What is the Unity of Consciousness?

Tim Bayne and David J. Chalmers

Department of Philosophy
University of Arizona
Tucson, AZ 85721.

[[Published in A. Cleeremans (ed.), *The Unity of Consciousness: Binding, Integration, Dissociation* (Oxford, 2003).]]

1 Introduction

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

This is a rough characterization of the unity of consciousness. There is some intuitive appeal to the idea that consciousness is unified, and to the idea that it must be unified. But as soon as the issue is raised, a number of questions immediately arise.

(1) *What is the unity of consciousness?* What does it mean to say that different states of consciousness are unified with each other, or that they are part of a single encompassing state? The idea of unity is multifaceted, and has been understood in many different ways by different thinkers. In some senses of "unity", the claim that consciousness is unified may be obvious or trivial; in other senses, the claim may be obviously false. So the first project in this area is to distinguish between varieties of unity, and to isolate those varieties that pose the most important questions.

(2) *Is consciousness necessarily unified?* Some thinkers (Descartes and Kant, for example) have argued that some sort of unity is a deep and essential feature of consciousness. On this view, the conscious states of a subject are necessarily unified: it is impossible for there to be a subject whose conscious states are disunified. On the other side, some thinkers (e.g. Nagel 1971) have argued that the unity of consciousness can break down. On this view, there are cases (especially neuropsychological cases, such as those involving patients with split brains) in which a subject's states of consciousness are disunified. Some (e.g. Dennett 1992) hold more strongly that consciousness is often or usually disunified, and that much of the apparent unity of consciousness is an illusion.

(3) *How can the unity of consciousness be explained?* If consciousness really is unified, and especially if it is necessarily unified, then it is natural to look for an explanation of this fact. What is it about consciousness that yields this unity? Is unity a primitive feature of consciousness, or is it explained by something deeper? Further, the unity of consciousness may put strong constraints on a theory of consciousness. If consciousness is necessarily unified, then a correct theory of consciousness should at least be compatible with this unity, and we can hope that it will explain this unity.

We can see these three questions as clustering around the status of what we can call the unity thesis (UT):

Unity Thesis: Necessarily, any set of conscious states of a subject at a time is unified.

The first question raises the issue of how the notion of unity in the unity thesis is to be understood: what is it for a set of conscious states to be unified? The second question raises the issue of whether the unity thesis is true. The third question raises the issue of how, if the unity thesis is true, its truth might be explained.

In this paper we will address all three of these questions. Our central project will be to isolate a notion of unity on which the unity thesis is both substantive and plausible. That is, we aim to find a more precise version of the unity thesis that is neither trivially true nor obviously false. With such a thesis in hand, we will look at certain arguments that have been made *against* the unity of consciousness, to determine whether they are good arguments against the unity thesis as we understand it. And finally, after fleshing out the unity thesis further, we will apply the thesis to certain currently popular philosophical theories of consciousness, arguing that the thesis is incompatible with these theories: if the Unity Thesis is true, then these theories are false.

We will not aim to conclusively prove the unity thesis in this paper, and indeed we are not certain that it is true. But we aim to suggest at least that the thesis is plausible, that it captures a strong intuition about the nature of consciousness, and that there are no knockdown arguments against it. If the thesis is true, it is likely to have strong consequences for a theory of consciousness.

2 Varieties of Unity

To start with, we need to distinguish different notions of unity. In particular, we will distinguish various different ways in which different states of consciousness might be said to be unified with each other.

(1) Objectual unity. We can say that two states of consciousness are *objectually unified* when they are directed at the same object. For example, when I look at a red book, I have an experience of redness, and an experience of rectangularity. The color experience and the shape experience here are unified in a particularly strong way. They are present in my consciousness as directed at a single entity: the book. The same goes for my experience of a blue car moving down the street. Here I experience color, shape, and motion, all of which are unified by being directed at the same object. I might even have an auditory experience of the car's engine, and also experience this as directed at the same object. So there can be objectual unity across different sensory domains.

For two experiences to be objectually unified, their object need not actually exist. If I hallucinate a red book, then my experiences of redness and of rectangularity will be objectually unified, despite the book's nonexistence. On the other side of the coin, two experiences can be experiences of the same object without being objectually unified. I might see a car's shape and hear its noise, without anything in my conscious state tying the noise to the car (perhaps I perceive the noise as behind me, due to an odd environmental effect). If so, the experiences are not objectually unified. For objectual unity, what matters is that two states are experienced *as* being directed at a common object.

The notion of objectual unity is closely tied to a central issue in cognitive psychology and neurophysiology. When I look at a red square, the color and the shape may be represented in different parts of my visual system. But somehow these separate pieces of information are brought together so that I experience a single red square, so that I can identity and report a red square, and so on. This phenomenon is often referred to as *binding*, and the question of how it is achieved is often referred to as the *binding problem*. The binding problem is in large part the problem of how objectual unity is possible. As we will see, this divides into two problems in turn.

Objectual unity is an important phenomenon, but it will not be central for our purposes. Where objectual unity is concerned, the unity thesis is almost certainly false. While some sets of experiences are objectually unified with each other, it seems that most sets are not. For example, my experience of the color of the book and of the shape of the car are not objectually unified: they are experienced as being directed at *different* objects. My experiences of a bird singing and of a sharp pain do not seem to be directed at the same object at all. If so, then objectual unity cannot unify all of a subject's conscious states. For such a notion of unity, we must look elsewhere.

(2) Spatial unity. A related notion of unity is that of *spatial unity*. We can say that two conscious states are *spatially unified* when they represent objects as being part of the same space. For example, my experiences of a book and of a car are not directed at a common object, but they represent both objects as part of the same visual space. More generally, all of my visual experiences seem to be spatially unified in this way: every visual experience represents something spatially, and everything that is represented is represented as part of a common space. Auditory experiences usually represent objects as part of the same space; such auditory experiences are spatially unified with visual experiences.

The notion of being "represented as part of the same space" can be fleshed out in various ways, but the crucial idea will be something like this: a set of experiences are spatially unified if (i) each has 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. So visual experience might represent a car as being near a tree, or behind a truck, or to the left of a building. Auditory experience might represent exhaust noise as coming from the same area as the car, or it might represent a siren as being much further away. This sort of comparability is endemic to visual experience and to much perceptual experience, and makes for a deep spatial unity in perception.

Like objectual unity, however, spatial unity does not yield a plausible version of the unity thesis. Some experiences seem to have no spatial representational content at all. An emotional experience such as that of melancholy does not obviously represent anything as located within space. A conscious thought about philosophy might have no spatial content at all. If so, these conscious states are not spatially unified with other conscious states. As before, to find a notion of unity that unifies all of a subject's conscious states, we must look elsewhere.

(3) Subject unity. Let us say that two conscious states are *subject unified* when they are had by the same subject at the same time. So all of my current experiences — my perceptual experiences, my bodily sensations, my emotional experiences and conscious thoughts — are subject unified, simply because they are all *my* experiences.

If we construe the unity thesis as involving subject unity, it is certainly plausible. If a set of conscious states are had by a subject at a time, then they will be subject unified by definition. The trouble with this version of the unity thesis is that it is *trivial*. It is true by definition, and tells us nothing substantive about consciousness. As such, it cannot capture the intuition that there is some nontrivial way in which consciousness is unified. So subject unity will not be our central focus here.

Still, the notion of subject unity is at least useful in articulating the unity thesis. As it was characterized above, the unity thesis holds that if a set of experiences of a subject at a time is subject unified, then that set is unified. So in effect, the unity thesis states that subject unity entails unity. Now we simply need to find a notion of unity for which this entailment is both plausible and nontrivial.

(4) Subsumptive unity. We started the paper by invoking the intuition that there is some substantial sense in which *all* of a subject's experiences — including at least perceptual, bodily, emotional, and cognitive experiences — can all be unified. This sense is not object or spatial unity, since these notions do not apply to all of the relevant experiences. And this sense is not subject unity, since the resulting unity holds trivially. Rather, it is the sense that all of these experiences are somehow subsumed within a single state of consciousness.

We can say that two conscious states are *subsumptively unified* when they are both subsumed by a single state of consciousness. The notion of one state being subsumed by another should be taken as intuitive for now; we will spell it out more shortly. To take an example: it seems plausible that all of my visual experiences are subsumed by a single encompassing state of consciousness, corresponding to my visual field. More generally, my visual and auditory experiences might all be subsumed by a single encompassing state of perceptual consciousness. And it does not seem unreasonable to suppose that there is a single encompassing state of consciousness that subsumes all of my experiences: perceptual, bodily, emotional, cognitive, and any others.

We can think of this last encompassing state of consciousness, for a given subject, as the subject's *total* conscious state. When it exists, a subject's total conscious state might be thought of as the subject's conscious *field*. It can be thought of as involving at least a conjunction of each of many more specific conscious states: states of perceptual experience, bodily experience, emotional experience, and so on. But what is important, on the unity thesis, is that this total state is not *just* a conjunction of conscious states. It is also a conscious state in its own right. If such a total conscious state exists, it can serve as the "singularity behind the multiplicity" — the single state of consciousness in which all of a subject's states of consciousness are subsumed.

It is worth pointing out certain sorts of unity with which subsumptive unity should not be confused. We are not talking about *gestalt unity*, where the conscious experiences of two different objects are deeply related in a way that transforms each of the experiences and produces a "gestalt" experience with a novel content. And we are not talking about *normative* unity, which requires some special coherence or consistency among multiple contents of consciousness. As we have characterized subsumptive unity, two conscious states might be subsumptively unified whether or not their contents stand in a special gestalt relation to each other, and whether or not they are especially consistent or coherent with one another. We are also not dealing with *neurophysiological unity*, which requires that conscious states involve a single area or mechanism in the brain. Finally, we are not dealing with *diachronic unity*, or the unity of consciousness across time. It might turn out that one or more of these notions is deeply related to the issues at hand, but none of them is our primary object of discussion.

To spell out the notion of subsumptive unity in more detail, we need to go into more detail about just what consciousness involves, and just what is involved in the idea of one conscious state being subsumed by another. This requires making some further distinctions.

3 Access unity and Phenomenal unity

What is it for a mental state to be a conscious state? There is no single answer to this question. As many have pointed out, the notion of "consciousness" is ambiguous, and is understood in different ways by different people. So to make progress we have to make distinctions. For our purposes, the most useful distinction is Ned Block's distinction between access consciousness and phenomenal consciousness (Block 1995).

A mental state is *access conscious* when a subject has a certain sort of access to the content of the state. More precisely, a state is access-conscious if by virtue of having the state, the content of the state is available for verbal report, for rational inference, and for the deliberate control of behavior. When I look at a red book, I can report the presence of the book ("there's a red book"), I can reason about it (e.g., concluding that I must have put it there when reading yesterday), and I can use its presence in deliberately directing by behavior (e.g., picking up the book and putting it back on the shelf). So my perception of the red book gives me the relevant sort of access to information about the red book. So my perceptual state here is access-conscious. One can also say that in such a case, the subject is access-conscious of the relevant object. So here, I am access-conscious of the red book.

