical look at these accounts.

3. Mereological accounts

3.1. Individuating experiences

Let us start with the following statement of the thesis of the phenomenal unity of consciousness – or ‘unity thesis’ for short –:

Unity Thesis: Necessarily, for any conscious subject of experience *S* and any time *t*, the simultaneous conscious states that *S* has at *t* will be subsumed by a single conscious state – the subject’s total conscious state.⁹ (Bayne & Chalmers 2003; Bayne 2010: 16)

As it has just been observed, mereological accounts of experience immediately raise the question of how a subject’s experiences should be individuated. Mereologists invoke the following ‘tripartite’ identity criterion:

I will assume that token experiences owe their individuality to three factors: their exact phenomenal character, their time of occurrence, and their physical basis... I will not speculate exactly what form this physical basis takes. (Dainton 2000: 25 and 189)

An experience is an instantiation of a phenomenal property, where a phenomenal property is a property that there is ‘something it is like’ to enjoy... A token experience can be equated with i) the instantiation of a phenomenal property ii) by subject iii) at a time. (Bayne 2010: 7)

Notice first that Dainton’s criterion differs from Bayne’s in that he favors a ‘physical basis’ factor over that of a ‘subject’. One of the reason for this preference is that ‘we do not need to know what kind of thing selves or subjects are – or even assume there are such things – to embark into a phenomenological inquiry into [the] relationships [between experiences]’ (Dainton: *Ibid*). I will simply observe for our present purposes that Bayne’s and Dainton’s criteria can be deemed equivalent since Bayne assumes, for the most part of his development at least, a biological conception of the self (Bayne 2010: 9). Anyhow, I will use Bayne’s afore quoted terminology because it better suits our linguistic habits: we ascribe experiences to *subjects*.¹⁰

Secondly, a little more should be said as to the phenomenal property (or phenomenal ‘character’) of an experience. Phenomenal properties are ‘properties that are individuated in terms of what it is like to have them’ (Bayne 2010: 70). According to Bayne, they ought to be distinguished from intentional states: intentional states are not individuated in terms of what it is like to instantiate them, but in terms of their attitude/mode (belief, desire, perception, etc.) and content (that *p*, that *q*, etc.)

⁹ For clarity’s sake, one should add that since the unity thesis concerns only conscious states, it concerns only *token* states; types of state cannot be conscious.

¹⁰ This terminology has the further advantage of accommodating, superficially at least, with anti-naturalist views of consciousness.

(*Ibid*). We will shortly see that this distinction is open to debate, and that some theorists actually *identify* phenomenal properties and intentional states (see sub-sections 4.3.2 and 4.3.3.). Finally, the instantiation of a phenomenal property, i.e. the undergoing of a phenomenally conscious mental state, can be called a *phenomenal event*, since events can be construed as instantiations of properties (Kim 1976, quoted in Bayne 2010: 25).

This last remark brings us to the delicate issue of the temporal structure of an experience. Here, the time t at which an experience occurs may be best understood as a ‘specious present’, a phrase coined by E. R. Clay (1882) and famously defined by James as a very short temporal interval, rather than a snap shot or ‘knife-edge’ instant (James 1890: 609). Indeed, one usually takes experiences (or at least *some* of them) to extend over time, ever so slightly (Dainton 2000: 23 and 171). I will not develop this issue any further in this work. I will simply assume this vague characterization of t , since I take it to be intuitive enough for the continuation of our discussion.^{11 12}

Here is how the tripartite criterion is readily justified. Let e_1 and e_2 be two experiences. If e_1 instantiates a different phenomenal property than e_2 , e_1 is obviously a different experience than e_2 . The same conclusion can be drawn if e_1 and e_2 happen at different times. But how are we supposed to distinguish between e_1 and e_2 if they happen simultaneously and have the exact same phenomenal character? In light of the fact that e_1 and e_2 are experienced by different subjects.

But consider the phenomenal property of seeing a certain shade of red. On the tripartite account, a subject S 's visual experience of the shade of red is individuated by its particular phenomenal feel, its occurrence at a certain time and its being S 's experience. Now suppose that instead, S is having at t a visual experience of multiple objects – at different locations in her visual field – that have the exact same shade of red. While there is a strong intuitive pull towards the claim that she has *multiple* experiences of red, the tripartite account entails that S has a single experience and is thus found wanting. It should therefore be added that the phenomenal property at play in the individuation of an experience must be *fine-grained*, or *maximally specific*. A phenomenal property is maximally specific if it has no determinates (Bayne 2010: 25; Bayne & Chalmers 2003: 45). Since, in our example, the phenomenal property of the shade of red has determinates, namely the phenomenal properties of red at specific locations, it is insufficient to individuate an experience; and since its determinates have themselves no determinates, they are maximally specific and are therefore individuating.

So far, so good. We turn now to the task of characterizing what it is, exactly, for a subject's experiences to be phenomenally unified. Mereological approaches to the unity of consciousness

¹¹ Also, it may very well be that the notion of specious present, or something akin to it, is necessary to account for the *diachronic* unity of conscious experience.

¹² For a quick overview of the problems involved with the notion, see Bayne 2001: 85-7.

address this issue by postulating a *unity relation* that ties experiences together. On Dainton's (2000) proposal, the relation that lies at the core of the unity thesis is that of *co-consciousness*. Thus put, the unity of consciousness is the thesis that all of a subject's experiences at *t* are co-conscious with one another. By contrast, Bayne and Chalmers (2003) submit that the unity relation is that of *subsumption*. On their view, for a subject's consciousness to be unified at *t* is for all her experiences to be subsumed by a further encompassing experience.

What is the difference between co-consciousness and subsumption? Although both relations entail a mereological approach to consciousness, the subsumption view is top-down in that it takes as its starting point the subsuming state, and then differentiate it into simpler – and *in fine* atomic – states that compose it. The co-consciousness view, on the contrary, is bottom-up: it starts from experiential parts and builds up an overall experience with them.¹³ Let us now inquire into the nature of these two relations.

3.2. Unity relations

3.2.1. Co-consciousness

First introduced by James (1909: 221), the term 'co-conscious' really became popular in Parfit's writings. For Parfit, for two distinct experiences to be co-conscious is for them to be 'the object of a single state of awareness' (Parfit 1984: 250). Similarly, on Dainton's introductory example of co-consciousness, when you snap your fingers, 'you see and feel a movement, and hear a sound. These three experiences — one auditory, one visual and one tactile — do not occur in isolation from one another, they occur together within your consciousness, you are aware of them all at once' (Dainton 2000: xiii).¹⁴

Since, on the unity thesis, phenomenal unity is a *constant* feature of our experiences, *S*'s awareness of her experiences must not be understood in a *focal* sense, as we have characterized reflective awareness as a punctual exercise. Moreover, even if *S*'s reflective awareness of her

¹³ That said, Dainton (2000: 91-5) concedes that an honorable case can be made in favor of the view – which he calls *primitivism* – that an entire stream is the basic experiential unit.

¹⁴ Importantly, the notion of co-consciousness invoked by mereological theorists such as Dainton and Bayne should not be conflated with James'. James had an anti-mereological approach to consciousness; consequently, he did not mean co-consciousness to be a relationship binding experiential parts. Rather he understood it as the relationship that stands between all the items of which the subject is conscious (Brook & Raymond 2010: §6.2); his conception is thus akin to Tye's, (see chapter 4).

experiences were constant, it would not likely target *all* of her experiences. These precisions laid out, we can now encapsulate Dainton's position as follows:

For a subject *S*'s experiences to be co-conscious at *t* is for *S* to be non-reflectively aware of her experiences conjointly at *t*.

While this definition sounds promising and provides a helping hand in roughly grasping co-consciousness, it leaves us wondering: what is it for *S* to be non-reflectively aware of her experiences, and what is it for *S* to be non-reflectively aware of *all* her experiences?

A good prospect would be to analyze *S*'s awareness of her mental state in terms of the particular *function(s)* that this state fulfills. Thus, one may hope to analyze the phenomenal relation that binds experiences together in functional terms. However, both Dainton's and Bayne and Chalmers' accounts resist this analysis.

Underlying the reluctance to reduce co-consciousness to another relation is Block's (1995) distinction between two kinds of consciousness: phenomenal consciousness ('P-consciousness' henceforth), on the one hand, and access-consciousness ('A-consciousness' henceforth), on the other.

A mental state is P-conscious if there is something it is like to be in this state. By contrast, a mental state is A-conscious 'if, in virtue of one's having the state, a representation of its content is (i) inferentially promiscuous..., that is, poised for use as a premise in reasoning, (ii) poised for rational control of action, and (iii) poised for rational control of speech' (Block 1995: 231).¹⁵ Note that the property of being A-conscious is purely dispositional: a mental state need not be actually accessed and used for reasoning, rational control and so on in order to be A-conscious; it need only be 'poised', *accessible* that is, for such purposes (Kriegel 2006b: 7-8).

According to Block, P-consciousness and A-consciousness often go hand in hand: a mental state that is P-conscious is A-conscious in most cases, and vice versa, which is why the two kinds are often conflated. However, Block argues, the two should be distinguished. A thorough justification of this distinction goes beyond the remit of my work. However, as it bears significant implications on the thesis of phenomenal unity, I shall briefly present the chief motivations for it.

Block's *conceptual* argument is itself inspired by the experiments conducted on subjects that suffer from blindsight ('blindsighters' for short). Due to brain damage in the occipital cortex, blindsighters have large blind areas in their visual field. Nonetheless some of them are able to make accurate judgments about some of the features of objects that are in their blind fields, such as presence, position, orientation and movement. They can for instance discriminate X-shaped objects

¹⁵ Nonetheless Block rightly notes that condition (iii) may not be necessary for A-consciousness, as it is plausible that 'very much lower animals are A-conscious' (*Ibid.*: 238).

from O-shaped objects. However, blindsighters issue reports about the objects that are in their blind fields only when forced to choose between multiple choices; moreover, they cannot identify even the objects that fall into categories with which they are familiar. For example, a blindsighter will fail to recognize, and reach out for, a glass of water even though he is thirsty.

From these data, it has been generally inferred that there is no conscious representation of the objects in the visual blind field. There is nothing it is like for a blindsighter to see the glass. Block claims that the fact that the patient does not use the representation of the glass for verbal report, reasoning and action only suggests that what he lacks is A-consciousness of the glass of water; the patient's inaction and lack of verbal report is not due to a lack of P-consciousness. He goes on to argue that even though the blindsighter most likely lacks both A- *and* P-consciousness in the cited cases, it is still conceptually possible that he lacks only the former but retains the latter (Block 1995: 242). Block takes this conceptual possibility to be empirically confirmed by Sperling's experiments, which I will leave to Bayne to present:

In [Sperling [1960](#) and Averbach & Sperling 1961's] experiments, subjects are presented with a matrix containing twelve or so alphanumeric figures for a brief period (say, 250 milliseconds). There are two kinds of conditions, a *full report* condition and a *partial report* condition. In the full report condition subjects are required to report the contents of the entire matrix. (p.81) Typically, subjects are able to correctly report only 4.3 of the twelve figures (on average). In the partial report condition, a tone is sounded immediately after the presentation of the matrix indicating which of the three rows the subject is to report; for example, a high tone indicates that subjects are to report only the four figures in the top row. On such trials, trained subjects are able to report three of the four figures in the cued row (on average). In other words, subjects are able to report more figures with respect to a row that has been cued (after display offset) than they are with respect to any arbitrary uncued row. (Bayne 2010: 80-1)

Block's and Bayne's interpretation of these data is that subjects are P-conscious of *all* the letters, jointly and at once, as they are flashed, but that they are not A-conscious of all jointly. There is a bottleneck that restricts the subject's access (access for report, for instance) to the entirety of the experiential, P-conscious that is, content.

Conversely, can a state be A-conscious without being P-conscious? Advocating for the possibility of zombies or 'superblindsighters' that function just like us without there being anything it is like for them to function so, Block answers in the affirmative (Block: *ibid.*).

What are the implications of Block's distinction on the thesis of phenomenal unity? Dainton understands the conjoint awareness of the subject's experiences in the *phenomenal sense* only. This means that even though S's experiences are phenomenally unified, in the sense that there is something it is like for her to undergo them *together*, they could fail to be A-conscious together: they could fail

to be conjointly accessible to S for rational reasoning and behavior. Hence, S's conjoint awareness of her experiences could not be analyzed in functional terms.¹⁶

Dainton postulates that co-consciousness is a 'basic experiential relationship', a *primitive* feature of consciousness of which one can spell out only the logical properties (Dainton 2000: 84 and 88 ff.; see also Lockwood 1989: 88). First, co-consciousness is a *material* rather than formal relation: its relata are experiences and experiences are not abstract entities. In other words, co-consciousness is a concrete relation that stands between concrete particulars, experiences. Secondly, there is logically no upper limit to the number of experiences that it can take as its relata. This means that a subject can theoretically entertain as many co-conscious experiences at *t* as one wishes. The lower limit of co-consciousness' adicity is one: an experience is co-conscious with itself. Therefore, co-consciousness is *reflexive*. Thirdly, it is *symmetrical*: if an experience *e1* is co-conscious with another experience *e2*, then *e2* is co-conscious with *e1*. Fourthly, co-consciousness does not come in degrees: it is an all-or-nothing relation. Project yourself again in your Céline Dion concert. You are fully enjoying the show with very rich visual and auditory experiences. Standing in a packed crowd, your arms are in direct contact with some of your fellow Céline Dion enthusiasts and you experience it, even if you pay very little – if any – attention to it. On Dainton's view, although you are certainly more aware of the co-consciousness that holds between the visual and auditory experiences than of the co-consciousness of these experiences with that of the tactile contact with your neighbors, the three experiences – together with all the others that you entertain at this instant – are co-conscious with each other to the same degree. There may be some kind of indeterminacy with regard to the phenomenal character of the experiences that you are peripherally aware of: for instance, you would not know whether it is leather or synthetic clothing that you are inadvertently touching. But the existence of this tactile experience is, *qua* phenomenal item, not indeterminate, and neither is its co-consciousness with your other, more salient, experiences. Finally, there is the question of whether co-consciousness is *transitive*. Could there be three experiences *e1*, *e2* and *e3* such that *e1* is co-conscious with *e2*, *e2* is co-conscious *e3* but *e1* and *e3* are *not* co-conscious? Paradigmatic cases suggest not, but there may be a few marginal cases of intransitive co-consciousness. I will get back to this issue in section 7.3., upon assessing the possibility that a subject's consciousness is, in particular circumstances, *partially unified*.

In addition to claiming that a subject's experiences are co-conscious at any given time, mereologists claim that an additional overarching experience *subsumes* these experiences. Hence they postulate another, 'vertical' unity relation of *subsumption*. Let us now characterize this relation and see how it complements that of co-consciousness.

¹⁶ Thus Schechter seems to misunderstand Dainton's 'conjoint awareness' in terms of the subject's 'simultaneous access to' her experiences (Schechter 2010b: 2-3).

3.2.2. Subsumption

Mereologists need 'a way of combining experiences that yields more than just an experiential aggregate, for a mere combination of experiences is not the experience of a combination' (Brook & Raymont 2006: 1059). This observation originates from James' following remarks:

Take a hundred [feelings], shuffle them and pack them as close together as you can (whatever that may mean); still each remains the same feeling as it always was, shut in its own skin, windowless, ignorant of what the other feelings are and mean...Take a sentence of a dozen words, take twelve men, and to each one word. Then stand the men in a row or jam them in a bunch, and let each one think of his word as intently as he will; nowhere will there be a consciousness of the whole sentence. (James 1890: 160, quoted in Bayne 2010: 34)

A restatement of the unity thesis should suffice to alleviate this worry. Given that the unity thesis is concerned with the putative unity of a *single* subject's experiences at a time, it is not threatened by James' challenge. A problem would arise only if the thoughts of the twelve words were entertained by a single subject. Let us suppose such a case (if it is ever possible): a single subject simultaneously entertains twelve thoughts of a single word each but has no experience of the twelve words unified into a sentence.

A crucial point, and one that is repeatedly stressed by theorists, is that the subject's phenomenally unified experiences have a 'conjoint phenomenology' (Bayne 2010:10-11) which is not tantamount to a mere conjunction of the phenomenal character of each experience. Rather, what the subject enjoys is a different, *total experience* which encompasses all her experiences (Bayne & Chalmers 2003: 23; Dainton 2000: 188).

According to Bayne and Chalmers, the relation that ties the subject's so-called total experience to each of her individual experiences is that of subsumption. Formally, when S entertains a set of experiences e_1 , e_2 , and e_3 (at t), S entertains a numerically distinct, multimodal experience e_4 which subsumes e_1 , e_2 and e_3 ; it is the joint occurrence, but not the mere aggregation, of e_1 , e_2 and e_3 ; e_4 is an experience in its own right.

The subsumption relation has been the object of many definitions. Bayne and Chalmers (2003: 39 ff.) repeatedly hold that e_1 and e_2 are unified by being *aspects* of a total experience e_3 . Alternatively, they try and define the notion in terms of the what-it-is-likeness of experiences: an experience e_3 subsumes other experiences e_1 and e_2 when what it is like to have e_3 , e_2 and e_1 simultaneously is the same as what it is like to have e_3 . However, they quickly renounce this definition because of the lack of consensus as to what is the what-it-is-likeness of an experience.

Bayne and Chalmers then state that an experience subsumes another (set of) experience(s) when the former entails the latter.¹⁷ According to Bayne (2010: 70), *e3* subsumes *e1* and *e2* when *e3* includes *e2* and *e3* as proper parts. These local characterizations are of assistance in understanding the notion, but subsumption is generally treated as ‘an intuitive primitive’; consequently, as Dainton does with co-consciousness, Bayne and Chalmers spell out its logical properties by way of definition. Subsumption, like co-consciousness, holds between token experiences and is therefore a *material* relation. It is *reflexive*: an experience can and in fact does subsume itself.¹⁸ It is *antisymmetrical*: for any two distinct experiences *e1* and *e2*, if *e1* subsumes *e2*, then *e2* does not subsume *e1*; another way of formulating this is that if it is both true that *e1* subsumes *e2* and that *e2* subsumes *e1*, then *e1* is numerically identical with *e2*. Finally, it is *transitive*: if *e3* subsumes *e2* and *e2* subsumes *e1*, then *e3* subsumes *e1* (Bayne & Chalmers: *Ibid.*).

Are the co-consciousness and the subsumption frameworks compatible? The two accounts appear to be largely coherent with each other and inter-translatable. It is true, as we have seen, that the two unity relations that they posit differ: co-consciousness ties experiential parts together whereas subsumption is the relation that holds between the total experience and each of its experiential parts. But on the subsumption account, experiential parts are regularly said to be ‘co-subsumed’ (Bayne 2010: 20), which hints at the relation of co-consciousness; similarly, on the co-consciousness account, Dainton claims that co-conscious experiences ‘compose’ a total experience (Dainton 2000: 95), which points to a part-whole relation much like subsumption.^{19 20 21}

Dainton (2004) has argued that co-consciousness presents the slight advantage of relating exclusively to mental states, whereas subsumption holds between other sorts of entity such as the ocean and its watery parts. Since co-consciousness is occasionally referred to as ‘togetherness’, a relation which too ties many other things than experiences, I think that this point is terminological rather than metaphysical. At any rate, this difference can be deemed negligible.

¹⁷ It is a simplification because there are very specific cases of gestalt unity where a phenomenal state entails another without subsuming it (Bayne & Chalmers 2003: 43. See also Bayne 2002: §2.6.).

¹⁸ A problematic implication of this claim will be highlighted in sub-section 3.3.3.

¹⁹ Amusingly, Dainton also employs a metaphor to name the relation between the experiential parts and the total experience: he says that experiences ‘fuse into a single unit of experience’ (2000: 105, my emphasis). Both Dainton and Bayne seem reluctant to explicitly parallel unity relations with (and thus completely reduce them to) part-whole relations.

²⁰ This is confirmed by the fact that the two unity relations face the same challenge in postulating a ‘total experience’, as will be demonstrated in sub-section 3.3.2.

²¹ In fact, Dainton and Bayne themselves emit doubt as to whether both accounts really are rivalrous (Bayne 2001: 82; Dainton 2006: 256-7).

I have now laid out the core tenets of the mereological accounts of phenomenal unity. A subject of experience entertains at any time t several 'atomic' experiences that are individuated by their respective phenomenal property, time of occurrence and subject. These experiences are co-conscious with each other at t . But this is not to say that they are a mere aggregate of experiences: they are together *subsumed* by a total experience.

I shall now present three objections to these views. The first bears on the epistemic grounds on which the unity relation is postulated. The second concerns the incoherence inherent to the notions of unity relations and total experience. The third has to do with the limited force of both mereological accounts.

3.3. Objections

3.3.1. From introspection

Current among unity theorists is the thought that we can have direct introspective access to the unity of our phenomenal consciousness:

The plausibility of the unity thesis derives largely from introspection. Consider the structure of your overall conscious state. I suspect that you will be inclined to the view that all your current experiences are phenomenally unified with each other – that they occur as components of a single phenomenal field; to put the same point in different terminology, that you enjoy a single phenomenal state that subsumes them all. (Bayne 2010: 75)

Unity judgment: by introspection I come to the conclusion that all my current states are phenomenally unified with one another. This unity is a permanent feature of consciousness, even when we don't introspect. (Bayne 2010: §4)

Our streams of consciousness do display a distinctive sort of unity, and this unity does not just consist in a relationship between certain experiences, it consists in a relationship between experiences that is itself experienced. (Dainton 2000: 4, 88)

Co-consciousness is such a primitive feature of experience, is it not plausible to think its basic phenomenal characteristics will be discernible through introspection? (Dainton 2000: 108)

One way to go about rejecting the unity thesis is to undercut its introspective justification. In order to do so, one could suggest that introspection is unreliable. This strategy, however, quickly proves self-undermining: since any inquiry about consciousness relies (at a personal, subjective level) on introspection – either directly via the subject's own introspection, or indirectly via other subjects' introspective report –, positing its lack of reliability precludes any judgment whatsoever about

consciousness. On pain of falling into general, paralyzing skepticism, and *faute de mieux*, one must accept that introspection is a basic and *bona fide* source of justification of judgments about the mental.

But then, how come no general consensus is reached about something so allegedly obvious as the phenomenal unity of our experiences? How is it that some people claim to introspectively notice this unity, while others profess the opposite? Here are some Humean reservations regarding the fruits of our introspective inquiries into our phenomenal experience:

If we are not aware of our experiences via introspection, we are not aware of them as unified. The unity relation is not given to us introspectively as a relation connecting experiences. Why, then, suppose that there is such a relation at all? (Tye 2003: 25)

I now feel that this view is wrong. It isn't possible to find this ghostly form of co-consciousness within one's experience. Hence, there is no reason to believe it exists. (Hill 1991: 239)

Although I am very much inclined to grant that a unity relation such as co-consciousness is phenomenally salient when I introspect, there are two problems with this epistemic ground.

The first problem is that the scope of introspection is not as wide as that of our phenomenal experience. Recall that introspection is a form of attention that a subject pays to some of her experiences. Recall moreover that attention has a limited scope vis-à-vis the contents of your experience: the subject is aware of more than what she pays attention to. Likewise, but at one level higher, introspection fails to grasp the entirety of the subject's experiences (Bayne 2010: 81). For instance, as Maria suddenly engages in an introspective reflection about how she feels about Sam while chitchatting with him, she reflects about some of her experiences, such as her attraction to him and the thought that she had never come to realize it before. But while so reflecting, she is surely entertaining other experiences about which she does not reflect, such as the audio-visual perception of the handsome man before her eyes, some bodily feelings, etc. Consequently, provided she experiences the unity relation as she introspects, she experiences it to hold only between the experiences that fall within the scope of this introspective glance. Introspection cannot in any direct fashion tell her whether the same relation binds these introspected experiences to the unintrospected ones; nor *a fortiori* can it tell her whether this relation binds the unintrospected experiences with each other.²²

I think that this first problem is surmountable. One could for instance hold, as Bayne does (2010: §4.4), that the fact that all experiences are co-conscious and thus phenomenally unified follows from a legitimate inductive generalization. Alternatively, it could be held (and I will later argue something along those lines) that one need not even introspect in order to experience the unity relation. In this line of thought, it may be argued that a subject is *aware*, in a non-reflective fashion and at all times, of

²² Following a similar line, Masrour suggests that introspection is 'subject to the limitations of working memory' (Masrour 2014: 513).

this unity relation. Hence the unity relations could be taken to hold as widely as the subject's awareness grasps. This would mean that regardless of whether the subject conducts an introspective exercise, all bits of phenomenal experience, including those in the 'penumbra', are experienced by the subjects *as being co-conscious with each other*.

Even if this is conceded, the introspective support for the unity relation is bedeviled by one of the mereological accounts' main presuppositions, namely Block's distinction between access (A-) and phenomenal (P-) consciousness. On this distinction, an experience could be P-conscious but not A-conscious. But then how could we take a P-conscious experience to be co-conscious with the subject's other experiences, if the subject has not even a remote or potential access to (is not even aware of) this experience?

Of course, what comes out of this last remark is not that the unity thesis is falsified. It only stresses that introspection falls short of providing direct and *positive* support for it. Assuming it is a *bona fide* access to our consciousness, the best that introspection entitles us to infer is that A-conscious experiences, i.e. experiences that are susceptible of falling within the introspective beam, are unified.

3.3.2. Just more content

Both subsumption and co-consciousness versions of the unity thesis posit that the subject's experiences compose a total experience. It is also naturally posited that this total experience is *maximal*, in the sense that there is no other experiences of which it is a part. The resulting problem is that mereologists face Russell's paradox: the very notion of a total experience is incoherent.

In the subsumption framework, a maximal phenomenal experience is an experience that is not subsumed by any other experience. Take again the set of unified experiences e_1 , e_2 , and e_3 . They are co-subsumed by the total experience e_4 , which is itself subsumed by no other experience. But as is stipulated by the unity thesis, e_4 must itself be unified with the states e_1 – e_3 that it subsumes. Therefore there must be a further (supposedly) maximal experience e_5 that subsumes e_1 – e_4 in order for them to be unified. But then e_4 is not maximal after all, and since e_5 is unified with e_1 – e_4 , there must be yet another subsuming state, and so on; a vicious regress of maximal experiences appears.

The problem arises *mutatis mutandis* in the co-consciousness framework: the total experience is 'maximal' in the sense that it is composed of 'a group of experiences which are all co-conscious with one another and which are not parts of a larger group of experiences which are all co-conscious with one another' (Dainton 2000: 95-6). But the total experience must itself be co-conscious with its parts (Bayne 2001: 84).

A first way of stopping the regress would be to reply that the subsuming experience e_4 need not be unified with e_1 – e_3 . The problem with this response is that it clashes with the unity thesis itself, according to which *all* of the subject's states at t must be unified. If the subsuming experience e_4 is indeed an experience as Bayne claims it is, it must be unified with the others; the fact that the latter are its parts changes nothing to it. This problem is equally salient in Dainton's formulation of the unity thesis. Dainton holds that 'there is no one right way to divide a subject's overall experience over a given interval into parts. As is plain, *no matter which division is considered*, all the relevant parts are related by co-consciousness...' (2000: 84, emphasis mine). Since there is one division according to which the total experience and all the co-conscious states are experiences in their own right, the former is co-conscious with the latter.

Another way of defusing the objection may be to reply that since subsumption, like parthood, is a reflexive relation – any experience (improperly) subsumes itself –, the subsuming state that unifies e_1 – e_4 is e_4 itself; so there need not be a state e_5 for e_4 to be unified with e_1 – e_3 . (Bayne 2010: 29). This is tantamount to replying, in the co-consciousness framework, that the set of co-conscious experiences e_1 – e_4 compose nothing other than e_4 itself. Ironically, Bayne is content with his own reply but expresses reservations regarding its co-consciousness version:

Dainton seems to regard co-consciousness as a *constructive* relation, a relation that enables one to build complex experiences out of (numerically distinct) simpler experiences. This constructive role for co-consciousness seems to be somewhat threatened if certain numerically distinct experiences, namely those that stand in the part–whole relation, can be co-consciousness without the product of this relation resulting in an experience which is greater than either constituent of the relation. (Bayne 2001: 84-5)

I think that this concern can be similarly applied to Bayne's own conception. The relation of subsumption has force in explaining the unity of consciousness only insofar as it is conceived as *proper* inclusion; once this restriction is waived, there is not much appeal left to it. In fact, it is explicitly stated by Bayne in his own account: an experience subsumes other experiences when it *includes* them *as proper parts* (2010: 70, quoted *supra*).

