

4

Motivating the Unity Thesis

In the first chapter I surveyed a number of things that might be meant by the claim that consciousness is unified. Of these, I singled out the unity thesis for special consideration. According to the unity thesis, the conscious states that any subject of experience enjoys at any one point in time will occur as the components of a single total phenomenal state—a unitary ‘phenomenal field.’ The unity thesis, I suggest, provides us with a conception of the unity of consciousness that is substantive, interesting, and plausible. But what exactly is the source of its appeal?

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 the 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. Call this claim *the unity judgement*. Although the unity judgement doesn’t entail the unity thesis, there is a line of argument from the unity judgement to the unity thesis. According to this argument, the unity of consciousness that is revealed to introspection is not a feature that consciousness possesses only when one attends to its structure but is a feature that it enjoys all the time—even when one doesn’t (and perhaps cannot) introspect. Moreover, the fact that one’s own experiences are unified gives one reason to think that other subjects of experience also enjoy conscious unity. That, in a nutshell, is the introspection-based argument for the unity thesis.

This chapter explores the tenability of this argument. My examination will proceed in two stages. The first stage (§4.1–§4.3) focuses on ‘the unity judgement’—the introspective judgement to the effect that each of one’s current experiences is unified with each other; the second stage (§4.4–§4.5) examines whether the route from the unity judgement to the unity thesis is a legitimate one. We will see that although the argument from introspection is far from decisive, it does provide us with a respectable case for the unity thesis.

4.1 The unity judgement

Perhaps the most general objection to the unity judgement concerns the fact that it relies on introspection. Worries about the trustworthiness of introspection are nothing new but date back to the earliest days of the science of consciousness. Indeed, the demise of the study of consciousness in the early years of the twentieth century was due in no small part to introspectively based disputes (Boring 1950; Lyons 1986). A century later, the science of consciousness is still grappling with introspective disputes and the sceptical challenges that they engender (Schwitzgebel 2008; Bayne & Spener 2010).

This is not the place to engage in a comprehensive examination of introspection's credentials, but in fact no such examination is needed. The question 'Is introspection reliable?' is not a good one to ask, for epistemic faculties are reliable only with respect to certain questions and under certain conditions. Vision is typically a good way of determining the colour of a nearby British telephone box on a bright summer's day, but a rather poor way of determining the identity of a small, rapidly moving animal on a dark winter's night. The question we need to ask is whether introspection is likely to be reliable *with respect to the unity judgement*.

Evidence that might directly bear on this question is hard to come by, for in order to have such evidence we would already need to know whether or not consciousness is unified, but if we knew that then we wouldn't need to invest any faith in the unity judgement. However, there might be less direct ways of putting pressure on the unity judgement. According to many, introspection is not to be trusted when it comes to determining the *capacity* of consciousness. Such claims might be taken to have an indirect bearing on the unity judgement, for if introspection is an unreliable witness with respect to the capacity of consciousness then we may have reasons to think that it is also unreliable with respect to its structure. Let us consider this line of thought in some detail.

Introspection, it is commonly said, suggests that we enjoy a 'rich' or 'lavish' stream of consciousness, according to which a single moment of consciousness will typically contain a rich, detailed, and multimodal representation of the world. There is, however, reason to doubt whether this introspective picture is accurate. Research on 'the span of apprehension' dating back to the earliest days of empirical psychology is often taken to show that this introspective judgement is radically wrong, and that the bandwidth of consciousness is vastly more 'austere' or 'sparse' than introspection suggests.

Let us begin with a selective overview of the evidence in favour of the sparse view of the capacity of consciousness.¹ One line of evidence involves research on numerosity judgements, in which subjects are required to determine the number of items presented in a display. A much-replicated finding is that subjects are able to identify the number of items in displays with four items or fewer with high degrees of confidence, rapidly, and with very low error rates, whereas levels of confidence, reaction times, and error rates worsen for displays containing five or more items (Atkinson et al. 1976). This is true not just for visual items but applies also to other modalities: within olfaction subjects are able to discriminate and identify only four items in mixtures containing multiple odours (Livermore & Laing 1996, 1998); within audition subjects are able to discriminate one speaker from two speakers but have great difficulty in segmenting an auditory stream that contains the voices of three or more speakers (Kashino & Hirahara 1996); and within touch subjects can reliably discriminate one tactile stimulus from two, but with three or more stimuli error rates exceed 30 per cent (Gallace et al. 2006). Small numbers of perceptual items can be subitized but larger groups of items need to be counted. Counting requires the deployment of attention, which in turn suggests that subjects can no longer draw on the contents of a single conscious state but must integrate the contents of successive experiences.

A second line of evidence for the sparse model derives from the work of Pylyshyn and colleagues on the ability of subjects to track visual objects in the context of distractors (Pylyshyn & Storm 1988; Sears & Pylyshyn 2000; Yantis 1992). In a typical experiment, a subset of items within a larger display set is cued, and subjects are required to track the cued objects whilst all the items—both targets and distractors—move randomly through the display. The items are then made stationary, and subjects are required to identify whether a randomly selected object is a target or a distractor. Results suggest that subjects are able to track only four items with any degree of reliability.