In a similar way, many of my perceptual states are access-conscious, and so are many of my emotional and cognitive states. Not all mental states are access-conscious, however. In some cases, such as those involving subliminal perception, blindsight, or unconscious belief, a mental state represents information without that information being reportable or usable in rational control of reasoning and behavior. The exact definition of access consciousness is somewhat flexible, and can be varied for different purposes. The most important point is that a state's being access-conscious is defined in terms of the causal role that the state plays within the cognitive system, and in particular in terms of the role that the state plays in making information available to other parts of the system.

A mental state is *phenomenally conscious* when there is something it is like to be in that state. When a state is phenomenally conscious, being in that state involves some sort of subjective experience. There is something it is like for me to see the red book — I have a visual experience of the book — so my perception of the book is phenomenally conscious. There is something it is like to hear the bird singing, and to feel the pain in my shoulder, so these states are phenomenally conscious. There is something it is like to feel melancholy, and there is arguably something it is like when I think about philosophy. If so, then these states are phenomenally conscious. Phenomenal consciousness is often taken to be the most important sort of consciousness, and to be the sort of consciousness that poses the most difficulty for scientific explanation.

There is a close empirical connection between phenomenal consciousness and access consciousness. It is arguable that the two almost always go together empirically: when a state is phenomenally conscious, it is access-conscious, and vice versa. That is, when there is something it is like to be in a state, a subject can usually report the contents of the state and use it to directly guide reasoning and behavior. And when a subject can report the contents of a state and use it to directly guide reasoning and behavior, there is usually something it is like to be in that state. So when I am phenomenally conscious of the red book, I am access-conscious of it, and vice versa.

Despite this empirical connection, there is plausibly a *conceptual* distinction between access consciousness and phenomenal consciousness. Access consciousness is defined in terms of the causal role that a state plays, whereas phenomenal consciousness is defined in terms of the way the state feels. It is arguable that we can at least *imagine* states that are access-conscious without corresponding states of phenomenal consciousness (the philosophers' zombie, which is functionally like a normal human being but without any conscious experience, would be one such imaginary case). And it seems that we can know about another being's states of access consciousness without knowing about their states of phenomenal consciousness: one might know what information is available for report and for behavioral control in a cognitive system without being in a position to know what it is like to be that system.

When there is something it is like to have a mental state, we can say that the mental state has a phenomenology, or a phenomenal character. Slightly more formally, we can say that such mental states have *phenomenal properties*, or *qualia*, which characterize what it is like to be in them. We can also say that *subjects* have phenomenal properties, characterizing aspects of what it is like to be a subject at a given time. We can then say that a phenomenal state is an instantiation of such a property. For example, the state of experiencing a certain sort of reddish quality is a phenomenal state.

When a subject is in a phenomenally conscious mental state, the subject will thereby be in a phenomenal state that reflects the phenomenology of being in the mental state. For example, if there is something it is like for a given subject to believe that Paris is in France, the subject will be in a corresponding phenomenal state. But the phenomenally conscious mental state and the phenomenal state may be distinct states. For example, it may be that there is now a certain phenomenology associated with my occurrent belief that P, while it is also possible for me to believe that P with a different phenomenology or no phenomenology at all. If so, the belief state (this instance of believing such-and-such) is a phenomenally conscious mental state, but it is not a phenomenal state. There is a special class of phenomenally conscious mental states such that the mental state and the corresponding phenomenal states are identical: phenomenal states themselves. Phenomenal states are at the core of phenomenal consciousness.

We can use the distinction between access consciousness and phenomenal consciousness to make a distinction between two corresponding notions of unity: access unity and phenomenal unity. Broadly speaking, two conscious states are *access-unified* when they are jointly accessible: that is, when the subject has access to the contents of both states at once. Two conscious states are *phenomenally unified* when they are jointly experienced: when there is something unified it is like to be in both states at once.

We can construct more precise versions of phenomenal and access unity by combining these distinctions with the distinctions outlined earlier between objectual unity, spatial unity, and field unity. These distinctions crossclassify each other, so that one can isolate notions of objectual phenomenal unity, objectual access unity, spatial access unity, subsmptive phenomenal unity, and so on. The distinction applies less clearly to the notion of subject unity, so we will set that notion aside here.

We can say that two conscious states are *objectually access-unified* when their contents involving attributing properties to a single object of representation, and when these contents are jointly accessible within the system. The contents will be jointly accessible when their *conjunction* is available for report and the rational control of reasoning and behavior. When I am conscious of a red square, I can report the presence of red, and the presence of a square, but I can also report the presence of a red square. Similarly, the presence of a red square can be used in guiding my reasoning and my behavior. So my perception of red and my perception of a square are not just individually access conscious: they are access-unified.

We can say that two conscious states are *objectually phenomenally unified* when they are experienced as representing a single object. When I am conscious of a red square, I experience the presence of red and I experience the presence of a square, but I also experience the presence of a red square. There is a distinctive sort of unity involved in what it is like to experience the redness and the squareness simultaneously here: the two states are unified by being experienced as aspects of a single object.

Objectual access unity and objectual phenomenal unity correspond to two distinct aspects of the binding problem. It has often been pointed out that there are actually two binding problems (see e.g. Revonsuo 1999). The first is that of how a system such as the brain manages to bring together two separately represented pieces of information (e.g., representations of color and shape in different areas of the visual cortex), so that these can play a joint role in the control of behavior (e.g. so that we can report the presence of a red square and a blue circle, rather than a red circle and a blue square). This is a sort of engineering problem concerning the design of the cognitive system; one can think of it as the neurophysiological or cognitive binding problem. This binding problem is the problem of explaining objectual access unity. The second binding problem is that of explaining how it is that we perceptually experience separate pieces of information as bound together in pertaining to the same object. This is the problem of explaining objectual phenomenal unity. On the face of it, these two problems are distinct: one could solve the neurophysiological binding problem, giving an explanation of how two pieces of information are brought together in the brain to be jointly accessible, while still having no explanation of why the jointly accessible information should be experienced. So objectual phenomenal unity and objectual access unity are at least conceptually distinct.

One can make a similar distinction between spatial phenomenal unity and spatial access unity. We can say that two conscious states are spatially access-unified when they have spatial representational contents that can be jointly accessed by the cognitive system, so that they can be spatially compared and so that the results of the comparison can be made available for report, reasoning, and behavioral control. For example, when I see a car and a tree, I do not just have access to their spatial locations individually; I also have access to the spatial locations jointly, in that I can report that the car is to the left of the tree. So these two perceptual states are spatially access-unified. Two conscious states are spatially phenomenally unified when they involve experiencing entities as part of the same space; as part of the same phenomenal space, one might say. I experience the car as being in the same space as the tree, and to the left of it; so these two states are spatially phenomenally unified.

The most important distinction is that between subsumptive phenomenal unity and subsumptive access unity. These notions apply to two arbitrary conscious states, as long as they are phenomenally conscious in the first instance and access-conscious in the second. Because these are the most important versions of unity, we will henceforth usually speak simply of "phenomenal unity" and "access unity", where it is understood that we are referring to phenomenal field unity and access field unity respectively.

We can say that two conscious states are *subsumptively access-unified* (or simply *access-unified*) if the *conjunction* of their contents is available for verbal report, reasoning, and the deliberate control of behavior. So if mental state A has content P and mental state B has content Q, these states will be individually access-conscious if the information that P is available for report and for control, and if the information that Q is available for report and control. They will be *jointly* access-conscious, or they will be access-unified, if the information that P∧Q is available for report and control. More briefly: two states A and B are access-unified if and only if the subject is access-conscious of the conjunction of their contents. In this case, there is an access-conscious mental state with the conjunctive content: this conjunctive mental state can be seen as subsuming the original states A and B.

For example, when I see a book and feel a pain, I can report the presence of the book and of the pain individually, but I can do more than that: I can report them simultaneously. I can also reason about the book and the pain jointly, and use information about both to jointly control my behavior (e.g. looking in the book for a remedy for the pain, or ceasing to read the book to help alleviate the pain). Because of the accessibility of this conjunctive content, the two states are (subsumptively) access-unified. Similarly, I can often jointly report or reason about an emotion and a sound: if so, the emotional state and the auditory state are access-unified. And so on.

It is worth noting that for a state to be access-conscious, it is not required that the content of the state actually be *accessed*, in the sense that it is directly used for report or for control. What matters is that it be *accessible*, in a certain direct sense, or "poised" for use in access, as Block puts it. The same goes for access unity. For two states to be access-unified, they need not be simultaneously accessed at any given moment. What matters is that they are simultaneously accessible, in that it would be possible for a subject to jointly report them, and to use them jointly in reasoning and behavior control. Typically, our conscious states are not jointly accessed, but they are much more often jointly accessible. It is joint accessibility that matters for our notion of unity.

We can use the notion of access unity to put forward a version of the unity thesis:

Access Unity Thesis: Necessarily, any set of access-conscious states of a subject at a time is access-unified.

This thesis appeals to the notion of a set of states being access-unified. This is a natural generalization of the notion of two states being access-unified. We can say that a set of states is access-unified if the contents of all of the states are jointly accessible.

It mght be objected that in requiring that *any* set of a subject's access-conscious states is access-unified, the thesis is highly implausible. A subject might have a large (possibly infinite) number of access-conscious states, and the conjunction of the contents of these states might be so complex that it is implausible that a subject could have access to this conjunction. The full conjunction would not be reportable or directly available to guide reasoning and behavior. To get around this, we could put forward a slightly weakened version of the thesis:

Pairwise Access Unity Thesis: Necessary, any two access-conscious states of a subject at a time are access-unified.

One might argue that the pairwise version is too weak to count as a full unity thesis (which requires unity of all states at a time), or that it suffers from the same problems as the full unity thesis (since it entails that conjunctions of conjunctions will be access-conscious, and so on). But none of this will matter for our purposes, since as we will argue shortly, even a weak version of the access unity thesis, limited to pairwise unity of relatively simple access-conscious states, is straightforwardly false.

We can say that two conscious states are *subsumptively phenomenally unified* (or simply *phenomenally unified*) if there is something it is like for a subject to be in both states simultaneously. That is, two states are phenomenally unified when they have a *conjoint phenomenology*: a phenomenology of having both states at once that subsumes the phenomenology of the individual states. When A and B are phenomenally conscious states, there is something it is like for a subject to have A, and there is something it is like for a subject to have B. When A and B are phenomenally unified, there is not just something it is like to have each state individually: there is something it is like to have A and B together. And the phenomenology of being in A and B together will carry with it the phenomenology of being in A and the phenomenology of being in B.

For example, when I look at the book while feeling a pain, there is something it is like to see the book (yielding a phenomenal state A), and there is something it is like to feel the pain (yielding a phenomenal state B). But there is more than this: there is something it is like to see the book while feeling the pain. Here there is a sort of conjoint phenomenology, that carries with it the phenomenology of seeing the book, and the phenomenology of feeling the pain. As in the discussion of field unity, we can think of the conjoint state here as involving at least the conjunction A&B of the original phenomenal states A and B. But importantly, the conjoint state is itself a phenomenal state: a single complex state of conscousness that subsumes the individual states of consciousness A and B. It is this encompassing state of consciousness that unifies A and B.