The so-called 'Just more content' objection, raised by Hurley (1994 and 1998: 70 ff., 97 ff.) and Tye (2003: 22), resembles the last in that it pinpoints yet another regress. As implied by the unity thesis, the unity relation that holds between experiences e_1 and e_2 is something that the subject experiences. If this relation was not experienced, there would be no phenomenal difference between a case where e_1 – e_2 are unified and a case where e_1 – e_2 are not; in other words, there would not be anything it is like for a set of experiences to be unified. The unifying relation is itself an experience, then. But this raises the question of how this experienced unity relation is itself unified with e_1 – e_2 . One must

postulate (so the objection goes) yet another unity relation that binds the first one with e_1 – e_2 . Consequently, mereologists are irresistibly wrapped up in a regress of unity relations.

To block this objection, Dainton's (2000: 215) move is to argue that the unity relation (co-consciousness) is an experience of a different nature than the others: it binds other states together and binds itself to them without requiring another relation. It is, in short, 'self-binding'. Similar to this line of response is the claim that although the unity relation has a certain phenomenal quality, it is not itself an experience but rather an 'experiential relation'. These replies have a very little appeal for they raise a whole new set of questions: what is the difference between an experience and an experiential relation? Why should some relations be self-binding while others are not? It looks like both distinctions are *ad hoc*.

Another way to defuse the threat of regress would be to reject Tye's charge according to which the unity relation must itself be an experience in order to make a phenomenal difference. The gist of this reply is that there is something it is like for the subject to undergo phenomenally unified experiences, but there is nothing it is like to experience this unity relation *in particular*. In Bayne's words, 'unity is not an *object* of experience but a *manner* of experiencing' (Bayne 2010: 31-2, emphasis in original).

In order to defend the phenomenal unity of our experiences without falling into regress, I will later develop a similar idea (see in particular 4.3.2., 4.3.3. and 5.2.). However, I must emphasize that this strategy is not *available* to the theorists under consideration. Indeed, the idea that unity is a manner of experiencing without being something that one experiences stands in contradiction with the tripartite individuation of experiences that is presupposed by the unity thesis: given that the unity relation makes a phenomenal difference, it bears a certain phenomenal property. And since this phenomenal property is instantiated by a subject at a time, it fulfills all three conditions for being an individuated experience in its own right (Tye 2003: 22). To sum up, holding that phenomenal unity is a manner of experiencing but not an experience is an option only insofar as the tripartite criterion is jettisoned, or at least refined.

3.3.3. Against Block

A last but not least concern about mereological accounts bears on the fact that co-consciousness, conceived as an irreducible phenomenal relation, lacks the informativeness that it should provide. This might reveal a problem with the definition of phenomenal consciousness that underlies it. Let me briefly motivate this diagnosis.

First, the conceptual distinction between A and P-consciousness relies on imaginary cases such as superblindsighters, as well as on the possibility that blindsighters *could* entertain P-consciousness of the glass of water despite the lack of A-consciousness. But on the other hand, Block concedes that the thirsty blindsighter lacks both A- and P-consciousness on the basis of the subject's denial that she has such an experience. While I do not take issue directly with these possibilities, they give rise to a methodological conundrum: if a mental state's being P-conscious is independent of whether the subject is aware of this property, then the blindsighter's denial that she has an experience of the glass cannot be taken as evidence that he lacks P-consciousness of it. So if the subject's testimony is to bear any weight in determining whether a state is P-conscious, then P-consciousness requires either that the subject be A-conscious of it, or at least that it be possible for him to be introspectively aware of it – for instance, when he is elicited to (Shoemaker 2003: 60-61; Dainton 2000: 32 ff. acknowledges this point).

Second, let us consider the empirical grounds that supposedly justify the distinction. We mentioned first Sperling's (1961) experiment: on Block's view, the visual experience of the whole matrix is P-conscious but not A-conscious. To put it another way, there is something it is like for subjects to experience the whole matrix, but they have access to (most of) the information of only *each row separately*; they are unable to use the *conjunction* of these row-specific contents for report.

There is however another interpretation of this discrepancy. When the matrix is flashed, subjects are both P and A-conscious of all of it, which explains why they are able to report the content of *any* row when prompted to. But we could hypothesize that the act of reporting either row 'deletes' both P-consciousness and A-consciousness of the whole matrix, which explains just as well why subjects are unable, after the report, to report any *other* row (Kobes 1995: 260). In other words, it may very well be that the subjects' complete P-consciousness of the matrix and their 'piecemeal' A-consciousness thereof obtain at *different* times.

Let me now turn to a quite different putative case of P-consciousness without A-consciousness:

Suppose you are engaged in intense conversation when suddenly at noon you realize that right outside your window there is - and has been for some time - a deafening pneumatic drill digging up the street. You were aware of the noise all along, but only at noon are you *consciously* aware of it. That is, you were P-conscious of the noise all along, but at noon you are both P-conscious *and* A-conscious of it. (Block 1995: 234)

This case parallels that of the absent-minded bike ride considered in section 2.1. Indeed, Block would say that during my bike ride, my perception of the surroundings is P-conscious but not A-conscious. But just as one has argued that my surroundings are accessible – though plausibly not *accessed* – for guiding reasoning and rational behavior, and thus A-conscious, one can too wager that

the pneumatic drill is A-conscious. So Block's picture of the pneumatic drilling is at odds with his own dispositional conception of A-consciousness.

It is undeniable, however, that there is a difference between the experience of the intense conversation and that of the pneumatic drilling (and similarly between my mind-wandering experience and that of my surroundings). This difference can be captured by means of the focal/peripheral awareness distinction previously canvassed (in 2.2.2.), and it can be even further illustrated by the following contrast: your experience of the intense conversation has a richer and more conceptualized content than that of the drilling outside (Dennett 1995: 253) (the same goes for the bike ride example). This conceptual difference explains why the former experience has an impact on reasoning and rational behavior that the latter lacks. But the mere fact that the experience of the pneumatic drilling *could* have such impact suffices for it to be A-conscious.

I should emphatically stress that I am not claiming that conceiving of co-consciousness as a primitive relation is erroneous just because it is premised on Block's distinction. Clearly, the above considerations do not show Block to be wrong. Rather, they point out that Block's conceptual distinction is more of a promissory note that is not likely to be ever cashed, since it is not at all obvious that P-consciousness *does* empirically obtain without A-consciousness. Moreover, the second example suggests that other distinctions within state consciousness more convincingly capture the nuances of our conscious experiences; indeed, it may be safer to conceive of all experiences in terms of representations that vary in richness of content and conceptualization. This means that the P/A-consciousness distinction may be more properly understood as one of *degree* and not of *kind*. Accordingly, there may be more to the phenomenal unity relation than its primitiveness.

3.4. Conclusions

It is time to recap. This chapter has focused on two compatible mereological defenses of the thesis that consciousness is phenomenally unified. Here are their four core tenets:

- (i) At any time, a subject undergoes several 'atomic' experiences that are individuated by their respective time of occurrence, subject and phenomenal property.
- (ii) All these experiences are unified by the primitive phenomenal relation of co-consciousness.
- (iii) Together they compose (are subsumed by) a total experience which differs from their mere aggregation.

- (iv) The phenomenal unity of our experiences is something that the subject experiences and to which she has an introspective access.

I hope that my development has sufficed to highlight that these accounts suffer problems on all fronts. On the theoretical front, claims (ii) and (iii) cannot be held together if claim (i) is true. For if (i) is true, experiential parts and the total experience are experiences *in the same sense*; and if they are experiences in the same sense, they are all co-conscious with each other in a way that entails a regress of total experiences. Similarly, claims (ii) and (iv) cannot be held together without entailing a painful regress of unity relations.

On the methodological front, a little more ought to be said about the mereologists' global argumentative strategy. Aside from appealing to the introspective salience of the unity relation, Bayne, Chalmers and Dainton adopt what may be called a 'wait-and-see' stance regarding the unity thesis: they take its truth for granted at the onset of their inquiry, and then address and reject potential objections, one at a time, in order to preserve it.

Motivation for this program stems from the fact that the unity thesis accommodates our strongest intuitions regarding our conscious lives: there is good sense in conceiving of our mental landscape as one, and to treat 'atomic' experiences – if such there be – that occur in its midst not only *in relation to* this unity, but in relation to one another. Thus Bayne and Chalmers (2003: 40) claim that there is a strong *prima facie* case in favor of the unity thesis and Bayne later states that '[the unity thesis] ought to be accorded a kind of default status – 'innocent until proven guilty'' (2010: 93). Transpiring from these claims is the clear intent to shift, onto the skeptics, the onus of the proof that consciousness is or can be *disunified*.

This wait-and-see strategy seems to me ill-advised in three important respects. First, it posits a fresh new set of entities that populates our mental ontology, namely new relations – subsumption and co-consciousness – and a further 'total' experience that encompasses all the others. This, I think, should be sufficient for the onus of the proof to befall the advocates of the unity thesis; and we have come to the conclusion that the resort to introspection falls short of providing substantive support for claim (ii) – that co-consciousness binds *all* the experiences that the subject undergoes at a given time –. Introspection does not capture the entirety of these experiences. Moreover, on Block's distinction, there could be P-conscious experiences that completely elude the reach of both introspection and awareness.

Secondly, and more importantly, the unity thesis lacks the informativeness that it ought to bear. The notion of 'phenomenal field' invoked by Bayne (2010: 75, quoted *supra*) is too metaphorical to be of much help in grasping that of 'total experience' (Van Gulick 2014: 492). Similarly, the irreducibility of the co-consciousness relation sheds as such very little light as to what it *is* exactly for a subject's

experience to be unified. Its logical properties are spelled out quite dogmatically and without other justification than the *prima facie* intuition of complete phenomenal unity. But decreeing that co-consciousness is a transitive relation is tantamount to little more than *restating* the unity thesis.

This brings me to my third point. It seems to me that the only legitimate reason for positing the phenomenal unity of consciousness would be its explanatory import. But thus stated, the unity thesis does not tell us at all what phenomenal unity concretely *achieves* in one's conscious life. I will show in chapter 7 that when confronted with certain neurophysiological conditions, it is completely otiose.

However, we have seen that the primitiveness of the unity relation is imputable to Block's dubious conception of phenomenal consciousness. I have argued that P-consciousness is likely to be more tightly linked to A-consciousness than Block would have us believe. If this idea finds convincing resonance, there might be hope after all: the unity relation may be enlightened by an analysis of phenomenology in terms of *representations* and/or their *functions*. In the following chapter, I evaluate Tye's development of this conception.

4. The One-experience account

As suggested by its name, Michael Tye's 'one-experience' view is that it is mistaken to couch the question of whether consciousness is unified in mereological terms, for there is simply no such things as experiential parts, understood as atomic experiences that are individuated by their phenomenal character, time of occurrence and subject. For instance, when you have a strong desire for a pint of beer and a sandwich, you do not have three distinct experiences of desire (a desire for a pint of beer, a desire for sandwich and a desire for both) but a single experience of desire that can be described in more or less rich ways. The same goes for other experiences, such as perceptions, and across modalities: 'sense-specific experiences... are the figments of philosophers' and psychologists' imagination. There are no experiences to be unified... There is a single multimodal experience, describable in more or less rich ways' but indecomposable in smaller or simpler conscious states (Tye 2003: 28-36). What this entails is that when you hear a loud sound and see a bright light at the same time, there is but a single multi-modal experience of a loud sound and bright light. This, again, is compatible with the claim that the sentences 'You are hearing a loud sound' and 'You are seeing a bright flash' are both (independently from each other) correct, albeit incomplete, descriptions of your experience.

4.1. Representationalism and transparency

Central to Tye's account of phenomenal unity is the thesis of representationalism. In rough outline, representationalism is the view that consciousness is best explained in terms of representations. Representations of 'information' and their functions were already at the heart of Block's definition of A-consciousness (Block 1995: 231, quoted *supra*). However, representationalism differs from Block's conception in a crucial respect: it ontologically binds P-consciousness with A-consciousness. The particular nature of this bond is the bone of contention among thinkers. Since the one-experience view is taken to be compatible with any version of representationalism (Tye 2003: xiv), I will simply sketch a disjunctive definition on which I need not adjudicate:

Representationalism: a conscious state's phenomenal character supervenes on, or is identical with, its representational content (Lycan 2015 : §2.1.; Bayne 2010: 52).

One can simply note that Tye's own version is of the strong variety: 'the phenomenal character of an experience *is one and the same as* its poised, abstract, non-conceptual, intentional content, or its PANIC for short' (Tye 2003: 176, emphasis mine). Importantly, though, Tye does not claim that an experience's content is wholly and *exclusively* PANIC ; rather, he takes PANIC to be a *part* of the experience's content, that is, the part 'which endows the experience with its phenomenal character' (Tye 2003: 36). For the time being, I need not specify what it is for an experience's content to be a PANIC. I shall simply mention that an experience's content is *non-conceptual* in the sense that it can represent properties of the world of which the subject lacks the concept. To simplify a bit, if you have for instance a visual experience of a tree, it will have a phenomenal 'tree-ish' feel even though you lack the concept of a tree.

A core tenet of strong representationalism *à la* Tye is the idea that our experiences are transparent. The argument, although never fully explicit in Tye's (2003), could be spelled out as follows. Experiences are transparent: the subject is aware of certain qualities of the world *through* them, but she is not directly aware of *them*, even when she introspects. To employ Tye's own metaphor, it is as though the subject were in a completely glass-walled room through which she sees the world. Never is she acquainted, even via introspection, with the qualities of the glass walls. Hence – the argument goes – one could not appeal to such qualities in order to individuate experiential parts. There are only representational contents, understood as the qualities of the world that are represented; and if there is to ever be an experience, there must be just the one (*Ibid.*: 25).

How does this picture of conscious experience concretely unfold? Immerse yourself once again into the Céline Dion concert. You are experiencing a great deal of perceptual qualities: the gloss of the singer's hair, the subtle tan of her skin, the smoothness of her voice, the smell of the sweaty unleashed crowd, etc. Furthermore, you are experiencing 'one single general bodily feeling', which is akin to a perceptual experience in that it gives you a 'body image' that represents its movements, its general location as well as the location of its part relative to the torso (*Ibid.*: 48). In our scenario, you experience a single complex bodily feeling that represent your whole body rocking to the rhythm, your hands clapping and your left foot frantically beating to the point where you are feeling a slight pain in it.²³ These experiences, moreover, are accompanied at this moment by an emotion, termed 'mood' in Tye's jargon, of sheer ecstasy. It consists partly in the representation of your bodily state and of its changes relatively to the moments when the mood was not experienced yet.

²³ The experience of pain has often been invoked as an objection to representationalism on the ground that it has a certain phenomenal character but does not seem to represent anything. Tye dissolves the problem by arguing that a pain in a particular body part *b* represents some tissue damage in *b* (*Ibid.*: 54-5).

Notice also that on this representational picture of phenomenology, *b* need not exist in order for a pain or sensation to obtain in *b*. All that is required is that *b*, and the relevant disturbance in it, be represented in the body image. Thus the reality of so-called 'phantom pains' or 'phantom sensations' in amputee is accounted for.

At last, you are entertaining at this instant a certain thought, for example a judgment whose content is ‘Boy am I lucky to witness such grandeur!’ Matters regarding thoughts are a bit tricky here. Tye holds that ‘the phenomenology of occurrent thoughts derives fundamentally... from the phenomenology of their associated linguistic, auditory images’ (*Ibid.*: 79). This auditory image consists in a sentence, with phonological, stress and even syntactic phenomenal features that you hear yourself speak in an ‘inner language’, concomitantly with your thought.

The claim that thoughts do not and cannot as such bear their own phenomenal character seems to follow directly from the combination of the PANIC view – a state’s phenomenal character just is its *non-conceptual* content – with the plausible claim that the contents of thoughts are essentially and wholly conceptual.²⁴ However, since Tye’s conservatism *à propos* cognitive phenomenology is by no means essential to his account of phenomenal unity, and since the broad version of representationalism that I laid out before is agnostic regarding the conceptual status of representational contents, I will liberally assume that thoughts, and cognitive states in general, have a phenomenology of their own.^{25 26}

4.2. Unity relation: closure

So far I have described the variety of representational contents that is experienced; so good. But the unification of these various experienced qualities is yet to be achieved. It is at this point that Tye’s view on phenomenal unity departs from the mereologists’. Since there are no experiential parts but only one experience, there is no need to posit a relation, akin to co-consciousness, that unifies *experiences*. If there is to ever be such a thing as a unity relation, it must hold between the various *contents* represented by the subject’s experience. The thesis that naturally follows from this is that for an experience to be phenomenally unified is for its various representational contents to be *representationally unified*. To start off, note that we examined two species of representational unity in section 2.3.: object unity and spatial unity. Recall that object unity obtains between experiences of the same object, and spatial unity obtains between experiences of spatially comparable contents. We

²⁴ It is more properly motivated by the following thought experiments (Tye 2003: 79–80). Get back to your concert and imagine entertaining the thought, a belief more specifically, whose content is ‘Boy am I lucky to witness such grandeur!’ Now imagine ridding this thought of its auditory image. There does not remain, Tye argues, a thought with a phenomenal character. For a thorough argumentation against cognitive phenomenology, see Tye & Wright (2011).

²⁵ I borrow the ‘liberal/conservatist’ jargon from Bayne (2010: 6) and Bayne & Montague (2011: ch. 1).

²⁶ The more phenomenal material an account can unify, the more robust the account!

have come to the conclusion that the problem of these unities is that they are too ‘local’ within our conscious experience; they are unable to account for the unity of consciousness across *all* modalities.

On Tye’s theory, representational unity is broadly deployed so as to account for the unity of representational contents *across the board*. Dive into your favorite Canadian artist’s concert once again. At a particular instant, you undergo a unique experience that represents the gloss of the singer’s hair *and* the smoothness of her voice *and* the foul smell of the crowd’s sweat *and* some tissue damage in your foot *and* body changes through which a particular emotion manifests itself, etc. These qualities are unified by virtue of entering the same ‘phenomenal content’; they are thus said to be ‘closed under conjunction’ (*Ibid.*: 36-7 and 76). To sum up, for you to entertain a phenomenally unified experience is for these various contents to be conjointly represented by a total experience.

Before evaluating this account, two crucial points should be stressed. First, the idea that qualities can be closed into a single content may lack intuitive resonance. If it is of any help, one could just as well understand experienced qualities in *propositional* terms, as Bayne does (2010: 51).²⁷ Instead of being a representation *of* a quality, an experience would amount to the representation *to the effect that p*, where *p* refers to the instantiation of this quality (at a certain space-time location). It is in this spirit that Hurley proposes the following ‘agglomeration principle’:

When there is a state in *i* at *t* with content *p*
 And there is a state in *i* at *t* with content *q*,
 Then there is a state in *i* at *t* with content *p* and *q*.

...Contents of consciousness agglomerate when the values taken by *i* for different contents match. (Hurley 1998: 117)²⁸

Secondly, as is suggested by the agglomeration principle, the closure of content is compatible with a mereological approach to consciousness: it could hold between contents of *several* experiences. In effect, other theorists characterize representational unity as ‘the relation that two [or any] experiences bear to one another when their representational contents are integrated in some way’ (Hill 2014: 500).

In what follows I consider three objections that could be made to Tye’s account. The first targets Tye’s anti-mereological stance on consciousness. Importantly, as the second point that I have just made indicates, this criticism does *not* entail that experiences are not unified by their representational

²⁷ A motivation for this interpretation is that Tye later addresses the issue of the semantic inconsistency of content, and inconsistency is a property of sets of propositions. However, for expository purposes, I shall indifferently talk in terms of propositions or properties, and I will take all that is said about one form of content to hold *mutatis mutandis* for the other.

²⁸ Let the placeholder *i* stand for a subject of experience, or more vaguely for ‘whatever it is that is the basis of agglomeration’. Note also that this principle ‘says nothing about what sort of thing *i* takes as a value or how to individuate the values of *i*’ (*Ibid.*).

content. I will in fact argue that while Tye's one-experience view is erroneous, representational unity could still hold; indeed, it could be the touchstone of phenomenal unity.

The second and third objections will then focus on two potential problems of representational unity, namely the alleged transparency of experiences and the idea that a single experience could 'represent it all'. Interestingly, then, the last objection will raise worries with regard to both Tye's notion of 'multi-modal experience' *and* mereologists' notion of 'total experience' (see chapter 3).

4.3. Objections

4.3.1. Against the one-experience view

The first problem of Tye's theory is simply that talk about a single experience is at odds with common sense (Bayne 2005: 497). Although I am inclined to share Bayne's intuition that we ought to conceive of consciousness in mereological terms, common sense arguments have limited force. In fact, Tye may even marshal a commonsense argument *in favor of* the one-experience view. After all, talk about a creature's consciousness, in the singular form, may stem from our pre-theoretical anti-mereological conception of consciousness. A creature's wakeful state may just be a neutral way of referring to a unique, multimodal but non-differentiated experience.

A slightly more convincing objection against an anti-mereological view is the fact that experiential modalities can represent things quite independently from each other. For instance, you could have heard a loud sound while seeing something completely different than a bright light, or vice versa. You could also have heard a loud sound while having no visual experience *at all*. Imagine that you block representations of all modalities but your audition, in such a way that your experience only represents the instantiation of an auditory property. It has a certain phenomenal property. I cannot see why any additional representation that you may experience, being independent of your auditory representation and having a phenomenal property of its own, should not equally count as an experience in its own right.

In fact, in developing his account of phenomenal unity, Tye seems unable to do without fragmenting the subject's experience. He starts off by characterizing multisensory perceptual experience as *the single* experience – so far so good – but then goes on to integrate into it, step by step, what seems to be yet other experiences, viz. bodily experiences, thoughts and moods. Of course, it is not clear how much weight such linguistic and methodological considerations should be afforded, but they might suffice to show that Tye's outspoken anti-mereological position is merely

terminological – I will get back to this point in the next sub-section. It is, at least, how rival theorists have understood it (Bayne 2005: 498; Dainton 2006: 253).

Another way of rejecting Tye's view is simply to rebut the arguments that are invoked in its favor. What are Tye's motivations for sustaining a one-experience conception? The first consists in the following argument from analogy. Consider a statue of clay or a cloud. The statue is naturally said to have lumps of clay as proper parts. But these lumps of clay are not themselves statues. Even in the special case where a clever sculptor carves a smaller statue out of the original statue, it would be odd to say that the original statue contained the smaller statue (as its part) before the sculptor's work. A cloud, likewise, has collections of water droplets as proper parts, but these collections are not typically clouds. Now a conscious experience, Tye argues, ontologically behaves much like the statue and the cloud: one can say that a subject's conscious experience at a given time is made of proper parts, but there is no reason for regarding these parts as *experiences* in their own right (Tye 2003: 30-3, 99).

This line of reasoning faces many problems. At the outset, we can remark that parts of statues and clouds *do* sometimes qualify as statues and clouds. So even if a subject's experience is indeed similar to a statue or a cloud, this analogy only stresses that it is not the case that *any* part of the experience qualifies as an experience. No mereologist would have trouble agreeing with this claim.

But we could more radically question the premise that an experience is like a cloud or a statue. Practically, distinguishing between parts of clouds and parts of statues does not prove as necessary as distinguishing bits of conscious experience (Bayne 2010: 23). One often invokes experiential parts in order to account for behaviors or even for other experiential parts; it seems that a total, coarse-grained experience would hardly fulfill such explanatory purposes. Of course, Tye does not deny that the sentence 'You are hearing a loud sound' would correctly, albeit incompletely, describe your unique experience and thus explain (say) that you shudder. But then it is not clear why this sentence could not refer to an experience in its own right.

In fact, Tye's own PANIC conception emphasizes this point. On PANIC, 'the phenomenal character of an experience is one and the same as its poised, abstract, non-conceptual, intentional content' (Tye 2003: 176, quoted *supra*). It is now that a clarification as to what 'poised' means becomes relevant:

The key idea is that experiences and feelings, *qua* bearers of phenomenal character, play a certain distinctive functional role. They arise at the interface of the nonconceptual and conceptual domains, and they stand ready and available to make direct impact on beliefs and/or desires. For example, how things phenomenally look typically causes certain cognitive responses—in particular, beliefs as to how they are if attention is properly focused. Feeling hungry likewise has an immediate cognitive effect, namely the desire to eat. In the case of feeling pain, the typical cognitive effect is the desire to protect the body, to move away from

what is perceived to be producing pain. And so on. States with nonconceptual content that are not so poised lack phenomenal character. (Tye 2000: 62)²⁹

On Tye's proposal, an experience is individuated by its phenomenal character, which means that it is individuated by its PANIC. So for the one-experience view to be consistent with a PANIC's approach to phenomenology, Tye is committed to admitting but *one* PANIC. But surely, many 'parts' of the content of one's total experience are (independently from each other) sufficiently 'poised' for fulfilling each a functional role, and therefore qualify as PANIC. Let me recapitulate: many parts of the one-experience's content are PANICs in their own right. But since a PANIC is what individuates an experience, there should be many experiences. Tye's turns out to be a mereological view in all but name.

The second motivation for the one-experience view is negative: it consists in rejecting mereological accounts of phenomenal unity. It involves the following line of thought: since phenomenal unity is a necessary feature of consciousness, and since postulating that the unity relation holds between experiences entails, as we have seen, a vicious regress of experiences, we must recuse the view that the unity relation holds between experiences. The only alternative, Tye argues, is the one-experience view on which the unity relation holds between *contents* of a single experience.

But this is a blatant *non sequitur* for there is a third alternative, and a quite obvious one, that we already envisaged when characterizing representational unity: there are several experiences, but they are unified by the integration of their representational content. So rejecting Tye's theory altogether may be throwing the baby out with the bathwater. One may argue that even though the anti-mereological conception of conscious experience collapses, Tye's account casts some interesting light on the issue of the unity of consciousness. Despite the plurality of experiences entertained by a subject at *t*, perhaps the thesis of representationalism can be put to good use; perhaps the phenomenal unity relation can be analyzed by the relation that holds between said experiences' representational content.

However, if my remarks undercut Tye's approach to consciousness as I suggest they do, the integration of representational contents, and thus the phenomenal unity of a subject's experience, has at this point no firm grounds upon which to be built. Indeed, on Tye's view, the unification of representational contents is achieved by their entering a single 'phenomenal content' (Tye 2003: 36); but one quickly realizes that the oneness of this phenomenal content is itself derivative of the oneness of *S*'s experience. Therefore contents are unified solely in virtue of the unity of the experience that represents them. To put this point differently, Tye has it that the things represented by your experience are unified by virtue of being closed under conjunction. But they are so closed only under the

²⁹ Notice that this definition of 'poised' is intended as defining a poised *state* and not a poised content. However, a poised content is just the content of a poised state. I will leave this nuance aside.

assumption that there is but one experience. Now that a shadow of suspicion has been cast on the claim the subject's experience is unique, the task of unifying representations remains to be done.

4.3.2. Against transparency

There is yet another motivation for the anti-mereological approach that I have deliberately not addressed in the last sub-section, namely the transparency of a subject's experience. Tye's argument is that since experiences are phenomenally transparent, nothing allows us to individuate and postulate several of them. I will now briefly argue against this assumption, which should put the final nail in the anti-mereologist's coffin (should any doubt have remained after the criticisms just made). The rejection of transparency will also serve to lay the grounds for my own view on phenomenal unity.

Recall that the claim that your experience is transparent is underlain by the fact that when you undergo an audio-visual experience of Céline Dion on stage, what you are aware of is only what is represented, namely that Céline Dion has a smooth voice, that she is gorgeous under the spotlight, and so on. You are not aware of vision, nor of audition, nor of any other modalities *per se*. While this claim is acceptable, it does not follow that these modalities are transparent, in a phenomenological sense. On the contrary, there is a certain 'feel' to each of them; and consequently, they contribute to the phenomenal character of your experience.

On Tye's view, endorsing the transparency of experience means that there are no such things as experiential parts; there are representational contents. But this view entails not only that the task of accounting for the phenomenal complexity of the subject's experience is left entirely to the complexity of these contents, but also that the similarity between two experiences must be accounted for only in terms of content. Is this picture right? A couple of simple cases conclusively show that it is not.