Change blindness provides a third motivation for the sparse conception of consciousness. This label refers to the surprising inability of subjects to detect large changes in naturalistic scenes across successive presentations.² In a representative experiment (Grimes 1996), subjects were shown an image of two cowboys sitting on a bench. Over 50 per cent of the subjects failed to notice when the heads of the cowboys were swapped during a saccade, despite the fact that they had been told to expect such changes. Kevin O'Regan (1992), amongst

¹ See also Cowan (2001); Huang et al. (2007); Irwin (1996); Luck & Vogel (1997).

² Nakayama (1990); McConkie & Rayner (1975); McConkie & Zola (1979); Rensink et al. (1997); Simons & Levin (1997).

others, has argued that subjects are unaware of such changes because visual experience fails to encode the information required for their detection. We have here the roots of the ‘grand illusion’ conception of visual experience, according to which the apparent richness of perceptual consciousness is an illusion generated by the fact that much of the world is immediately and effortlessly *available* to awareness.

Taken as a whole this research certainly provides *some* support for the sparse conception of consciousness. After all, the very fact that these results are surprising suggests that they are at odds with our naïve conception of the capacity of consciousness (Dennett 2001). However, some degree of caution would not be inappropriate here. First, each of these experimental paradigms requires the exercise of capacities that go beyond those required for consciousness as such. This is perhaps most obvious with respect to change blindness. It is one thing to be aware *of* features in a scene that have changed, and it is another to be aware *that* they have changed. Detecting a change requires not only that one represent the features that have changed but also that one integrate the representations of those features, and it is possible that the failure to detect changes often arises not because the features themselves haven’t been consciously represented but because the representations of those features haven’t been appropriately integrated (Dretske 2004; Henderson & Hollingworth 2003; Hollingworth et al. 2001; Simons et al. 2002). Similar concerns can be raised about the numerosity and multi-object tracking experiments. In each case, reliable performance requires not merely that subjects are aware of the presented items but that they also bring them under certain kinds of representations—either number categories (numerosity judgements) or temporally stable object files (multi-object tracking).

A second ground for caution is that much of this work is concerned with identifying the capacity of accurate visual *perception* rather than that of visual *experience* as such. In other words, these results may show not that visual experience is austere, but that its content owes a great deal more to top-down expectation and ‘filling-in’ than we tend to assume. Perhaps part of the surprise that we experience in response to these experiments is a reaction to discovering how poor we are at tracking the world rather than a reaction to discovering how impoverished the contents of consciousness are.

Finally, we should keep in mind the fact that these experiments are typically unimodal. Even if there is a four-item limit on experience within any one modality the *overall* capacity of consciousness may be somewhat larger. One could argue that although these experiments are unimodal in structure they actually tap the overall capacity limits of consciousness, for in the relevant experimental contexts the subject’s experience is restricted to a single modality,

but it is hard to believe that the stream of consciousness is ever restricted to a single perceptual modality, even in the context of (say) a visual-tracking experiment. These impressionistic remarks can be buttressed by research on intermodal integration, which strongly suggests that the various perceptual modalities do not function as autonomous streams of processing but are highly interdependent (see § 10.3). All in all, the evidence for the sparse conception of the capacity of consciousness is rather less compelling than it is often taken to be.

We should also note that it is not entirely obvious how much support introspection actually does provide for the ‘lavish’ conception of consciousness. Although introspection appears to lead some subjects to embrace a rich account of consciousness, others take introspection to support a rather more modest picture of the capacity of consciousness (Schwitzgebel 2007). I myself incline towards some degree of modesty here. As best I can tell, my typical phenomenal field involves a small band of focal experience surrounded by an experiential penumbra. This focal experience is usually dominated by at most two or three modalities at a time, with only a few objects and features represented in any detail across these two or three modalities. The penumbra surrounding this focus might include a background sense of affective tonality, an awareness of the general orientation of my body, and perhaps also various fragments of fringe and cognitive phenomenology, such as tip-of-the-tongue experiences or stray thoughts. This sketch might not qualify as an endorsement of the ‘sparse’ view of consciousness, but it is certainly some distance from the lavish model. Naïve introspection—or, perhaps better, our naïve theorizing about consciousness on the basis of naïve introspection—might lead us to overestimate the capacity of consciousness somewhat, but the case for thinking that it is radically unreliable has not been made.

Let us return to the unity judgement. Even if introspection *were* an unreliable witness with respect to the capacity of consciousness, it is a further question whether it is also unreliable with respect to the unity judgement. Taking their lead from the ‘grand illusionists’, the critic might suggest that the sense of experiential unity provided by introspection is an illusion that arises from mistaking features of the world for features of consciousness. The world itself may be unified but our experience of it is fragmentary and piecemeal, built up over successive experiences rather than contained within a single unified conscious state.