More generally, we can say that a *set* of conscious states is phenomenally unified if there is something it is like for a subject to have all the members of the set at once, and if this phenomenology subsumes the phenomenology of the individual states. As a special case, we can say that the set consisting of all of a subject's conscious states at a given time is phenomenally unified if there is something it is like for the subject to have all these states at once, where this phenomenology subsumes the phenomenology of the individual states. If so, then the subject has a *total phenomenal state* that encompasses all of the subject's phenomenal states. One can think of a total phenomenal state as capturing what it is like to be a subject at a time. If a subject has a total phenomenal state, there is a clear sense in which all of a subject's phenomenal states are unified within it.

We can put forward a phenomenal version of the unity thesis, as follows:

Phenomenal Unity Thesis: Necessarily, any set of phenomenal states of a subject at a time is phenomenally unified.

This is not quite the same as the unity thesis that falls out of our original definition. Such a thesis would hold that any set of *phenomenally conscious mental states* of a subject at a time will be phenomenally unified. But the two theses are clearly equivalent. The first version (regarding phenomenally conscious mental states) entails the second version (regarding phenomenal states) as a special case. In reverse, the second version entails that for any set of phenomenally conscious mental states, their associated phenomenal states will be phenomenally unified. So there will be a phenomenal state that subsumes each of the original phenomenal states. So there will be something it is like to be in all the original mental states simulaneously that subsumes what it is like to be in them individually. So the original mental states will be phenomenally unified, and the first version of the unity thesis holds.

One can also put forward slightly weaker versions of the phenomenal unity thesis:

Pairwise Phenomenal Unity Thesis: Necessarily, any two phenomenal states of a subject at a time are phenomenally unified.

Total Phenomenal Unity Thesis: Necessarily, the set of all phenomenal states of a subject at a time is phenomenally unified.

The original phenomenal unity thesis clearly entails the pairwise unity thesis and the total unity thesis. The pairwise thesis does not obviously entail the first version. It is plausible that subsumption is transitive, so that necessarily, if A subsumes B and B subsumes C, then A subsumes C. If so, the pairwise unity thesis will entail the phenomenal unity thesis for any finite set of phenomenal states, as any pair of these will be subsumed by a single phenomenal state, and any pair of those in terms will be subsumed by a single phenomenal state, and so on. But the pairwise thesis does not obviously entail the original thesis where infinite sets of phenomenal states are concerned. The total unity thesis entails the original phenomenal unity thesis, however: if there is a state that subsumes each phenomenal state of the subject, that state will also subsume each member of an arbitrary set of phenomenal states of the subject, so that set will be phenomenally unified. So the total unity thesis and the original phenomenal unity thesis are equivalent.

The total unity thesis arguably captures the central intuition behind the unity of consciousness. This thesis suggests that there is always a single phenomenal state that subsumes all of the phenomenal states of a subject at a time. That is, it suggests that any conscious subject at any time has a total phenomenal state. If a subject has a total phenomenal state, subsuming every specific phenomenal state of the subject, then the subject's consciousness will be unified in a deep way.

It might be objected that when a subject experiences a number of phenomenal states at once, the original phenomenal states will be transformed. For example, it might be phenomenally different to see a red book in the context of a moving car than to see a red book on its own, and the phenomenal state that was present when one saw the book on its own might not be present at all. This may be so, but it is no objection to the unity thesis. The unity thesis says that the phenomenal states had by a subject *at a time* are subsumed by a complex phenomenal state. So the experience of a red book and a moving car at a given time should subsume the experience of the red book at that time and the experience of the moving car at that time. It is not required that the complex experience should subsume the experience of a red book as the subject might have it at a different time, in a different context. If the experience of the book is itself transformed by the context of the car, then it is the transformed experience that will be subsumed by the complex state.

It might also be objected that these unity theses are *trivial*. If a subject has a set of phenomenal states, there will automatically be a phenomenal state that subsumes them: the conjunction of the original states. But this is not a trivial claim. It is trivial that if a subject is in a number of phenomenal states, the subject will be in the conjunction of those states. But it is nontrivial that this conjunction will itself be, or be subsumed by, a phenomenal state. That is, it is nontrivial that there will be *something it is like* to be in the conjunctive state. This can be seen from the fact that some philosophers deny the total unity thesis, or at least entertain its denial. For example, when Hurley (1998) discusses the possibility that the unity of consciousness could break down and that consciousness could be "partially unified" (so that two phenomenal states are each unified with a third state, but not with each other), she says:

> Therefore, we cannot imagine what it is like for there to be partial unity. That doesn't show partial unity is unintelligible, because being partially unified isn't the sort of thing there *could be* anything it is like to be. We shouldn't expect to be able to imagine what it is like.

In general, it seems that a case in which the unity of consciousness breaks down would be precisely a case in which there is no total phenomenal state of the subject: that is, there is nothing it is like to be the subject at that time, or at least there is no single something-it-is-like that captures all the phenomenal states of the subject. Such a subject would have states with a local phenomenal character, but there would be no global phenomenal character involved in having these states. It is certainly very hard to see how this could be the case. Indeed, one might suspect (as we do) that such a scenario is impossible and perhaps incoherent. But to say this is not to say that the unity thesis is trivial: it is a substantive thesis about the nature of consciousness. This is reflected by the fact that (as we discuss later in the paper) certain theories of consciousness entail that the unity thesis is false. If so, then the thesis puts substantive constraints on a theory of consciousness.

4 When Access Unity Breaks Down

The access unity thesis holds that necessarily, any two access-conscious states are access-unified. This entails that whenever a subject is access-conscious of P and is access-conscious of Q, the subject will be access-conscious of P∧Q. This thesis is clearly false.

To see that the thesis is false, we need only note that it is possible for a subject to be access-conscious of P and access-conscious of Q, without being access-conscious of P∧Q. For this to happen, it should be the case that P is reportable and available for guiding reasoning and behavior, and that Q is reportable and available for guiding reasoning and behavior, but that P∧Q is not reportable and not available for guiding reasoning and behavior. This can happen in a quite straightforward way. All that is required is that there be an *access bottleneck*. This will be a pathway of information access through which only a limited amount of information can pass at one time. If P and Q are both accessible only through the bottleneck, and if each carry an amount of information that is near the capacity of the bottleneck, then P and Q will be individually accessible, but the conjunctive content P∧Q will not.

This is not merely a hypothetical description of an imaginary case. Such access bottlenecks can occur in real cognitive systems, and are revealed by a number of experiments in the psychological literature. Perhaps the clearest example of such a bottleneck is given by a famous experiment by George Sperling (1960).

In Sperling's experiment, a subject is presented with a matrix consisting of three rows with four letters each. The matrix is flashed only briefly, for 250 milliseconds. After the matrix vanishes, a tone sounds, indicating whether the subject is to report the contents of the first, second, or third row. When subjects are required to report the contents of the top row, on average they correctly report 3.3 of the four letters in that row. The same goes when they are required to report the contents of the middle row, or of the bottom row. But when subjects are asked to report the contents of the entire matrix, on average they correctly report 4.5 of the twelve letters. So, to simplify a little, it seems that the subject has access to the information in any single row, but the subject does not have joint access to the information in all three rows.

In this case, it is natural to hold that the subject (just after the matrix disappears, before the tone sounds)is access-conscious of the contents of any individual row. Recalling that access consciousness requires accessibility for report and for reasoning and behavior: the contents of each row are available for report (individually), and could presumably be used to guide reasoning about those contents and to guide behavior. But it also seems that the subject is not access-conscious of the conjunctive contents of the whole matrix, or of any two rows. The conjunctive contents of more than one row are not available for verbal report, and presumably are not available to guide reasoning and behavior. If so, then a subject can be access-conscious of P (one row) and of Q (another row), without being access-conscious of P∧Q (both rows). So two access-conscious states of a subject at a time can fail to be access-unified, and the access unity thesis is false.

We do not claim that the Sperling experiment alone *proves* that the access unity thesis is false. There are other possible interpretations of the experiment: for example, one could hold that the subject has some sort of internal access to the conjunctive content, but that the process of report destroys this access. But the interpretation we have suggested is a natural one: on the face of it, the conjunctive content does not seem to be available for any sort of reasoning or control, although the individual contents are available taken one at a time. And importantly, whether or not this interpretation is correct of the actual case, it seems to be a perfectly coherent interpretation, one that describes a perfectly reasonable way for a cognitive system to function.

Indeed, given a natural design for cognitive systems with limited resources, we would expect certain restrictions on the flow of information in access and control, and we would expect access bottlenecks to arise in some cases. It may be that most of the time, when a subject has access to P and to Q, the subject has access to P∧Q. But this sort of joint access clearly cannot hold *necessarily*. So even if there is a reasonably high degree of access unity in ordinary conscious states, this sort of access unity cannot hold across the board.

This breakdown of access unity does not entail a breakdown of phenomenal unity. This can be seen by examining the Sperling case. It is difficult to know exactly what is going on in the phenomenology of the subject who is undergoing the Sperling experiment, before being asked about the contents of a row. Perhaps the details of all nine letters are present in the subject's phenomenology (as some subjects report); perhaps these details are not present, and there is merely an indeterminate patch in each cell of the matrix; or perhaps there is something in between. But whatever the exact phenomenology here, there is little reason to suppose that phenomenal unity breaks down.

No matter what it is like for a subject to experience each individual cell of the matrix in the Sperling case, it is plausible that there will be something it is like for the subject to see the entire matrix. And it is plausible that the phenomenology of seeing the matrix will subsume the phenomenology of seeing the individual cells. If the phenomenology of seeing a cell involves just a hazy patch, then the phenomenology of seeing the matrix will plausibly involve nine hazy patches. If the phenomenology of seeing a cell involves a detailed shape, then the phenomenology of seeing the matrix will plausibly involve nine detailed shapes. Either way, the individual phenomenal states are subsumed by the overall phenomenal state. So there is no reason to deny phenomenal unity here.

At most, this sort of case suggests that a subject does not always have simultaneous access to the contents of all of the subject's phenomenal states. If the subject is indeed experiencing the details of all nine letters, then the subject is in a position where the contents of these experiences can be accessed and reported only a few at a time, and not all at the same time. There is nothing paradoxical or contradictory about this. It simply suggests that a subject's access to a total phenomenal state is sometimes piecemeal. But this is just what we might expect.

One consequence of this is that access consciousness and phenomenal consciousness can come apart. We have seen that the subject is access-conscious of the individual letters but not of their conjunction. And it is natural to hold that either (i) the subject is phenomenally conscious of neither the individual letters nor their conjunctions, or (ii) the subject is phenomenally conscious of both the individual letters and their conjunction. In case (i), a subject is access-conscious of an individual letter but not phenomenally conscious of it. In case (ii), a subject is phenomenally conscious of the conjunction but not access-conscious of it. Either way, access consciousness and phenomenal consciousness of a given content can come apart. Our own view is that description (ii) is somewhat more plausible. If this is so, we can still hold that access consciousness and phenomenal consciousness are correlated with each other for *simple* contents. But access consciousness and phenomenal consciousness will not always be correlated for *complex* contents.

