



Combining Minds: How to Think about Composite Subjectivity

Luke Roelofs

Print publication date: 2019

Print ISBN-13: 9780190859053

Published to Oxford Scholarship Online: February 2019

DOI: 10.1093/oso/9780190859053.001.0001

Conscious Subjects, Conscious Unity, and Five Arguments for Anti-Combination

Luke Roelofs

DOI:10.1093/oso/9780190859053.003.0002

Abstract and Keywords

This chapter looks at five arguments that have been advanced to show that minds cannot combine (under the heading of the “combination problem for panpsychism”) and considers the options for addressing them. They are the subject-summing argument, the unity argument, the privacy argument, the boundary argument, and the incompatible contexts argument. All of these arguments, under scrutiny, turn out to rest on assumptions either about the metaphysics of subjects of experience or about the unity of consciousness, so this chapter contains some in-depth examination of these two topics. For both topics, there is room for a range of plausible but conflicting views, and so the chapter outlines a plan to sketch three different theories of mental combination, starting from different assumptions about subjects and unity.

Keywords: subject of experience, personal identity, unity of consciousness, combination problem, metaphysics

USUALLY WHEN SOMEONE explicitly discusses the idea I’ve called “composite subjectivity,” it’s to say that it’s demonstrably impossible—that “subjects cannot combine” (Coleman 2012, 1), that “experiences don’t sum” (Goff 2006, 53), that “planning for tomorrow or feeling pain are not activities that a lot of simples can perform collectively” (Van Inwagen 1990, 118).

This is held to be something that applies to all forms of composite subjectivity, or at least to all that are of any theoretical interest, and it is held to be knowable *a priori*, by reason alone. To defend the possibility of composite subjectivity, I need

to engage and refute the arguments advanced for this extreme denial—the arguments alleging an internal problem with combinationism. So in this chapter I identify five arguments against the possibility of composite subjectivity; in section 2.3 they will be discussed in detail, but they can be briefly outlined as follows:

The subject-summing argument: It seems possible for a whole made of conscious parts not to be conscious itself, not to be a further subject of experience. So how does consciousness in the parts help to explain consciousness in the whole?

(p.35) The unity argument: Each of us has many experiences, but they seem to all be connected, “unified” in various ways. Experiences spread across many subjects, by contrast, seem separate, “disunified.” How can experiences spread across many parts provide the whole with unified consciousness?

The privacy argument: It seems to be an important fact about experiences that each of us can have only our own—others can be aware of ours only indirectly. But then if each experience is “private,” available exclusively to a single subject, how can the part’s experiences (which are private to those parts) do anything for the whole?

The boundary argument: If minds can combine, what stops my mind and your mind, or even everybody’s minds, from combining into some sort of giant mind which would absorb and do away with each of us as individuals? What, in general, stops component minds from “vanishing” into the whole they form?

The incompatible contexts argument: Experiences seem to get much of their character from context, from the other experiences they are unified with. But then any experience shared by part and whole must reflect both the relatively limited context of “what else the part is experiencing” and its larger context of “what else the whole is experiencing”; how can it reflect both?

These arguments are closely related: they all deal either with the metaphysics of the subject or with the unity of consciousness. That is, they all say that composite subjectivity is ruled out either by something about the way that experiences “belong to” certain beings or by something about the way that experiences “hang together.” But both of these—conscious subjects and conscious unity—are multifarious notions, often given conflicting definitions and prompting conflicting intuitions. I believe that part of the force of the above arguments comes from this multifariousness: responses that work for one way of conceiving unity, or subjecthood, come across as missing the point on a different conception. Combinationists need multiple responses, reflecting the multiple ways that unity and subjecthood can be conceived, which is why this book contains outlines of three different combinationist theories. Each theory will

have to address the five arguments above, though they will do so in different ways. So before evaluating the arguments themselves, and the options for responding to them, I need to examine these two connected notions: the “I” that undergoes experiences and the “unity” that ties all my experiences together.

(p.36) 2.1. What Is a Subject of Experience?

We can define a subject of experience as a being which it is like something to be. If I ask what it is like to be this dog, or this chair, or something else, I am thereby presupposing that the dog, chair, or whatever is a subject. This definition does not tell us much, though it tells us that certain things are subjects (unless we are all very deeply mistaken), such as me, my friends, the people I walk past in the street, etc. And it tells us that whatever subjects are, they are the things we mean when we use personal pronouns, standard human names, and so on. But this does not provide a deep understanding of what subjects are.