I don’t find this line of thought convincing. It is certainly true that the world is unified (whatever that might mean), and it is also true that the representational contents of our experience of the world are also built up over successive experiences. But there is nothing in these platitudes that might undercut or ‘explain away’ the force of the unity judgement. Not only is the world itself

unified, so too are our experiences of it. Perceptual *content* is often ‘fragmentary’ and ‘piecemeal’, but no matter how partial one’s take on the world may be it is invariably contained within a single overarching experiential state. I submit that the unity judgement has little to fear from attacks on the reliability of introspection.

4.2 Beyond the reach of introspection

Introspection itself might be reliable when it comes to the unity judgement, but what about experiences that elude the reach of introspection? Even if each of the experiences of which one is introspectively aware were phenomenally unified with each other, those experiences might constitute only a subset of one’s overall set of experiences. The possibility of introspectively inaccessible experiences casts a shadow of suspicion over the unity judgement.

In examining this objection it is useful to distinguish between two types of introspective inaccessibility. One type of inaccessibility obtains when experiences occur within parts of cognitive architecture that are sealed off from the mechanisms of introspection. We might call such states ‘deeply inaccessible’. Experiences immured within Fodorian modules (Block 2007) or sub-personal homunculi (White 1987)—if such there be—would be deeply inaccessible. I am going to set the question of deeply inaccessible states to one side here. I do this not because I regard deeply inaccessible experiences as conceptually or metaphysically impossible, but because we have no evidence of such states. It seems to me that deeply inaccessible experiences ought to be treated in the way that we treat radically sceptical scenarios: although they may be of interest to certain philosophical projects, they ought not constrain sober exercises in theory-building.

Another class of inaccessible experiences cannot be dismissed quite so easily. If introspection has a more restricted ‘bandwidth’ than consciousness itself, then we might expect that complex experiences would not be introspectively accessible (as such). Let us call such states ‘superficially inaccessible’.

Unlike deeply inaccessible experiences, there may be reason to think that superficially inaccessible experiences occur. The data in question involves Sperling’s well-known experiments on the reportability of briefly presented visual stimuli (Sperling 1960; Averbach & Sperling 1961). In these 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.

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. This is known as the *partial report superiority effect*.

Arguably, the most natural explanation of this effect is that there is a bottleneck on reportability—and, perhaps more fundamentally, on introspection—that prevents subjects from gaining access to the full sweep of their experiential content (Block 2007). Subjects in the experiment enjoy a rich visual experience whose content ‘outstrips’ that to which they have introspective access. According to this interpretation, the visual phenomenology of subjects is not limited to a generic representation of the matrix as (say) containing twelve alphanumeric figures, but includes a detailed representation of the specific identity of each—or at least most—of the figures in the matrix. There are other interpretations of the Sperling data (see the commentaries on Block 2007; Phillips forthcoming), but this is perhaps the most plausible one.

What implications might this picture have for the unity judgement? Well, if subjects have a detailed, fine-grained experience of the matrix as a whole to which they lack introspective access, then they are hardly in a position to have introspective warrant for thinking that each of their current experiences is unified with each other. Subjects might be able to tell whether or not their experiences of each of the *rows* and/or *columns* of the matrix are unified with each other, but they will not be able to tell whether those experiences are unified with their experience of the entire matrix. And, the objection continues, there is no reason why it shouldn’t be possible for experiences of the parts of the matrix to be unified with each other without also being unified with a non-generic experience of the matrix as a whole. (As an aid to intuition imagine a subject with one stream of consciousness in its right hemisphere and another in its left hemisphere: its experiences of the parts of the matrix might be located in one hemisphere and its experience of the matrix as a whole might be located in the other hemisphere.) What should we make of this objection?

It is, I think, an open question whether the scenario just outlined is possible, for one might argue that there are deep constraints between phenomenal unity on the one hand and representational unity on the other, such that experiences of the parts of an object cannot be phenomenally unified with each other without also being unified with an experience of the whole of that object at

least for subjects who do have an experience of the whole of the object (although see §3.3). But even if it is *possible* for a subject to have an experience of the matrix as a whole that is not itself phenomenally unified with experiences of the various parts of that matrix, this would surely be a highly unusual state of affairs. If there is one place where we should expect relations between the contents of experience to constrain relations between the experiences that carry those contents it is here. A cognitive architecture in which experiences of a matrix as a whole were not phenomenally unified with experiences of its various parts would surely have little to recommend it. In sum, it seems reasonable to assume that if the experiences of the parts of an object are phenomenally unified with each other then they will also be unified with an experience of the object as a whole, at least when that object is indeed experienced as a whole.

But the critic is not yet done. Consider experiences e_1 and e_2 , which are such that each represents a single column of the Sperling matrix. Although the contents of e_1 and e_2 are *individually* accessible to introspection, they are not *conjointly* accessible. And, the critic might continue, the fact that a pair of experiences is not conjointly introspectable suggests that they are not phenomenally unified with each other. Indeed, it is not implausible to suppose that it is part of the functional role of phenomenal unity that phenomenally unified experiences are conjointly available to introspection.