The moral of all this is that a breakdown of access unity does not entail a breakdown in phenomenal unity. There is a sense in which a breakdown of access unity is a "disunity" in consciousness, but it is a relatively shallow sense. Such a breakdown is quite compatible with an underlying phenomenal unity. Of course we have not demonstrated that no breakdowns of access unity involves a breakdown of phenomenal unity. But this discussion does strongly suggest that one cannot *infer* a breakdown of phenomenal unity from a breakdown in access unity. To accept a breakdown of phenomenal unity, one would need some quite distinct reason.

An opponent might try to argue that the Sperling case is a case where phenomenal unity breaks down. For example, the opponent might argue that the phenomenology of seeing each individual cell involves a detailed letter, but that the phenomenology of seeing the whole matrix does not, and that any "global" phenomenology here involves only hazy patches. Such a response would seem unmotivated and implausible, on the face of it, at least in the absence of much supporting argument. If the phenomenology of each letter is detailed, then there seems to be good reason to hold that this phenomenology is present in a global phenomenal state. And even if it is *coherent* for an opponent to hold this, it is equally coherent (and seemingly more plausible) to deny this, and to hold that the experience of the letters is phenomenally unified. The mere coherence of the denial is enough to show that one cannot *infer* a breakdown in phenomenal unity from a breakdown of access unity.

5 Can Phenomenal Unity Break Down?

We think that there is a strong prima facie case that the unity thesis is true. This prima facie case is brought out by the fact that there seems to be something *inconceivable* about phenomenal disunity. It is difficult or impossible to imagine a subject having two phenomenal states simultaneously, without there being a conjoint phenomenology for both states. And there is a sense that there is something incoherent about the suggestion. This prima facie inconceivability — whether it takes the form of unimaginability or apparent incoherence — gives at least some reason to believe that cases in which phenomenal unity break down are impossible, so that the unity thesis is true.

But this is only a prima facie case. There are some possible scenarios that humans cannot imagine, and there are arguably some possible scenarios that no being could imagine. And the judgment of incoherence in this case is not so strong that it could not be incorrect. So the prima facie case for the unity thesis needs to be balanced with the case *against* the unity thesis. A number of philosophers and scientists have argued that the unity of consciousness can break down. So to assess the unity thesis, one needs to examine these arguments in order to see what force they have against the unity thesis as we have understood it.

By far the most common reason for holding that the unity of consciousness can break down is grounded in neuropsychology. It is widely held that patients in various unusual neuropsychological states have a disunified consciousness. The paradigm case here is that of a *split-brain* patient, whose corpus callosum has been severed for medical purposes, preventing the left and right hemispheres of the cerebral cortex from communicating directly (although there is still some connection through lower areas of the brain). Such a patient behaves in a surprisingly normal fashion much of the time, but in certain circumstances they behave quite unusually. For example, when presented with different pictures in different halves of their visual field (e.g., a cat on the left and a dog on the right), and asked to report the contents, the patient will report seeing only a dog, since the left hemisphere, which dominates speech, receives input from the right visual field. When asked to write down what they see with their left hand (which is controlled by the right hemisphere), such a patient may slowly write "CAT"; with the right hand, the patient may write "DOG". If a patient writes with her left hand in her right visual field, a conflict may occur when the patient sees what is written, and in some cases the right hand scratches out what the left hand has written.

It is often held that in cases like this, consciousness is disunified. On one interpretation (e.g. Puccetti 1981), there are two distinct subjects of consciousness, corresponding to each hemisphere. Such an interpretation is actually compatible with the unity thesis, since the unity thesis requires only that every subject have a unified consciousness. More threatening to the unity thesis are intepretations on which there is a single subject with a disunified consciousness. Some (e.g. Marks 1980) hold that the subject has two separate streams of consciousness, at least under experimental conditions. Others (e.g. Lockwood 1989) hold that the subject has a fragmented consciousness with nontransitive unity between the states: for example, the experiences of "CAT" and of "DOG" might each be unified with some background emotional state, but not with each other. Others (e.g. Nagel 1971) hold that our conceptual framework in speaking of subjects may simply break down in this area.

Adjudicating this question requires a very detailed examination of both the empirical details and the philosophical analysis of these phenomena, which we cannot provide here. Here, we will simply note that given what we have said so far in this paper, the advocate of a phenomenal unity thesis has a natural line of response.

It is plausible that in split-brain cases, there is some sort of breakdown of *access* unity. If we assume that there is a single subject, then it seems that the subject in the case above has at least a weak sort of access both to the presence of a cat and to the presence of a dog, and can use each in reasoning and in the control of behavior. But it seems that the subject has no access to a conjunctive content involving both the cat and the dog. The conjunctive content is not reported, and plays no apparent role in reasoning and in the control of behavior. So this may well be a case in which access unity fails. In this case, it seems that two accessed contents are not jointly accessible, because of a disconnection between the relevant access mechanisms.

But as we have seen, a breakdown of access unity does not entail a breakdown of phenomenal unity. So the possibility remains open that split-brain subjects have a unified phenomenal field, with some sort of conjoint phenomenology subsuming each of the separate contents. It is just that the subject has pathologies of access, so that the contents of the field are accessible only singly and not jointly. If so, the subject in the experiment described has a phenomenal field that includes experiences of both "CAT" and "DOG". The subject simply has no conjoint access to these contents. Of course this implies that the subject has highly imperfect knowledge of her conscious states: she will believe (in both "halves" of the brain) that she is experiencing only one word, when in fact she is experiencing two. But it is plausible for many other reasons that knowledge of consciousness is fallible, and it is not unreasonable to suppose that in cases of brain damage, this fallibility might be quite striking.

Of course nothing here proves that this interpretation is correct. It does suggest, however, that we should not be too quick to conclude that these cases involve a breakdown of phenomenal unity. Most of those who have discussed these cases have not carefully distinguished the relevant notions of unity and consciousness (an exception is Marcel 1994, who distinguishes "reflexive consciousness" from "phenomenal experience" and argues that the disunity concerns the former), and have often discussed things in terms of access and related functional notions. Once we distinguish access unity from phenomenal unity, it becomes clear that the direct evidence concerns access disunity, not phenomenal disunity. To establish phenomenal disunity requires substantial further argument. It may be that such arguments can be given, but the case is far from clear.

One might say something similar about other disorders of consciousness, such as dissociative identity disorder (multiple personality disorder). In this case, it seems that there are pathologies of access between different parts of a cognitive system. But it seems quite tenable to hold that nevertheless, there is a single field of consciousness at any given time, subsuming the conscious states of the subject, even if they are in certain respects mutually inaccessible. Of course as in the split-brain case, the subject may well have various false beliefs about her own consciousness (e.g. that the various states belong to different subjects), but again this is not unexpected. Indeed, there are other cases involving "alien thoughts" in schizophrenic patients where it seems most plausible to hold that a subject incorrectly attributes some of her conscious states to a subject other than herself.

To completely assess this thesis requires much further analysis. But for now, we conclude that the empirical case against the phenomenal unity thesis is at best inconclusive. Given the strong prima facie positive case for accepting the phenomenal unity thesis, this suggests that the unity thesis remains quite plausible.

6 Formalizing the unity thesis

6.1 Subsumption and entailment

For further analysis, we need to clarify the phenomenal unity thesis, and the corresponding notion of phenomenal unity. We have said that a set of states is phenomenally unified when there is something it is like to be in all those states at once. When this is the case, the subject will have a phenomenal state (corresponding to the conjoint what-it-is-like) that *subsumes* each of the states in the original set. So phenomenal unity can be seen as a sort of subsumptive unity, and the phenomenal unity thesis on the table is a sort of subsumptive unity thesis.

> **Subsumptive Unity Thesis**: For any set of phenomenal states of a subject at a time, the subject has a phenomenal state that subsumes each of the states in that set.

There are also closely related total and pairwise subsumptive unity theses, requiring subsumptive unity only for pairs of phenomenal states or only for the complete set of a subject's phenomenal states at a time, but we can focus on the thesis above for now.

As it stands, the notion of subsumption is something of an intuitive primitive. There are some things we can say about it. It is a relation among token phenomenal states. It is plausibly reflexive (a state subsumes itself), antisymmetric (if A subsumes B and B subsumes A, then A = B), and transitive (if A subsumes B and B subsumes C, then A subsumes C). Note that reflexivity eliminates any apparent problem of regress in the unity thesis (if A and B are subsumed by C, there is no need for a further state to subsume A and C, since C subsumes itself).

The paradigm case of subsumption is the relation between a complex phenomenal state and a simpler state that is intuitively one of its "components". One might think of subsumption as analogous to a sort of mereological part/whole relation among phenomenal states, although this should be taken as an aid to intuition rather than as a serious ontological proposal, at least at this point. It is also useful to stipulate that subsumption holds between a phenomenal state and less specific states that intuitively correspond to the same experience: for example, that the state of experiencing a sharp pain subsumes the corresponding state of experiencing a pain. This sort of subsumption is required in order for it to be possible that a highly specific total phenomenal state can subsume all of a subject's phenomenal states, including unspecific states.

It should be noted that there are alternatives to analyzing phenomenal unity in terms of subsumption. Often, phenomenal unity is analyzed in terms of an intuitive relation of co-consciousness, where this relation is taken as primitive. We think that the analysis in terms of subsumption runs deeper in certain respects than a primitive analysis in terms of co-consciousness, and offers the promise of further analytic tools, as discussed below. But the exact relation between these notions is an open question. (Dainton (2000) gives a thorough and insightful analysis of the unity of consciousness in terms of a primitive co-consciousness relation; Bayne (2001) discusses the relationship between the different accounts.)

The notion of subsumption is connected to the notion of "what it is like" in at least the following sense: when A subsumes B, what it is like to have B is an aspect of what it is like to have A. Of course this appeals to the unexplained notion of an "aspect". One might try to go further by *defining* subsumption wholly in terms of notion of "what it is like" as follows: A phenomenal state A subsumes phenomenal state B when what it is like to have A and B simultaneously is the same as what it is like to have A. This seems to capture the connection articulated above, and it also can ground the connection between subsumptive unity and the original definition of phenomenal unity. If there is something it is like to be in a set of states (as the original definition requires), then this phenomenology will correspond to a phenomenal state A of the subject, and it is clear that this state will subsume the states in the original set in the sense defined above. It is arguable that the defined notion if subsumption goes beyond the intuitive notion in certain respects (someone might hold that the "what it is like" locution can be read such that what it is like to have A and B differs from what it is like to have A, even when A subsumes B), and we will not rely on it in what follows, but nevertheless it can serve as a useful aid to the understanding.