What is the common feature of my experience of a red book on my desk today and my experience of a huge tree in my backyard tomorrow?³⁰ I am not sure how these experiences' contents alone could satisfactorily account for this similarity. The proper answer surely is that both are tokens of a single psychological mode type, vision, and this mode is pinpointed in both experiences by means of a proper phenomenology. Again, Tye's phrasing betrays him, as he repeatedly talks of the 'multi-modal' character of the subject's single experience. And again, I do not think that this point is merely terminological. Talk about modes, and the multiplicity thereof, commits Tye to individuating mode types; and this cannot be done without ascribing to the tokens of each mode a proper phenomenology.

One can reach this conclusion from the other direction, as Crane does:

³⁰ I locate both experiences at different times to thwart the reply that they are part of a single diachronically unified stream of consciousness. Classically, and on Tye's proposal, a unified stream is taken to extend from wake-up to the last instant before you fall asleep.

The content of a state is part of what individuates that state: that is, what distinguishes it from all other states. The other thing which distinguishes an intentional state is whether it is a belief, or a desire, or hope or whatever. Again following Searle, I call this aspect of the state – the relation which relates the subject to the content – the ‘intentional mode’. (Crane 2003: 8)

It is indeed very likely that there are several ways of representing one and the same content: I can for instance see, know, believe, intuit, hope, desire, imagine that p , and these all seem to be *different* experiences, albeit of the same content. Endorsing Tye’s view entails one of the following two claims:

- (i) These experiences are just one and the same experiential item (under different descriptions).
- (ii) These experiences’ content is wrongly referred to as a common p , as each experience has in reality a distinctive content that individuates it from the others.

Claim (i) is simply unpalatable. As for (ii), it is far from clear how a new description of each experience’s content could do a proper job of individuating them. It is not clear either what is to be gained from (ii). Parsimony of type-experience comes with great expense of contents and, above all, it falls short of doing justice to the diversity of representing that our conscious experience accomplishes (or more precisely, consists in). It is pretty straightforward that your fear that it may rain does not feel the same, phenomenally speaking, as your belief that it may rain.

At this point, it is worth noting that the rejection of the transparency just sketched puts pressure first and foremost on Tye’s anti-mereological inclinations; it can therefore serve as a powerful additional reason for endorsing a mereological approach to consciousness. Importantly, though, the rejection of the transparency assumption does not entail the rejection of representationalism (Lycan 2015: §3.3). In fact, my sympathy lies with representationalism and I will take the forthcoming refinement of our stance on phenomenology to be a simple *variation* of it:

Representationalism₂: A conscious state’s phenomenal character supervenes on, or is identical with, its representational content *and* its psychological mode.

Now that this phenomenal character has been illuminated, the tripartite criterion described in chapter 3 can be refined as follows: experiences are individuated by their respective representational content, psychological mode, time of occurrence and subject.

Above all, the phenomenal input of experiential modes are of particular interest in the debate regarding phenomenal unity. Recall that since the unity relation is posited by mereologists to make a phenomenal difference, we regarded it as an experience that must be unified with the other experiences and thus involves a nasty regress of unity relations. We have observed that two strategies could be used in order to block it. The first consists in denying that the unity relation is something that the subject actually experiences; but in this case, the grounds on which the unity relation is postulated

simply evaporate. The second is to argue that the unity relation is not experienced *in the same way* as other experiences, like that of Céline Dion on stage. Underlying this claim is the idea that there is something that it is like for the subject to enjoy a unified set of experiences, but there is nothing it is like to experience this unity relation in particular. It seems to be in this spirit that Bayne concludes: ‘unity is not an *object* of experience but a *manner* of experiencing’ (Bayne 2010: 31-2 quoted *supra*, emphasis his).

This ‘manner of experiencing’ could very well hint at what I characterized as a modality, like vision or imagination. The resulting idea, tentative and provisional at this point, would be that the representational contents are integrated and thus unified by means of one (or perhaps several) mode(s). In the next sub-section, I explore the difficulties pertaining to this idea.

4.3.3. From the plurality of modes ³¹

With the mode/content distinction in hand, let us briefly re-consider the core tenets of the theses of phenomenal unity that have been discussed heretofore. On mereological theories of phenomenal unity, the subject undergoes at every instant *t* a total multimodal experience that supposedly subsumes all its experiential parts but is not just the conjunction thereof. ³² On Tye’s conception, likewise, the subject’s experience consists in but one multi-modal experience.

Now, connecting with the view that experiences are partly individuated by their modality, the question that I am getting at should be clear: what is the *nature* of this total, ‘multi-modal’ experience? In other words, how are our numerous experiences and their respective mode unified into one experience, and thus one mode, in any meaningful sense?

It seems that a natural strategy to unify modes of experiences is to argue that all of a subject’s experiences, including her multisensory perceptions, bodily feelings, emotions and thoughts, share a common mode. One of the most promising such attempt is *perceptualism*, the view that ‘all consciousness is perceptual’ (Prinz 2007, quoted by Bayne 2010: 60).

I take Tye to have something akin to perceptualism in mind, at least in regard to bodily sensations and emotions. On his conception, your experience of bodily sensations is a perception of your ‘body image’ (Tye 2003: 49, quoted *supra*) and the various changes that it undergoes – recall for example the particular case of pain, which is the representation of some tissue damage –, and your ‘moods’ are perceptual representations of bodily states and their changes. As for your thoughts, as we have seen,

³¹ This chapter significantly draws on Bayne 2010: ch. 4.

³² For simplification purposes, I choose here to systematically adopt only one of the two terminologies considered in chapter 3. My choice of Bayne & Chalmers’ is purely arbitrary; so again, what is said of their frameworks equally applies to Dainton’s, *mutatis mutandis*.

they bear phenomenal character only insofar as they are accompanied by auditory images, so one may charitably conceive of them as special, indirect perceptions.³³ To recapture Tye's view in a nutshell, the subject's total experience at a given time consists in various qualities, qualities of a perceptual nature, entering the same phenomenal content.

The significant upshot of treating all experiences as kinds of perception is that it allows you to claim that your total experience's mode is perception. This, in turn, seems to go some way towards explaining their unity: if the things that you experience are perceptual qualities, then they are unified by a *binding* of some sort. Recall that when you see a blue car, your visual apparatus not only 'assembles' the different car parts that you see into one unified object, but binds to it various features such as its color, motion and so on in an unifying fashion (see 2.3.). Can analogous binding processes account for the unity of experiences of different modes? To answer this, let me begin by considering the following sets of multisensory experiences, drawn from Bayne (2010: 61-3):

- (a) You are simultaneously seeing a big coin and touching a small coin.
- (b) You are simultaneously seeing a small coin and touching a big coin.

We further assume that all other things are equal in (a) and (b). Now perceptualism handily states that the total experience that unity theorists claim you have in both (a) and (b) is perceptual. Without this assumption, it is not clear what mode this total experience should involve. One could not hold it to be exclusively visual, for this would fail to capture the tactile dimension of the experience, and vice versa. This total perception-like experience, moreover, will have as its content the conjunction of the contents of its experiential parts, i.e. the properties of being a small coin and of being a big coin. But this picture entails that your total experience is the same in (a) and (b): it represents in both cases the properties of being a small coin and of being a big coin in a perceptual manner. This, however, is a terribly unwelcome result, for there certainly is a significant difference in phenomenal quality between the two cases.³⁴

Two strategies can be deployed in order to block this wrong conclusion. The first is 'Fregean'. It consists in stressing that the experience of sizes of coins involves different modalities. This means that in (a) you do not only experience the bigness of a coin and the smallness of the coin, but you experience the bigness *in a visual manner* and the smallness of a coin *in a tactile manner*; and vice versa in (b). This response, however, brings us back to the initial problem: how can these visual and tactile modalities be meaningfully combined into one? The generic perceptual mode proposed by perceptualism simply does not do the job.

³³ Because there are deeper problems to discuss, I deliberately choose not to question the link that would tie a thought with such image.

³⁴ And this is true whether you are enjoying such a thing as a 'total experience' or the mere conjunction of your visual and tactile experiences.

The second reply, by contrast, is ‘Russelian’: it does not admit of ‘modes’, and entails that there is no phenomenal difference whatsoever between touching a big coin and seeing it. Rather, in order to distinguish the phenomenal experience of (a) from that of (b), the strategy is to appeal to qualities other than shape that are represented only by one of the two senses. For instance, the touching of the coin represents its temperature. And since the coin that you touch in (a) is smaller than the one you touch in (b), the amount of temperature that will be represented by this tactile experience will not be the same as in the other. Another, related way of telling the two experiences apart is that although vision and touch both represent the shape of coins, they do so with varying degrees of precision: vision arguably represents shape more clearly than touch.³⁵ Consequently, you will more clearly experience the bigness of the coin in (a) than in (b). We can momentarily sigh with relief: the obvious phenomenal difference between (a) and (b) is salvaged.

This account of phenomenal contrast, however, hardly works when extended to other, non-sensory modes of experience. Ever since Anscombe’s (1963) notorious shopping-list analogy and Searle (1983), it is now widely accepted that beliefs, perceptions and hypotheses on the one hand, and conative states such as desires, willings, hopes and intentions on the other, have different *directions of fit*. Here is a nice synthesis of this fundamental difference:

The distinction is in terms of the direction of fit of mental states to the world. Beliefs aim at being true, and their being true is their fitting the world; falsity is a decisive failing in a belief, and false beliefs should be discarded; beliefs should be changed to fit with the world, not vice versa. Desires aim at realization, and their realization is the world fitting with them; the fact that the indicative content of a desire is not realized in the world is not yet a failing in the desire, and not yet any reason to discard the desire; the world, crudely, should be changed to fit with our desires, not vice versa. (Platts, 1979: 257, quoted in Humberstone 1992: 59).

In other words, belief-like states and desire-like states are said to have respectively a ‘mind-to-world’ and a ‘world-to-mind’ direction of fit. On other terminologies, the former states are said to be descriptive or ‘thetic’, while the latter are said to be directive or ‘telic’. Note that other experiences seem to have neither direction. Consider for instance the state of idly imagining what a flying pig would look like. Arguably, evidence that there is no such thing would not likely make you change this experience; and nor would this experience cause you to try and change the world so as to fit it (Bayne 2010: 67).

In light of this varying feature among experiences, it becomes increasingly hard to conceive of a total experience that would faithfully capture (or ‘subsume’) the following pair of simultaneous experiences:

(c) I desire that *p* and I believe that *q*.

³⁵ Indeed, you will be more apt to detect a default in the roundness of the coin by seeing it than by touching it.

Indeed, what would be the direction of fit of (c)'s total experience? Advocating for either one would surely prove unsatisfactory. Replying that this unified total experience is a-modal, or that it is perceptual as perceptualism has it, would not take us very far either. Here too, proceeding by way of phenomenal contrast should drive this point home. Consider this other set of experiences:

(d) I desire that q and I believe that p .

If perceptualism is true, then (d) is a perception just like (c). It is arguably the same perception, since it consists in the same conjunction of states: a desire and a belief. Moreover, (c) and (d) would have the exact same content: $\langle p \ \& \ q \rangle$. On our revisited version of representationalism, (c) and (d) are therefore phenomenally identical. But this is an undesirable conclusion: (c) and (d) sure do not *feel* the same.

It must be concluded, therefore, that perceptualism falls short of capturing fundamental structural features of experience, and I think it safe to claim that no category of mental state could possibly have such heteroclitic experiences as thoughts, perceptions, affective and conative states as determinates; accordingly, the endeavor of unifying experiences by way of a single mode appears to be doomed.

4.4. Conclusions

It is time for an interim summary of my inquiry. In the last two chapters, I proceeded to critically evaluate in turn Bayne and Chalmers' and Dainton's accounts, on the one hand, and Tye's account, on the other. This critical survey allowed me to separate the wheat from the chaff, so to say.

From the mereological accounts, I retained the commonsensical and more practical idea that experiences are best construed in mereological terms. But at the same time, I rejected the different relations that were proposed by these accounts. More specifically, I highlighted that the notion of 'total experience', which is supposedly at the core of the unity of phenomenal consciousness, is incoherent. I further stressed that the unity relations of subsumption and co-consciousness, understood as *primitive*, do not clarify the idea that our experiences are phenomenally unified.

From Tye's view, I retained an unspecified version of representationalism, namely the idea that an experience has a representational content, and that it is this content that endows it, somehow, with its phenomenal character. In a nutshell, this content is what makes the experience *conscious*.

Particularly enticing in representationalism is the prospect of analyzing phenomenal unity in terms of representational unity. The idea, roughly, would unfold as follows: for one's experiences to

be phenomenally unified (for there to be something it is like to entertain them together) is for their representational contents to be somehow unified, or integrated.

This claim, however, calls for refinements. Indeed, your experience of content p is individuated by both p and its mode of representation. You can imagine, believe, desire, hope or fear that p , and each of these modes of representation of p yields a distinct experience. This has led us to the conclusion that the unification of representations could not satisfyingly be achieved by closure of content *alone*, on pain of wrongly identifying the experience of believing that p and desiring that q with the experience of desiring that p and believing that q .

Consequently, we should conceive of closure of representations in a way that does justice to our different modes of representation. In an attempt to unify all of a subject's experiences into a single experience as Tye purports to, we briefly explored the idea that our perceptions, thoughts, bodily feelings and emotions are determinates of a single mode, and jettisoned it on grounds of these experiences' deep structural differences, notably with regard to their direction of fit.

5. The Self-Representational account

5.1. Self-awareness

My inquiry into phenomenal consciousness as led me to endorse representationalism, the popular view that experiences are representations, construed as *contents* represented in certain *modes*.

Let me start off by claiming that there is something amiss in this representational picture of phenomenal consciousness. Indeed, there seems to be more to the qualitative character of my seeing a tree than the properties of the tree that are represented. In particular, it seems that if there is something it is like *for me* to undergo this particular visual experience then I must be somehow aware of having it. Were I completely unaware of having it, it would make little sense to say that this visual experience is conscious 'for me' (Kriegel 2005: 25).³⁶

This subjective character can be captured as follows: when having an experience of a tree, I am acquainted not only with certain qualities of the tree but with the fact that these qualities are represented *to me*, that the representation is mine and no one else's. This subjective character is alternatively called 'for-me-ness' (Kriegel 2005), 'quality of mineness' or 'first-personal givenness' (Zahavi 2006: 16).

The view that any experience bears a certain quality of mineness is not new. Husserl, already, characterized our conscious experience as a 'Für-sich-selbst-erscheinens', which means literally 'self-manifestation'. Brentano claimed that 'in the same mental phenomenon in which the sound is present to our minds, we simultaneously apprehend the mental phenomenon itself' (Brentano 1874: 127). In a similar line of thought, Heidegger held that in any experiencing 'I am always somehow acquainted with myself' (*Gesamtausgabe* 58: 251), while Sartre (1943) later took the only 'mode of being' of conscious states to be *pour soi*, that is, 'for-oneself'.³⁷

In spite of its deep-rooted historical foundations and intuitive resonance, this dimension of consciousness has often been neglected, notably by representationalism; it is only recently that it has started to enjoy increasing popularity. In contemporary terms, one can synthesize what transpires (in perhaps peculiar terms) from these theorists by the general idea that *consciousness presupposes self-awareness*.

³⁶ For this reason, Kriegel takes the claim that only states that we are aware of count as conscious to be 'something of a conceptual truth' (*Ibid.*: 27).

³⁷ These quotes are drawn from Zahavi (2006: 11-2) and Kriegel (2004: 188).

What is self-awareness? Let us start with a rough definition according to which for a subject *S* to be self-aware of her experience *E* is for *S* to be *acquainted with* or *conscious of* herself having *E*.³⁸ To put it differently, it is for *S* to be aware of herself as being the subject or owner of *E*. To turn it in yet another way, it is for *S* to be aware of *E* as being her own (Zahavi 2006: 15; Kriegel 2004: 189).

Notice that there may initially appear to be actually two different claims here: when *S* is self-aware of *E*,

- (i) *S* is aware of *E*;
- (ii) *S* is aware of *herself* as being the one having *E*.

The reason why I press this point is that (i) and (ii) may be initially thought to be *partly independent* claims. Indeed, it sounds plausible that *S* being self-conscious of *E* entails only (i) but not (ii) – though (ii) seemingly entails (i). However, a simple argument provided by Rosenthal conclusively shows that this is misguided:

- (1) One cannot think about a particular (as opposed to a *type* of) mental state without thinking about the subject of this state.
- (2) But *S*'s awareness of her conscious state *is* an awareness of a particular mental state.
- (3) So *S*'s awareness of her conscious state comes with an awareness of the subject of her conscious state.
- (4) *S*'s awareness of a state as belonging to *another* subject would not make it conscious.
- (5) Therefore, *S*'s awareness of her conscious state entails *S*'s awareness of *herself* as the subject of her conscious state.³⁹

In short, claims (i) and (ii) are equivalent. Accordingly, in what follows, even though I only speak of self-awareness in terms of (i) for ease of presentation, (ii) shall always be implied. When needed, I shall emphatically get back to the dimension of mineness of our experiences.

Thus spelled out, the view according to which consciousness presupposes self-awareness will certainly raise some eyebrows: at first glance, some mental states of *S* can be conscious without *S* being aware of them. But this criticism ensues from an underspecification of the notion of self-awareness. There are actually two senses of self-awareness, one on which consciousness does plausibly entail self-awareness.

Recall our preliminary distinctions between attention, awareness and introspection. Awareness can be understood as the generic notion of attention: to be aware of something is to pay attention to

³⁸ Note that accordingly the terms 'self-consciousness' and 'self-awareness' are used completely interchangeably across the literature. I will likewise freely juggle the two terms.

³⁹ The argument in Rosenthal 1997: 741 is only partial, for it infers (5) from (1) and (2) only; the later addition of premises (3) (4) (in Rosenthal 2003: 330) facilitates this inference.

it. It has two essential features. The first is that it comes in degrees: I can be more or less aware of the cars on the road when riding my bike to work. We have simplified this graduation by saying that awareness can be either focal or peripheral – although there no doubt exist limit cases. As we have noted, in my initial case of absent-minded bike ride my awareness of the cars is peripheral to a point where it does not sound odd at all to say that I am unaware of them. But this is a figure of speech: I am not *wholly* unaware of them. The second feature of awareness, understood as attention, is that it can be directed either outwards, at worldly objects and properties, like cars and trees, or ‘inwards’, at one’s own experiences.

Just as one’s awareness of the external world can be either focal or peripheral, so too can one’s awareness of one’s experiences. For instance, Maria is focally aware of her attraction to Sam. What is meant by this is that she pays careful attention to her experience; she reflects on it.

Now, to say that any experience involves self-awareness in this focal sense is surely wrong. One simply does not scrutinize any experience that we enjoy the way that Maria does her attraction to Sam. Mowgli, for instance, is not focally aware of his visual experience of the tiger that he is fighting.

On the self-awareness view just introduced, however, Mowgli is *peripherally* aware of his visual experience of the tiger – and of *any* other experience that he may entertain at this instant, for that matter –. Here is a nice presentation of this idea:

[Consider] the case of [self-consciously] thinking about x or attending to x. In the process of [self-consciously] thinking about x there is already an implicit awareness that one is thinking about x. There is no need for reflection here, for taking a step back from thinking about x in order to examine it... When we are thinking about x, the mind is focused on x, not on our thinking of x. Nevertheless, the process of thinking about x carries with it a non-reflective self-awareness. (Goldman 1970: 96 quoted in Kriegel 2004: 188-9)

As these lines show, the distinction between focal and peripheral self-awareness is a fairly spread one, though the terminology varies among authors. Goldman, for instance, contrasts ‘reflection’ with an ‘implicit’ or ‘non-reflective’ self-awareness. Borrowing from Sartre, Zahavi frames this distinction in terms of ‘reflective’ and ‘pre-reflective’ self-awareness (Zahavi 2006: 21); Flanagan, in terms of ‘high-level’ and ‘low-level’ self-awareness (Flanagan 1992: 194); Kriegel, in terms of ‘transitive’ and ‘intransitive’ self-awareness (Kriegel 2004: 189).

To summarize, the view proposed here is that consciousness *necessarily* involves self-awareness of the second kind. Again, this means that even though the subject does not engage in any reflection about her experiences, she is necessarily *non-reflectively* aware of them.

Having laid out this missing piece of our phenomenal puzzle, I shall shortly venture and defend the thesis that for S’s experiences to be phenomenally unified (at a given time) is for S to be non-reflectively aware of her experiences conjointly ; it is for S’s experiences to be closed under a single act

of non-reflective self-awareness. Before elaborating any further, the following twofold question needs to be answered: what is the *nature* of this self-awareness, and what *relation* does self-awareness bear to our experiences? Answers to this question can be regrouped into two main conceptions of self-awareness. On one conception, proposed by Higher-Order Theories (HOT), awareness of one's experiences is a higher-order mental state. On the other conception, proposed by Same-Order Theories (SOT), awareness of one's experiences is somehow *imbedded*, at the same level, into each of these experiences.

I what follows I critically consider these two conceptions in turn. HOT will quickly appear to be rivals to our representationalist account of consciousness. I will argue that in lieu of renouncing representationalism, one should reject HOT on account of the various problems that bedevil them. This rejection will immediately entail that SOT must be endorsed; nevertheless, I will provide positive argumentation in its favor and show not only that it fits nicely into our representational picture of consciousness, but that it lends itself to a powerful account of phenomenal unity.

5.1.1. Higher-order theories

On Higher-Order Theories (HOT), a subject *S*'s mental state *M* is conscious when and only when *S* has second mental state *M** that *represents* *M* in an appropriate way. This means in other words that *M* is conscious insofar as it is the object of an appropriate higher-order (HO) representation *M**.

HO theorists disagree as to the nature of this HO state *M**. While Armstrong (1968) and Lycan (1997) think of it as a *perception-like* state, Rosenthal (1997) thinks of it as a HO *thought*, more specifically a thought to the effect that *S* is in *M*. I need not enter into further details for the present discussion, and will treat both views indifferently in regarding this HO state as a state of *awareness* – I join Kriegel (2006a: §5) in adopting this broad formulation. Accordingly, HOT's slogan may be the following:

A conscious state is a mental state whose subject is aware of being in. (Lycan 2001: 3)

The representation of *M* by a HO state *M** is *appropriate* insofar as it is *non-inferential* and *de se*. The representation of *M* must be non-inferential in the sense that *S* must be aware of *M* in an immediate (or at least apparently immediate) manner. You may come to notice an anger that you harbor against your father only after your therapist presented to you a convincing amount of evidence that you are in this state. This sort of inferential awareness does not endow your state with consciousness (Rosenthal 1997: 737-8).

The second requirement for M's consciousness, the one that will be of particular interest for my discussion, is that your awareness of M be *de se*: it must be an awareness of M *as being had by yourself*. In other words, for your state M to be conscious it does not suffice that you be aware that M is occurring; your awareness of M must be the awareness that *you are in M* (Rosenthal 1997: 741).

The weaker, non-reflective sense in which we have defined self-awareness handily neutralizes two worries that are generally elicited by this view.⁴⁰ The first has already been put to bed: it is that S is not systematically aware of her experience. The second is that a self-awareness account of consciousness denies consciousness to non-conceptual (nonlinguistic) creatures like children and animals. The objection is that for S to be self-aware of her mental states, S must be equipped with the concepts of thought and of self, which non-conceptual creatures certainly lack.⁴¹ But on a non-reflective conception of self-awareness, non-conceptual creatures need only have a 'minimal' (admittedly very minimal, as will be argued in 6.1.2. and 6.3.) concept of self and of states in order for their experience to be conscious (Rosenthal 1997: 741). It does not seem beyond the pale to grant them these competences. Although more radically, Hurley militates for the same idea:

If consciousness of pain is possible without the concept of pain, or consciousness of the sun is possible without the concept of the sun, why not self-consciousness without the concepts of oneself, of conscious states, or of the objects or properties of which one is conscious? (Hurley 1998: 136)

Consequently, HO theorists' claim that consciousness implies self-consciousness is secure with regard to what they intend by the notion of self-consciousness.

HOT stand in tension with representationalism. According to the latter, your visual experience of the tree is conscious in virtue of *representing* the tree in a certain way; according to the former, your visual experience of the tree is conscious in virtue of *being appropriately represented by* a HO state of awareness.

Since HOT account for the apparently essential subjective character of experiences, and since I have admitted that representationalism neglects this dimension of consciousness, the latter view is put in jeopardy and should perhaps be jettisoned in favor of the former. However, HOT accrue vexing problems that considerably reduce their appeal, not only comparatively with representationalism but *tout court*. I consider here two of them.

⁴⁰ For a demonstration that the reflective/non-reflective awareness distinction holds as much for a perceptual as for a cognitive approach to higher-order awareness, see Kriegel 2004:189-92.

⁴¹ It is worth remarking that certain theorists (notably Carruthers 1996a: 155-8 quoted in Zahavi 2006), which advocate for a reflective self-awareness picture of consciousness, are willing to bite this bullet and argue that while non-conceptual creatures are 'creature conscious' (cf. section 2.1.), their mental states are not conscious.

A first problem concerns the epistemology of our conscious experiences. On HOT, M is conscious only when it is appropriately represented by a HO state M*. It follows that M* is not itself conscious unless it is itself appropriately represented by a further HO state. Rosenthal argues that this typically happens in case of introspective consciousness; but since we do not introspect on our experiences most of the time, M* is unconscious most of the time (Rosenthal 1997: 742-6).⁴²

Accordingly, in order to know that we are in a conscious state M, we need to know that we are in the unconscious state M*. But since M* is unconscious, we have no direct awareness of it. Rather, our knowledge of M* must be theoretical (inferential) and third-personal. Since our knowledge of M* is necessary for our knowledge of M, this entails that our knowledge of M is theoretical and third-personal. But this falls afoul of the axiomatic view that our knowledge of our own conscious states is necessarily first-personal and non-theoretical (Goldman 1993; Shoemaker 1996).

In fact, this conclusion is also at odds with our previously accepted conception of self-awareness, as well as with HOT's own arguments in favor of higher-order awareness. It seems that we have first-personal knowledge not only of our (first-order) experiences, but of *the very fact that we are aware of our experiences*. Indeed, it is on first-personal epistemic grounds that HO theorists themselves sustain that conscious states are states that we are aware of. The only way that HOT could salvage our first-personal knowledge of both our experiences and our awareness thereof would be to hold that the representing state M* is itself conscious. But since, on HOT, M* would need to be the object of yet another HO state in order to be conscious, this reply entails an infinite regress of HO states (Kriegel 2006a: §5.3.).

A second line of problem of HOT concerns the sort of property that they take consciousness to be. Since, on HOT, a state M's being conscious is a matter of being appropriately represented by HO state M*, M is not conscious in and of itself; rather M is conscious in virtue of the two obtaining together in the relevant way.⁴³ Since, on this picture, M and M* are two numerically distinct and logically independent states, M's consciousness is a nonintrinsic, relational property that obtains between M and M* (Rosenthal 1997: 736-7). The problem is that thus construed, M's consciousness is causally inert:

If what makes E (some experience) conscious is the fact that S (the person in whom the experience occurs) is, somehow, aware of E, then it is clear that E's causal powers (as opposed to S's causal powers) are unaffected by the fact that it is conscious. The causal powers of a rock (as opposed to *my* causal powers) are not changed or enhanced by my observing the rock or having thoughts about it. Why should the causal powers of a thought or an experience be any

⁴² Which is why, according to Rosenthal, the fact that that we have *many* higher-order state like M* at any time (for as many conscious states) is perfectly plausible.

⁴³ The 'in virtue of' ought not to be understood in a causal sense. The relevant representation of M by M* does not *cause* M to be conscious; rather, M's being conscious *consists in* or is *constituted by* M* appropriately representing M (Kriegel 2006a: note 58).

different? If the consciousness of mental states and processes comes down to higher-order experiences of them, or higher-order thoughts about them, then consciousness is epiphenomenal. (Drestke 1995 : 117, emphasis in original)

Unless we are willing to bite this bullet, HOT are considerably undermined by this charge.