Although this argument has some force, there is another line of argument that must also be taken into account. The fact that e_1 and e_2 are individually introspectible surely gives us some reason to think that they are phenomenally unified with each other, for experiences that are not phenomenally unified are unlikely to be available to the same consuming systems. So we have two competing considerations: the fact that e_1 and e_2 are not conjointly introspectable suggests that they are *not* unified, but the fact that each is introspectable suggests that they *are* unified. Which of these lines of argument has most weight?

It seems to me that the case for unity trumps the case for disunity. The reason for this is that we have an account of why e_1 and e_2 are not conjointly introspectible—namely, that any experience that subsumes both e_1 and e_2 would be ‘too large’ to make it through the bottleneck of introspection. Just as the experience of the entire matrix is too complex to make it from the visual buffer to working memory ‘in one piece’, so too the experience of any two columns is also too complex to make it through this channel. If, however, it turned out that the content of the conjunction of e_1 and e_2 was ‘smaller’ than that of an experience that could be introspected, then we would have reason to think that e_1 and e_2 are not phenomenally unified. But, to the best of my knowledge, it doesn’t and so we don’t.

Let us take stock. I began this section by distinguishing two forms of introspective inaccessibility: deep inaccessibility and superficial inaccessibility. I then set to one side the potential threat from deeply inaccessible experiences in order to focus on the more pressing—not to mention more tractable—challenge posed by superficially inaccessible experiences. Although it is contested, we noted that Sperling’s data provides us with some evidence of superficially inaccessible experiences. Do such states call the unity judgement into question? In principle they might, for the fact that experiences are not conjointly available for introspection gives us some reason to deny that they are phenomenally unified. But such reasons must be weighed against the fact that such states are individually introspectible. Introspective support for the unity judgement derives not only from the fact that one is aware of one’s experience *as* unified, but also from the fact that experiences that are available to introspection—whether conjunctively or individually—are likely to be unified with each other.

4.3 Hurley’s objection

Hurley’s treatment of the unity of consciousness provides us with the materials for a quite different objection to the argument from introspection. According to this objection, thinking that introspection might provide support for the unity thesis would involve something akin to a category mistake, for introspection provides us with access only to the *contents* of consciousness and not to its *structure*. This objection should be familiar, for we examined a version of it in Chapter 2 in connection with the question of partial unity. Let us begin by briefly revisiting that discussion.

One of the main objections to partial unity concerns the challenges one confronts in projecting oneself into a partially unified perspective. As we noted, Hurley argues that it would be wrong to think that such difficulties show that the notion of a partially unified perspective is incoherent, for the subjective perspective has access only to the contents of experiences. There are, she suggests, partially unified perspectives that would be subjectively indistinguishable from certain fully unified perspectives.

To appreciate the force of Hurley’s argument consider *Fully Unified* and *Partially Unified* (see Figure 4.1).³ Despite the fact that *Fully Unified* has a fully unified consciousness and *Partially Unified* has only a partially unified conscious-

³ Because Hurley’s cases involve unnecessary complications I have changed their details whilst preserving their spirit (1998: 108–12; see also Hurley (2003)).

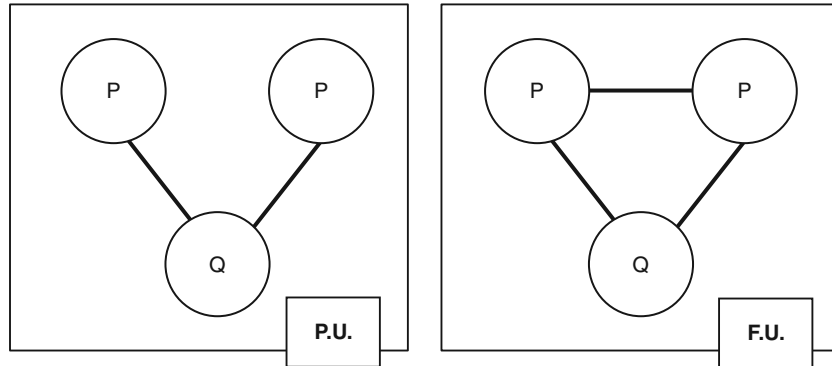


Figure 4.1 A partially unified perspective (P.U.) and a fully unified perspective (F.U.)

ness, there will be no subjective ('what it's like') difference between them. Both subjects will be aware of experiencing $\langle p \rangle$ and $\langle q \rangle$ together, in the context of a single experiential state. Introspection might lead both subjects to endorse the unity judgement ('I have a unified consciousness'), but only in *Fully Unified's* case would that unity judgement be correct.