(Extending this line of thought, one could say that a state A *precisely subsumes* a set of states S what what it is like to be in A is the same as what it is like to simultaneously be in the members of S. Then if A precisely subsumes S, A subsumes each of the members of S, but the reverse entailment does not hold. For example, a subject's total state of consciousness subsumes each of the subject's visual experiences, but it does not precisely subsume the set of them. One could then articulate a *Correspondence Thesis* holding that for any set of phenomenal states of a subject at a time, there is a corresponding phenomenal state that precisely subsumes that set. The correspondence thesis is formally stronger than the original subsumptive unity thesis: the existence of a total phenomenal state suffices for the truth of the original thesis, but it does not suffice for the truth of the correspondence thesis. The correspondence thesis nevertheless has some intuitive plausibility, and one could argue that this thesis, rather than the subsumptive unity thesis, best captures the idea articulated in the original phenomenal unity thesis. The difference between these theses will not be important for our purposes, however.)

There is a close relation between subsumption and *entailment*. Let us say that a state P entails a state Q when it is impossible (logically or metaphysically impossible) for a subject to instantiate P without instantiating Q. Then it seems clear that when a phenomenal state P subsumes a phenomenal state Q, P will entail Q. For example, if P involves the phenomenal character as of seeing a red book and hearing a bird singing, and if Q involves the phenomenal character as of seeing a red book, then it is impossible to have P without having Q. The same goes with any case of subsumption: by its nature, the subsuming state carries with it the subsumed state.

Note that strictly speaking, entailment is a relation among state-types, while subsumption is a relation among state-tokens. For present purposes, we can regard entailment as derivatively a relation among state-tokens, so that one state-token entails another when there is entailment between the corresponding state-types (although see below). We will generally pass over this nicety in discussion, acknowledging it where it is relevant. Note also that a phenomenal state A entails a phenomenal state B if necessarily, a subject in A is also in B — *not* if the content of A entails the content of B.

The close relation between entailment and subsumption raises an interesting possibility: perhaps we can simply *define* subsumption in terms of entailment? That is, perhaps we can hold that phenomenal state A subsumes B when A entails B? If this were possible, instead of relying on a novel primitive relation, we could analyze unity in terms of a well-understood relation that allows the use of standard logical tools. To help assess this possibility, we can define a corresponding notion of unity and a corresponding unity thesis:

> A set of phenomenal states of a subject at a time is *logically unified* when the subject has a phenomenal state that entails each of the phenomenal states in that set.

> **Logical Unity Thesis**: For any set of phenomenal states of a subject at a time, the subject has a phenomenal state that entails each of the states in the set.

This gives an attractively simple formulation of the unity thesis, and one that has some intuitive force. Unfortunately there is an obstacle to replacing subsumption by entailment. We know that when A subsumes B, A entails B. But the reverse is not obviously the case. In fact, there are two ways in which it may seem that A could entail B without subsuming B.

First, A and B might correspond to intuitively distinct experiences that share a type. For example, a subject might have two pains at the same time, or two experiences of red, and so will have two distinct phenomenal states of the same type. In this case, one state-type will entail the other, so if entailment among tokens is derivative on entailment among types, one state-token will entail the other. In this case, it is not plausible to hold that one state subsumes the other. (What it is like to have A and B simultaneously is quite different from what it is like to have A.) One might instead refine the definition of entailment among state-tokens, requiring that it is impossible for one token to exist without the other, in addition to the requirement that one type cannot exist without the other. But one can also deal with this case by a strategy discussed below.

Second, A and B could be intuitively distinct phenomenal states that do not share any simple type, but that are nevertheless necessarily connected. This would involve a sort of *gestalt unity*, which involves constraints on the co-occurrence of distinct phenomenal states. For example, perhaps there are cases where feeling a pain in one's shoulder while also experiencing a splitting headache produces a unique sort of pain that could not be experienced in the absence of the headache. Or perhaps seeing a certain person in the middle of a crowd produces a unique sort of visual experience of that person that could not be had in the absence of the experience of the crowd. Or perhaps (to use an example from Dainton 2000) the experience of the boundaries of a Kanisza triangle is of a special sort that could not be had in the absence of the circles in which the triangle is embedded. In this sort of case, we can say that the pain is gestalt unified with the headache; the experience of the person is gestalt unified with the experience of the rest of the crowd; and the experience of the boundaries is gestalt unified with the experience of the circles.

Whether there are really any cases of gestalt unity is arguable. One could argue that in the above cases, it would be possible to experience the pain without the headache, or have the visual experience of the person without that of the crowd, or to have that of the boundaries without that of the circles, perhaps in some very different context. But it is not implausible that at least some experiences put *some* constraints on concurrent experiences, and that one cannot mix and match experiences arbitrarily. If this is so, then there is at least a weak sort of gestalt unity, since the presence of one phenomenal state puts constraints on the nature of concurrent phenomenal states. In this case one can even say that the presence of one phenomenal state entails the existence of another phenomenal state, where the second is understood as an instantiation of a sufficiently unspecific phenomenal property.

If there is gestalt unity, then there will be cases in which one phenomenal state entails another phenomenal state without the first subsuming the second, at least in any intuitive sense. For example, the experience of the boundary of a Kanisza triangle might entail something about the experience of the nearby objects, but the experience of the nearby objects does not intuitively subsume that of the boundary of the triangle. Similarly, the experience of the shoulder pain might entail the experience of the headache, but it does not intuitively subsume the experience of the headache. (Intuitively, what it is like to have the pain and the headache goes beyond what it is like to have the pain, even if the former is entailed by the latter.)

So: If there is gestalt unity, then subsumption cannot be understood in terms of entailment. But this does not mean that we must give up on the logical unity thesis. Even if subsumption cannot be understood in terms of entailment, one can make a case that the logical unity thesis entails the subsumptive unity thesis.

To see this, we can first note that not *all* phenomenal states are gestalt unified. Even if some pairs of phenomenal states are gestalt unified, it seems very unlikely that all pairs are, and it seems much more plausible that most pairs are not. Given a typical pair of phenomenal states had by a subject such that neither subsumes the other, it usually seems to be straightforwardly possible that a subject could have an instance of the first state without the second. When I see the red book and hear the bird singing, there seems to be no good reason to deny that I could have a visually identical experience without hearing the bird singing, and so on. (Dainton 2000 gives a more extended argument for the conclusion that gestalt unity is not universal and is in fact rare.)

If there can be pairs of states that are not gestalt unified, it also seems that there can be subjects none of whose states are gestalt unified. One simply needs a subject all of whose basic phenomenal states are independent in the above way: each of them could occur without any of the others. There seems to be no obstacle in principle to such a subject, and one could even argue that our own phenomenal states are often like this. Let us say that such a subject is gestalt-free. In gestalt-free subjects, the gestalt cases of entailment without subsumption will not arise. So (setting aside for a moment any other cases of entailment without subsumption), we can say that if the logical unity thesis holds, the subsumptive unity thesis holds at least when restricted to gestalt-free subjects.

Now let us assume that the subsumptive unity thesis holds for gestalt-free subjects: for any set of phenomenal states of a gestalt-free subject, there is a subsuming phenomenal state. If so, it is very plausible that the subsumptive unity thesis holds for all subjects. If there is always a subsuming state in gestalt-free cases, there will plausibly always be a subsuming state in gestalt cases. There is nothing about gestalt unity that makes the existence of a subsuming state in such cases *less* likely. If anything, the situation is the reverse. In a case of gestalt unity, the experiences will be connected in such a way that the existence of a subsuming state will be more likely, not less. So if there are cases in which gestalt unified states are not phenomenally unified, there should equally be cases in which gestalt-free states are not phenomenally unified. So the subsumptive unity thesis for gestalt-free subjects plausibly entails the subsumptive unity thesis for all subjects.

We are close to establishing a connection between the logical unity thesis and the subsumptive unity thesis in general. But we still need to deal with the other case of entailment without subsumption discussed above, in which a subject has distinct simultaneous experiences that share a type. We can deal with this in an analogous way. Let us say that a subject has *duplicate* experiences when the subject has two intuitively distinct experiences that share a maximally specific phenomenal type (two pains or two color experiences with exactly the same quality, say). It is not entirely obvious that duplicate experiences are possible; but in any case, let us say that a *duplicate-free* subject is a subject without duplicate experiences. It is plausible that *if* the subsumptive unity thesis is true when restricted to duplicate-free subjects, it is true also of subjects with duplication: if it is possible for duplicate experiences not to be subsumed by a common experience, it will be equally possible for non-duplicate experiences not to be so subsumed. As with gestalt phenomena, there is nothing about duplication per se that contributes to a breakdown of phenomenal unity. So the subsumptive unity thesis for duplication-free subjects plausibly entails the subsumptive unity thesis for all subjects.

Combining the last two cases, we can say that the subsumptive unity thesis restricted to gestalt-free, duplication-free subjects plausibly entails the subsumptive unity thesis for all subjects. But it is also clear that the logical unity thesis entails the subsumptive unity thesis for gestalt-free, duplication-free subjects. In such subjects, a phenomenal state T that entails all phenomenal states will also subsume all phenomenal states, since we have removed the relevant gaps between subsumption and entailment. One might worry that one gap remains: by ruling out duplication, we have ruled out the possibility of entailment without subsumption for maximally specific phenomenal states, but two sufficiently nonspecific states of the same type might entail each other without subsuming each other. Nevertheless, since T entails maximally specific versions of each of these nonspecific states, T will subsume these maximally specific states, and so T will subsume the nonspecific states. So T subsumes all the subject's phenomenal states.

We have established that the logical unity thesis entails the subsumptive unity thesis for gestalt-free subjects, duplication-free subjects; and we have established that the latter thesis plausibly entails the subsumptive unity thesis for all subjects. So the logical unity thesis plausibly entails the subsumptive unity thesis. In reverse, the subsumptive unity thesis clearly entails the logical unity thesis. So it is plausible that the subsumptive unity thesis holds if and only if the logical unity thesis holds. The only obstacle to this equivalence will arise if there are breakdowns of phenomenal unity that are solely due to gestalt unity or to duplication, but there seems to be little reason to take that possibility seriously.

If this is correct, we can assess the truth of the subsumptive unity thesis by assessing the truth of the logical unity thesis. This latter tasks is in some respects more straightforward, since we no longer have to deal directly with the primitive notion of subsumption. This also allows the possibiliy of using familiar logical tools to formulate and assess versions of the unity thesis. We will look more closely at some versions of the thesis in the following section.

6.2 Logical unity and conjunctive closure

There are three versions of the subsumptive unity thesis: the pairwise version, the general version, and the total version. There are correspondingly three versions of the logical unity thesis, holding either that there is logical unity among either any two states of a subject at a time, any set of states, or the complete set of states. Or more directly:

> **Pairwise Logical Unity Thesis**: Necessarily, for any two phenomenal states had by a subject at a time, the subject has a phenomenal state that entails both original states.

> **General Logical Unity Thesis**: Necessarily, for any set of phenomenal states of a subject at a time, the subject has a phenomenal state that entails each state in the set.