Wrapping up, we have hypothesized as our point of departure that a conscious state necessarily comes with self-awareness, understood as the subject's non-reflective awareness that she is in that state. Since HOT initially promised to account for this feature, it was worth pausing to consider them and perhaps recast the issue of phenomenal unity in higher-order terms, as Rosenthal (2003) for instance does. However, HOT of self-awareness do not sit well with our apparent first-person epistemology of both our conscious states and our awareness thereof, nor with the causal powers with which consciousness is uncontroversially endowed. I take HOT to succumb to these problems.

I will now sketch the view according to which the subjective character of our experiences is somehow *imbedded in them*; it is the view, dubbed *self-representationalism*, that our experiences represent themselves to us. I will argue that this account not only plausibly captures the central claim that we are self-aware of our experiences while incurring neither of HOT's pitfalls, but has the significant merit of being compatible with a representationalist approach to consciousness.

5.1.2. A same-order theory: self-representationalism

To start off, it is worth stressing that the traditional phenomenologists quoted at the beginning of this chapter do not construe the 'sense of mineness' that they postulate as an additional, contingent state that targets other mental states. Rather, they treat it as an *intrinsic* feature of our experiences; it is in other words located *within* them.⁴⁴

The idea is that our awareness of our mental states is non-objectifying: it does not take these mental states as its objects. An objectifying awareness entails a *phenomenal distinction* between the awareness and its targeted object. To illustrate this phenomenal distinction, one can note that it is certainly salient in the first-order experiencing of world-features: there is a feeling of the duality between the *world features*, on the one hand, and our experiencing them, on the other; we feel a clear demarcation between the tree and our visual experience of it. However, it is dubious that this sort of phenomenal distinction is at work with regard to self-awareness. The impression is rather that we entertain our mental states and non-reflective self-awareness as an *experiential unit*. Thus, talking about self-awareness *of* our experiences should not mislead us into thinking of self-awareness as being

⁴⁴ Of course, this is not to say that self-awareness, and thus consciousness, is an intrinsic feature of *all* mental states; there surely exist non-conscious mental states (Zahavi 2006: 20).

distinct from the experiences. This ‘of’ associated with self-awareness is not the objectifying particle that is appropriately used in talking about a visual experience *of* the tree; it is merely a grammatical constraint. It is in this spirit that Sartre cleverly refers to self-awareness by the phrase ‘conscience (de) soi’ (*sic*), the parentheses underlining the ambiguity that attaches to the preposition (Sartre 1943: 22, quoted in Zahavi 2006: 23). Another way to remove self-awareness’s intentional connotations is to treat it in an adverbial form; thus the idea *S*’s self-awareness of her visual experience of the tree would be best captured by saying that *S* is ‘*self-awarely*’ or *self-consciously seeing* the tree (Kriegel 2005: 48 and 2004: 188).

The views that transpire from these authors could be synthesized as follows: self-awareness is a *constitutive* (as opposed to *extrinsic* on HOT) *feature* of our experiences. This claim has given rise to many positions which are nowadays championed under the label of ‘Same-Order Monitoring Theories’ (Kriegel 2007). For simplicity, I will hereafter generically refer to them as the ‘Same-Order Theories’ (SOT).

Besides their commonsense appeal and respectable lineage, SOT follow from the apparent untenability of HOT:

- (1) A mental state *M* of subject *S* is conscious when, and only when, *S* is aware of *M*;
- (2) Awareness of *X* requires mental representation of *X*; therefore,
- (3) *M* is conscious when, and only when, *S* has a mental state *M**, such that *M** represents *M*.
- (4) Either *M** and *M* do not entertain a constitutive relation, or they do;
- (5) There are good reasons to think that it is *not* the case that *M** and *M* do not entertain a constitutive relation; therefore,
- (6) There are good reasons to think that it is the case that *M** and *M* do entertain a constitutive relation. (Kriegel 2006a: §6.2.)

It should be noted that this argument goes through only if the present view is not vitiated by the same problems as HOT. In order to ascertain that it is not, a crucial issue must be addressed: what is meant exactly by the claim that *M** is *constitutive of M*?

I first argued at the beginning of the chapter that a mental state *M* is conscious only on the condition that the subject *S* is non-reflectively self-aware of *M*. Recall that for *S* to be non-reflectively self-aware of *M* is for *S* to be in a mental state *M** that represents *M*. Consequently, *M** being constitutive of *M* seems to mean that *M* *represents itself*. Call this version of SOT ‘Self-representationalism’ (*Ibid.*).

I take my introductory remarks to have sufficiently motivated the claim that self-awareness is a necessary feature of our experiences. However, it may be worried that on the present self-representational reading, self-awareness is not *sufficient* for consciousness. Consider for instance the following sentence: ‘This very sentence contains exactly seven words’. While this sentence represents itself, it is certainly not conscious. But one should distinguish between *extrinsic* and *intrinsic* self-representations. This sentence represents itself only to the extent that someone reads it and endows

it with meaning. In fact, it does not represent *anything* in and of itself; hence its representational and self-representational content is extrinsic. By contrast, a mental state is conscious by being *intrinsically* self-representing (Kriegel 2005: 39).⁴⁵

More fundamentally however, as Kriegel repeatedly notes, the idea that a mental state can represent itself may come across as mysterious, to say the least. In particular, it falls afoul of the project of naturalizing consciousness. Naturalists aim at reducing phenomenal consciousness to *natural* relations. Representationalism *à la* Tye is a good example of such reduction. It construes mental representations as relations between world states and brain states. These relations, as opposed to logical or conceptual ones, are *natural* in the sense that they are based on causal processes: world states cause brain states. So if representations are natural relations and if phenomenal properties are representations, phenomenal properties are natural as well. It is in this spirit that Tye famously declares: ‘phenomenology ain’t in the head’ (Tye 1995: 151).

On self-representationalism, an experience is both the object of representation *and* the representation thereof – since it represents itself. This means that the self-representation of an experience consists in a reflexive relation. But the causal relation is anti-reflexive: something cannot cause itself. Therefore, self-representation is not a causal relation. This means that no naturalist account of representation leaves room for self-representing experiences (Kriegel 2005: 43-5). We have to postulate a further, ‘super-natural’ relation accounting for self-representation.

We are here at a crossroads. Theorists can take either one of two significantly divergent paths. The first consists in fully embracing the anti-natural character of self-awareness. Along Burge and Crane, it can be argued that naturalization is only one of many philosophical issues surrounding consciousness, and that philosophers would benefit from emancipating themselves from it when pursuing other projects (Burge 1997: 433, quoted in Crane 2003: 27).⁴⁶

However, I want to follow Kriegel (2005) in pursuing the more ambitious path of reconciling the naturalist project with the enticing proposal that self-awareness is at the heart of consciousness. Indeed, it seems that if this is successfully done, self-awareness, as well as the account of phenomenal unity that I will shortly build upon it, will gain in appeal and robustness.

⁴⁵ See also Kriegel 2006a: §6.3. for the same distinction in terms of *derivative* and *original* self-representation

⁴⁶ I must emphatically stress that taking this path would by no means runs counter to the critics that I made about Bayne and Chalmers’ and Dainton’s accounts. Indeed, I have not taken issue with the seemingly unnaturalizable Blockean phenomenal consciousness *per se* that they presuppose. Rather, I have charged them with the quasi absence of informativeness and explanatory power of their accounts. In couching phenomenal properties in terms of representations, *even if some of these representations are super-natural*, self-awareness would still go a long way towards explaining what it is for a mental to be conscious and, consequently, towards making the thesis of phenomenal unity more substantial.

5.1.3. Naturalizing self-representation

The impossibility of naturalizing self-awareness is only apparent and is due to the somewhat naive understanding that our conscious mental states literally represent themselves. Fortunately, there are other SOT that account for the idea that self-awareness is constitutive of an experience.

I have previously alluded to the famous *binding* phenomenon, which consists in the integration of certain representations into a simple representational content. For instance, when I see a tree, this unitary visual experience involves the integration of several features, among which its shape of colors, and so on. We further know that these features are processed in different parts of the brain. The representational content emerges from the integration of these features (Kriegel 2005: 47, n. 59).

Now, it can be ventured that my self-awareness of this experience involves an additional binding phenomenon. When I see the tree, a representation of this visual representation takes place elsewhere in the brain. What makes my visual experience conscious, then, is the *integration* of its representational content with the representation of this representation as being mine.⁴⁷

Admittedly, the *integration* of a representation with a higher-order representation of it may sound just as mysterious as the notion of self-representation. Carruthers gives a plausible account of how consciousness may emerge from this integration. Upon seeing the duck-rabbit figure for the first time, a subject may experience nothing more than a non-conceptual combination of lines and curves. But once she comes to see it as a rabbit (or a duck), what happens is that the conceptual content 'rabbit' is integrated with the non-conceptual representations of the lines and curves, thus giving rise to a single perceptual state. It is plausible that self-awareness involves a similar process: the non-conceptual representation to the effect that S is having this experience comes to be integrated with the other visual contents so as to yield the visual *experience* of the rabbit, which comes with S's non-reflective awareness that she is undergoing it (Carruthers 2016: §6).

Strictly speaking, then, the visual experience does not represent itself. 'Self-representation' may be a useful proxy or *façon de parler*, but it is not essential to the S's awareness of her experience; rather, what seems essential is that S's self-awareness be *integrated with* her experience. On the present proposal, a psychological process integrates a representational content with a higher-order representation of this representation, in a way that yields a single self-conscious experience. And since this psychological process is a causal one, the integrative relation that gives rise to phenomenal consciousness is natural.⁴⁸

⁴⁷ A quick terminological point: Kriegel (2005) coins this view the 'Cross-Order Monitoring Theory' because it states that an experience involves a higher-order representation which is yet constitutive of the experience. However, for the sake of simplicity, I will not burden myself nor confuse the reader with yet another acronym.

⁴⁸ Admittedly, a zombie argument could be ultimately cooked up: unconscious creature may have states that represent themselves. But since the conceivability of these creatures is a bit of an open question, I shall not

With this characterization in hand, let us now see if our picture satisfactorily rises to the epistemological and causal challenges that beset HOT. Consider the first. Because HOT construe self-awareness as an unconscious state, it leads to the unwelcome conclusion that our awareness of our conscious states could be known only on a third-personal and theoretical basis. By contrast, since SOT regard self-awareness as a constitutive feature of our conscious states, self-awareness *must* be conscious as well; self-awareness is therefore first-personally knowable as it ought to be. Thus one could treat the first-personal knowledge constraint on experiences as an independent, positive argument for SOT, irrespective of HOT's limitations (Kriegel 2004: 196).

What about the second challenge? On HOT, consciousness is an extrinsic property which makes it an epiphenomenon devoid of any causal powers. One might worry that on SOT, consciousness is as much as an extrinsic property as it is on HOT.

Self-representationalism, the view that an experience literally represents itself, offers a ready-made reply to this worry. It is true that self-awareness is relational in the sense that it stands in a relation to something, namely itself. However, a relational property ought not to be conflated with an extrinsic property. On a self-representational approach, self-awareness is *not* extrinsic for it stands between overlapping (indeed, *identical*) entities, namely M* and M. Just as my having an arm is a relational but intrinsic property, so too is self-awareness of one's experiences. Consequently, since M's representing itself is what makes M conscious, consciousness is an *intrinsic* property of M (Kriegel 2007: 380). However, this intrinsic character cannot deal with the causal powers of consciousness for it rests on the reflexive relation of self-representation.

On the naturalist account just presented, self-awareness is technically a *part* of, rather than identical with, the experience. Hence it stands in relation to the *other non-overlapping* representational part of the experience. However, this relation is causal, as opposed to *conceptual* as a couple of SO theorists claim it is. In effect, since this consciousness emerges from the integration of the two parts, and since this integration is achieved by a real, temporally extended psychological process, consciousness is endowed with the causal potency of this process.⁴⁹

In light of this, SOT of self-awareness, and in particular the naturalist variation that I endorsed, wins the day. But there may still be something disturbing about the idea that a representation of the tree is integrated with a representation of this representation, for they seem to be very *different*

address this concern any further and will simply concede that this account of self-awareness, as any other for that matter, irremediably remains exposed to it.

⁴⁹ A theorist who takes the relation between self-awareness and first-order representations to be *conceptual* is Gennaro (1996, 2002), presented in Kriegel (2005: 48-9; see also Kriegel 2007: 364 ff. for more technical details) A worry about this view is that its difference with HOT is merely verbal (Kriegel 2007: 367). Kriegel's conception does not face this charge.

objects. However, it is possible to construe the integrated information as concerning the same object, namely the tree. The visual experience of the tree would amount to the representation of the tree i) as bearing such-and-such features (a particular color, shape, etc.) and ii) as being visually represented (to me) as bearing such-and-such features (Kriegel 2005: 49).

If this construal still fails to find intuitive resonance, another way of understanding this integration consists in treating self-awareness as part of the *attitude* that the subject takes toward the intentional content of her experience. So understood, self-awareness functions as an additional intentional mode.⁵⁰ This means that when I see a tree, there are two modes at work: the visual mode that represents the tree under a certain light, and a self-representational mode that represents this representation as mine.

Kriegel seems to adumbrate something along the same lines when he claims that the subject's self-awareness 'modifies' her experiences and should be understood as 'the way in which it is had by the subject' (Kriegel 2004: 188, emphasis mine). However, he later harbors doubt as to whether this treatment of self-awareness fits our construal of modes of experiences as having a certain direction of fit. His worry is that self-awareness is neither mind-to-world (as are doxastic states) nor world-to-mind (as are desire-like states) directed, and that therefore it cannot qualify as a mode of representation (Kriegel 2005: 40-1). Two replies can be offered to put this worry to bed. The first is that a mode need not have a direction of fit – recall our earlier example of idly imagining what a flying-pig would look like –; thus the claim that self-awareness lacks a direction of fit would seem insufficient to deny that it is a mode. A second reply is that self-awareness does have a direction of fit. Though it would be odd to say that it mind-to-world directed, as it targets representations, one could hazard the idea that it is *mind-to-mind* directed, in a thetic sense: it aims at describing the way the subject's mental states are. Accordingly self-awareness has satisfaction conditions too: it is satisfied when it truthfully represents the representation of the tree.⁵¹

The major upshot of this picture of self-awareness is that it incorporates the apparently essential subjective character of our experiences within a representationalist approach to consciousness at the only cost of endowing our experiences with an additional modality, namely self-awareness. I will now

⁵⁰ This is inspired by Smith (1986: 156). Importantly, Smith terms this intentional mode 'modality', for he distinguishes between 'mode of presentation' and modality of presentation. When you see a frog, the *mode* of presentation is that of the frog, whereas the *modality* of presentation is that of vision (*Ibid.*:150).

⁵¹ Perhaps ironically, this proposal is encouraged by a weakness of the naturalist SOT, namely the fact that it leaves room for mismatches between self-awareness and the representation with which it is integrated: nothing prevents that the representational content 'blue car' be integrated with a higher-order awareness that I am visually experiencing a *green* car. It admittedly cries for an explanation of why, in such mismatches, the car would feel blue or green to me; I cannot delve into these problems here (but see Carruthers 2016: §6 for promising solutions). All that matters now is that such mismatches sits well with the construal of self-awareness as a thetic mode.

present and motivate a compelling account of phenomenal unity that draws on this conception of self-awareness.

5.2. Phenomenal unity

5.2.1. Closure revisited

How does the above account of self-awareness play out for the more specific issue of phenomenal unity? I have endorsed the view that any experience consists in the psychological integration of a representation with the representation of this representation. This higher-order representation is the now-familiar 'self-awareness', which is constitutive of the experience in a naturalizable way. Furthermore, one has not lost sight (I hope) of the following tenets of my view: the constitutive self-awareness of a subject *S*'s experiences is *non-reflective*, and it consists not only in *S*'s awareness of *her experiences*, but in *S*'s awareness of *her undergoing them*.

Consider *S*'s desire that *p* and belief that *q* at *t*: *S*'s desire is self-consciously experienced, and so is *S*'s belief. Since they are both experienced as being *S*'s own, they are all indexed, so to say, to the same bearer, namely *S*. This is ensured by the fact that *S* can only be acquainted first-personally with *one* subject of experience, and that this subject of experience must be *S* herself.

My contention, adapted from Rosenthal (2003: 332), is this: in lieu of positing two numerically distinct modes of self-awareness, one for each experience, there is a *single* self-awareness that represent *both S*'s desire that *p* and *S*'s belief that *q* as *S*'s own experiences. Once generalized, my view is then the following: for *S*'s experiences of all modalities (perceptions, bodily feelings, thoughts and emotions) to be phenomenally unified is for them to be conjointly experienced as *S*'s own. In short, self-awareness provides *S*'s experiences with a common denominator, namely the subject that is elusively experienced as undergoing each of them; and this common denominator is what naturally integrates, or closes under conjunction, *S*'s experiences together.

Our evaluation of Tye's account in the last chapter laid important groundwork for our conception of phenomenal unity. In particular, it has drawn us to the idea that for several experiences to be phenomenally unified is for the representations they involve to be integrated together. Furthermore, it has led us to the conclusion that a satisfying account of phenomenal unity should be able conciliate these three core claims:

- (i) The subject's experiences are integrated in such a way that it makes sense to speak of a 'total experience'.

- (ii) This total experience should be able to meaningfully integrate both content *and* mode of each experience that is to be unified.
- (iii) The subject has some kind of access to the unity of his experience.

The conception on which our experiences are unified by self-awareness results from an inference not only to the best explanation, but to the *only explanation available* that satisfies (i), (ii) and (iii) together. First, it makes sense of the claim that there is such a thing as a ‘total experience’. S’s total experience is nothing other than S’s elusive ‘self-awarely’ way of undergoing all of her experiences at any time; in other words, it consists in the representation of all S’s experiences as her own. Furthermore, since these experiences are individually self-conscious but yield together a unique integrative self-awareness, i.e. the total experience, phenomenal unity does not consist in a mere aggregation of experiences as James feared.

Secondly, since self-awareness consists in the higher-order representation of S’s first-order representations, and since representations are construed as contents *and* modes of representation, the closure that self-awareness accomplishes does not leave out the variety of S’s intentional modes. Remember that at any instant of your concert, your experiences are incredibly diverse. Some, like your desire to have your shirt signed after the show, are world-to-mind; others, like your perceptions of the glowing stage, are mind-to-world; still others may lack a direction of fit at all. On my view, these representations are unified in virtue of being *wholly* integrated, and not just with respect to their content (the signed shirt, the glowing stage) but by the awareness that they belong (as conative and visual experiences) to you.

Thirdly, self-awareness is on par with the idea that the subject has an access, and a privileged one at that, to this unity. Recall Bayne’s ‘Unity judgment’:

Unity judgment: by introspection I come to the conclusion that all my current states are phenomenally unified with one another. This unity is a permanent feature of consciousness, even when we don’t introspect. (Bayne 2010: §4, quoted *supra*)

The present conception accounts for this feature: S’s self-awareness is this elusive and yet constant sense that her experiences are hers, even when she does not engage in introspection. Moreover, I hypothesize that the very fact that S can at will, immediately and first-personally introspect about several of her mental states at a time *requires* that she be non-reflectively (or more precisely *pre-reflectively*) aware of them. One can look at this plausible claim in the following way: were it not the case that S is non-reflectively aware of her experiences, she would not be capable of introspecting into any of them *immediately*. And were it not the case that S is non-reflectively aware of them *together*, she would not be capable of introspecting into some of them *together* immediately.

5.2.2. No more content

Another major upshot of my proposal is that it is immune to the just-more-content objection. On the present proposal, one's experiences are unified by the elusive and common self-awareness. As I have argued, self-awareness is best regarded as an additional *mode* by which one's experiences are represented as being one's. In light of this, self-awareness is unified with one's experiences by a mode-content relation.

Is this unity *phenomenal*, as the account that we are after would require? It certainly is. On my view, your visual experience of Céline Dion on stage has three phenomenal components: a representational content (Céline Dion on stage), a mode of representation (vision) and a mode of self-representation (*you are* visually experiencing of Céline Dion on stage). Each component is endowed with a distinctive phenomenal feel, which is revealed by simple contrasts. Seeing Céline Dion on stage does not feel the same as seeing your feet; seeing Céline Dion on stage does not feel the same as remembering Céline Dion on stage. As for the sense of mineness, it could not be brought forth by comparison with an experience that lacks it simply because it is what makes a mental state an experience. But elusive though it may be, we have argued that it is definitely there and that it is phenomenally different from the visual feel and the Céline-Dionish feel of your experience.

Although they each bring their own flavor to the phenomenal recipe of your experiences, these three ingredients are ontologically and experientially bound: you could not represent Céline Dion on stage without representing in a certain way, and without representing the representation as belonging to you. This in turn entails that you simply could not experience the qualities of Céline Dion on stage without thereby experiencing the phenomenology attached with its mode (here, vision). They are like two sides of the same phenomenal coin, so to say: either one cannot obtain without the other. The same reasoning applies to self-awareness: the visual representation could not be experienced at all without being integrated with a representation of this representation as being yours; it is an integral part of it. Thus they are not felt in piecemeal manner: they are experienced as a single experiential unit.

Summing up, our experiences are phenomenally unified by being represented together as belonging to a single subject, and the total experience of self-awareness is phenomenally unified with our experiences simply by a mode-content relation or, more simply put, by being a *part* of them. These relations are much more intelligible than the primitive relations of co-consciousness and subsumption: they explain phenomenal unity in terms of conjointly indexed representations, which is in my opinion a familiar and fathomable enough notion; thus I take them to actually *clarify* the otherwise vague notion of a 'conjoint phenomenology'. Consequently, holding self-awareness to be an essential feature

of consciousness should be considered an explanatory gain, rather than an ontological cost, when compared with the other accounts on the market.

My job, however, is far from completed yet. Tye's view and the one that I have just developed attempt to analyze phenomenal unity in terms of representational closure. I have rejected Tye's on the grounds that it could not satisfyingly close experiences; my view has aimed at improving the idea of representational closure so as to make it *viable*.

However, many theorists may still rest discontent with the idea of representational closure on any conception, Tye's just as mine. Two questions may be raised in particular:

- 1) Can *any* set of experiences be the object of a conjoint representations achieved by self-awareness?
- 2) Is self-awareness really *sufficient* for two experiences to be representationally and thus phenomenally unified?

Thinkers of a skeptical persuasion may consequently bring into doubt the claim that self-awareness integrates our representations *in a meaningful sense*. They may take self-awareness as a purely *ad hoc* move intended to accommodate the plurality and putative incommensurability of modes. They may even venture the claim that integrated experiences are in some cases completely artificial or arbitrary. For instance, what would the conjoint representation of the taste of strawberry flavor in my mouth with my visual experience of Geneva's Jet d'Eau concretely amount to? What would it do? What about the conjunction of the auditory experience of *Pour que tu m'aimes encore* with the imagining of the number pi? Are they really *unified* by being both represented to me? Supposedly integrated representations such as these are strongly evocative of mereological sums,⁵² a concern that has been voiced about Bayne's conception of a total state:

Of course, if you buy classical mereology, it is trivially true that any two experiences are parts of a single state. This follows from the Leonard–Goodman axiom that says that for any set of entities, there is another entity of which all of the members of the set are proper parts – an entity called the 'fusion' of the set (see Leonard and Goodman 1940). But the relation being parts of a Leonard–Goodman fusion won't serve Bayne's purposes, because the Leonard–Goodman fusions of pairs of entities usually aren't unified in any meaningful sense. Thus, it is true of an experience in your mind and an experience in mine that they are both parts of a Leonard–Goodman fusion. (Hill 2014: 504-5)

In order to alleviate this concern, I will contend in the remainder of this chapter that the integration of our experiences is *constrained* by certain norms. I will moreover argue that this normativity most likely stems from the self that is elusively sensed to be the owner of all experiences.

⁵² Or as Kevin Mulligan calls them, 'Polish sums'.

That is, I will argue that self-awareness monitors our experiences by seeing to the respect of these norms.

5.3. The monitoring role of phenomenal unity

Against the view that for a subject's experiences to be phenomenally unified, they must be representationally integrated together, arguments generally unfold *ad absurdum* as follows:

- (1) For *S*'s experiences to be phenomenally unified is for them to be representationally integrated together as being *S*'s.
- (2) If *S*'s experiences are representationally integrated together in any meaningful sense, they must (at least) exhibit some kind of coherence.
- (3) But some of *S*'s phenomenally unified experiences can be (and sometimes are) incoherent.
- (4) So (1) is false.

In order to block this argument, one can reject either (2) or (3). Premise (2) is what makes representational unity a substantive claim rather than a mere mereological sum of representations, but there are also practical reasons for defending it. Central in Block's and representationalist accounts of consciousness is the idea that experiences, understood as representations, play certain functional roles in the rational and agentic economy of the subject. But a lot of experiences do not fulfill these functions individually. For instance, desires are usually taken to be functionally inert, absent relevant beliefs. In order to be functionally efficacious together, they must abide by certain norms of coherence. Drawing on these lines of thought, I will accept *and* motivate (2) in the following development. Consequently, if (1) is to be sustained, one should argue against (3) that phenomenally unified experiences *cannot* be incoherent.

But what is it exactly for two experiences to be *coherent*? One could at least name four coherence norms: semantic consistency, inferential integration, means-end rationality and agentic unity.⁵³

Before characterizing these norms, three crucial points must be stressed. First, bear in mind that for these norms to relevantly apply to closure, representational contents must be construed in propositional terms and will be so in the following development: an experience takes as its object a

⁵³ The first three are proposed by Hurley (1998: 124), the last by Schechter (2010b: 4), although Schechter joins Hurley in citing 'semantic coherence' and 'disposition for rational thought' as well.

proposition, and for two experiences e_1 and e_2 with respective content $\langle p \rangle$ and $\langle q \rangle$ to be closed is for there to be an experience that represents them together, i.e. that has as its representational content the conjunction $\langle e_1(p) \ \& \ e_2(q) \rangle$.

Secondly, it should be clear that the normative coherence constraints can only serve as a *necessary* condition for the representational unity of experiences. The fact that two experiences e_1 and e_2 (or two sets of experiences E1 and E2) are coherent *does not suffice* for these experiences to be unified. They could very well be respectively entertained by two different subjects, Sam and Maria, in which cases e_1 and e_2 (E1 and E2) would obviously not be phenomenally unified (Hurley 1998: 120). Nevertheless, if these constraints turn out to efficiently enforce coherence within our representations, this will still be a significant result because cases of blatant incoherence will constitute evidence of phenomenal disunity (*ibid.*). This said, I will push Hurley's view a step further in arguing not only that the subject's experiences are coherent, but that the subject somehow *strives* for this coherence; and while coherence is not strictly sufficient for phenomenal unity, the subject's striving for it is.

Finally, for these constraints to put pressure on representational integration so as to make it substantive, they must be normative together. In other words, it is necessary that representations respect them *all*, and not just either one of them, to be unified. Upon scrutiny, this will turn out to follow naturally from the characterization of these norms. Indeed, we will see that they work hand in glove, so much so that compliance with one necessitates compliance with the others. While this requirement makes representational unity a fairly substantive claim, it also makes it all the more exposed to failure: a single case of infringement of either norms will suffice for the objector to conclude to the truth of (3), and consequently to the rejection of our view that for experiences to be phenomenally unified is for them to be representationally closed.

5.3.1. Semantic consistency

Be M1 and M2 modes of representation, and $\langle p \rangle$ and $\langle q \rangle$ representational content. On the view just canvassed, for a subject S's mental states M1(p) and M2(q) to be representationally integrated is for S to be aware of $\langle M1(p) \ \& \ M2(q) \rangle$. With the integration of experiences thus construed, does the semantic consistency require that representational contents $\langle p \rangle$ and $\langle q \rangle$ be consistent? It does certainly not so. For example, you frequently believe that the weather is not sunny (*not-p*) and desire that the weather be sunny (p). Here, even though the contents of your experiences are semantically inconsistent, the co-representation of these two experiences is not incoherent, and this is because the two intentional modes do not have the same direction of fit: the belief is *thetic*, while the desire is *thelic*. There cannot be semantic inconsistency between the two kinds of attitudes, for they fulfill

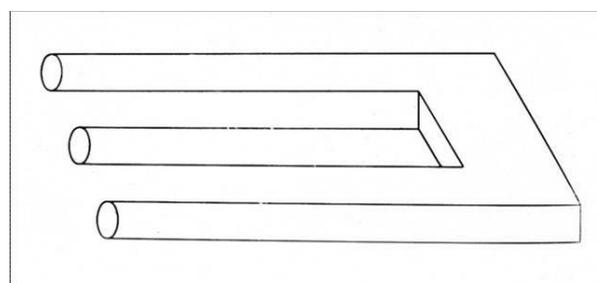
different and relatively independent goals: the former aims at the *truth* whereas the latter aims at *realization* (see sub-section 4.3.3.).