Although Hurley's point focuses on the question of partial unity, parallel considerations apply to the question of disunity more generally. Contrast *Fully Unified* not with *Partially Unified* but with *Disunified* (Hurley 1998: 110; see Figure 4.2). *Disunified* and *Fully Unified* have experiences with the same contents, but whereas *Disunified* has two streams of consciousness *Fully Unified* has only a single stream of consciousness. Again, the thought is that there would be

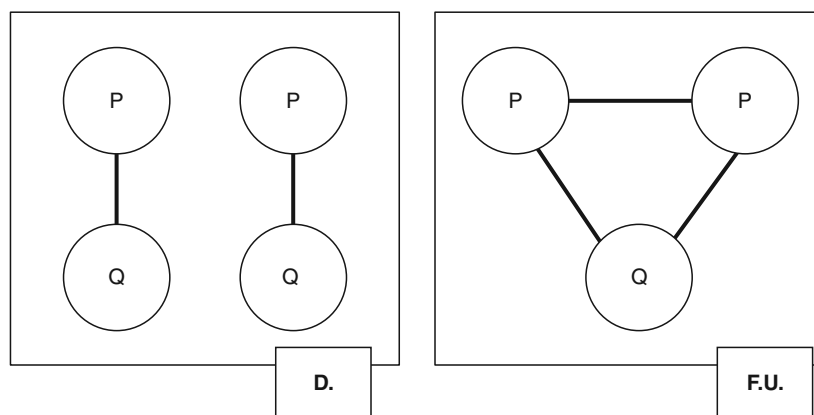


Figure 4.2 A disunified perspective (D.) and a fully unified perspective (F.U.)

no subjective—and hence no introspectively accessible—difference between these two perspectives. *Disunified* would have as much reason to endorse the unity judgement as *Fully Unified* has, but in her case it would be false. Call this *Hurley's objection*.

In response to Hurley's objection one might argue that because consciousness is necessarily unified then neither *Disunified* nor *Partially Unified* is possible, and as a result we needn't worry that our treatment of the unity of consciousness cannot distinguish their phenomenal perspectives from each other. Although tempting, this response is dialectically problematic. Perhaps it is true that neither *Disunified* nor *Partially Unified* is possible, but it would be illegitimate to dismiss Hurley's objection on that basis given that the very point of the objection is to undermine the main argument for the unity thesis. To appeal to the unity thesis here would be placing the cart before the horse, to put it mildly.

But there are responses to Hurley's objection that do have some force. Hurley's cases make essential appeal to the notion of *phenomenal duplication*, the idea that a subject can have multiple tokens of a single experiential type at one and the same time. If phenomenal duplication is not coherent then neither is her objection. And we have seen that the coherence of phenomenal duplication is very much up for grabs (see §2.4). Phenomenal duplication might be possible if we were to individuate experiences in sub-personal or vehicular terms, but I argued that experiences should be individuated in tripartite terms—that is, by appeal only to their subject, time, and phenomenal character. And the tripartite account rules out the possibility of phenomenal duplication, for it requires that numerically distinct experiences differ from each other in terms of either their subject of experience, time, or phenomenal character, and—by definition—phenomenal duplicates differ from each other in none of these three ways.

Might we be able to reformulate Hurley's objection without appealing to phenomenal duplicates? No, for were the subjects in question to have experiences with different contents then there would also be subjective (and hence introspectable) differences between their perspectives. Suppose that we replace one of *Disunified's* <p> experiences with (say) an experience with the content <r>. In that case, *Disunified* would have conjoint experiences with contents <p&q> and <p&r> respectively, but no experience that subsumes both of those experiences. As a result, *Disunified* would have introspective reason to reject the unity judgement. In short, Hurley's objection cannot be formulated without appealing to phenomenal duplication.

What if we were to individuate experiences in vehicular terms, and thus allow for the possibility of phenomenal duplication? Even then, I doubt that we should be overly worried by Hurley's objection. The unity judgement would

be under real threat only if it were to turn out that we are subject to phenomenal duplication on a widespread scale. Do we have reason to believe that that is the case? Some theorists have suggested that we do. Puccetti (1981) and Bogen (1977) argue that since split-brain patients have two streams of consciousness with duplicate contents, we have reason to suppose that we too enjoy two streams of consciousness with duplicate contents. On their view, the left and right hemispheres of a normal human brain sustain distinct streams of consciousness—with essentially the same contents—that operate in parallel. The split-brain procedure doesn't create two streams of consciousness but merely reveals the presence of the two streams of consciousness that were already in place prior to the operation.

This proposal is best rejected. For one thing, the two-streams interpretation of the split-brain data is far from irresistible, as I will argue in Chapter 9. But even if we had good reason to suppose that *split-brain* patients have two streams of consciousness, we should not read the phenomenal structure of the split-brain syndrome back into that of normal cognition. The two-stream model of the split-brain is motivated by the striking cognitive disunities that split-brain patients manifest, but these disunities are striking only against the backdrop of the unity exhibited by normal human subjects.

In attempting to motivate the claim that normal human beings are subject to phenomenal duplication Puccetti and Bogen don't just appeal to the supposed parallels between normal subjects and split-brain subjects, they also appeal to the bilateral structure of human consciousness.⁴ They argue that because each hemisphere has areas devoted to processing the same kinds of stimuli, we should think of consciousness as taking the form of two streams, one of which is located in the subject's left hemisphere and the other in the right hemisphere. What should we make of this argument?