> **Total Logical Unity Thesis**: Necessarily, for any conscious subject at a time, the subject has a phenomenal state T such that for any phenomenal state A of the subject at that time, T entails A.

As before, it is clear that the general thesis entails the pairwise thesis and the total thesis as special cases. The total thesis also entails the general thesis and the pairwise thesis, since a state that entails all phenomenal states of a subject will also entail any pair or any set of states. Arguably the pairwise thesis does not entail the other two theses, because of the formal possibility that there might be entailing states for any finite set of states, but not for infinite subsets.

We can start by focusing on the total logical unity thesis, since this corresponds most closely to the total phenomenal unity thesis, which arguably captures the central intuition behind the unity of consciousness. Intuitively, we can think of T, the entailing state in the thesis, as the subject's total phenomenal state, capturing what it is like to be the subject at that time. If such a state exists, it will fulfill the requirement of the total logical unity thesis.

One can also approach the matter in logical terms. Let us say that the *conjunction* of a set of states is a state C such that necessary, a subject is in C if and only the subject is in each of the states in that set. (Like entailment, conjunction is fundamentally a relation among state-types, and derivatively a relation among state-tokens. Note also that the conjunction of states is quite different from conjunction of the *contents* of states.) This identifies C at least up to mutual entailment. For present purposes, it is useful to assume that when two states A and B mutually entail each other (i.e. when necessarily, a subject is in A if and only if the subject is in B), then the two states are identical. If so, then C is identified uniquely. Nothing that follows rests essentially on this assumption — one could rephrase things in terms of equivalence classes of states — but this makes the discussion easier.

We can then propose a natural candidate for T: the conjunction C of all of a subject's phenomenal states at a time. It is clear that if T exists, T entails C (since T entails each of the conjuncts of C). And it is clear that if T exists, C entails T (since T is itself a phenomenal state). So if T exists, then T is identical to C (and C is therefore a phenomenal state), by the criterion for state identity above. It is also clear that if C is a phenomenal state, then C will satisfy the total logical unity thesis with T=C. We can therefore say that an appropriate T exists if and only if C is a phenomenal state.

Let us say that a set of states is *conjunctively unified* when the conjunction of the members of that set is itself a phenomenal state. Then from the discussion above, it follows that the total logical unity thesis is equivalent to the claim that the set of a subject's phenomenal states is conjunctively unified:

> **Total Conjunctive Unity Thesis**: If C is the conjunction of all of a subject's phenomenal states at a time, then C is itself a phenomenal state.

As before, someone might think that a thesis of this sort is trivially true, but this would be incorrect. It is trivial that for any set of phenomenal states of a subject at a time, there will be a conjunctive state C that entails each of the original states. But it is nontrivial that C will itself be a phenomenal state. That is, it is nontrivial (although very plausible) that there will be something it is like to be in C: some global phenomenal character that a subject will have if and only if the subject is in C. Those who deny the original unity thesis will deny the existence of such a phenomenal character, and so will deny that C is itself a phenomenal state.

In effect, we have seen that the original phenomenal unity thesis is equivalent to a thesis about the *conjunctive closure* of co-instantiated phenomenal states (where co-instantiated states are states had by the same subject at the same time): certain conjunctions of states in this class must also be states in this class. This is very useful, since conjunctive closure is amenable to relatively straightforward analysis.

One can also formulate conjunctive closure theses that are closely related to the other versions of the logical unity thesis. There is a pairwise version, and a general version:

> **Pairwise Conjunctive Unity Thesis**: For any two phenomenal states of a subject at a time, their conjunction is a phenomenal state.

> **General Conjunctive Unity Thesis**: For any set of phenomenal states of a subject at a time, their conjunction is a phenomenal state.

These theses are not quite formally equivalent to the corresponding versions of the logical unity thesis. To see this, note that it is at least a formal possibility that two states might be logically unified but not conjunctively unified. For example, it is at least formally possible that the conjunction of *all* of a subject's phenomenal states might be a phenomenal state, but that the conjunctions of certain pairs and subsets might not be. If so, then these pairs and subsets will be logically unified but not conjunctively unified. In this case the pairwise and general conjunctive unity theses will be false, but the the pairwise and general logical unity theses will be true.

However, it is clear that these conjunctive unity theses *entail* the corresponding versions of the logical unity thesis. And they are interesting and plausible theses in their own right. The first says that for any two phenomenal states A and B of a subject at a time, there will be something distinctive it is like to be in A and B: that is, a distinctive conjoint phenomenal character that a subject will have if and only if the subject is in both A and B. The second says the same thing for arbitrary sets of co-instantiated phenomenal states. These theses are not formally trivial, but they are highly plausible theses about phenomenal consciousness. (These theses are closely related to the correspondence thesis discussed in the previous section.)

All three theses are simple and elegant. The pairwise conjunctive unity thesis says that the class of phenomenal states is closed under pairwise co-instantiated conjunction: the conjunction of two co-instantiated phenomenal states is a phenomenal state. The general conjunctive unity thesis says that the class of phenomenal states is closed under general co-instantiated conjunction: the conjunction of any set of co-instantiated phenomenal states is a phenomenal state. And the total conjunctive unity thesis says that the class of phenomenal states is closed under maximal co-instantiated conjunction: the conjunction of a maximal set of co-instantiated phenomenal states is a phenomenal state.

The total conjunctive unity thesis remains the core version of the unity thesis, but all of these theses are plausible and useful. Each of them can be used as a tool in assessing the status of the unity of consciousness, in assessing its consequences, and in assessing its compatibility with various theories of consciousness.

6.3 Hurley, Shoemaker, and what it is like

It should be noted all of these unity theses are stated simply in terms of the notions of phenomenal state, of co-instantiation, and of conjunction. And the notion of a phenomenal state is tied constitutively to the notion of there being something it is like to be a given subject, or to be in a given state. So we have an account of unity that requires little more than the existing "what it is like" conception of phenomenal states.

This stands in tension with a claim in a very interesting analysis by Hurley (1998; this volume). Hurley (1998, pp. 165-66) argues that the unity of consciousness cannot be characterized "subjectively", and that suppositions about the structure of consciousness are not captured by the "what it is like" test, so that we need to appeal to further "objective" properties to give an account of unity. This claim is grounded in the claim that in a case where unity breaks down, there is no "what-it-is-like" that captures the structure of a subject's consciousness. Hurley backs up this claim by considering two cases: (i) two subjects, one experiencing red and hot, the other experiencing red and dizzy; and (ii) a partially unified single subject, in whom red and hot are unified, red and dizzy are unified, but hot and dizzy are not. Hurley argues that no "what it is like" facts can distinguish these two cases.

But from the claim that there is no what-it-is-like that characterizes a disunified subject, it does not follow that one cannot characterize unity in what-it-is-like terms. Indeed, following Hurley's own claim, one can hold that unity breaks down precisely when there is nothing it is like to have all a subject's conscious states simultaneously. We can distinguish case (i) from case (ii) above by noting that in case (ii), both subjects have a phenomenal state that subsumes all their phenomenal states, whereas in case (i), the subject has no such phenomenal state. Of course, our characterization of unity appeals to something more than phenomenal states themselves: it appeals to subsumption, and to co-instantiation in a subject. Perhaps Hurley would count these notions as in some sense "objective". There is no point arguing over terminology here, but we can at least note that subsumption is a phenomenal relation, fixed by phenomenology alone: if A subsumes B, then the phenomenology of A guarantees that it subsumes B. And subjects are simply the bearers of phenomenal states. So we are staying quite close to home in characterizing unity this way.

Hurley might extend her argument by suggesting case (iii): a bifurcated subject with two different (but indistinguishable) tokens of red in separate streams. In this subject, red1 is unified with hot, red2 is unified with dizzy, and no state in either pair is unified with a state in the other pair (Hurley 1998, p. 166, seems to point toward such a case). If (iii) is possible, one could argue that it could not be distinguished from (ii) by talk of subjects and their phenomenal states alone. We would need to appeal to the identity of phenomenal states: a single "red" experience is involved in both complex experiences in (ii), but not in (iii).

There are a number of things one could say in response. One might concede that "what it is like" talk cannot distinguish the two different cases of disunity (ii) and (iii), but hold that it can nevertheless distinguish unity from disunity, which is the most important work we need it to do. If the unity thesis is true, then cases of disunity will be impossible, and distinctions among impossible cases will not matter to characterizing the structure of consciousness. More deeply, one can suggest that Hurley's argument shows at best that one cannot distinguish the cases in terms of the distibution of phenomenal state-*types*. If we appeal to facts about the distibution of phenomenal state-*tokens*, things are straightforward: there is a token experience that is subsumed by two different complex experiences in (ii), but not in (iii). It may be that (ii) and (iii) will be introspectively indistinguishable, so that the structure of consciousness is not *transparent* to a subject. But nevertheless, a characterization of the structure of consciousness in terms of phenomenal relations among phenomenal state-tokens is still, in a deep sense, a characterization in subjective terms.

Our characterization of unity in phenomenal terms also stands in tension with a claim by Shoemaker (this volume). Shoemaker suggests that if a conscious state is understood as one with a phenomenal property (i.e. one such that there is something it is like to be in it), this leads to "consciousness atomism": the view that the factors that make a state conscious are independent of the factors that make two states unified. Our discussion here suggests that this is false. What it is for two conscious states to be unified can be understood in terms of the existence of a more complex conscious state, where both the simple state and the complex states are states characterized by what it is like to be in them. So the factors that enter into unifying conscious states are the same sort of factors as those that enter into those states being conscious in the first place.

At one point, Shoemaker characterizes "consciousness atomism" differently, as the view that "whether a state is conscious will be independent of whether there are other conscious states with which it is co-conscious". The account here is neutral on this claim. For all we have said here, it may be possible for there to be a subject with a single conscious state. This claim does not seem to us to be obviously objectionable, and it is compatible with the more important view that the factors that enter into conscious states are the same as those that enter into co-consciousness.

In fact, the definitions of unity that we have given here suggest that any account of what it is to be a phenomenal state will automatically yield a theory of what it is for two such states to be unified. We need simply to apply the theory to the relevant conjunctive states, in order to determine whether they are phenomenal states. In this way, it seems that any substantive theory of phenomenal consciousness can yield unified definitions of consciousness and of co-consciousness. It is precisely because of this that the unity thesis (if it is true) puts strong constraints on a theory of phenomenal consciousness, as we will see.

7 Applications of the Unity Thesis

We have already mentioned the objection that the conjunctive versions of the unity thesis are trivial: that is, that it is trivial that the the conjunction of a set of co-instantiated phenomenal states is itself a phenomenal state. It is clear that the thesis is not *formally* trivial, in that there are many classes of states that are not closed under co-instantiated conjunction: e.g., states of the sort "talking with X", where X is an individual. Closer to home: there are also many classes of *mental* states that are not closed under co-instantiated conjunction.