Let us therefore hypothesize that semantic consistency holds (at least) within experiences that have the same direction of fit – I shall characterize these experiences as ‘uni-directed’ –. It unfolds roughly as follows: for two uni-directed experiences $M1(p)$ and $M2(q)$ to be representationally integrated, their content $\langle p \rangle$ and $\langle q \rangle$ must be consistent with each other. Therefore, bear in mind at all times that when talking about (in)consistent experiences hereafter, I will focus exclusively on contents of uni-directed experiences.

It is frequently claimed that it is possible, and sometimes the case, that a subject undergoes phenomenally unified experiences that represent inconsistent contents. The arguments against closure generally take the form of examples of such experiences. I will now put the consistency constraint to the test by briefly discussing some of these examples.

One can distinguish between two kinds of inconsistencies: inconsistent content of experiences of the same mode (unimodal inconsistencies) and inconsistent content experienced across modes (intermodal inconsistencies). Bear in mind that I am restricting the discussion to uni-directed experiences.

So-called ‘impossible figures’ may be adduced as examples of unimodal inconsistency (Tye 2003: 38; Bayne 2007: 204). Consider the devil’s pitchfork: it has three prongs at its left end, but only two at the other end. Closure would imply an experience whose integrated contents would be inconsistent, viz. $\langle V(p \ \& \ \text{not-}p) \rangle$, where ‘ V ’ and ‘ p ’ respectively stand for ‘visually experiencing’ and for ‘the fork is three-pronged’. Another putative case of visual inconsistency is the waterfall illusion. Right after looking at a waterfall for a long enough time, the motionless object at which one looks appears both to be moving and to stay still. There are at least two ways of explaining away these inconsistencies.



The Devil’s Pitchfork

First, it could be argued that the properties of having three prongs and of having two prongs are not ascribed to the same object: the first is ascribed to the pitchfork *with respect to its left end*, whereas the second is ascribed to the pitchfork *with respect to its right end*. Hence, when you visually

experience the pitchfork, your experience is phenomenally unified in virtue of the binding of two objects with different properties. (As an analogy, think of a car whose trunk is fat and round and whose hood is slim and sharp.) Hence no inconsistent content is actually experienced.

Second, the claim that these cases even involve a single and phenomenally unified experience could be challenged. When you look at the devil's pitchfork, it rather seems that you see it as having not simultaneously but *alternatively* two and three prongs. What this means is that at t , you experience $V(p)$, but it is only after shifting your gaze onto the fork's handle at $t + \Delta$ that you experience $V(\text{not-}p)$. Thus, there are no less than *two consecutive* experiences of the pitchfork. Similar explanation holds with a fair amount of plausibility with regard to the waterfall illusion: you see the motionless object not moving at a given time, only to see it moving at the next moment, and these impressions switch back and forth several times. Hence those experiences hint at most at *diachronic* representational disunity. This result is completely innocuous. One often revises or abandons beliefs; the unity of consciousness over time cannot plausibly be constrained by the norm of semantic consistency.

What about doxastic experiences? Tye (2003: 116) argues that a subject can simultaneously believe that the sunny weather is wonderful and believe that the sunny weather is not wonderful. But this certainly stems from an equivocation in the propositions: for instance, the first occurrence of 'wonderful' means *esthetically pleasant*, while the second means *healthy for my skin*.

What about thetic representations such as desires and intentions? *Pace* Kobes (1998: 534), I argue that it cannot be the case that Hamlet simultaneously desires to be and not to be. There is not, nor will there ever be, anything that it is like to experience these two desires *together*, in a phenomenally unified manner. A very likely explanation of Hamlet's experience is that he hesitates: he alternatively has a desire to be and a desire not to be; therefore his famous phrase is yet another example of inconsistent representations over time.

I submit that all these cases of unimodal inconsistencies have been satisfactorily dismissed. For lack of more convincing cases, I consequently take the consistency constraint to hold *within any* modality, like this:

Semantic consistency: for any two unimodal experiences $M1$ and $M2$ of subject S at t , if $M1$ and $M2$ are phenomenally unified, their closure cannot take the form $\langle M(p \ \& \ \text{not-}p) \rangle$.

Let us now turn to alleged inconsistent experiences of different (uni-directed) modalities. Classically, such cases involve inconsistency between the content of a sensory experience and that of doxastic thought. Consider for instance a half-immersed stick. Although you visually experience it as being bent, you simultaneously believe, and rightly so, that the stick is straight. Yet no matter how hard you remind yourself of the belief that it is straight, you cannot shake off the visual impression of

its bentness. In other words, it is said, you experience $\langle V(p) \ \& \ B(\text{not-}p) \rangle$ (where V and B respectively stand for visually experiencing and believing). Similar conflicting representations occur in the Müller-Lyer illusion.

A first understanding of these inconsistencies is simply that they do involve a phenomenally disunified experience. But this is implausible and makes phenomenal unity too fragile – whichever of these two considerations grounds the other –. I will assume, quite uncontroversially I hope, that they are phenomenally unified – that there is something it is like to undergo both experiences *together*. Some will try and deny that these are real semantic inconsistencies at all cost. Others, on the contrary, will bite these bullets and face the implications that they bear on representational unity.

To begin with, it must be acknowledged that the strategy of separating both experiences in time in order to account for these inconsistencies is not available in those cases. They no doubt both happen simultaneously. Another strategy for avoiding phenomenally unified inconsistent experiences of different modalities would be to differentiate the objects to which properties are ascribed. Thus, it could be argued that the bentness is ascribed to a half-immersed stick, that is, to a stick composed of an above-water half and another reflected half, while the straightness is ascribed to a completely above-water stick. But things do not plausibly occur in this manner: you do not take the half-immersed stick to be a different object than the above-water stick. What is more, the fact that a subject would be able to ascribe the two inconsistent properties to different objects presupposes that she would be well aware that her visual experience is illusory. Hence this explanation could not plausibly account for cases where the subject does not have this prior knowledge.

A perhaps more promising strategy, championed by Glüer, would be to argue that sensory experiences, contrary to doxastic ones, do not ascribe sensible properties such as bentness, but instead ascribe 'phenomenal properties' (Glüer 2009: 311), i.e. properties such as *looking-bent*. Since the experience of seeing the bent stick and that of believing it to be straight would not involve ascriptions of the same kind of properties, there could not be anything like a semantic inconsistency between sensory and doxastic experiences.

Although this strategy initially elicits sympathy, at least with respect to the experience of the bent stick, it quickly proves riddled with problems. In addition to overpopulating our ontology with a fresh kind of properties, it raises the question of why representations of phenomenal properties normally lead to representations of sensible properties – typically beliefs. How can the latter be inferred from the former? The problem is that if there is some kind of principle bridging the two kinds of properties, it presupposes that phenomenal and sensible properties share a common *genus*. It may then be argued that beliefs and perceptions actually consist in ascriptions of this common genus of properties, and that consequently the experience of the bent stick does involve an inconsistency.

Besides these theoretical problems, Glüer's distinction of properties is at odds with the way subjects appear to deal with inconsistent experiences. Consider for instance the inverting-goggles experiment, wherein subjects wear special goggles that make them see the world upside down:

The subjects reported that at times they had the impression that they themselves had been turned upside down, but that they knew that this was not the case. But no subject claimed that he had regained upright vision or that his visual image matched his body sense at any point during the experiment, not even when the subjects were exploring the visual scene by touch. (Linden et al. 1999: 475, quoted in Bayne 2010: 56)

The puzzlement exhibited by the subjects indicates that they *experience* the inconsistency of their representations, which in turn indicates that these representations involve the same kind of properties. Furthermore, although upright vision was not 'restored' over time, the subjects were able, after a certain time, to 'coordinate' their movement (their tactile experience) with their inverted view. And this phenomenon, known as 'perceptual adaptation', necessitates that the different perceptual modalities represent a shared pool of properties:

When observers look through a prism, the visual image of the world appears laterally displaced from where it really is. A (point on an) object that is grasped will appear to be in one position through vision but in a different position through touch. The spatial discrepancy leads to adaptation (e.g., Wallach, 1968; Welch, 1978), in which one modality will change to realign with the other. (Bedford 2004: 908)

I think that the lesson to be learnt from these considerations is that the cross-modal (uni-directed) inconsistencies should not be denied at all cost. I think that the reasonable thing to do is to fully acknowledge these inconsistencies, as do Hurley (1998: 115) and Tye (2003: 38). However, they by no means undermine my account. *Au contraire*, they bring water to my mill. In the next subsection, I show that the way in which the subject deals with inconsistent representations suggests that said representations are indeed normatively integrated.

5.3.2. The normative 'I'

Let me begin by giving a thumbnail sketch of the three other norms that likely constrain our representations:

Inferential Integration: For any two different thetic experiences e_1 and e_2 of subject S at t , S must believe in the implications of e_1 and e_2 's contents, unless these implications are semantically inconsistent (as per *Semantic Consistency*).⁵⁴

Means-End Rationality:

⁵⁴ Note that the 'implications' of beliefs include their conjunction: if S believes that p and believes that q , S 's conjunctive belief that $p \& q$ is an implication of the first two beliefs.

If S believes at t action ϕ to be impossible, then S cannot intend at t to ϕ .⁵⁵

If S intends at t to ϕ and believes at t ψ to be a means to ϕ , then S incurs at t a commitment to intending to ψ . (Millar 2004: 109)

Agentive Integration: for any subject S, and any set of experiences E of S at t , necessarily E commit S at t to performing coherent and compatible actions.⁵⁶

Most of the time, we are not reflectively aware of our experiences. For instance, you do not typically entertain a reflective awareness that you want a cup of coffee and believe that there is coffee at the coffee place. Yet these experiences comply with the norms just sketched, without requiring neither effort nor active coordinating on your part. If they conjointly comply with a norm (here, *Means-End Rationality*), then they must be somehow integrated. But as such, the fact that this integration constrains them to yield appropriate reasoning and behavior without neither actual decisions nor reasoning on our part is highly puzzling. Where does the normativity of integrated representations come from?

My contention is that if our experiences are so often efficiently integrated in this functional manner, it is because they are integrated by the awareness that they are ours. I hypothesize in other words that the normativity of this integration stems from this very self-awareness. More specifically I take the elusive 'I' to which they are non-reflectively ascribed to fulfill a monitoring role: it monitors our experiences in such a way as to make them coherent and functionally integrated for certain purposes. Indeed, our experiences adjust to each other (Shoemaker 1996: 182-6). For instance, your intention to go to the coffee place will vanish along with the belief that they sell cups of coffee there. Likewise, your desire to rest your left foot during the Céline Dion concert will likely supplant other desires (for instance, the desire to stay until the end of the show) if the pain grows stronger. Far be it from me to advocate for this hypothesis solely on grounds of a couple of examples. These are just supposed to get the gist of my proposal across. Yet it may still be objected at this stage that my proposal amounts to little more than guesswork. And rightly so: how could this view be positively defended? I have argued that we normally do not experience in any reflective manner that our representations are unified. Nor do we experience the self's normativity in any direct fashion.

As a first motivation for my idea, I should stress that it fills a great explanatory void. The normativity of our experiences *has* to be imported from some entity. The elusive 'I', to which a subject's experiences are non-reflectively ascribed, seems a good candidate for the source of our experiences' normativity. It may be in just this spirit that one regards all conscious creatures to be

⁵⁵ 'The sense of impotence inhibits the volition', argues James (1890: 560n1).

⁵⁶ These norms would have to be properly specified of course, but a rough characterization will suffice for our inquiry.

somehow 'driven' towards some *telos*, something broadly akin to self-preservation. But admittedly, it is too thin a thread to sew a whole theory together.

Ironically, it is the cases of inconsistent experiences such as those presented above which most strongly suggest that there is a pre-reflective integration of our experiences. Something along those lines is already hinted at by Hurley's observation: 'intermodal incoherence tends to be unstable, especially when the subject is active, and... this tendency in itself reveals something essential about the unity of consciousness' (Hurley 2000: 560). What does Hurley mean by this instability?

Let me venture the plausible-sounding view that a subject's global phenomenal experience, though it sometimes admits of inconsistent representations, *strives* for coherence and efficient functional integration. Concretely, it means that when inconsistent representations arise in the subject's experience, they are immediately put under pressure by coherence norms. In other words, they are monitored by the subject's awareness that they are both hers. But let us not rush to this conclusion. First, let us make good sense of the idea that experiences are integrated by their constitutive self-awareness.

Imagine your visual (and tactile) experience of a straight stick, and then, after half-immersion, your visual experience of a bent stick. For the sake of the argument, let us assume that you are a child, and that it is the first time that you enjoy such an experience. Once the stick is half-immersed, you simultaneously entertain these two experiences:

- (a) I visually experience that the stick is bent. $V(p)$
- (b) I believe that the stick is straight (and thus not-bent). $B(not-p)$

The experience (b) naturally follows from your perception of the stick before its immersion. It is rather straightforward that (a) and (b) are *phenomenally* unified. Yet something else occurs in your experience. Since you are not acquainted with the impact of water reflection on vision, you initially take your visual experience to be trustworthy. But then an inconsistency arises, since representations (a) and (b), beingthetic, together suggest that it is the case that both p and $not-p$.

I submit that a feeling of unease or puzzlement will pop up in your conscious experience, probably accompanied with something akin to the following reflective question: 'how can I experience $V(p)$ and $B(not-p)$ together (implicitly : given that they are both supposedly truthful)?' I take these additional experiences to be conclusive evidence of the integration of experiences (a) and (b) by self-awareness. It would seem odd to claim that the self-ascription of your experiences is merely a result of your reflective question. Since the feeling of unease is spontaneous, and since the reflective question that arises following the inconsistent experiences is not *deliberate*, it seems as though the non-reflective self-awareness of both representation is *prior to* and *responsible for* the unease and the arising of the

reflective judgment. To put it differently, it is *because* your experiences necessarily represent themselves as being yours that they sometimes spontaneously bring about *reflective* self-awareness.

It is now that the monitoring role of self-awareness is put to the fore, and encourages the idea that inconsistencies are *unstable*. The inconsistency of the integrated representation $\langle V(p) \text{ and } B(\text{not-}p) \rangle$ will immediately tell you that something is wrong. And this, in turn, will likely give rise to the following rationale:

A visual experience that p generally leads to a belief that p . But in the present case, such inference would lead to the belief that p and $\text{not-}p$. It cannot be the case that p and $\text{not-}p$: a stick cannot be both bent and not-bent! I cannot believe such a result. Therefore, one of my two representations must not be what it looks to be. Either it is not the case that I see truthfully a bent stick, or my belief that the stick is straight is false, in which case I ought to stop entertain it. I must either take my perception that p to be erroneous, and thus block the inference that would normally lead me to the belief that p , or take the evidence on which was inferred my belief that $\text{not-}p$ to be erroneous, and thus jettison this belief.

Being accompanied by grown-ups that eventually tell you all about light refraction, you end up modifying your experience by considering your representation an '*illusory* vision that p' – $IV(p)$. The idea is not that the immersed stick *looks different* than it looked before you operated said modification, for there is little doubt that the visual experience, *qua* visual experience, will feel exactly the same. Rather, it is that you adopt a cautionary stance towards this visual experience that *prevents you from inferring a belief in its content*. The inconsistency thereby vanishes. It is in this sense, and *to this precise end*, that the two representations are integrated.

It is worth noting that this is neatly on par with our general view on phenomenology, for the modification of representation $\langle V(p) \text{ and } B(\text{not-}p) \rangle$ into $\langle IV(p) \text{ and } B(\text{not-}p) \rangle$ will be naturally be accompanied with a modification of the phenomenal feel of the whole experience. Your introspective judgment to the effect that you are experiencing something abnormal will disappear, as well as the sense of unease and puzzlement.

Of course, the piece of reasoning just sketched is conscious and verbalizable only the *first time* that you experience such inconsistency. Chances are that your subsequent experiences of the half-immersed stick will automatically bring about the perfectly inferentially consistent experience of $\langle IV(p) \text{ and } B(\text{not-}p) \rangle$, without you even thinking about it.

In fact, inconsistencies between thetic modes of representations may be more frequent than what the development of the last section might have us believe. But this is only mildly problematic, since their being co-represented as belonging to the subject ensures that the inconsistency signals itself by means of reflective judgment, which in turn triggers a rectification. I must emphasize that the rectification does not entail a change in how things *look* or are represented; rather it entails a change in the way that we treat the representation in our system of belief and inference.

Importantly, it should be noted that inconsistencies are not the only experiences in which the monitoring self-awareness reveals itself. The monitoring role of self-awareness in agentic unity is put to the fore by well-known thought experiments involving head-body separations. First introduced by Dennett (1981), numerous versions of the thought experiments have since circulated in the literature.⁵⁷ The version that I propose here is a variation on Tye's.

Imagine that one night, during your sleep, your eyes and ears are unbeknownst to you separated from your body in such a fashion that your brain still processes, as though no separation had occurred, the stimuli coming from your audio-visual system and the stimuli coming from the body – the technological advancements in neuro-tele-transmission enable such procedure –. Imagine then that your body is completely buried in sand, and that your eyes and ears are placed on a roller-coaster ride. Upon waking up, you instantly experience sensations that represent your body in sand, and an audio-visual content to the effect that you are on a roller-coaster.

I follow theorists in agreeing that these experiences are phenomenally unified (again, I insist on the fact that neuro-transmission of your audio-visual stimuli to your brain is perfectly seamless). There is something it is like to enjoy the vista from a roller-coaster perspective while feeling completely buried in the sand.⁵⁸ Moreover, your experiences are strictly *consistent* with one another. In general, bodily sensations are indexed to specific body parts; for instance, when you experience a particular pain, you represent some tissue damage *in a particular location of your body*. The same goes in the present scenario: you tactilely experience the sand relatively to almost all body parts, but there is no tactile experience of the sand relatively to your eyes and ears; in its stead, there is something like a tactile experience of a breezy air relatively to these body parts, given that they are located on a roller coaster. Therefore, there is no tactile experience whose content is $\langle p \ \& \ \textit{not-p} \rangle$.

Still, this integration of representations is highly peculiar and disturbing, since you normally do not experience such sensations *together*. When riding a roller coaster up in the air, you usually do not feel buried in sand; when letting your buddies bury you in sand on vacation, you usually do not enjoy the vista provided by a roller coaster ride. This uncustomary character aside, your experiences leave you puzzled because their integration plausibly leads to very different, if not incompatible, beliefs and behaviors. Consequently, upon awakening from this nightly transplant, you will likely undergo the following reflective thought: 'I am currently feeling buried in sand while audio-visually experiencing a roller-coaster ride'. You will moreover ask yourself: 'How could I be undergoing these experiences?!' Indeed, this integration will cause you trouble deciding *what to infer* ('Should I believe that I am in the

⁵⁷ See *inter alia*: Bayne 2004: 223 ff., Dainton 2000: 67 ff. and 2003; Schechter 2010b: 3; Tye 2003: 78.

⁵⁸ As Bayne notes, a potential problem with this thought experiment is that it assumes that a subject's audio-visual sense field can function in independence from her body-centered sense field, while there is no guarantee that it actually can (Bayne 2004: 224). Hence we should assume this independence with a pinch of salt.

sand, or that I am riding a roller-coaster?’ and *how to act* (‘Should I wait till the roller coaster ride stops before I try to unbury my head of the sand?’).

5.4. Conclusions

Let us take stock. In this chapter, I have argued that the phenomenal unity of the subject’s experiences consists in their being represented together by a unique state of self-awareness that non-reflectively ascribes them to the subject. When you desire that Céline Dion signs an autograph to you, this experience inextricably comes with the representation that this desire is yours. When you contemplate Céline Dion on stage, this experience inextricably comes with the representation that this visual experience is yours. When undergone simultaneously, both experiences are represented *together* as being yours.

Experiences are of varying directions of fit that do not likely ‘merge’ into one. The total state of self-awareness that integrates representations together not only respects and preserves this diversity; it yields seamlessly fruitful and coherent coordination in the subject’s reasoning, decision and action. In short, representation of her experiences as being hers is what makes them normative. Some experiences have individually no inferential or behavioral impact. It is in virtue of being brought together by their relation to the subject that they give rise to rationales and actions. Thus self-awareness can be seen as fulfilling substantial *monitoring* functions. In other words, the representational integration achieved by self-awareness is what makes phenomenal unity explanatory powerful rather than a mere epiphenomenon. And this easily puts to bed the worry that the integration of representations may be no more than an artificial sum. Let me then capture my account:

Self-Representational Account: For *S*’s experiences *E* to be phenomenally unified at a time *t* is for *E* to conjointly represent themselves as being *S*’s.

S’s self-awareness of *E* monitors *E* by putting them under the pressure of coherence norms.

In the last two sections, I have demonstrated that incoherent experiences are unstable, and that this instability actually constitutes *positive evidence* that they are represented as being undergone by the subject in an integrated and monitoring fashion. Indeed, the spontaneous reflective awareness that normally accompanies incoherence is best explained by a non-reflective awareness of these experiences. It is what signals to the subject that something is wrong in her experiences, and this is what moves her to adjust her inferences, or adopt different courses of action, so as to make her

experiences cohere again. It is in this particular sense that incoherent experiences are unstable: the elusive 'I' to which they are ascribed makes the subject strive for coherence.

From this proposal, it naturally follows that *stable, persistent* breaches of any of the coherence norms listed above is constitutive of phenomenal *disunity*. Indeed, if they are stable, then they are not pressured by the normative constrains. If they are not so pressured, then they are not monitored by the normative 'I'. Which means that *there is nothing it is like* for the subject to undergo inconsistent experiences *together*.

Moreover, as the examples considered above suggest, phenomenal disunity can be 'diagnosed' by the following bridging principle between behavior and phenomenal consciousness: a persistently incoherent behavior is evidence of a lack of integration of representations. And since for experiences to be phenomenally disunified is for them to be representationally disunified, a persistently incoherent behavior is evidence of phenomenal disunity.

An important implication of this conception is that the subject of phenomenal disunity cannot herself *experience* this disunity (as is observed by Dainton 2000: 109). For if she was able to experience it, it would thereby elicit behavioral adjustments in order to abide by coherence norms. It is for this reason that the subject's verbal report to the effect that she experiences no phenomenal disunity ought not to be taken as evidence that her experience is phenomenally unified. Thus the defense of my account of phenomenal unity differs significantly from Bayne and Chalmers' and Dainton's. I did not *infer* my account *from* the fact that a subject is aware (non-reflectively or reflectively) of her experiences' unity; I only built it in such a way that it *accounts for* this epistemic access.

These last considerations will come in handy upon considering neuro-pathological split-brain cases. But before putting my self-representational account of phenomenal unity to the test of empirical data, I shall ensure that it is theoretically sound by fending off potential objections that could beset it.

6. Objections and replies

6.1. Against the coherence norms

6.1.1. The norms' scope

A first criticism against the self-representational account concerns the plausibility of the claim that the norms sketched in the last chapter constrain a subject's experiences *across the board*. Indeed, these norms are very 'local' and *too* neat, in the sense that they constrained only a few representations of specific modalities, viz. doxastic and conative states. But one should not lose sight of the fact that the subjects' experiences at any time are numerous and of various modalities. Our favorite example of the Céline Dion concert involves bodily feelings, among which a global proprioceptive feedback and a mild pain in the foot, the thought that you are lucky, the elation, the audio-visual experiences of the stage, the olfactory experience of your fellow fans and so on. Hence the following concern may be voiced: can there really be norms constraining other experiences than beliefs and conative states? Are emotions or bodily feelings, for instance, so constrained? I have said (in section 5.3.) that for norms to apply to experiences, experiences must be construed as propositional attitudes, i.e. representational states whose contents are propositions. I must admit that I may have been a bit too expeditious in assuming this. This is not the time to discuss this structural aspect, but even if it is conceded that some experiences such as emotions and bodily feelings are 'raw', in that they are non-propositional and maybe even entirely non-*conceptual*, there are at least two reasons why this poses no threat to the claim that they are constrained by norms.

A first reason is that it is extremely plausible that so-called 'raw' experiences are provoked by and give rise to beliefs, and not only reflective beliefs to the effect that one is undergoing a raw experience but beliefs about how the world is. For instance, fear is generally accompanied by a belief that something dangerous is around;⁵⁹ should evidence defeat this belief, the fear will likely (or perhaps ought to) cease to be experienced. Thus emotions are no doubt constrained by certain norms, starting with *Inferential Integration*, even though I have not explicitly spelled them out.

This brings me to the second reason. Coherence and functional coordination are concepts that by far surpass in complexity those of semantic consistency, inferential integration, means-end rationality and agentic unity. So the norms proposed earlier should be understood as a few rules of thumb, if one will; they comprise an oversimplified sample of much more complex norms. There is good sense

⁵⁹ I deliberately remain agnostic as to whether this belief is constitutive of, responsible for or triggered by the fear.

in holding our raw experiences to be integrated with our thoughts. For instance, it is empirically averred that your proprioceptive feedback, no matter how discreet it feels in your phenomenal horizon, is integrated with your perceptual and conative representations in the way that you move. Just because you are not reflectively aware of it does not imply that they are not integrated together. If this lack of reflective awareness tells you anything, it is precisely that your experiences *are indeed* exhaustively monitored by being non-reflectively experienced by yourself.

6.1.2. Non-conceptual creatures and coherence

I have argued that since self-awareness of our experiences is non-reflective it does not require any capacity other than a non-conceptual sense of self and of what the subject experiences. Hence, there is no reason why non-conceptual (nonlinguistic) creatures should not be self-aware of their experiences in the same unifying way as conceptual subjects are. Moreover, since I have argued that the normativity of a subject's experiences is constitutive of their phenomenal unity, the fact that the non-conceptual creatures appear to behave in a coordinated fashion should be taken as evidence that their experiences are constrained just as those of conceptual subjects are, and thus that there is something it is like for them to enjoy their experiences *together* as theirs.

What does, however, appear to require conceptual skills is the having of reflective thoughts that arise in cases of incoherent experiences. Therefore, it may be argued that non-conceptual creatures lack the ability to *strive* for coherence, which on my self-representational account constitutes evidence of phenomenal *disunity*. But this strongly runs counter to our pre-theoretical intuitions; a lot of theorists would take issue with the claim that non-conceptual creatures cannot enjoy phenomenally unified experiences.

One could nip this concern in the bud by observing that the very notion of *incoherence* seems to demand that the subject's representations be conceptual; this would *eo ipso* make incoherent representations impossible within animals' and children's experiences.

But still, there could be a weaker, non-conceptual sense in which these creatures might entertain incoherent experiences. Let me take a simple example. Don is a toddler and desires an ice cream. He sees an ice cream on the kitchen table but he cannot reach it despite desperate attempts to do so. Yet he puts his heart and soul into trying to reach for it. So, as per the *Means-end Rationality* norm, there is arguably something like incoherence here: Don believes that (or more likely *feels as though*) it is impossible for him to reach the ice cream, and yet persistently keeps trying to do so for quite some time. Should we conclude from this apparent lack of striving toward coherence that there is nothing it is like for baby Don to undergo both experiences together?

I think that the incompatibility of Don's experiences manifests itself within his self-awareness. Of course, since Don is incapable of reflecting about his own thoughts (if his experiences can be so labeled), it does not take the form of a reflective thought to the effect that he intends to ϕ even though he believes that he is incapable of ϕ -ing. However, just as subjects in the experiences of wearing inverting-goggles and of seeing a half-immersed stick for the first time, Don likely *feels* that something is wrong with his experiences: something akin to frustration accompanies his incoherent experiences. This, in turn, triggers a certain behavior: he moans (if you are lucky) or cries. I wager that this is conclusive evidence that his proto-intention and proto-belief are phenomenally unified. His non-conceptual representations are integrated *and* monitored by his self-awareness in a way that signals proto-incoherence by the arising of a further (equally non-conceptual) *experience* of this incoherence. Similar stories can be told about animals' experiences.