The first point to note is that this argument could undermine the unity judgement only if a great deal of neural processing were bilaterally duplicated. If bilaterality held only occasionally—say, for only a few select types of conscious states—then it would leave the unity judgment unscathed even if there were a direct inference from bilaterality to phenomenal duplication. How much bilaterality is there within consciousness? We don't really know. We do know that although a good deal of neural processing is lateralized to one or other hemisphere—for example, early visual processing of events in each visual hemifield is carried out in the contralateral visual cortex—some neural processing seems to be bilaterally duplicated. Perceptual processing of certain high-level categories

⁴ Arguments against the unity of consciousness based on the bilateral structure of the brain have a long history. See Harrington (1987) for a fascinating account of their 19th-century roots.

appears to be carried out by each hemisphere in parallel (see e.g. Marsolek et al. 2002; Andresen & Marsolek 2005; and Rousselet et al. 2002), as does the representation of objects that are located near the centre of the visual field (where the two visual fields overlap). However, we do not yet have a clear conception of the degree to which consciousness involves neural processing that is inter-hemispherically duplicated.

Even where bilaterality does obtain it may not support claims of phenomenal duplication. In order to move from bilaterality to phenomenal duplication we need to consider how the homologous areas in each hemisphere are related to each other. Do they function as a single unit or as autonomous causal nodes? Does the left-hemisphere neural state subserve (say) verbal behaviour and the right-hemisphere state subserve (say) non-verbal grasping behaviour, or do both right- and left-hemisphere states conspire together to subserve all forms of behavioural control that are based on this experiential state? Homologous neural events might underpin or realize distinct experiential tokens if they operate autonomously, but if they operate as an integrated unit then it might be best to think of them as jointly underpinning or realizing a single experiential token—a token whose realizer is distributed between the two hemispheres (Marks 1981; Tye 2003)—rather than two tokens of the same experiential type. So even on a vehicular conception of experiences the route from bilaterality to phenomenal duplication is not straightforward, and may turn on questions about the architecture of consciousness to which we do not yet have the answers. In sum, the bilaterality-based case for thinking that normal subjects of experience have two streams of consciousness has not yet been made.

At this point it might be useful to review the various steps that we have taken in our examination of the unity judgement thus far. I began in §4.1 by examining an objection to the unity judgement based on the claim that introspection is an unreliable guide to the capacity of consciousness. According to this objection, the introspectively based sense we have of enjoying a rich stream of consciousness is illusory. I suggested both that the evidence in favour of the austere conception of consciousness is far from decisive, and that introspection might provide us with a reliable guide to the structure of consciousness even if it should turn out to be unreliable with respect to its capacity. In §4.2 I examined a threat to the unity judgement provided by introspectively inaccessible experiences. Putting to one side the *recherché* possibility of deeply inaccessible experiences, I focused on the challenge of superficially inaccessible experiences. I argued that we might be able to explain the inaccessibility of such experiences by appealing to their ‘size’ rather than the fact that they are not phenomenally unified with the rest of the subject’s (introspectively accessible) experiences. A final challenge to the unity judgement focused on Hurley’s

attack on the ‘what it’s like’ conception of phenomenal unity. Despite its ingenuity, I argued that Hurley’s objection is vulnerable at two points. First, it requires that phenomenal duplication is possible, and there is reason to doubt whether that is so. Secondly, even if phenomenal duplication is possible, Hurley’s objection undermines the unity judgment only if we have good reason to think that we are subject to phenomenal duplication on a widespread scale, which we don’t. I leave it to the reader to judge where these considerations leave the unity judgement, but my own view is that it is in quite good shape.

4.4 From the unity judgement to the unity thesis: I

The unity judgement might be secure, but how do we get from the unity *judgement* to the unity *thesis* given that the former is a claim about the structure of one’s own consciousness and the latter is a claim about the structure of human consciousness in general? One might be forgiven for thinking that evidence about the structure of one’s own consciousness—indeed, evidence about the structure of one’s own consciousness that is limited to contexts of introspective attention—is rather too thin a reed on which to ground the unity thesis.

There is clearly some risk in adopting the unity thesis on the basis of the unity judgement, but—as Hume pointed out—a certain amount of risk attends each and every inductive generalization. The question is whether the amount of risk involved in this generalization is *excessive*. Let us begin by asking how an inference from the unity judgment to the unity thesis might go wrong. Two possibilities spring to mind. One way in which the unity judgement could be true but the unity thesis false is if the structure that consciousness possesses when the subject is introspecting is not representative of its structure at other times. I examine that challenge to the argument from introspection in the following section. In this section, I focus on a prior challenge to the argument: that the subject making the unity judgement might be unrepresentative of human beings in general.

First-year logic students are typically taught that induction on the basis of a single case is not to be trusted. There is some truth to this claim, for one-shot induction is often unreliable. The fact that I adore Vegemite gives me little reason to think that you too adore Vegemite, and even less reason to think that Vegemite is universally adored. But what every first-year logic student is told needs to be taken with a grain of salt, for there are contexts in which one-shot inference is perfectly legitimate. The fact that a certain substance is poisonous to

me might make it reasonable for me to infer that it will also be poisonous to you, and even that it will be poisonous to human beings in general.