For example, the class of *belief* states does not seem to be closed under conjunction. Let us say that a belief state is the state of believing some proposition. Then it is not the case that the conjunction of any set of belief states is a belief state. For example, if A is the state of believing that P, and B is the state of believing that Q, there is plausibly no belief state that a subject will be in precisely when they are in A and B. The only tenable candidate for such a belief state is the state of believing P and Q. But there are well known reasons to believe that a subject can believe that P and believe that Q without believing the conjunction P∧Q. For example, P and Q might be believed in different "compartments" of a compartmentalized mind. It may even be that for some P, a subject can believe that P and separately believe that ~P, without believing the contradiction P∧~P. And it seems quite possible that a subject can have many different beliefs without accepting the massive conjunction of the contents of all of those beliefs. If this is right, then the conjunction of co-instantiated belief states will not in general be a belief state. So the class of belief states is not closed under co-instantiated conjunction.

It may seem plausible or even obvious that the class of *phenomenal* states is closed under conjunction. But if so, this is a substantive thesis about the class of phenomenal states, and their difference with other classes of mental states. It may even be a *conceptual* truth, in some sense, that the class of phenomenal states is closed under co-instantiated conjunction. But if so, this is again a substantive thesis about the *concept* of a phenomenal state, and a way in which it differs from the concept of a belief state, and of other sorts of states.

The substantive nature of the thesis is revealed by the fact that the conjunctive unity theses put a strong constraint on potential theories of consciousness. We have seen that the unity thesis is *prima facie* plausible, and there there seem to be no strong arguments against it. If this is right, then the unity thesis puts a *prima facie* constraint on theories of consciousness: they must be compatible with the unity thesis. And in particular, any account of phenomenal states must be compatible with the total conjunctive unity thesis. Whatever phenomenal states are, according to a given account, the class of phenomenal states must be closed under total co-instantiated conjunction. A number of prominent theories of consciousness appear to be incompatible with this constraint.

One example is the higher-order thought theory of consciousness, put forward by David Rosenthal (1997) and others. Not all higher-order thought theorists intend the theory as an account of phenomenal consciousness (e.g. Lycan 2000 explicitly rejects the idea), but we are only concerned with versions of the theory that are aimed at phenomenal consciousness. The central idea of these theories is the following:

> **Higher-Order Thought Thesis**: A mental state M is phenomenally conscious if and only if a subject has a higher-order thought about M.

Here, a higher-order thought about M should be understood as a thought by the subject with the content "I am in M". The thesis will usually be modified and qualified in some ways. For example, Rosenthal holds that for M to be conscious, the higher-order thought must be brought about in the right sort of way, and in particular must be a noninferential thought. Rosenthal also holds that only sensory states can be phenomenally conscious, so that we would have to insert a rider to that effect in the definition above. This is arguably a mere terminological difference, however, since Rosenthal holds that there will be something it is like to be in a state whenever it is the object of the right sort of higher-order thought, whether the state is sensory or not. In any case, for our purposes we will take the thesis in the simple form above. Our arguments should apply straightforwardly to most modified versions.

Is the higher-order thought thesis compatible with the unity thesis? It is easiest to approach this question by considering the conjunctive versions of the unity thesis. The conjunctive versions say that the class of phenomenal states are closed under conjunction. So we can ask: on the higher-order thought theory, is the class of phenomenal states closed under conjunction?

We can start by thinking about phenomenally conscious mental states. If A and B are phenomenally conscious mental states, is A&B necessarily a phenomenally conscious mental state? Assuming the higher-order thought thesis, this translates into the following: if a subject has a higher-order thought about A and a higher-order thought about B, does the subject necessarily have a higher-order thought about A&B? That is, if the subject has a thought "I am in A" and a thought "I am in B", does it follow necessarily that the subject has a thought "I am in A and B"?

It seems not. It is surely possible for a subject to think "I am in A" and "I am in B", without connecting these into a thought "I am in A and B". We can take a case like those discussed above, in which a subject has contradictory beliefs, knows that she has each belief, but never puts the two together. She might have the thought "I believe P" and the thought "I believe ~P" without ever putting these two together into a thought "I believe both P and ~P". This might be strange or unusual, but there is nothing contradictory about it. There would only be something contradictory here if the beliefs of a subject are necessarily closed under logical consequence; but of course no subject's beliefs are closed under logical consequence.

The same is even clearer where total conjunctivity is concerned. On the higher-order thought theory, if a subject has a number of phenomenally conscious mental states, is their conjunction a phenomenally conscous mental state? That is, if a subject is has mental states A_1, ..., A_n, and has the thoughts "I am in A_1",..., "I am in A_n", does the subject necessarily have the thought "I am in A_1&A_2&...A_n"? Again, it seems not. One might reasonably argue that this entailment does not even hold *typically*, let alone necessarily. That is, it is arguable a *typical* subject with these higher-order beliefs would not have the complex conjunctive belief. Whatever one says here, it is hard to dispute that it is *possible* for a subject to have the individual higher-order beliefs without the complex conjunctive belief.

So it appears that if the higher-order thought view is true, the class of phenomenally conscious mental states is not closed under co-instantiated conjunction. This already contradicts the central intuition behind the unity thesis: that necessarily, if there is something it is like to be in each of a set of states, there is something it is like to be in all the states at once. On the higher-order thought view, this thesis will clearly be false.

The official version of the unity thesis is stated in terms of phenomenal states, not phenomenally conscious mental states. The analysis of phenomenal states is slightly tricker, since advocates of the higher-order thought view have not usually talked about phenomenal states and phenomenal properties directly. But given that higher-order thought theorists hold that there is something it is like to be in a mental state when the subject has a higher-order thought about it, they presumably hold that what it is like to be in that state is determined by the content of the higher-order thought. If so, it seems that phenomenal properties will be the properties of having higher-order thoughts with certain contents, and phenomenal states will be the states of having such higher-order thoughts.

Do phenomenal states, understood this way, satisfy the unity thesis? It seems not, for much the same reason as before. Here it is useful to take the entailment version of the unity thesis: that necessarily, when a subject has a set of phenomenal states, the subject has a phenomenal state that entails each of the individual states. When a subject has a set of higher-order thoughts H1, ..., Hn, does the subject necessarily have a higher-order thought HH such that being in HH entails being in H1, ..., Hn? It seems not, for the usual reasons. A subject might think "I am in A" and "I am in B", without any higher-order thought (e.g. "I am in A&B") such that having that thought entails having the original thoughts.

The problem is not that the higher-order thought theory provides no way to understand phenomenal unity. It can do so in a natural way. Two phenomenally conscious mental states A and B are unified when the subject has a higher-order thought about them not just singly but jointly. And two phenomenal states, the states of having higher-order thoughts "I am in A" and "I am in B", are phenomenally unified when there is a complex phenomenal state that entails them: that is, if there is a complex higher-order thought such that having the complex thought entails having the specific thoughts. This requirement will arguably be satisfied when the subject has a complex higher-order thought such as "I am in A&B".

The problem is rather that on this account, there is no reason to believe that phenomenal states, or phenomenally conscious mental states, will always be unified. Certainly it will not be necessary that they be unified, and it seems plausible that in a typical case, they will not be unified. So the higher-order thought thesis is incompatible with the unity thesis. It is clearly incompatible with the conjunctive and logical versions of the unity thesis. It is therefore also incompatible with the subsumptive versions, since any failure of logical unity automatically entails a failure of subsumptive unity. So if the higher-order thought thesis is true, the unity thesis is false. And if the unity thesis is true, the higher-order thought thesis is false.

Proponents of the higher-order thought thesis might reply in a number of ways. Most straightforwardly, they might reply by denying the unity thesis. This is a tenable response, since the truth of the unity thesis cannot be taken for granted. But still, there is a strong intuition that the unity thesis is true, so the incompatibility is at least a cost of the higher-order thought thesis. Proponents might also embrace a more limited version of the unity thesis, arguing for example that unity holds typically but not necessarily, or that it holds given contingent facts of human psychology, but not for all possible beings. Here there would still be the cost of denying the intuition of necessary unity, and there would be the added difficulty of defending the claim that unity holds in the relevant range of cases, when there seems to be no obvious reason why complex conjunctive thoughts about all the objects of our higher-order thoughts should typically exist.

A higher-order thought theorist might also respond by finding fault with the argument for incompatibility: they might hold, for example, that it is necessary that the class of mental states that are objects of higher-order thoughts is closed under conjunction. This would be a difficult case to make, in face of the apparent possibility of failure of this principle, and in face of the general phenomenon that beliefs are not closed under logical consequence.

Finally, a proponent might modify the higher-order thought thesis to make it compatible with the unity thesis. To do so, they must modify the definition of a phenomenally conscious mental states. It could be held, for example, that a mental state is phenomenally conscious when either (i) it is the object of a higher-order thought, or (ii) it is the conjunction of states that are the objects of higher-order thoughts. This sort of disjunctive account would be contrary to both the letter and the spirit of existing higher-order thought views (which hold that a conscious state is one that the subject is conscious of). One could also raise questions about whether this thesis delivers any *substantive* unity of consciousness, or merely a stipulated sort of unity of consciousness that holds trivially. And so far as the unity of consciousness seems to be a substantive fact about consciousness, one could argue that this modified version of the higher-order thesis does not really account for it.

Of course all of this is debatable and could lead to fruitful further discussion. But the prima facie incompatibility between the two theses is at least interesting. It is worth noting that the incompatibility extends straightforwardly to other "higher-order" views of consciousness, including views on which a conscious state is an object of a higher-order perceptual state, or the object of some other sort of higher-order representational state. The existence of a set of higher-order perceptual states does not entail the existence of a complex conjunctive higher-order perceptual state, and the same goes for other sorts of representational states. So if the unity thesis is true, these theses are false, and vice versa.

The unity thesis is also incompatible with many *representationalist* views of consciousness. According to representationalist views (e.g. Dretske 1995, Tye 1995), all phenomenally conscious mental states are representational states (that is, states with representational content). This is commonly allied with a further functional criterion to yield:

> **Representationalist Thesis**: a mental state is phenomenally conscious if and only if it is a representational state that plays an appropriate functional role.

We will focus on this broadly functionalist variety of representionalism. The details of the relevant functional role different between representationalists, but it is typically held to involve some sort of access and control. One can then say that what it is like to be in a mental state is determined by the content of the representional state, on the condition that it plays the relevant functional role. On this sort of view, then, a phenomenal state is a state of having a certain sort of representational state play the appropriate functional role, where distinct phenomenal states are individuated by distinct representational contents.

Two phenomenal states P1 and P2 are conjunctively unified when there is a phenomenal state P that entails each of the original states. On the representationalist account, two phenomenal states P1 and P2, corresponding to representational states A1 and A2 (with contents C1 and C2) playing the relevant functional role, will be conjunctively unified when there is a phenomenal state P, corresponding to representational state A (with content C) playing the relevant functional role, such that P entails P1 and P2. This will occur if and only if the existence of A playing the role entails the existence of both A1 and A2 playing the role. The only reasonable way to satisfy this is for the content of A to entail the content of A1 and the content of A2: that is, for C to entails both C1 and C2, or for C to entail C1&C2. For example, if A1 has content "red to the left", and A2 has content "green to the right", P1 and P2 will be conjunctively unified if there is a state A (playing the role) whose content entails "red to the left and green to the right". So two phenomenal states, corresponding to two representational states, will be conjunctively unified if and only if there is a conjunctive representational state (playing the appropriate role) whose content entails the conjunction of the contents of the original representational states.