Although the experience of incoherence shows the subject's experiences to be unified, it must be granted that in non-conceptual creatures self-awareness is not accompanied with the striving for coherence. So this calls for a refinement of my view into two different claims.

- (i) *Self-Representational account* : For a subject's experiences to be phenomenally unified is for them to conjointly represent themselves as being hers in such a monitoring way that if they came to be incoherent, an additional experience of this incoherence would arise in the subject's experience.
- (ii) *Strive for coherence*: A rational subject normally undergoes this experience of incoherence together with a reflective thought that makes her strive for coherence.

Up to this point I have more or less tacitly assumed that the coherence norms under which integrated representations were constrained were norms of *rationality*, and that subjects followed them when summoned to reflect on their experiences. In other words I took claim (i) to systematically work hand in glove with claim (ii). But baby Don's case has demonstrated that coherence norms cannot be reflectively followed by a-rational subjects. Moreover, it is quite commonsensical to occasionally peg a rational subject's behaviour as *irrational*, thereby suggesting that their experiences somehow failed to reflectively comply with the norms. Consequently, my view ought to be restricted to claim (i) in order to apply to all conscious creatures. But then, how do a-rational and occasionally irrationally-behaved subjects deal with the experience of incoherence?

Put yourself in baby Don's tiny shoes for an instant, and imagine simultaneously entertaining the intention to reach for this ice cream and the belief that the ice cream is irremediably out of reach. These representations being pressured by the *Means-end Rationality* norms, this incoherent integration is unstable, in the sense that you should feel there is something wrong with the fact that *you* are entertaining it. On my account, this feeling is conveyed by a reflective thought that represents

these experiences together as yours. The rational thing to do is to abandon one of your representations. Provided the belief is robust enough, you will abandon the intention to reach for the ice cream. Now consider baby Don's situation again. The frustration triggered by his incoherent experiences will not entail such a wise abandonment, since he is incapable of reflecting about them. Rather, the frustration will linger as long as he keeps entertaining them. It is only when Don wearily stops entertaining his intention or his belief (or both) that the feeling will fade off.

A similar denouement may occur in the case wherein a rational subject behaves irrationally. Instead of resolutely abandoning one of the incoherent experiences, the subject may choose to distract himself from the puzzling integration and momentarily stop undergoing the intention and think about something else completely. The immediate phenomenal result is the same as with a rational modification: the feeling (and reflective thought) that something is wrong with her experiences disappears, to her relief. But this is a dilatory measure, for the intention is likely to arise later in her conscious experience and to provoke yet again desperate gestures to reach it, as well as the thought that something is wrong with his experiences.

So, this should dissipate any concern that the self-representational account denies phenomenal unity to non-conceptual and/or a-rational subjects of experience. The conclusion that there is something it is like for these creatures to enjoy their experiences *together* swiftly follows from the integrated behavior that they exhibit most of the time and the fact that incoherent experiences are accompanied with unease or frustrations (which transpires through cries, barking, sometimes fist-pounding). Again, this does not require any reflecting nor any conceptual competences on their part; it is unreflective, and thus seamless and effortless. It is true that the conceptual shortcomings of these creatures do have consequences on how they deal with incoherent representations. But this by no means undermines my conception.

6.2. Against self-awareness

6.2.1. Humean concerns

Another worry with regard to my self-representational account bears on the conception of self-awareness that it presupposes. It is that we do not ‘find’ self-awareness in our phenomenology and that we should therefore restrain ourselves from positing it.

It is not clear how one should go about defusing this eliminativist attitude vis-à-vis self-awareness. Indeed, there may be no *positive* demonstration of its existence. This argumentative lack is symptomatic of any statement of existence. For instance, how can I positively prove that there *is* a watch on my wrist to someone who claims there isn’t? It seems that the only strategies available to the realist are to show anti-realist arguments to be weak, to pinpoint the factors which may have led the anti-realist into error, or to invoke the explanatory advantages of the realist view (Kriegel 2007: 27). I shall briefly mention three factors misleadingly underlying anti-realism about self-awareness, and one explanatory advantage of realism.

First, anti-realism about self-awareness may stem from immoderately high expectations regarding the phenomenal impression that our experiences are supposed to make on us. Perhaps anti-realists take color experiences as the ‘standard’ for phenomenal character. But it is pretty straightforward that any other experience is bound to be less phenomenally impressive than those. As has already been pointed out at the beginning of last chapter, our *de se* representations are characteristically elusive. This idea could not be better put than by Chalmers:

One sometimes feels that there is something to conscious experience that transcends all these specific elements [visual experiences, auditory, olfactory, tactile, and taste experiences, experiences of hot and cold, pain and other bodily sensations, and conscious thoughts]: a kind of background hum, for instance, that is somehow fundamental to consciousness and that is there even when the other components are not. This phenomenology of self is so deep and intangible that it sometimes seems illusory, consisting of nothing over and above specific elements such as those listed above. Still, there seems to be something to the phenomenology of self, even if it is very hard to pin down. (Chalmers 1996: 10; quoted in Kriegel 2004: 194)

A first explanation of this discreetness is that it is a *constant* feature of our conscious life. Consequently we have no way of making it more phenomenally salient. Since we could not entertain a conscious mental state without representing it as ours, and since it has been a ubiquitous feature of our conscious life ever since we were born, we cannot contrast this particular subjectivity by entertaining (or imagining) an experience that lacks it. As Kriegel rightly notes, we easily fail to notice some stimuli that have been constant for a relatively short time (think of a refrigerator’s hum). Why, then, would our non-reflective awareness of ourselves not be so discreet?

The second related reason why self-awareness is so intangible is that it completely eludes our control. Remember that self-awareness is a *peripheral* awareness. But think about other kinds of peripheral awareness, in vision or audition for instance. Whilst we certainly choose to direct our *focal* awareness at certain world features, we have no direct control of what we peripherally see and hear. Likewise, we do not deliberately choose to be peripherally aware of ourselves, and we cannot stop at will to be so peripherally aware. Which leads to the same conclusion as the first observation: we could not contrast our peripheral awareness of ourselves by putting a stop to it.

These remarks unsurprisingly apply to my account of phenomenal *unity*. There has arguably not been a time of in our conscious life when our experiences have not been co-represented as being ours. Thus we simply would not know what it feels to experience different things (this piece of paper, the chair on which we sit) in different ways (visually, tactilely) without experiencing them conjointly *as ours*. It is so embedded in our consciousness that we tend to neglect its phenomenal character; we take it for granted.

Finally, the existence of peripheral self-awareness may well be bolstered by its evolutionary and functional advantages. In effect, some theorists regard peripheral awareness as a cognitively economical way of keeping representational contents readily available. In the case of visual peripheral experience, for instance, it seems that the peripheral content of our visual field is too functionally important to be excluded from our phenomenology altogether (Mangan 2001, quoted in Kriegel 2005: 27 n. 12). Hence it is kept in something akin to the Jamesian ‘fringe of consciousness’ – what we poetically termed *experiential penumbra* earlier (see sub-section 2.2.2.). Likewise, it is not unreasonable to posit self-awareness in light of its apparent functional benefits. Indeed, I take my conception to have sufficiently advocated in favor of the integrative and monitoring advantages of self-awareness.

6.2.2. Cartesian implications⁶⁰

Another potential criticism that could be made against self-awareness is that it may imply an overly Cartesian picture according to which we have a special, *transparent* knowledge of our mental states. For our present discussion, let us restrict this privileged epistemic access to our *conscious* mental states. Starting from a common view on which knowledge entails (at least) justified true belief, Descartes’ view could be synthesized as the conjunction of the following claims:

Infallibility: whenever a subject S undergoes an experience E, S has a true belief about E.

⁶⁰ This objection is forestalled by Kriegel (2005: 26-7, n. 10).

Incorrigibility: whenever S undergoes E, S has a justified belief about E.

As Kriegel points out, the thesis that S is necessarily aware of E entails neither *Infallibility* nor *Incorrigibility*. In the elaboration of my account I have repeatedly stressed the peripheral character of self-awareness. A good example of what peripheral self-awareness exactly represents is by analogy with another kind of peripheral awareness, peripheral vision. Right now you are peripherally aware of the table on which your computer (or your draft of this work) is set. Contrary to your focal visual experience of the computer, this peripheral experience consists in a representation with a rather poor and non-conceptual content. Likewise, peripheral self-awareness amounts to a non-conceptual representation of you entertaining a first-order representation. This lack of conceptual content comes with the faintness of this self-awareness, and it permits that non-conceptual creatures also be aware of their mental states – and thus that these mental states be conscious.

The fact that I regarded self-awareness as an additional thetic mode should not mislead us into regarding self-awareness as a belief. Self-awareness is thetic in the same sense that proto-belief states of non-conceptual creatures such as baby Don are thetic: it has certain satisfaction conditions with regard to how the world, and more specifically the mind, is. However, from the non-conceptual character of the representation in which self-awareness consists, it naturally follows that self-awareness is *not* a belief, and *a fortiori* not a justified and/or true one. Beliefs differ from self-awareness in two crucial respects: they always have a propositional (and thus conceptual) content, and they entail the subject's endorsement of this content. Both these features are not exhibited by self-awareness, understood in the peripheral sense.

In short, the most that a self-awareness account entitles us to is the claim that whenever S undergoes E, S is *first-personally* and *immediately* aware of E and of undergoing E. This privileged access is completely non-committal with regard to its 'epistemic quality'; in fact, I have conceded that S's awareness of E is not immune to erroneously representing E (see n. 51, sub-section 5.1.3.)

Admittedly, self-awareness guarantees a weaker transparency with regard to our experiences. First, it cannot represent an inexistent representation. Second, it cannot wrongly represent a representation as being *another subject's*. If it represents a representation, and if the integration of the two yields an *experience*, the higher-order representation is perforce that of the relevant subject's – it would not yield an experience otherwise. In any case, this weaker transparency is completely unpresumptuous and quite uncontroversial, contrary to the 'strong' transparency posited by Descartes.

6.3. Phenomenal unity and the self

6.3.1. The self and the 'I'

At the very beginning of my inquiry I invoked the deep-rooted philosophical problem of the self and its identity as a prime motivation for investigating the unity of phenomenal consciousness. Since I have hitherto remained silent on this matter, it is time for me to take stance on the relation that binds the self to phenomenal unity. This issue goes well beyond the remit of this work, so I shall sketch my view in broad outline only; I do not doubt that it ought to be refined and worked out in more detail.

I concur with Bayne in adopting the following approach to the self: rather than assume a fixed conception of the self and assess whether it must be phenomenally unified, it seems more judicious to start from the claim that the self must have phenomenally unified experiences and then construct a conception of the self accordingly (Bayne 2010: 281).

The received view, which will be my working hypothesis, is that phenomenal unity is constitutive of a self. More specifically, I argue that what constitutes a single self is a single *stream* of consciousness. The following definition will suffice, at least for the time being, to get a handle on the notion:

Stream of consciousness (Generic): a maximal set of phenomenally unified experiences.

Before elaborating on this conception of the self, let me forestall a confusion to which it may give rise. The idea that the self is constituted from a stream of consciousness ought not to be understood literally as the claim that the self is *identical* to a stream. Indeed, this would be at odds with our commonsense way of treating the self as the *owner* (or 'entertainer') of a stream. Rather, the relation of constitution between a stream and a self means that the self is defined by, or abstracted from, a stream.

Now, it appears at first glance that this conception of the self is circular: the self is defined by a stream of consciousness, but the stream is defined by the agglomeration of experiences to the same self. This circularity is only apparent, however; it is due to a confusion between the *self*, on the one hand, and *the subject of experience*, on the other.

The entity that I claimed is at the core of our experiences' unity is the subjective character that is experienced with each of our conscious mental states. It is the object of the *de se* reference, i.e. the 'I' that is referred to in the implicit, non-reflective awareness that *I* am seeing the computer screen. This 'I' is, to borrow Rosenthal's phrase, a 'raw bearer' of experiences (2003: 330). This is not to say that the 'I' is what metaphysicians call a *bare particular*. Bare particulars need not bear any property in order to exist. By contrast, the 'I' to which we implicitly ascribe all our experiences would not exist absent the experiences that are self-consciously ascribed to it. This 'I' is an 'essential indexical' in the

sense that it refers only to ‘the very individual who is entertaining this experience’, nothing more. It makes no reference to the representation in which consists this experience.

Pace Descartes, we do not *identify ourselves* as mere substances, understood as raw bearers of mental states (Rosenthal 2003: 334). The self with which we identify ourselves is rather the bearer of a broad collection or ‘bundle’ of interconnected mental states, memories, character traits, locations and so forth. We generally take this bundle to be *contingent*. For instance, it could have been the case that you did not go to the Céline Dion concert and that consequently you did not have beautiful memories of it; more dramatically, it could have been the case that Maria never made the delightful acquaintance of Sam. Accordingly, you and Maria could have identified your respective self differently than how you actually do (Bayne 2010: 282).

The difference between the two notions, the ‘I’ and the ‘self’, is vividly illustrated by Perry’s (1979) example of the sugar spilling cart (presented in Rosenthal 2003: 338). Imagine that I see a trail of sugar apparently spilling from somebody’s cart; if I am the one actually spilling sugar unbeknownst to me, and if I think that the person who is spilling sugar is making a mess, it does not imply that *I*, myself, am making a mess. This suggests that the particular way in which the subject of experience refers *de se* is independent from the contingent collection of properties (for example mental states, memories, character traits and so forth) that she invokes in order to describe herself.

The ‘I’ and the self differ from each other in that they each enjoy a distinct mode of being. The self, like a center of gravity of an object, can be regarded as a fictional object, ‘abstractum’ or ‘fictum’, in the sense that they result from a theorist’s abstraction; you cannot locate it anywhere (Dennett 1992: 276).⁶¹ To illustrate the self’s mode of existence, Dennett takes the example of Sherlock Holmes. Sir Conan Doyle wrote stories featuring a person with certain properties such as bearing the name of Sherlock Holmes, having an incredible sense of deduction and an addiction to cocaine, and so on. This person is the novelist’s fiction; and as the novelist’s stories are written and more properties and events are recounted, we can ascribe more and more properties to this fiction. As a result, Sherlock Holmes, this thing that we abstract as the bearer of these properties, becomes more and more ‘determinate’. Likewise, Dennett invites us to regard the self with which the subject identifies as a virtual object to which a collection of mental states, memories and so on, is ascribed. Although I am inclined to adopt Dennett’s view, I wish to inject a dose of realism into this picture.

Bayne prescribes that we treat the ‘I’ which is referred to by *de se* reference as a virtual or ‘intentional’ entity as well (Bayne 2010: 289). He holds this view simply because he *identifies* the self

⁶¹ Of course, you could say that they are ‘located’ in the brain. But this would be an abuse of language, for you could not pick out a special region of the brain (or all of it, for that matter) and say that this is the thing that is referred to when we say ‘self’.

to the 'I'. As I have just mentioned, I cannot indulge in such identification on pain of circularity. However, I wish here to prove Bayne's position to be flawed regardless of whether my account is true.

Bayne takes experiences to be undergone by a merely fictional entity, and although he first expresses some reservations vis-à-vis this anti-realist stance, he ends up soothing them by claiming that the 'I' so virtually conceived is 'as real as [it] gets' and 'provides all the reality that we *need*' (Bayne 2010: 292-3, emphasis his). I very much doubt that it does, for the fact that a fictional entity could be the subject of experiences that are presumably *real* is very puzzling (Van Gulick 2014: 497). It seems instead that it *must* be an existing entity. Another way of motivating this claim is to rebut the idea that it *could* be a virtual entity. *De se* reference is so imbedded in the subject's experiences that it is not only concomitant with their occurrence but completely unintentionally experienced. Thus, one could hardly construe it as the result of a theorist's fiction in any sense. Yet another way of putting it is this: were there no *real* subject of experience, there would be *no experiences at all* from which to abstract a virtual self in the first place.

The reason why the I/self distinction may be overlooked is that they bear epistemic resemblance to each other: they are both first-personally accessible. What is implied by this is that when you see Céline Dion on stage, you are immediately and non-inferentially aware of you (raw bearer) seeing Céline Dion on stage. Similarly, when you are asked (say) about who it is who is currently seeing Céline Dion on stage and was gifted his (her) first Céline Dion mixtape at age nine, you can immediately and non-inferentially refer to *yourself* as the bearer of these experiences, along with many others. By way of describing yourself you can just invoke on the spot your past and present experiences without engaging in any reasoning.

However, there is a fundamental difference between the raw bearer and the self. Your non-reflective awareness of yourself as the bearer of your experience is *essentially* first-personal. I could not be acquainted in any sense with the subjective character of *your* experience. By contrast, the self is first-personally *and third-personally* accessible, and this straightforwardly follows from its ontological status of abstracta. Just as we are entitled to abstract Sherlock Holmes from Conan Doyle's stories, we could abstract your self as the bearer of particular collections of mental states, memories, character traits and so on, were you to tell us about it.

With this crucial distinction in hand, we can now make sense, without incurring circularity, of the intuitive claim that selves should be 'constructed out of streams of consciousness' as Bayne says (2010: 281). The 'I' is what entertains such-and-such experiences at a given time. It is the subject which is non-reflectively self-aware. As such, its existence involves no reference whatsoever to what these experiences represent; it is a raw bearer of experiences *qua* experiences. And since there is but one raw bearer of all the subject's experiences, this raw bearer is the entity around which the stream of

consciousness is centered. Having initially endorsed a realist view of consciousness, we are also committed to realism about its raw bearer; this elusive 'I' is the real deal.

Summing up, non-reflective self-awareness is what unifies experiences into a single stream of consciousness. This stream of consciousness determines the representations that the subject reflectively and narratively takes to define herself. Thus, this self is limned by the boundaries of the subject's stream. It is in this precise sense that selves are made out of streams of consciousness.

It is worth observing that it is for this reason that we tend to conflate the self with the subjective character of an experience. When the subject's experiences cohere from the perspective of the raw bearer 'I', the subject's abstraction of her own self is normally automatic and seamless, which is why she can readily refer to it in an immediate, non-inferential manner. An implication of this is that generally, the abstraction of the self as the bearer of a stream *coincides* more or less with the raw bearer of the experience in the sense that they both bear the same coherent experiences (Shoemaker 1996: 186; Van Gulick 2014: 497).

However, this automaticity is sometimes hampered by unfortunate integrations of experiences such as those involved in the inverting-goggles experiment. In those cases, the subject of experience is momentarily invited to reflect on what experiences she takes herself to be the bearer of. What happens is that the normative 'I' refuses that a self be abstracted from incoherent experiences, and therefore elicits a modification of the experiences of which the self is supposed to be the center of gravity.

Arguably the most extreme illustrations of the ontological independence and discrepancy between the 'I' and the self are supplied by patients suffering from schizophrenia, multiple personality disorders (MPD) and depersonalization. I submit that the schizophrenia or MPD patient has a single experiential perspective, that is, a single *de se* reference, but that she somehow abstracts *several selves* from several bundles of coherent experiences (Dennett 1992: 282; Bayne 2010: 290). The subject suffering from depersonalization, by contrast, will fail altogether to complete her conception of self with certain experiences that she yet experiences *as hers*. Put otherwise, the depersonalized subject undergoes an experience that she cannot ascribe to her narrative self; but the fact that she experiences it entails, on our view, that it is experienced *from a certain point of view* – 'I' –. *Pace* Hume, there is no such thing as an unborne experience.

6.3.2. The continuity of the self

An important challenge of an account of the self is to ensure its unity and continuity over time. If the self is constructed out of a stream of consciousness, it entails that the self ceases to exist during ‘phenomenal gaps’, typically nightly hours of dreamless sleep. This is a considerably unwelcome result given the unwavering intuition that you are the same person, or self, that you were yesterday and at 2 years old, and that you have existed until now without discontinuity.

A first obvious response to this challenge is to bite the bullet and claim that the self ceases to be during your sleep, and comes back into existence upon awakening. However, this is only partly satisfactory for it leaves the question of how is it that the self that you were yesterday is numerically the same as the one you are today. One must, in other words, bridge the gap between yesterday’s and today’s stream.

I have argued in the last section that the primary and real subject of experiences is the raw bearer to which the experiences refer. But since the raw bearer’s existence is ontologically dependent on, and thus concomitant with, the occurrence of experiences, it also ceases to exist during the subject’s sleep. Hence, the raw bearer cannot plug the two temporally gapped streams. And to argue that you are the same person as yesterday because the subject of experiences that you are today closely resembles that which you were yesterday does not take us any further, for such a reply stands in serious need of explanation: in virtue of which feature do yesterday’s and today’s subjects resemble each other, or are identical?

The solution that I provide only *tentatively* here is to ground the raw bearer in the *physical structure* that generates it. This position originates from Mackie:

Unity of consciousness is, as it were, the nominal essence of personal identity...But the real essence of personal identity will be whatever underlies and makes possible the unity of consciousness. (Mackie 1976: 200, quoted in Bayne 2010: 287)

Thus, insofar as this physical structure continues to exist, the raw bearer is somehow preserved (albeit dormant) in such a manner that when you wake up, you have the exact same experiential perspective as you did before your sleep, which means that you can continue to be yourself by undergoing a stream of experiences. In short, it could be argued that the self is not made out of a stream of consciousness *per se* – saying this may be considered a proxy – , but rather out of its underlying physical structure.

Bayne (*Ibid.*: 288) resists this proposal on the grounds that there is no *a priori* guarantee that a single consciousness-generating structure will produce a single stream of consciousness. If a physical structure can theoretically give rise to different streams of consciousness and since our working

hypothesis is that a self is constructed out of a stream, then the continuity and unity of a physical structure is insufficient for the unity of the self.

These *a priori* matters are really touchy and I cannot presume to issue any definitive verdict on them. By way of closing my discussion of the self, I will simply point out that, provided that the consciousness-generating mechanism is more specifically circumscribed, there may be *empirical* guarantee that one such mechanism cannot give rise to several streams and several selves.

What is the physical structure of consciousness? On my account, consciousness is the result of the binding of first-order representations, perhaps via the higher-order representations with which they are bound. This binding is a causal process, and it is sufficiently quick to go unnoticed – it is quasi-instantaneous. Consequently, I hazard that the physical structure of consciousness is whatever unit – typically the brain but not *necessarily* (as the next chapter will show) – which can bind representations by what could be approximately called ‘immediate’ causal processes. And there may be *a posteriori* motivations for the claim that one such unit cannot give rise to more than one stream of consciousness.

An implication of this view is that it leaves open the possibility that a subject’s experiential perspective literally splits into two perspectives, into two raw bearers that is. In effect, if the unit and continuity of the ‘raw bearer’ is derivative of the causal unity of the mechanism that generates it, then the *complete* fission of this mechanism into two would appear to entail the fission of this perspective. An imaginary case of such fission is provided by Parfit’s (1984: 246) notorious thought experiment of the physics problem. If a subject could, in order to solve a physics problem by two completely different methods, block any causal path linking his two hemispheres by a remote control, then two mutually closed off causal units would be created. Accordingly, such subject, if it deserves to be so named, might ‘host’ not one but two experiential perspectives.

As is customary among all unity theorists, I will now put my account of phenomenal unity to the test of empirical data. I am here referring to the experiments that were conducted on cognitively damaged subjects, and most notably split-brain patients. My demonstration will be two-staged. I will first argue that in addition to their theoretical failures, the accounts of phenomenal unity that have been evaluated so far fall short of providing insights as to *how* and *why* of these patients’ conscious life differs from ours. Then, I will finally show that my self-representational account of phenomenal unity is well equipped to yield a plausible interpretation of these pathological cases.

7. Split-brains

7.1. The data

The human brain divides into two more or less symmetrical hemispheres. It is now well-known that the removal of one of them does not deprive the subject of consciousness, that is, of the capacity of undergoing conscious mental states. One also knows that the two hemispheres directly communicate – in a sense that one shall attempt to define – by a large transverse band of nerve fibers called the corpus callosum (Nagel 1971: 398).

Split-brain patients ('split-brains' or 'split-brainers' for short) are people whose corpus callosum is damaged, either congenitally or surgically. The procedure consisting in cutting connections in the corpus callosum is called callosotomy, and was first performed by Roger Sperry in the 1960's on epileptic patients. It was designed to prevent their epileptic seizures to spread from one hemisphere to the other, and proved successful in doing so.

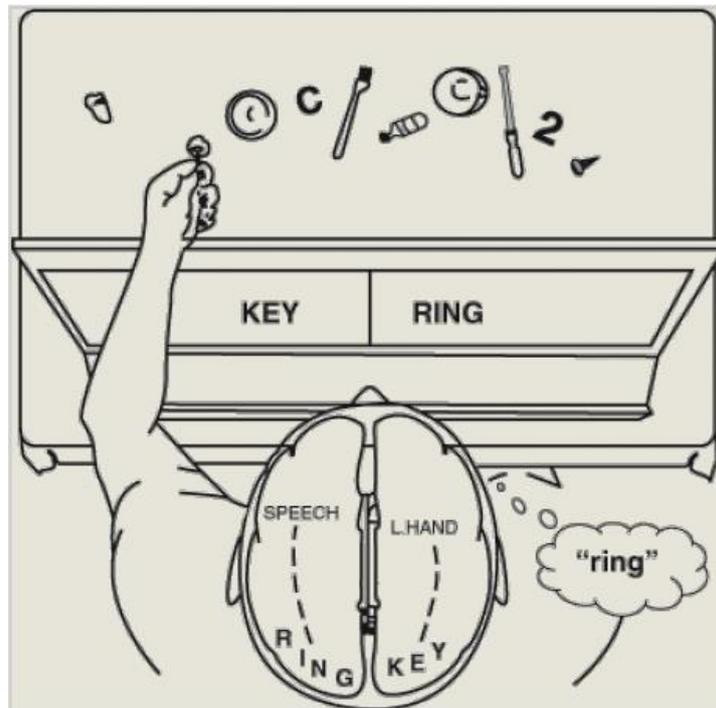
What may come off as a surprise is that split-brains are able to behave in as functionally and socially adequate a manner as cognitively intact (or more simply 'normal') subjects. They can have a job, drive a car, play the piano, casually interact with others, etc., so much so in fact that one could hardly tell a split-brain and a normal subject apart. Indeed, until recently, the corpus callosum had long been conjectured to serve no function whatsoever: not only do split-brains live normal life, but they do not report any sensation of reduction of their visual field, or of division whatsoever. What is more, they typically *act* as if they were single persons, who enjoy a unique and unified stream of consciousness; accordingly, their relatives naturally treat them as single individuals (Nagel 1971: 401, 406).

In specific and carefully controlled experimental conditions, however, they exhibit behavioral disunities that have been taken by many as evidence that their conscious experiences somehow differ from normal subjects'. The paradigmatic format of experiments on split-brains is the following. Information is perceptually presented to the patients in such a way as to 'reach' and be processed by only one of the patient's two hemispheres. As a result, the behavior dictated by the *other* hemisphere subsequently hints at a lack of representation (access) to said information in this hemisphere.

In one of the most notorious such experiments, the patients are asked to stare at a fixed point at the center of their visual field while words are flashed on both sides of the fixed point. The words 'Key' and 'Ring' are respectively flashed within the patients' left and right visual field (hereafter, respectively LVF and RVF); they are flashed during just 1/10 of second, which is long enough to allow each eye to

register the word that falls within its half of the visual field, but too quick for the right eye to move and register what is flashed in the LVF, and *vice versa* – a method called ‘tachistoscopic stimulation’ (Nagel 1971: 399). Since the visual system has a contralateral structure, information flashed in the LVF is processed by the brain’s right hemisphere (RH), and information flashed in the RVF is processed by the left hemisphere (LH).

When a patient is asked what he sees, he can only report having seen the stimulus flashed on the RVF, ‘Ring’, because this stimulus has been processed by the LH which is also responsible for speech; he verbally denies having seen the word ‘Key’. However, when asked to pick with his left hand, among several items, the object whose name has been flashed, he picks a key, for both procession of the stimulus ‘Key’ and movement of the left hand are performed by the RH. Symmetrically, when asked to pick the appropriate object with his right hand, he picks the object whose name was flashed in the RVF and processed in the LH, namely a ring.



Source: Sperry 1974, copied in Bayne 2010: 191.

Other experiments have studied sensory experiences of other modalities. In one of them, a smell of garlic is presented to the patient’s right nostril; the patient denies verbally that he experiences any distinctive smell, while *at the same time* appropriately picking a garlic clove with his left hand.