Whether or not one-shot inference is warranted depends on how the property in question is distributed in the relevant population (or, perhaps, on how it is believed to be distributed). Suppose that we are dealing with a property that is uniformly distributed in the relevant population: either every member of the population has it or no member of the population has it. For such a property, one-shot induction will be highly reliable (indeed, it will be infallible). If you want to know whether human beings are mortal, you need a sample of only one. Other properties, of course, are not homogeneously distributed. (Liking Vegemite would be one such property.) So, in order to know whether the unity thesis might be justified on the basis of one-shot inference (the unity judgement) we need to ask how likely it is that human consciousness has a uniform structure.

Let us approach this question by stages. I think we have good reason to suppose that there is a high degree of homogeneity in how the broad features of consciousness are realized in the members of our species. In fact, the very ascription of consciousness to other people depends on such an assumption. My warrant for thinking that you are conscious is based primarily on an inductive inference from my own case. I know that I am conscious, and I know that you and I are similar in ways that are highly germane to the possession of consciousness. Not only do we have various behavioural dispositions and physical properties in common, we also share a common evolutionary heritage (Sober 2000). These facts make it reasonable for each of us to engage in a one-shot inference when it comes to ascribing consciousness to our fellow citizens. (Indeed, not only is this one-shot inference warranted, it would be positively irrational to harbour serious doubts about it.)

One-shot inference is not just warranted when it comes to the ascription of consciousness *per se*, it is also warranted when it comes to the ascription of fine-grained conscious states. It is reasonable for me to suppose that the phenomenal states that I enjoy when hearing a trumpet are much like these which you enjoy when hearing a trumpet. Of course, such inferences *can* be derailed by the presence of individual differences. Some of these differences are obvious (those who are profoundly deaf lack auditory experience); others perhaps slightly less so (musical training can alter the phenomenology of auditory experience); still others are extremely surprising (individuals with synaesthesia might have musical experiences in contexts that are completely unknown to the rest of us). But esoteric cases to one side, one-shot inference is generally secure when it comes to the ascription of particular conscious states.

It seems reasonable to suppose that one-shot inference will be at least as reliable when it comes to the structure of consciousness as it is with respect to other features of consciousness—indeed, it may even be *more* reliable here than it is elsewhere. This is because the structure of consciousness is likely to be fixed by deep and stable features of our cognitive architecture, and thus to exhibit little variation within the normal range of adult human experience. Whether that homogeneity extends to (say) infants or those who have experienced cerebral insult is more of an open question. Here, I think, we can't expect the argument from introspection to have much force. In such cases, we will just have to determine the structure of consciousness as best we can from the agent's behaviour. (I return to this issue in the following section.)

There is a further point worth noting here. Although I don't have introspective access to your experiences, I do have access to your introspective reports. And if introspection can provide me with evidence for thinking that my current experiences are unified, then it can also provide you with evidence for thinking that your current experiences are unified. So in this sense my introspective evidence for the unity thesis is not limited to the results of my introspective endeavours but can include your introspective endeavours too. And if that is right, then the inference from the unity judgement to the unity thesis is not just a *one*-shot inference, but can draw on as many 'shots' as there are individuals who have reflected on the structure of their experience and have endorsed the unity judgement.

4.5 From the unity judgement to the unity thesis: II

If the unity judgement is correct then one's consciousness is unified when one introspectively attends to its structure. However, such occasions are surely highly unusual. Introspection, a critic might claim, is a pretty uncommon presence in ordinary waking life, and it is even less common for subjects to introspectively enquire into the structure of their own consciousness. With these thoughts in mind, one might wonder whether the unity judgement is representative of the normal structure of consciousness. Perhaps consciousness is unified only during periods of introspective attentiveness. And if that were so then the unity argument would be fatally flawed, for the unity judgement would provide a misleading picture of the typical structure of experience.

What should we make of this line of thought? First, I'm not so sure that introspection *is* that uncommon. True, we don't often explicitly attend to our own conscious states—our concern is generally with the world and not with our experience of the world—but arguably a background sense of one's own

experience pervades normal waking life. Everyday experience often contains an implicit awareness of the fact that one is conscious, and also of various facts about the generic character of one's experience—that one is (e.g.), perceiving, or daydreaming, or thinking hard about a particular problem (although see Schooler 2002). And, it seems to me, we are also implicitly aware of the fact that our conscious states are unified with each other. Awareness of this form doesn't involve focal attention, but it is nonetheless a genuine form of introspection insofar as it takes as its objects (facts about) one's own conscious states.

'Fair enough', the critic might reply, 'but there is a further problem here. In order to be justified in thinking that the structure of introspected experiences is representative of the structure of unintrospected experiences, we need to reckon with the "refrigerator light" problem. How do you know that introspection (or the possibility thereof) isn't responsible for phenomenal unity in the way that opening the refrigerator door is responsible for switching on its light? If introspection (or the possibility thereof) were responsible for the unity of consciousness, then far from providing us with a reliable picture of the structure of consciousness introspection would provide us with exactly the wrong picture.'