The unity thesis is true if and only if necessarily, every set of phenomenal states is conjunctively unified. On the representationalist view, is this the case? It seems not. It seems at least possible to have a state with content C1 and a state with content C2, each playing a certain role, without having a state with content C1&C2 that plays the role. We saw this earlier in the case where the relevant role involves accessibility: it is possible that C1 is accessible and C2 is accessible with C1&C2 being accessible. Something similar will hold for any functional role involving access and control. If this is so, then representationalist thesis in the relevant class are incompatible with the unity thesis.

As before, representationalists could respond in a number of ways. They could deny the unity thesis, at cost of denying a strong intuition. They could modify it to apply to a more limited range of cases, at cost of some intuition and perhaps some empirical constraint. (For example, in the Sperling case, this representationalist may have to deny that the subject has a phenomenally unified visual field.) They could modify the representationalist thesis to allow a disjunctive definition on which is stipulated that conjunctions of phenomenal states are phenomenal states, at cost of endangering the substantive status of the unity thesis. Alternatively, they could move to a different sort of representationalism which is not so closely tied to functionalism: for example, it might be held that phenomenally conscious states are representational states whose content is represented *phenomenally*, or that they are representational states with some other property that is not functionally defined. The resulting version of representationalism might be compatible with the unity thesis (as well as being independently more plausible than the previous versions), at cost of giving up the reductive aspirations of many representationalist views.

One might also argue that other nonrepresentationalist forms of functionalism are incompatible with the unity thesis, on the general grounds that there will not be the relevant conjunctive property among states playing the functional role. The details will depend on the details of the functionalist theory, and in particular on the account that is given of phenomenal states and properties. These accounts can vary between functionalist theories, and are often not clearly articulated, so it is difficult to give a general analysis of such theories with respect to the unity thesis. But it is clear that it will be at least highly nontrivial for a functionalist account to satisfy the unity thesis.

If what has gone above is correct, then the unity thesis is incompatible with higher-order thought (and other higher-order representation) views of consciousness, with many representationalist views of consciousness, and with many functionalist views of consciousness. So the unity thesis is clearly nontrivial. Nevertheless, it has strong independent plausibility as a thesis about phenomenal states. So the incompatibility of the unity thesis with these views of consciousness should be seen as at least a *prima facie* argument against these views.

8 Explaining the Unity Thesis

If the unity thesis is true, how is its truth to be explained? We do not know the answer to this question. But in this concluding section, we will explore some possibilities.

One common strategy is to try to explain unity in functional terms. For example, one might try to explain unity in terms of some sort of informational integration, or in terms of serial processing in the brain, or something along those lines. One obvious problem with this sort of strategy is that it is not clear why this sort of functioning should yield phenomenal unity, as opposed to something like access unity. But an equally deep problem is that for reasons similar to those discussed above, it seems inevitable that this sort of functioning will be present *contingently*, and that it will be possible for conscious states to exist that do not stand in the relevant functional relations. If so, unity (on these analyses) will obtain only contingently, and the unity thesis will be false. If unity is to obtain necessarily, as the unity thesis suggests, we must look elsewhere.

Much of the reason for accepting the truth of the unity thesis comes from the fact that its denial seems to be inconceivable, and perhaps incoherent. This suggests that the unity thesis may be at some level a *conceptual* truth, although perhaps a deep conceptual truth, whose roots are revealed only by a deep analysis of our concepts. The central concepts involved in the unity thesis are that of a phenomenal state and that of a subject, along with various additional notions such as subsumption, entailment, conjunction, and so on. So one might hope that some light could be shed by attention to the concept of a subject, or by attention to the concept of consciousness.

One natural suggestion is that our concept of a subject of experience is somehow premised on unity. For example, one could suggest that ascriptions of subject-hood require as a precondition that subjects correspond to unified phenomenal fields. In the spirit of a sort of bundle theory of the subject, one could argue that we have a prior notion of a phenomenal field, and that we then associate subjects with phenomenal fields. If this is the case, we would expect that every subject would have a unified consciousness. A subject with two distinct phenomenal fields, for example, would be ruled out as a conceptual impossibility: where there are two phenomenal fields, there will automatically be two subjects.

How might this work? Our articulation of the notion of a phenomenal field in this paper appeals to subjects and co-instantiation, but one might argue that these can be bypassed. For example, one might appeal to a primitive relation of subsumption (or of co-consciousness) among phenomenal states that makes no presuppositions about subjects of those states, and then define a phenomenal field as a maximal phenomenal state: a phenomenal state that is not subsumed by any other phenomenal state. But even if something like this works, there is a deeper problem. This strategy might explain why distinct phenomenal fields correspond to distinct subjects, but it cannot explain why states of consciousness come packaged into unified phenomenal fields in the first place. For example, nothing in this strategy explains why a phenomenal state cannot be subsumed by two different phenomenal states such that no further phenomenal state subsumes both of these in turn. More generally, nothing here explains why the subsumption relation does not hold in quite unsystematic and fragmented manner. It is possible that an analysis of subsumption itself could do some work: for example, one could argue that subsumption is conceptually akin to a mereological part-whole relation, and so must hold reflexively, antisymmetrically, and transitively, and perhaps in a way that allows no overlap. But this conceptual stipulation does not really make the problem go away. It simply raises the question of why conscious states come packaged as parts and wholes.

One might then take a different approach. Instead of focusing on the concept of a subject, one can focus on the concept of consciousness itself. It could be argued that our *basic* concept of consciousness is not the notion of a simple phenomenal state — what it is like to such-and-such at a time. Rather, our basic notion of consciousness is that of a total phenomenal state: what it is like to be a subject at a time. This yields a holistic rather than an atomistic view of consciousness. On this approach, we do not start with basic atomic states of consciousness, and somehow glue them together into complex states. Rather, we start with a basic *total* state of consciousness, and then differentiate it into simpler states, and ultimately into atomic states.

If this were truly our basic notion of consciousness, then it might explain why the unity thesis is true. On this view, any non-total phenomenal state is *derivative* on a total phenomenal state that subsmes it. On this view, it is to be expected that any phenomenal states of a subject at a time are all simply aspects of what it is like to be that subject at that time. As such, it is to be expected that for any set of co-instantiated phenomenal states, there will be a subsuming state. On this view, the most basic problem with the theories of consciousness discussed in the last section is that they are atomistic rather than holistic, starting with simple states rather than total states. If this view is right, then any such analysis of consciousness will be a misanalysis from the start.

It is not obvious that this sort of conceptual claim on its own yields a substantive unity thesis. But one might tie naturally tie this analysis to a corresponding view of the metaphysics of consciousness. In nature, it may be that the most basic sort of conscious state is the total phenomenal state, or the phenomenal field, or even the phenomenal world. These total states are basic, but they are not featureless: they come with a complex structure from which one can differentiate many aspects. (As an analogy, one can think of a quantum wavefunction, which is a basic state in physics but which nevertheless has a complex structure.) So metaphysically, simple conscious states might be derivative on total conscious states. If so, we would have a clean explanation of why a substantive unity thesis is true.

This sort of suggestion is highly speculative, and much needs to be worked out. For example, it is far from obvious that our basic concept of consciousness is that of a total state of consciousness, and one needs to make a direct case for this. And the corresponding metaphysics needs to be worked out in much more depth. But there is at least some plausibility in the idea that the concept of concsiousness, and states of consciousness, are fundamentally holistic rather than atomistic. And this squares well with our intuition that consciousness is necessarily unified.

In any case, whether the substantive claims that we have made in this paper are correct or incorrect, we hope to have helped to pin down some of the crucial issues. It is clear that there is much need for further work in analyzing the notion of unity, in assessing the truth of the unity thesis, and in seeking an explanation of its truth. It is likely that such work will be philosophically fruitful.

Acknowledgments

Versions of this paper have been presented at the ASSC conference on the unity of consciousness in Brussels, at a Pew conference on the metaphysics of human beings in Syracuse, at a conference on neurophenomenology in Boulder, and in talks at Alabama, Arizona, Barcelona, University College London, and Wake Forest. We thank audience members on all those occasions for helpful comments. Special thanks to Barry Dainton, Bernie Kobes, Graham Macdonald, and Paul Studtmann.

References

Bayne, T. 2001. Co-consciousness: Review of Barry Dainton's *Stream of Consciousness*. Journal of Consciousness Studies 8:79-92.

Block, N. 1995. On a confusion about a function of consciousness. Behaviora and Brain Sciences 18:227-47.

Dainton, B. 2000. *Stream of Consciousness: Unity and continuity in conscious experience*. Routledge.

Dennett, D.C. 1992. The self as a center of narrative gravity. In (F. Kessel, P. Cole, & D. Johnson, eds), *Self and Consciousness: Multiple Perspectives*. Hillsdale, NJ: Erlbaum, 1992.

Dretske, F. 1995. *Naturalizing the Mind*. MIT Press.

Hurley, S. 1998. Unity, neuropsychology, and action. In *Consciousness in Action*. Harvard University Press.

Hurley, S. 2002. Action, the unity of consciousness, and vehicle externalism. In (A. Cleeremans, ed) *The Unity of Consciousness: Binding, Integration, Dissociation*. Oxford University Press.

Lockwood, M. 1989. *Mind, Brain, and the Quantum*. Blackwell.

Lycan, W. 2000. Representational theories of consciousness. In (E. Zalta, ed) *The Stanford Encyclopedia of Philosophy*.

Marcel, A.J. 1993. Slippage in the unity of consciousness.

Marcel, A.J. 1994. What is relevant to the unity of consciousness? In (C. Peacocke, ed) *Objectivity, Simulation, and the Unity of Consciousness*. Oxford University Press.

Marks, C. 1980. *Commissurotomy, Consciousness, and Unity of Mind*. MIT Press.

Nagel, T. 1971. Brain bisection and the unity of consciousness. Synthese 22:396-413. Reprinted in *Mortal Questions* (Cambridge University Press, 1979).

Puccetti, R. 1981. The case for mental duality: Evidence from split-brain data and other considerations. Behavioral and Brain Sciences 4:93-123.

Revonsuo, A. 1999. Binding and the phenomenal unity of consciousness, Consciousness and Cognition 8: 173-185.

Rosenthal, D.M. 1997. A theory of consciousness. In (N. Block, O. Flanagan, and G. Güzeldere, eds) *The Nature of Consciousness*. MIT Press.

Shoemaker, S. 2002. Consciousness and co-consciousness. In (A. Cleeremans, ed) *The Unity of Consciousness: Binding, Integration, Dissociation*. Oxford University Press.

Sperling, G. 1960. The information available in brief visual presentations. *Psychological Monographs* 498:1-29.

Tye, M. 1995. *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. MIT Press.