In yet another experiment (described in Nagel 1971: 400-1), the split-brainer holds in his left hand, but out of sight, a pipe and is then asked to write down with the left hand (controlled by the RH) what he was holding before. He writes ‘P’, then ‘I’, but then changes the ‘I’ to ‘E’ and rushes to write ‘PENCIL’

instead, suggesting that the LH, responsible for speech, takes over and ‘makes a guess’ based on the first two letters that were written. But then the subject deletes the word ‘PENCIL’ and sketches a pipe in its place, in turn suggesting that his RH regains control.

Experimental results of this sort are fascinating and abundant. They all uncontroversially point to dissociations and disunities of some sort in the split-brainers’ behavior. They raise two important questions: how is it that they exhibit highly integrated and thus functionally successful behavior in everyday life? What conclusions do behavioral disunities bear on the question of the unity of consciousness? Many interpretations of the experimental data have been challenged throughout the years in order to provide answers to these questions.⁶² Before examining them, I shall briefly stress important terminological and methodological points.

First, it must be warned that experiments on acallosal patients often anarchically elicit talk about the number of *streams of consciousness*, the number of *minds* and the number of *subjects* that supposedly obtain in such or such case. I understand the term ‘mind’, used among others by Nagel (1971) and Lockwood (1989), as referring in the present discussion to the same thing as the phrase ‘stream of consciousness’, which I broadly defined in the last chapter as follows:

Stream of consciousness (Generic): a maximal set of phenomenally unified experiences.

One should then distinguish at this point between two questions regarding a split-brainer SB’s experiences:

- 1) Does SB entertain phenomenally disunified experiences?
- 2) Does SB entertain more than *one* stream of consciousness?

Most theorists who answer ‘yes’ to 1) are committed to answering ‘yes’ to 2). However, one will remark that this inference may be too quick for it runs on a particular assumption regarding the logical properties of the relation of phenomenal unity. At any rate, I will handle these two questions separately in order to avoid any confusion.

Finally, in the last chapter we regarded the self as a theorist’s abstraction that corresponds to the bearer of a stream of consciousness. Consequently, the answer to the question 2) will bear implications for the number of *subjects*, understood as persons or selves, that SB is supposed to be.

⁶² For complete references of these positions, see Nagel 1971; Bayne 2010: ch. 9; Lockwood 1989: ch. 6; Tye 2003: ch.5; Hurley 1994 and 1998: ch. 3.

7.2. Is there phenomenal disunity?

In the key-ring experiment, the discrepancy between SB's verbal report and his left-hand's grasp indicates that the following two representations are simultaneously entertained by SB:

- (a) The word flashed before me is 'Ring'.
- (b) The word flashed before me is 'Key'.

I consider here three responses that these data have prompted to the question 1) of whether SB undergoes phenomenally disunified experiences.

7.2.1. The automaton interpretation

On a first interpretation, only the representations processed in SB's LH, such as (a), are conscious. SB's behavioral responses produced by the 'non-speaking' RH are those of a non-conscious automaton, or 'zombie mechanism'. Thus the representation (b) is not conscious.

This interpretation is riddled with problems of both empirical and philosophical nature. First, cases have been reported of apparently conscious human beings whose LH has been completely destroyed. These cases indicate that the RH alone is capable not only of generating consciousness, but of performing tasks that were originally to be performed by the LH; for instance, infants who suffer this sort of LH destruction grow up to speak normally (Lockwood 1989: 81-2).

More generally, underlying this automaton interpretation is the assumption that consciousness requires the faculty of verbally reporting conscious states. But as Nagel observes, taking the split-brains' incapacity of verbally reporting the RH's representations as evidence that these representations are not conscious is begging the question, since this verbal capacity is assumed by the LH. Furthermore, accepting the verbalizability as a precondition on consciousness has the very unwelcome consequence of denying consciousness to pre-linguistic children and aphasics. Above all, the kind of behavior generated by the RH is 'too elaborate, too intentionally directed and too psychologically intelligible to be regarded merely as a collection of unconscious automatic responses' (Nagel 1971: 403; see Bayne 2010: 193 for a similar point). In other words, SB seems *aware* of entertaining representations that trigger his behavior, which means that these representations are conscious.

Summing up, the automaton interpretation must be rejected on account of the wealth of empirical evidence that there are conscious representations in both LH and RH. But the question of interest is yet to be answered: are conscious experiences generated by one hemisphere phenomenally unified with experiences that are generated by the other? Are (a) and (b) phenomenally unified?

7.2.2. The phenomenal unity interpretation

The second interpretation of the key-ring experiment, championed by Bayne and Chalmers (2003), is that although SB's uncoordinated behavior reveals some disunity or 'breakdown' in his experiences, SB's *phenomenal* unity remains unscathed.

Assuming Block's (1995) phenomenal/access-consciousness distinction, Bayne and Chalmers naturally claim that there are two kinds of unity at play in consciousness. A subject's experiences are phenomenally unified (P-unified) when there is something it is like to undergo them together. By contrast, a subject's experiences are access-unified (A-unified) when they are conjointly available to the subject for verbal report, rational inference and deliberate control of behavior (c.f. sub-section 3.2.1.; see also Bayne & Chalmers 2003: 31).

It is worth noting that A-unity is roughly another terminology for 'representational unity': for experiences to be conjointly accessed is for them to be closed under conjunction, or 'integrated', in the normatively constrained and functional sense that has been characterized earlier (in section 5.3.). In spite of small differences between the two formulations, bear in mind at all times that I will take A-unity and representational unity as equivalent theses in the development that follows.

It is undeniable that SB's representations (a) and (b) are not conjointly accessible: were they so accessible, SB would most likely grasp both objects with either hands, and would report seeing them both – alternatively, were he to name just one object, he would pick the same –. So SB's representations (a) and (b) are A-disunified. This said, Bayne and Chalmers' move is to argue, in line with Block, that since A-consciousness is conceptually distinct from P-consciousness, it should not be inferred from this A-disunity that (a) and (b) are P-disunified.

I have already taken issue at length with Block's distinction (see section 3.4.3.) However, I shall now hammer down an additional (and perhaps final) nail in Bayne and Chalmers' coffin.

Bayne and Chalmers' interpretation allows them to preserve their unity thesis, according to which any subject's consciousness is *necessarily* phenomenally unified. But a little more must be said about the modal status of this thesis. Bayne denies that this necessity is of a conceptual or metaphysical nature, and that it is grounded in the laws of nature (Bayne 2010: 17). Rather, he intends to argue that 'we *never* have disunified experiences' (*Ibid.*, emphasis in original), and since his defense of this claim largely consists in answering empirical challenges to it, his thesis must be an *empirical* one, as Bayne himself observes (2010: 94 and 281; see Shoemaker 2011 for a similar point).

Recall that Bayne and Chalmers partly justify their endorsement of the unity thesis by its *prima facie* appeal. This endorsement is not as such problematic, for it could be understood as a starting hypothesis, something like an 'educated guess' that is then to be tested against empirical data. Yet

they subsequently motivate the *prima facie* case for the unity thesis by the ‘inconceivability – whether it takes the form of unimaginability or apparent incoherence – [of] having two phenomenal states simultaneously without there being a conjoint phenomenology’ (Bayne & Chalmers 2003: 40); and this in turn premises the claim that a phenomenal unity break down is ‘impossible’ (*Ibid.*).⁶³ Although in slightly different terms, Dainton deploys a similar strategy. He advocates for the transitivity of the co-consciousness relation on the grounds that non-transitive co-consciousness is ‘unimaginable and unconceivable’ (Dainton 2000: 108).⁶⁴ In short, these theorists seem to motivate their *prima facie* case by an inconceivability case.

On Block’s distinction, a subject’s verbal report and behavior can only shed light on whether her experiences are A-unified. This is symptomatic of a general ailment that vitiates Block’s A/P distinction. With the A-consciousness shouldering the entirety of the functional load, P-consciousness appears to be devoid of any functional significance, and to be nothing more than an epiphenomenon (Kriegel 2006b: 11-12; Chalmers 1997: §3; Tye 2003: 125). Consequently, it is not clear why *any* empirical data should be considered *at all* in the evaluation of the thesis of phenomenal unity; at the simple evocation of such thesis, Popper must be turning over in his grave.

These remarks are not intended to show Block’s distinction to be erroneous. Rather, they show that thus assumed, it comes to grips with the very modal nature of the mereologists’ unity thesis. In other words, champions of this thesis face the following dilemma. Either they take the unity thesis to be a conceptual truth, in which case split-brain experiments, or any cognitive-scientific studies for that matter, are simply *off point* – since they all target A-consciousness as Block concludes –; or they take the unity thesis to be an empirical truth, in which case they should envisage that other kinds of disunity entail phenomenal disunity (Tye 2003: 125) or that their thesis actually concerns A-consciousness and not P-consciousness.

If they take phenomenal unity to be a conceptual truth, one should note from the get-go that the case that they make for it is a poor one. Indeed, one of the lessons that is to be drawn from Nagel’s seminal (1974) is that the fact that we are incapable of imagining or conceiving what is like to be some creature, be it a bat or a subject with phenomenally disunified experiences, should not mislead us into concluding that there is nothing it is like to be this creature.

If phenomenal unity is an empirical truth, and if Block’s distinction is to be retained, it ought to be refined so as to parry the threat of epiphenomenalism. An attractive solution proposed by Kriegel

⁶³ Interestingly, the unimaginability of phenomenal disunity does not meet universal consensus. See Tye (2003: 120) and Lockwood (1989: 92) for a demonstration of how one could imagine undergoing two phenomenally disunified (sets of) experiences at the same time.

⁶⁴ As will be shortly developed, this logical property has important implications on the conception of a stream of consciousness.

(2006b) consists in treating the property of P-consciousness as the *categorical basis* of the property of A-consciousness. The argument unfolds roughly as follows. A-consciousness is a dispositional property, but dispositional properties cannot exist on their own: they must be grounded by other non-dispositional properties. For instance the fragility of a glass is a dispositional property that is grounded by non-dispositional physico-chemical properties of this glass. This means that the physico-chemical properties *explain why* it is fragile (or *constitute reasons for* its fragility). So the glass's physico-chemical properties are the categorical basis of its fragility. Similarly, a state's P-consciousness is the categorical basis for its A-consciousness: it is in virtue of its phenomenal character that it is endowed with its functionality. And consequently, concludes Kriegel, we should construe A-consciousness as the *functional role* of P-consciousness.⁶⁵

This conception could be adapted to Bayne and Chalmers' ends; they could still hold that P-consciousness is necessarily unified, even though it is incapable of carrying out its functional tasks in split-brain cases. But as a result, their thesis would be greatly weakened, for split-brains' unified P-consciousness would come off as cheap, not to say completely otiose. More generally, Kriegel's solution leaves one wondering whether Block's distinction is really worth salvaging, at least in the present discussion. I will now sketch the interpretation that naturally follows from my self-representational account of phenomenal unity.

7.2.3. The phenomenal disunity interpretation

As has been rightly observed by Bayne and Chalmers, SB's representations (a) and (b) are representationally (or 'access-') disunified. Indeed, the verbal report and behavior that are prompted by the experimenter's instructions suggest that SB more precisely entertains the following two doxastic representations:

- (a) The only word flashed before me is 'Ring'.
- (b) The only word flashed before me is 'Key'.

Not only does the conjunction of representations (a) and (b) fail to comply with the norm of semantic consistency, it seems stable and unprone to rectification (judging from SB's behavior), which indicates that (a) and (b) are not conjointly monitored. This means that representations (a) and (b) are not represented *together* as being SB's, which in turn means that there is nothing it is like for SB to undergo them together.

⁶⁵ Bayne seems to hold a similar view (2010: 279).

But how can this be? On my account, a mental state is conscious in virtue the subject's being aware of undergoing it. Accordingly, SB is aware of undergoing (a) and of undergoing (b); how come the two representations are not integrated by being indexed to the same experiential point of view, SB's?

Three hypotheses can be hazarded. The first is to deny conscious status to either (a) or (b), by claiming that either representation is not represented as being entertained by SB. This gets us right back to the automaton interpretation, which we have deemed highly contentious for reasons I need not recall.⁶⁶ A crucial implication of this interpretation is that experiences are necessarily unified. Indeed, it seems to run on the assumption that if two representations are consciously entertained, then they are necessarily integrated by the subject's awareness that they are her own. However, I have *not* advocated for the view that the subject's experiences are *necessarily* integrated. All I have claimed is that for the subject's experiences to be phenomenally unified is for the representations in which they consist to be integrated. My view does not rule out the failure of integration of two representations. But before elaborating on this idea, let me dismiss a second hypothesis that some may be tempted to infer from the necessity of integration.

The second hypothesis is to accept that both (a) and (b) are conscious while arguing that the failure of their integration stems from the fact that they each involve a different *de se* reference. In this vein, it may be argued that S is aware of (a) as being (say) S1's but is aware of (b) as being (say) S2's. This would mean that there are actually *two* subjects of experiences. This hypothesis stands in tension with our intuitions: it is metaphysically unintelligible that a subject's *de se* reference involved in the self-awareness of his experience splits into two numerically distinct such references. If it is *de se* then there must be just one entity. Moreover, provided SB would regain a single perspective after the experiment, it is all the more obscure how experiential perspectives may come in and out of existence. The advocate of this hypothesis may pugnaciously retort that my reservations are due only to the limits of my imagination. I plead guilty to this charge, but at any rate I think that this hypothesis might be worth exploring only as a last resort, that is, if it were the only one available.

Luckily, my naturalist self-representational account readily motivates a third hypothesis. Before developing it, inquiring into SB, a little more must be said about how integration of representations could work in cases of phenomenal unity. Following Kriegel (2005), I have defended the view that an experience consists in a psychological integration, akin to binding, of a first-order representation with a higher-order (HO) representation of this representation. I have further argued that the HO representation is processed by a different part of the brain. Now take two experiences *e1* and *e2* who are respectively composed of HO representations. At the phenomenal level, *e1* and *e2* are both

⁶⁶ Even if we are willing to grant that either (a) or (b) is actually unconscious, which one is it? One has seen that each hemisphere is individually capable of assuming consciousness. This epistemic indeterminacy considerably undercuts the automaton interpretation.

experienced by S as her own experiences – by being both ascribed to the same indexical ‘I’ –. In other words, the *de se* reference common to both HO representations somehow brings about one common HO representation of *e1* and *e2* being together experienced by S.

At the neural level, two different phenomena could generate this integration. The first is that *e1* and *e2* involve higher-order representations (HO1 and HO2) that are both processed in the *same* brain state *N*. *N* would then be what integrates the first-order representations *e1* and *e2*, both at the neural and at the phenomenal level. But this is both unlikely and empirically disproved.⁶⁷

Rather, our experiences *e1* and *e2* involve HO representations that are processed by *different* brain states *n1* and *n2*. Accordingly, the plausible phenomenon responsible for the integration of *e1* and *e2* would be that *n1* and *n2* somehow ‘communicate’ with each other so as to share, and thus integrate, their content – the first-order representations –. In other words, in addition of the occasional binding that is at work in a first-order representation (notably, in visual representation) and of the binding that is involved in the integration of this first-order representation with a HO one, there may be binding between first-order representations via some connections between their respective HO representations.

Let us now get back to the key-ring experiment. I suggest that both SB’s representations (a) and (b) yield conscious mental states by being integrated with HO representations processed respectively in *n1* (in the LH) and *n2* (in the RH). These HO representations (trivially) involve each a *de se* reference: SB is aware of experiencing (a) and is aware of experiencing (b). But the lack of phenomenal unity between (a) and (b) need not be explained by a difference in experiential perspective. Rather, it can be explained by the fact that binding between first-order representations of the key and ring is hampered by the lack of connectivity between *n1* and *n2* that is due to the damage of the corpus callosum.

Of course, for cognitively intact subjects like you and me, this idea is as unimaginable as the second hypothesis according to which SB experiences (a) and (b) are those of two different subjects. As far as I am concerned, I must concede that I cannot fathom that there is something it is like to experience (a), and that there is something it is like to experience (b), but that there is nothing it is like to experience (a) and (b) *together*. Thus I can settle to a draw with the second hypothesis with regard to their conceivability. Nevertheless, since mine postulates only one raw bearer of experiences, it should prevail on grounds of ontological parsimony.

⁶⁷ While this phenomenon may be at work for small sets of representations, it simply cannot account for the integrations of *all* the subject’s experiences, including thoughts, perceptions, bodily feelings and emotions for it is unlikely that the higher-order representations that they involve are *all* processed by one and the same brain state *N*. Moreover, the contralateral structure of the visual system, for instance, is such that a visual representation acquired via the right eye is processed in the LH and represented by a brain state in LH, and vice versa for a visual representation acquired via the left eye.

This said, my hypothesis may be encouraged on more than purely theoretical grounds. The fact that split-brain subjects display such a high level of behavioral coordination in everyday contexts is explained by inter-hemispheric cross-cuing, a strategy by which split-brains use environmental cues to pass information from one hemisphere to the other. For instance, a subject can transfer to the RH the representation of an object perceived by the LH by fixating a related object located in a part of the visual field that is commonly accessible by the LH and RH. Similar transfer of content can be achieved by means of movements of the neck, tongue and facial muscles (among other cues) since these can be controlled by both hemispheres (Nagel 1971: 401-2; Bayne 2010: 196). In short, the two hemispheres continue to cooperate and exchange representations despite the lack of corpus callosum. The difference with normal subjects is that they do so by means of external vehicles.

It is worth noting that this cooperation is not the result of an *intentional* communication, for there is good evidence that neither hemisphere takes itself to be a distinct agent from the other. This is confirmed by the fact that in experiments where the patient displays uncoordinated behavior, he does not think of himself as several subjects that would be each associated with one (set of) experience(s). (Bayne 2010: 196 and 204).

Two interpretations can legitimately be drawn from this cross-cuing phenomenon. The first is our initial assumption: *de se* references involved in experiences of the LH refer to the same subject of experience, the same 'I', as *de se* references involved in experiences of the RH (Bayne 2010: 204). Indeed, one could take cross-cuing as indicating that there is a unique subject of experience 'trying', at a sub-personal level, to patch up representations. Confirming this interpretation, other experiments have shown it to be impossible for the two hemispheres to *compete* in a game (MacKay & MacKay 1982, quoted in Lockwood 1989: 86).

The second interpretation is that it does not matter that the subject's experiences be integrated by external vehicles, such as muscular cues, rather than by 'internal' ones such as impulses between neurotransmitters. As long as this passing yields the integration of representations, it does not seem far-fetched to venture that there is something it is like for the split-brainers to undergo, in everyday contexts, these experiences together. Note that this is on par with my tentative hypothesis that the continuity and unity of a 'raw bearer' is warranted by the continuity and unity of a physical basis, understood as a unit that supports the binding of representations by 'immediate' causal processes.

Admittedly, I am exposed to the reproach that my little story about neural correlates is completely mistargetted: it can only go so far as to account for neural integration, and cannot as such premise any phenomenal claim whatsoever. But in reality, the humble purpose of my development was to illustrate that on my naturalist account, the integration of a subject's experiences by their agglomeration to a single index is no conceptual necessity. Since this integration is akin to binding, it is as much a

psychological process as is the integration of a first-order representation with a higher-order representation. And this, in turn, makes the phenomenal unity of experiences likely to break down when brain parts are severed or missing, as in the cases of split-brains.

In short, on my account the stable incoherent behavior exhibited by split-brain patients show that their experiences of 'Ring' and of 'Key' are phenomenally disunified. The fact that both these perceptions involve a reference to the same subject of experience, inconceivable though it may be, is not only compatible with my account but empirically encouraged.⁶⁸ I now turn to the second question of whether this phenomenal disunity entails that split-brains undergo more than one stream of consciousness.

7.3. Is there more than one stream?

As previously alluded to, the view that split-brain patients entertain phenomenally disunified experiences has often gone hand in hand with the view that they entertain *two* streams of consciousness. However, because it has led theorists to a variety of postulates regarding the structure of a normal subject's consciousness, on the one hand, and the mechanisms underlying this duality, on the other hand, the two-stream interpretation shades off in several versions. In the next two subsections I consider three of them in turn, and infer from their shortcomings that a one-stream interpretation is preferable.

7.3.1. The two-stream interpretation

On a first, naïve version, *all* subjects undergo two streams of consciousness at all times. The high degree of behavioral integration that we exhibit is explained by the fact that the two streams, 'like an elderly and devoted married couple, have had a lifetime's practice in mutual cooperation' (Lockwood 1989: 81; see also Bayne 2010: 196). Less metaphorically, the gist is that they share their representations in such a way that the two hemispheres run on a single 'cognitive-behavioral workspace'. On this conception, the callosotomy procedure merely reveals this duality by sealing off the sets of experiences associated with each hemisphere from each other in such a manner that contents associated with one hemisphere can only be integrated with other contents of the same

⁶⁸ I could phrase my proposal differently by saying that (a) and (b) are integrated by their common subjective character, but that this subjective character fails to *monitor* them in conjunction. However, this difference would be merely verbal, for this integration would resemble a mereological sum.

hemisphere. The resulting picture is that of a subject with two internally unified streams, one relative to the RH and the other relative to the LH, which cannot share representational content with each other.

What is wrong with this picture? First, the fact that the two streams run on a single cognitive-behavioral workspace is simply incoherent with the conjunction of two of our previous assumptions:

- (i) *Self-representational account*: For a set of experiences to be phenomenally unified is for them to conjointly represent themselves in such a way as to be integrated and normatively monitored.
- (ii) *Stream of consciousness (Generic)*: a maximal set of phenomenally unified experiences.

In effect, the suggestion that representations are pooled into a single cognitive-behavioral workspace means in all likelihood that they are integrated and thus phenomenally unified. Then, in accordance with our generic definition of a stream, they are part of *a single* stream. The image of a stream is supposed to capture the idea of a total state, or more neutrally, of a unit; it would be absurd to conceive of 'sharing' or 'overlapping' streams, *a fortiori* if they are completely sharing.

To put it a bit differently, postulating two streams by default instead of one seems a dispensable ontological expense, and an *arbitrary* one at that. As Nagel points out, if the two hemispheres are associated each with a distinct stream of consciousness, even when they are linked by the corpus callosum, what is there to prevent us from making additional fragmentations within each hemisphere, and from ascribing a distinct, further stream to each of the resulting brain parts? In short, phenomenally unified experiences must be part of the same stream.

But what about phenomenally disunified experiences? I have advocated for the idea that split-brain patients undergo phenomenally disunified experiences in certain laboratory conditions such as those of the key-ring experiment. Does this entail that they undergo more than one stream of consciousness? The final two sub-sections consider two different understandings of a stream of consciousness which yield opposite answers to this question.

7.3.2. The momentary split interpretations

To have a clear view of the challenge on hand, let us lay out the three crucial pieces of data that must be taken into account when interpreting the split-brain cases. The first has already been mentioned: it is the fact that SB normally exhibits normal, seamlessly integrated behavior, which on our account suggests that his experiences are phenomenally unified: there is something it is like for him to undergo his experiences together.

The second, resulting from my development in the last section, is that SB undergoes phenomenally disunified experiences under experimental conditions, which shows through his stably uncoordinated behavior. The representations of 'Key' and 'Ring' are an example of such phenomenally disunified experiences.

The third piece of data, which is most startling, is that under experimental conditions, SB exhibits both behavioral dissociation and coordination *at the same time*. During the key-ring experiment, for instance, SB displays a completely coordinated behavior with respect to other environmental stimuli: he adapts his posture, follows instructions about what to stare at, acts normally vis-à-vis the experimenter and the experimental set-up, etc. (Nagel 1971: 408; Lockwood 1989:84). On our account, this indicates that SB undergoes some phenomenally unified experiences as well.

The first datum invites us quite naturally and uncontentiously to regard SB as normally entertaining a single stream of consciousness, just like a cognitively intact subject. Accordingly, the fact that SB's experiences exhibit phenomenal disunity under experimental conditions has been understood by theorists as a *momentary split* in SB's stream of consciousness. Resulting from this split are *two* numerically distinct and mutually sealed off (and thus non-overlapping) streams, each associated with a hemisphere, that remain two for the lapse of the experiment. Bayne dubbed this view 'contextualism'.

The momentary split interpretation is motivated by the assumption that the relation of phenomenal unity is *transitive*. This means that for two experiences e_1 and e_2 , if e_1 and e_2 are phenomenally disunified, there is no e_3 such that e_3 is phenomenally unified with *both* e_1 and e_2 . The transitivity of the relation of phenomenal unity yields the following definition of a stream of consciousness:

Stream of consciousness (Strong): a maximal set of experiences wherein *any* experience is phenomenally unified with *all* the others.

This idea, however, does not sit well with other phenomenally unified representations which appear to be associated with *different* hemispheres. For if the contextualist is right, each hemisphere is associated with its own stream during the experiment. But if representations of different hemispheres are phenomenally unified, they must be part of the same stream.

In order to accommodate simultaneous phenomenal unities and disunities, contextualists postulate that in normal contexts, both hemispheres of any subject have access to more or less the same information from the environment. Contrary to the first two-stream interpretation, this sameness is not numerical – this would clash with the accepted conception of a stream – but *qualitative*. Contextualists, in other words, hold that all the representations of each hemisphere are

'mirrored' in the other. For instance, when you see a tree, there is a representation of the tree in the LH and another representation of the tree in the RH.

Contextualists then argue that when SB exhibits phenomenal disunity in the key-ring experiment, it is because the representation of the key in the RH is not mirrored in the LH, and because the representation of the ring in the LH is not mirrored in the RH. By contrast, unified representations that are entertained *during this momentary split* are held to be processed in the *same* hemisphere in virtue of the mirroring. Thus contextualists can claim, on par with the experimental data and with their conception of a stream, that no representation processed in a hemisphere is actually integrated with a representation of the other. Phenomenally unified experiences therefore belong to the same stream of consciousness, be it the one associated with the LH or the one associated with the RH.

This said, theorists disagree about exactly what the mirroring of representations amounts to, and particularly what it implies for a normal subject S's experiences. On a first interpretation, S has *experiential* 'duplicates' in the sense that mirroring representations in LH and RH are both *conscious*. This means that S has a token experience *e1* in one hemisphere, and another numerically distinct token experience *e2*, which is *phenomenally identical* to *e1*, in the other hemisphere. In other words, when S sees a tree, she has two numerically distinct, and yet phenomenally identical, visual experiences that represent the tree.

However, accepting the possibility of duplicates entails either rejecting the tripartite method of individuation of experience (according to which two experiences that instantiate the same phenomenal property and are entertained by one and the same subject at the same time must be numerically identical) or postulating not one but *two* subjects of experience so as not to contradict the tripartite account – to each token experience its subject –. Neither option is desirable. The tripartite method is fairly adequate and one should not reject it unless a better method of individuation is proposed. As for the second option, I provided sufficient motivation against it at the beginning of this section. So the first interpretation must be discarded.

A second interpretation, proposed by Tye, appeals to experiences' physical basis. It tells us that all subjects have *neural* duplicates, each being in a distinct hemisphere and causally isolated from the other. However, each pair of neural duplicates realizes a single token experience. This means that S has a neural event *n1* in the LH, and another neural event *n2* in the RH, and either one event could have realized the experience *e* had the other not occurred. Contrary to the first interpretation, this does not entail that S has two numerically distinct but phenomenally identical experiences when *n1* and *n2* both occur. The idea, rather, is that *n1* and *n2* work redundantly: they realize the *same* token experience *e*. To get this point across, Tye invokes the following analogy. Imagine that two projectors

project the same movie at t on exactly the same screen. There are, so is posited, two projections – one of which surely is redundant – but only one image (Tye 2003: 127). Accordingly, split-brainers' phenomenal unity is due to the mirroring of the representations in the two hemispheres; their disunity is due to a difference of representations between the two hemispheres.

It must be granted that this story, contrary to the first, does not conflict with our accepted method of individuation of experiences, since the tripartite criterion is neutral with regard to their neural basis. But on a naturalist approach to consciousness, two causally isolated neural correlates simply cannot realize together one single token experience. There is no honest and non-arbitrary sense in which $n1$ and $n2$ together 'compose', or 'co-realize' *one* neural event that bears causal powers as a unit. Moreover, if Tye's construal is correct, what is there to prevent that either $n1$ or $n2$ co-realizes an experience e with a further neural event $n3$ that occurs in *another* subject's brain?