Sceptical worries of this sort are most pressing when one has reason to think that the conditions of epistemic access to the scenario in question might themselves play a role in bringing it about. The refrigerator light problem is a problem only because our background knowledge of refrigerators makes it plausible to suppose that they might be designed with lights that are triggered by opening the door. We would have reason to take the refrigerator light objection seriously if—and perhaps only if—we had reason to suppose that either introspection itself or the mechanisms underlying introspective accessibility were responsible for the unity of consciousness. Do we have any such reason?

I am inclined to think not. Introspection, as I am conceiving of it, is a faculty whereby the phenomenal character of experience becomes available to the mechanisms of judgement (more specifically, to the mechanisms of self-ascription). It is, I think, more plausible to suppose that introspection reveals the unity within consciousness that exists independently of it than it is to suppose that phenomenal unity is a product of introspective accessibility (see also Dainton 2006: 34–41). I suspect that the mechanisms responsible for introspection (and conscious thought more generally) are relatively late additions to the architecture of consciousness. They are grafted onto those mechanisms responsible for the creature's awareness of its perceptual and bodily environment, and it is these more primitive mechanisms that ensure that the creature's

experiences are unified. On this view, the unity of consciousness is not imposed on it by introspection but is a feature of consciousness that can be found even in creatures who enjoy only primitive forms of sensory experience.

Of course, even if phenomenal unity is independent of introspective capacities *per se* it might nonetheless depend on the kinds of integrative capacities whose presence is highly *correlated* with the presence of such capacities. Suppose, for example, that the unity of consciousness depends on working memory. If so, then it is likely to be compromised precisely in those conditions—such as delirium, dreaming, and the minimally conscious state—in which the mechanisms of working memory are disrupted. And if the unity of consciousness depends on the capacity for first-person thought (to take another possibility), then it is likely to be compromised in those syndromes—such as schizophrenia, dementia, and the dissociative disorders—in which the mechanisms underpinning first-person thought are also disrupted. In other words, even if the unity of consciousness doesn't depend on introspection *per se*, there may be reason to think that it is most likely to break down in precisely those conditions in which introspective access to the structure of consciousness is least reliable.

In my view this objection poses the most serious challenge to the argument from introspection. However, it is very difficult to say just *how* serious this challenge is, for we are largely ignorant of the mechanisms that might be implicated in the unity of consciousness, and thus we have relatively little grip on the conditions under which that unity is most likely to be lost or impaired. (Indeed, we are in something of a methodological bind here, for in developing such an account we will need to make assumptions about when the unity of consciousness might be lost and when it is retained.) My hunch, however, is that phenomenal unity is a rather basal phenomenon, one that requires little in the way of either working memory resources or the capacity for first-person thought. But to the extent that this is merely a hunch, the inference from the unity judgement to the unity thesis is certainly open to question.

4.6 Conclusion

This chapter has examined the first-person motivation for the unity thesis. As we have seen, providing introspective justification for the unity thesis turns out to be a rather challenging affair. Let us review the central points at which the argument from introspection is vulnerable—or at least apparently so.

The first half of this chapter examined three challenges to the unity judgement. An initial challenge involved the charge that introspection might be an unreliable guide to the structure of consciousness. A second challenge to the unity judgement came in the form of unintrospectible experiences. Inspired by Hurley's treatment of partial unity, a third challenge focused on the claim that the contrast between unity and its absence is not the sort of thing that might be introspectively discernible, even in principle. I argued that there are plausible responses to each of these challenges.

In the second half of this chapter I examined the inference from the unity judgement to the unity thesis. We saw that there are two gaps here: a gap between oneself and others, and a gap between contexts in which one enjoys introspective access to consciousness and contexts in which one doesn't. In §4.4 I argued that the first of these gaps can be adequately bridged, for inductive inferences from claims about the structure of one's own consciousness to that of others' are warranted even if they are one-shot inferences (which they needn't be). The most serious challenge to the argument from introspection involves the second of these two gaps. If the unity of consciousness were dependent on either introspection itself or the kinds of integrative capacities that are correlated with the presence of introspection then claims about the structure of consciousness based on introspection would not be a reliable guide to its structure in the context of non-standard background states. Is it likely that the unity of consciousness does depend on the kinds of integrative capacities that are correlated with introspection? My own view is that although it is not likely it is certainly an open possibility. Given our ignorance about the mechanisms underpinning consciousness, it would be foolish to suggest otherwise. Perhaps what we ought to say is that the 'pre-theoretical' case for thinking that a target individual has a unified consciousness will be a function of its background state of consciousness: the further we stray from normal waking consciousness, the weaker the case provided by the unity judgement for supposing that the individual in question retains a unified consciousness.

Where does that leave the unity thesis? It would be premature to put too much faith in it, but I do think that the foregoing considerations show that it ought to be taken seriously. Not only does it provide us with a respectable focus for third-person investigations into the structure of consciousness, we might even go so far as to say that it ought to be accorded a kind of default status—'innocent until proven guilty'. Let us turn now to the question of how we might go about establishing whether or not it is guilty.