It appears that Tye is guilty of a token-type fallacy: given causal isolation between hemispheres, and assuming supervenience of the phenomenal experiences on the neural events, $n1$ and $n2$ cannot but cause two different token experiences of the same type (Schechter 2010a: 11). The projectors analogy can be readily understood this way: there are *two* images, albeit superimposed, of *one* type.

One has assessed the potency of contextualism without questioning the assumption that representations mirror each other in the two hemispheres. But it should be noted that empirical evidence undermines this very assumption:

The hemispheres are no doubt subject to more similar contents outside of experimental situations than they are inside of them, and this no doubt offers at least a partial explanation for the fact that split-brain subjects behave differently in the two types of circumstance. Nonetheless, since the hemispheres have different patterns of perceptual access to the world, and since they also have different processing styles and capacities, and appear to experience emotions somewhat differently, and to have access to a somewhat different store of long-term memories—not to mention the fact that one hemisphere can presumably generate a fairly normal stream of inner speech and one hemisphere probably can't—it would be a stretch to imagine that the hemispheres are associated with highly similar, much less identical, conscious contents. (Schechter 2010a:4-5)

More generally, contextualists fall short of explaining how a person's stream of consciousness can undergo such prodigious changes as a momentary split and a reunification (Nagel 1971: 409; Lockwood 1989: 84; Hurley 1994:70). That experimental conditions elicit local disunities seems intelligible enough; but it seems quite a stretch to posit that they bring about a clear-cut bisection of the subject's workspace into two mutually sealed off streams.

What is more, the experiential or a physical redundancy that is invoked to deal with the apparent dominance of phenomenal unity seems extremely costly and conceptually shaky when compared to its explanatory gains. These proposals leave us under the impression that contextualists are willing to go to incredible extremes in order to preserve the *prima facie* claim that phenomenal unity is transitive.

But maybe this claim ought not to be retained after all, as Nagel already foreshadowed: for apparent lack of better explanation of SB's simultaneous integrations and disunities, he concluded that there may be no definite answer as to the number of streams of consciousness that a split-brain patient undergoes at a given time (Nagel 1971:410).⁶⁹ In a similar spirit, I think it reasonable to hypothesize that our conception of a stream is too strong. In the final sub-section, I suggest that a stream's unity is not an all-or-nothing matter but rather admits of degrees of unity.

7.3.3. The partially unified stream interpretation

Underlying the two-stream models was the assumption that the relation of phenomenal unity is *transitive*. The transitivity of phenomenal unity has been decreed only *prima facie* (cf. sections 3.1. and 3.2.) and has been motivated neither conceptually nor empirically. Moreover, we have just seen that the strong understanding of a stream of consciousness to which it gives rise does not square with the claim that SB displays both phenomenal unity and phenomenal disunity at the same time.

From the implausibility of neural and experiential duplicates, I now champion the view that the relation of phenomenal unity is *non-transitive*. If this is true, answering 'yes' to the question of whether SB undergoes phenomenally disunified experiences does not entail answering 'yes' to the question of whether SB undergoes more than one stream of consciousness. In effect, a stream may be only weakly or partially unified in the following sense:

Stream of consciousness (Weak): a maximal set of experiences wherein

- i) there are at least three experiences *e1*, *e2* and *e3* such that *e1* is phenomenally unified with *e2*, *e2* is phenomenally unified with *e3* but *e1* is not phenomenally unified with *e3* (Hurley 1994: 67-8);
- ii) any experience is phenomenally unified with *at least one* other experience.

There is no watertight argument for the interpretation that SB enjoys a single but partially unified stream of consciousness. I follow its advocates, and most notably Hurley, in endorsing it by the rejection of all other alternatives.

Our development has so far provided a propitious framework for partial unity. I have suggested that phenomenal unity could be understood as a psychological phenomenon that binds, by cerebral or external paths, experiences together. Accordingly, in the key-ring experiment, the phenomenal

⁶⁹ To be more precise, Nagel emits doubt primarily as to the number of *subjects* that split-brains are supposed to be. But since he equates number of subjects with the number of 'minds', i.e. streams, which are exhibited by a split-brain patient, the indeterminacy of the number of subjects holds *eo ipso* for the number streams. Hence this shortcut is legitimate and innocuous.

disunity of the representations of 'Ring' in LH and of 'Key' in RH can be explained by some default in the binding paths, namely the absence or damage of the corpus callosum and the hindrance of any visual cross-cuing. As a result, two experiences are individually, but not conjointly, experienced as being SB's. Elaborating on this picture, it is plausible that different integrations of representations involve different specific paths, and that the severing of only a portion of these paths could not *completely* obliterate phenomenal unity.

Lockwood seems to have this in mind when he engages in the following thought experiment: if a normal subject's corpus callosum were to be gradually severed, one fiber at a time, it would yield at some point a partially unified stream of consciousness (Lockwood 1989: 90ff.) Although this thought experiment alone has limited force, it is overwhelmingly confirmed by a wealth of experiments on patients with *partial* callosotomy. In one of them, a patient's visual experiences appeared to be disunified for touch but not for vision:

Tactile information in the left hand and right hemisphere remained isolated from the right hand and the left hemisphere. Yet, when a visual stimulus, such as the picture of an apple, was lateralized to either hemisphere, either hand could manually retrieve the apple, unaided by visual exploration. (Gazzaniga & LeDoux 1978: 10, quoted in Bayne 2010: 206)

In this experiment, the patient's visual representation 'straddles' both hemispheres and appears to be phenomenally unified with both his right hand tactile experience of the apple and his left hand tactile experience of the apple; but the left and right tactile experiences are phenomenally disunified with one another.

Even though the key-ring experiment studies fully acallosal patients, it is arguable that other binding paths such as external cues are taken so much so that some experiences are phenomenally unified with both the experiences of 'Ring' and that of 'Key'.

Admittedly, this little story of a partially unified stream does not bolster in any way the case for partial unity; it merely shows it to be compatible with the key-ring experiment. However, there are other experimental data for which partial unity may be the only plausible and satisfying explanation.

In Sergent's (1990) experiments (reported in Hurley 1994: 52-3), split-brains patients were shown numbers '6' and '7' by means of the standard and now-familiar tachistoscopic stimulation, so that the number '6' was shown exclusively to the LH, and the number '7' was shown exclusively to the RH. In three different experiments, respectively three questions were asked to the patients:

- 1) Are the numbers the *same*, or *different*?
- 2) Which one of the two numbers is *higher*?
- 3) Are the two *equal*? Which one of the two is *higher*?

Note that for the experimenters to ascertain that responses were not due to hemispheric specialization,⁷⁰ the patients had to give responses with each hand.

In experiment 1), patients were unable with either hand to make the adequate comparisons: the percentage of accurate comparisons amounted to *chance*, indicating that their responses were sheer guesses.

In experiment 2), patients were supposed to give their response by pressing a lever on the side of the higher number. Here, by contrast, they found the higher/lower comparison ‘very easy’ and showed ‘no hesitation’ in their answers, which were highly accurate. What comes off as a surprise is that the same/different response that failed to be accurately given in 1) can be directly (or so it seems) inferred by the higher/lower responses given in 2).

In experiment 3), finally, patients were able with high accuracy, but not as high as in 2), to determine whether the numbers were equal as well as whether one was higher than the other. When asked, in another projection of the numbers an hour later, to decide whether these were the *same* or *different*, they were unable to do so, just like in experiment 1).

To sum up, in contexts with no mention of the same/different, and with only quantitative higher/lower judgments, responses about the equality of the projected numbers could be accurately given; in contexts where the question is asked in terms of same/different, judgments of sameness could not be made. These data thus yield the following interpretation:

[Sergent] suggests that the disunity affects certain categories of information but not others, even though there may be inferential connections between them.

...there is unity with respect to awareness of comparative quantity but disunity with respect to awareness of the identities of the digits. To spell this interpretation out: awareness of the number on the left as lower is co-conscious with awareness of the number as a 6 and also co-conscious with awareness of the number of the right as higher, which in turn is co-conscious with awareness of the number on the right as seven, even though awareness of the number on the left as a 6 is not co-conscious with awareness of the number on the right as a 7.⁷¹ (Hurley 1994: 53, 70)

With these data comes the end of the empirical testing of my account of phenomenal unity. I intended them to show that once phenomenal disunity is accepted as a possibility and an empirically motivated phenomenon, nothing commits us to the endorsement of the transitivity of the relation of phenomenal unity, and therefore nothing commits us to the endorsement of a two-stream model. Not only do two-stream models fail to account to SB’s phenomenally unified experiences, but they are also less suited than a partially unified stream to account for certain experimental results such as Sergent’s.

⁷⁰ For instance, speech is associated with the LH; the motion of the right hand is initiated by the LH, and vice versa (see section 7.1.).

⁷¹ The property of being ‘co-conscious with’ is here devoid of any connotations that Parfit or Dainton might have endowed it with. It should be understood here as a neutral shorthand for ‘phenomenally unified with’.

Now, since I suggested that a self is made out of a stream of consciousness, does it follow that a split-brain patient who entertains a partially unified stream of consciousness is (or has) a partial self?

It certainly does not. My account of the self is such that a stream wears the trousers: it is an ontologically more fundamental notion than that of the self: selves are derived from streams (and not the other way around). Consequently, it is compatible with the fact that selves do not admit of degrees whereas streams do. A neat illustration of this difference is the relation that ties values to norms. It is generally agreed that values constitute a more fundamental ontological category than norms, and that norms are derived from values. Values admit of degrees, whereas norms do not. Likewise, streams admit of degrees of unity whereas selves do not. So SB is as much of a person as a normal subject of experience is.

However, I have defended the view that we construct selves, and typically ourselves, out of *coherent* streams of consciousness by a process akin to autobiographical writing. How, then, can SB 'narrate himself' in the presence of inconsistencies? What will most likely happen in cases of inconsistencies is that SB will treat *a posteriori* his experiences of 'Key' and 'Ring' (provided he remembers them both) as a *spot of indeterminacy*. We can distinguish between spots of *ontological* indeterminacy and spots of *epistemic* indeterminacy. Spots of ontological indeterminacy are what characterize, according to Dennett (1992), fictional entities such as selves. If you were to ask yourself whether Sherlock Holmes has a mole on his left shoulder, your question would remain forever unanswered because there is no such fact of the matter; it is simply not part of Sir Conan Doyle's script. There is ontological indeterminacy for it is neither true nor false that such mole exists.

In SB's case, the spot is rather of epistemic indeterminacy. What this means is that SB cannot adjudicate on whether he really visually experienced 'Key', or 'Ring', or both. He will most likely let this local disunity slide, just like normal subjects intentionally or forgetfully leave some blanks in their narrative. Did I wear a blue t-shirt for my sixth birthday? Don't know; don't care. And while there certainly is a fact of the matter regarding this precise question, it would seem absurd to say that the replies 'yes' and 'no' would each engender a distinct self. So SB can arbitrarily choose between either of the three states of affairs and add it to his personal narrative; or he can simply leave this detail out of it. Rather than plain contradictions, SB's local phenomenal disunities should be considered harmless spots of epistemic indeterminacy.

This said, a self is constructed out of a stream insofar as the stream endows it with global coherence and functional unity. But if stably incoherent experiences were to spread within the stream, they would impede such unity. Thus, just as there is a degree of value at which an action is no longer permissible but forbidden, there may be a degree of partial disunity at which one should hesitate to

speaking of a single self or two. That is, there may be disunities that elicit two significantly distinct narratives.

8. Conclusions

In this work I have developed an account of phenomenal unity on which, at that particular instant of your Céline Dion concert, all your experiences, including your perceptions of the singer on the glowing stage, your bodily feelings and notably this slight pain in the left foot, your emotions of joy, your thoughts and so forth are phenomenally unified in virtue of the fact that they conjointly represent themselves as being *your experiences*. This thesis more or less boils down to the conjunction of the following core theses:

- (I) *Mereology*: experiences are individuated in terms of their phenomenal character, time of occurrence and subject S.
- (II) *Representationalism₂*: all experiences are representations, and their phenomenal character supervenes on, or is identical with, its representational content and psychological mode.
- (III) *Self-awareness*: an experience necessarily comes with S's non-reflective awareness that S is having it.
- (IV) *Normativity*: S's exhaustive set of experiences at *t* are conjointly constrained by certain norms in such a way as to yield coherent reasoning and agency.

From (I) and (II), I have inferred that for S's experiences to be phenomenally unified is for the representations in which they consist to be integrated in a way that does justice to both their representational contents and modes. (IV) provides positive motivation for this integration: if S's experiences were not somehow integrated, they simply could not conjointly be under the range of norms. (III) has led me to claim that these representations are all unified by agglomerating around a single subjective character. This self-awareness can be regarded as a mode of representation that is shared by all of S's experiences and represents them *as being S's*. Hence my self-representational account of phenomenal unity:

Self-representational account: For S's experiences E to be phenomenally unified at a time *t* is for E to conjointly represent themselves as being S's; that is, it for S to be self-aware of E. S's self-awareness of E monitors E by putting them under the pressure of coherence norms.

The advantages of identifying self-awareness as a sort of total state that encompasses and integrates our experiences together are numerous. Phenomenally, it makes good sense of the claim that we never cease to *experience*, ever so elusively, the unity of our experiences. Epistemically, it explains why several of our experiences can be immediately conjured up, at will or spontaneously,

under the beam of our reflective glance. Functionally, the indexical 'I' which is the object of self-awareness is most likely the primary source of our experiences' normativity. Indeed, there is a gloss of plausibility to the idea that our most primitive sense of individuality, this sense of mineness, is what drives us, conscious creatures, to strive for coherence, coordination and efficiency in order to attain some ends. In a nutshell: the normativity of consciousness' essential subjectivity is what gives our experiences their *raison d'être*.

I have found the rival accounts of phenomenal unity wanting. Bayne and Chalmers and Dainton endorse a conception of phenomenal consciousness that says very little, if anything, about what it is for a mental state to bear a phenomenal property. Consequently, claiming that a subject entertains at any time a total state or 'conjoint phenomenology' of her experiences is more of a restatement of their thesis than a *case for* it. Their wait-and-see strategy of defense of the unity thesis is testament to this ailment: they only grasp on a pre-theoretical intuition which cannot be specified otherwise than by a set of logical properties. These accounts fall short in *principled* ways too: the notions of total experience and of unity relation are both incoherent within the authors' mereological framework.

Tye's one-experience conception went some way towards analyzing phenomenal unity in terms of representations. But his account proved equally lacking: we have quickly realized that his anti-mereological inclinations were superficial and poorly founded. His account of phenomenal unity therefore rests on quicksand: once the so-called 'one experience' unravels, there is nothing left to unify representations in any meaningful sense.

Contrary to other theorists, I have not presumed to provide a thesis of phenomenal unity that holds as a *necessity*. Rather, I have characterized *what it is* for a subject's experiences to be phenomenally unified. I have done so on both phenomenal and naturalist levels. On the phenomenal level, my account may be rephrased in a counterfactual statement: a subject S's experiences are phenomenally unified if and only if, if some of these experiences failed to comply with the coherence norms, S would non-reflectively experience this infringement.

On the naturalist level, I have argued that an experience consists in the binding of a first-order representation (of a worldly property) with the representation of this representation. Since binding is a causal relation, as opposed to a conceptual one, it endows the property of 'being conscious' with all the causality that it needs. By analogy, I have ventured that the integration of *several* experiences may consist in yet another binding-like phenomenon.

Having dressed my account in this physical clothing, I can indulge in the following statement: a subject's experiences are *normally* phenomenally unified. This 'normally' ought to be understood as an 'other things being equal' or perhaps 'if the subject is normally physically constituted' clause. This

is certainly not as strong as a necessity claim; but it is the most that we are entitled to hold. It also provides what we could hope for, explanatory-wise: since phenomenal unity is at the core of the subject's agency and rationality, the thesis of phenomenal unity allows for safe predictions as to how the subject feels like on the face of certain experiences (notably perceptual), and how she will reason and act upon these phenomenally unified experiences.

A propos the problem of the self, I have argued that we should not mistake the unity of a stream with that of a self, nor think that each is circularly derived from the other. The unity of a stream is derived from an experiential perspective 'I'. Streams are real things composed of real experiences. Hence the perspective from which streams are undergone must be more than just virtual objects. Selves, by contrast, are *a posteriori* and potentially third-personal fictions that serve as centers of narrative gravity of coherent streams. Babies and cats do not have selves other than those which we 'novelists' endow them with; but they have a first-personal sense of their own individuality from their birth onward. By contrast, a subject suffering from Multiple Personality Disorder may abstract several selves from a yet unique experiential perspective.

Finally, the functional relevance that I claim to be at the core of phenomenal unity enjoys the virtue of partially 'bridging' (though certainly *not* reducing!) what it is like for a subject to undergo her experiences together with how she behaves. I argued that stable incoherent behavior as that exhibited by acallosal patients is a conclusive sign that some of their experiences are (temporarily at least) phenomenally disunified. However, the fact that other representations are still integrated by binding mechanisms invites us to posit that they entertain one partially unified stream rather than two separated streams. We should not let us be held back by some pre-theoretical intuition that phenomenal unity is a transitive relation; nothing commits us empirically to this claim. In line with the view that the self is narratively constructed out of a coherent stream of consciousness, I have naturally concluded that we should not regard split-brain patients as each being several selves *unless they regard themselves as such* – which is not the case.

Although I have claimed my self-representational account to be naturalist, I do not doubt that there remain problems with the naturalization of our experiences' subjectivity, that is, this first-personal 'raw bearer' around which they inevitably gravitate.

Construing a conscious mental state as a causal integration of a representation with a higher-order representation thereof provides us with a respectable account of how S is aware of her *experiences*. But where and how exactly does the *for-her-ness* dimension of experiences arise? If each experience is conscious by the integration of a higher-order representation that is generated by a different brain state as I suggested, it seems that there is no single neural correlate associated with

and responsible for the subjective dimension of consciousness; it spreads all over the brain, so to say. Accordingly, I have tentatively suggested that the 'I' of our experiences is grounded in a certain causal unit that binds representations. But saying this does not advance us any further with regard to the question just posed; the explanatory gap is yet to be bridged. I take solace in the fact that this is a problem with which thinkers unanimously grapple, and in the fact that it will prompt many mind-stimulating philosophical and neurobiological inquiries to come.

Acknowledgments

I would like to thank Professors Fabrice Teroni and Julien Deonna, whose great pedagogical skills and interesting lectures and seminars on the philosophy of mind drew my interest to this field. I more specifically thank Fabrice Teroni for the supervision of this work, as well as my brother Maxime, whose stylistic corrections significantly contributed to the improvement of the original draft. At last, my thanks and gratitude go to Galit Bengui for the love and moral support that she gave me throughout the redaction of this work.

I also wish to acknowledge my indebtedness to the works of Bayne (2010), Bayne and Chalmers (2003), Dainton (2000), Hurley (1998) Kriegel (2004, 2005, 2006a., b.) and Tye (2003) which inspired many of the good ideas that are presented in this work.

References

- ARMSTRONG, D. M. 1981. What is consciousness? In ARMSTRONG, D. M. (ed.), *The Nature of Mind and Other Essays*. Ithaca, NY: Cornell University Press: 55–67.
- BAYNE, T., 2001. Co-consciousness: Review of Barry Dainton's *Stream of Consciousness*. *Journal of Consciousness Studies*, 8:79–92.
- , 2004. Self-consciousness and the unity of consciousness. *The Monist*, 87: 219–36.
- , 2005. Divided brains and unified phenomenology: A review essay on Michael Tye's *Consciousness and Persons*. *Philosophical Psychology*, 18(4):495–512.
- , 2010. *The Unity of Consciousness*. Oxford: Oxford University Press.
- , and CHALMERS, D., 2003. What is the unity of consciousness? In CLEEREMANS, A. (ed.), 2003, 23–58.
- , and MONTAGUE, M. (eds.), 2012. *Cognitive Phenomenology*. Oxford: Oxford University Press.
- BEDFORD, L. F., 2004. Analysis of a Constraint on Perception, Cognition and Development: One Object, One Place, One Time. *Journal of Experimental Psychology: Human Perception and Performance*, 30: 907–12.
- BLOCK, N., 1995. On a confusion about the function of consciousness. *Behavioral and Brain Sciences*, 18: 227–47.
- BROOK, A. and RAYMONT, P., 2006. Unity of consciousness. *Proceedings of the 28th Annual Conference of the Cognitive Science Society*, New York: LEA: 1056–61.
- , 2010. The Unity of consciousness. In ZALTA, E. (ed.), *The Stanford Encyclopedia of Philosophy*.
- CARRUTHERS, P., 2016. Higher-Order Theories of Consciousness. In ZALTA, E. (ed.), *The Stanford Encyclopedia of Philosophy*.
- CHALMERS, D., 1995. Facing Up to the Problem of Consciousness. *Journal of Consciousness Studies*, 2: 200–19.
- CLEEREMANS, A. (ed.), 2003. *The Unity of Consciousness: Binding, Integration and Dissociation*. Oxford: Oxford University Press.
- CRANE, T., 2003. The Intentional Structure of Consciousness. *Consciousness: New Philosophical Perspectives*. In JOKIC, A. AND SMITH, Q. (eds.) Oxford and New York: Oxford University Press (online version 1–27).
- DAINTON, B., 2000 (2nd ed. 2006). *Stream of Consciousness: Unity and continuity in conscious experience*, London: Routledge.
- , 2007. Coming Together: The Unity of Consciousness. In VELMANS, M. & SCHNEIDER, S. (eds.), *The Blackwell Companion to Consciousness*. Oxford: Blackwell: 209–22.
- DENNETT, D. C., 1981. Where Am I? In HOFSTADTER D., and DENNETT D. (eds), *The Mind's I*, New York: Basic Books.
- , 1991. *Consciousness Explained*. Boston: Little, Brown and Company.

- , 1992. The Self as a Centre of Narrative Gravity. In KESSEL, F., COLE, P. and JOHNSON, D. (eds.), *Self and Consciousness: Multiple Perspectives*, Hillsdale, NJ: Erlbaum Publishers
- , 1995. The path not taken. *Behavioral and Brain Sciences*, 18: 252–3.
- DESCARTES, R., 1647/1992. *Méditations métaphysiques*. Garnier Flammarion (Méditation VI : 175–213).
- DRETSKE, F. 1995. *Naturalizing the Mind*. Cambridge, Mass: MIT Press.
- GLÜER, K., 2009. In Defence of a Doxastic Account of Experience. *Mind and Language*, 24: 297–327.
- HILL, C. 1991. *Sensations: A Defense of Type Materialism*. Cambridge: Cambridge University Press.
- , 2014. Tim Bayne on the Unity of Consciousness. *Analysis Reviews*, Vol. 74, 3: 499–509.
- HUMBERSTONE, I. L., 1992. Direction of Fit. *Mind*, 101: 59–83.
- HURLEY 1994. Unity and Subjectivity. *Proceedings of the British Academy*, 83: 49–77
- , 1998. *Consciousness in Action*. Cambridge, MA: Harvard University Press.
- , 2000. Clarifications: Responses to Kobes and Kinsbourne. *Mind & Language*, Vol. 15 No. 5: 556–61.
- , 2003. Action and the Unity of Consciousness. In CLEEREMANS, A. (ed.), 2003.
- JAMES, W., 1890. *The Principles of Psychology*. New York: Henry Holt and Company.
- KOBES, B. W., 1995. Access and what it is like. *Behavioral and Brain Sciences*, 18: 260.
- , 2000. Unity of Consciousness and Bi-Level Externalism. *Mind & Language*, Vol. 15 No. 5: 528–44.
- KRIEGEL, U., 2004. Consciousness and Self-consciousness. *The Monist*, Vol. 87 No. 2: 182–205.
- , 2005. Naturalizing Subjective Character. *Philosophy and Phenomenological Research*, Vol. LXXI (1): 23–57.
- , 2006a. Philosophical Theories of Consciousness: Contemporary Western Perspectives. In MOSCOVITCH, M., THOMPSON, E. and ZELATO, P.D. (eds.), *Cambridge Handbook of Consciousness*. New York: Cambridge University Press.
- , 2006b. Consciousness: Phenomenal Consciousness, Access Consciousness, and Scientific Practice. In THAGARD, P. (ed.), 2006. *Handbook of Philosophy of Psychology and Cognitive Science*. Amsterdam: North-Holland: 195–217 (online version 1–33).
- , 2007. The Same-Order Monitoring Theory of Consciousness. *Synthese Philosophica*, 44: 361–84.
- LEVINE, J., 1983. Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly*, 64: 354–61.
- LOCKWOOD, M., 1989. *Mind, Brain and the Quantum*, Oxford: Blackwell Publishers, Ch. 6.
- LYCAN, W. G., 2001. A Simple Argument for a Higher-Order Representation Theory of Consciousness. *Analysis* 61: 3–4.
- , 2015. Representationalism. In ZALTA, E. (ed.), *The Stanford Encyclopedia of Philosophy*.
- MASROUR, F., 2014. Unity, Mereology and Connectivity. *Analysis Reviews*, Vol. 74, 3: 509–20.

- MILLAR, A., 2004. *Understanding People. Normativity and Rationalizing Explanation*. Oxford: Oxford University Press.
- NAGEL, T., 1971. Brain bisection and the unity of consciousness. *Synthese*, 22: 396–413.
- , 1974. What is it like to be a bat? *Philosophical Review*, 83: 435–56.
- O'BRIEN, B. and OPIE, J., 1998. The Disunity of Consciousness. *Australasian Journal of Philosophy*, 76: 378–95.
- PARFIT, D., 1984. *Reasons and Persons*. Oxford: Clarendon Press.
- PUCETTI, R., 1981. The case for mental duality: evidence from split-brain data and other considerations. *The Behavioral and Brain Sciences*, 4: 93–123.
- ROSENTHAL, D., 1997. A theory of consciousness. In BLOCK, N., FLANAGAN, O., AND GUZELDERE, G. (eds.). *The Nature of Consciousness*. Cambridge, MA: MIT Press: 729–53.
- , 2003. Unity of consciousness and the self. *Proceedings of the Aristotelian Society*, 103: 325–52.
- SEARLE, J., 2002. Consciousness. In *Consciousness and Language*, Cambridge, UK: Cambridge University Press. (Originally published in *Annual Review of Neuroscience*, 23 (2000): 557–78.)
- SCHECHTER, E., 2010a. Individuating mental tokens: The split-brain case. *Philosophia* 38: 195–216.
- , 2010b. Two Unities of Consciousness. *The European Journal of Philosophy* 21: 197–218.
- , 2013. The Unity of Consciousness: Subjects and Objectivity. *Philosophical Studies* 165: 671–92.
- SCHWITZBEHEL, E., 2014. Introspection. In ZALTA, E. (ed.), *The Stanford Encyclopedia of Philosophy*.
- SHOEMAKER, S., 1996. Unity of consciousness and consciousness of unity. In *The First-Person Perspective and Other Essays*, Cambridge: Cambridge University Press:
- , 2003. Consciousness and co-consciousness. In CLEEREMANS, A. (ed.), 2003. *The Unity of Consciousness: Binding, Integration and Dissociation*. Oxford: Oxford University Press: 59–71.
- , 2011. Review of Tim Bayne's *The Unity of Consciousness*. *Notre-Dame Philosophical Reviews*.
- SMITH, D. W., 1986. The Structure of (Self-)Consciousness. *Topoi*, 5: 149–56.
- STEIN, B. E. et al., 2002. Multisensory Integration. In RAMACHANDRAN, V. S. (ed.), *Encyclopedia of the Human Brain*. New York: Academic Press: 227–41.
- TYE, M. 1995. *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.
- , 2000. *Consciousness, Color and Content*. Cambridge, MA: MIT Press
- , 2003. *Consciousness and Persons*. Cambridge, MA: MIT Press.
- and WRIGHT, B. 2011. Is There a Phenomenology of Thought? In BAYNE, T. and MONTAGUE, M. (eds.) *Cognitive Phenomenology*. Oxford: Oxford University Press.
- VAN GULICK, R., 2014. Getting it All Together: Phenomenal Unity and the Self. *Analysis Reviews* Vol. 74, 3: 491–8.
- ZAHAVI, D., 2006. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge, Mass: MIT Press, Ch. 1.