Unfortunately, all available philosophical accounts of subjecthood are highly controversial. Indeed, there has been far-reaching disagreement on this topic for centuries. It would be cumbersome in the extreme to discuss every plausible option that is available, but committing to one in particular would inappropriately narrow the interest of my subsequent discussion. Moreover, I am inclined to think (following Rovane 1998, 11, 33ff.; cf. Schechter 2009, 2012a, 583) that no single analysis fully captures our intuitive notion of a subject. However, I do think that two rough clusters of accounts can be distinguished, and that we can be relatively confident that the right account is either somewhere within one cluster or the other, somewhere intermediate between them, or some combination of them. Thus by considering combinationism in terms of both conceptions of subjects, I can perhaps ensure that my arguments will be relevant to almost all reasonable views.

2.1.1. Two Conceptions of Subjects—Metaphysical and Psychological

I will call these two clusters the “psychological” and the “metaphysical” conceptions of subjects. On the psychological conception, a subject is some sort of integrated psychological structure—a self-image, a point of view, a stable personality, an autobiographical narrative, an ongoing stream of causes and effects, or something of that sort. What individuates a subject is usually their distinctive psychological traits, like their episodic memories, their cognitive capacities, or their values and life projects. Crucially, on this conception we have to identify the states of a subject—experiences, beliefs, memories, or whatever—prior to identifying the subject themselves. They are individuated, even constituted, by those states.

By contrast, on the “metaphysical” conception, a subject is whatever being “underlies” those states and their continuities—an immaterial soul, a brain, a

(p.37) human being, or some other kind of enduring concrete substance.¹

Rather than starting from the experiences and constructing something to “have” them, we ask “What entities does the world contain?” and try to identify the one “in which” certain experiences are going on. We may still attend to complex psychological structures, but we look for the entity whose nature and operations *in fact explain* them—whereas on the psychological conception we define subjects to *logically guarantee* them. Even if, like Descartes, we decide that no detectable physical object can be the right thing to have experiences, we are still employing the metaphysical conception if our response is to *add something*, some purely mental substance, to our worldview. For on the psychological conception subjects would be nothing in addition to their experiences, no further thing to be added.

The primary examples of the psychological conception of subjects would include all of what are broadly called “Neo-Lockean” theories of personal identity (e.g., Shoemaker 1970, 1997; Parfit 1984; Noonan 2003; Lewis 1976; Rovane 1998), who trace their intellectual ancestry to John Locke’s (1836) “Of Identity and Diversity.” There Locke argues that we have the same person at two times wherever consciousness at one time can “reach back” to consciousness at the other; interpreters disagree as to whether this “reaching back” relation is just episodic memory, but Locke’s crucial point is that since this relation can connect states of consciousness in different substances (whether brains, souls, or something else), a “person” is not the same thing as any substance. Later writers have often changed the formulation, but the core idea (discussed in more detail in chapter 7) remains: persons are individuated by psychological relations (often grouped under the umbrella term “psychological continuity”) that have no essential connection to the substances that underlie thought. Since it is nearly platitudinous that “persons” are conscious subjects, we are clearly dealing here with a psychological analysis of subjects.

One need not be a Neo-Lockean, however, to hold a psychological analysis of subjects. One might be a functionalist, holding that a subject is individuated by the pattern of causal relations that hold among mental states (Shoemaker 1997; Schechter 2015). One might, like Hume (1888, 252), regard the subject as a “bundle,” collection, or stream of experiences; one might, like Schechtman (2009), regard it as being, or as constituted by, a series of narratives that are told to oneself; one might, like Nedelsky (2012), regard it as constituted by some kind of social process **(p.38)** of care or enculturation; or one might, like Dennett (1992) and Bayne (2010, 289), regard it as a sort of fictional character or intentional object, defined by the representations that posit it.² Or one might have no specific theory of the subject, and simply be inclined to think that what really defines someone is something about how exactly their mind *works*.

The metaphysical conception of a subject is less unified, since adherents will differ over what a subject is in proportion as they differ over what kind of substance underlies thought and experience. Many contemporary philosophers

focus on *biological* substances, in particular human organisms (e.g., Van Inwagen 1990, 142ff.; Olson 1997), and they can with only mild fear of anachronism associate themselves with the views of Aristotle. Others, particularly the earliest critics of Lockean theories, posit nonphysical souls and take subjects of experience to be these (Descartes 1985; Butler 1860; Swinburne 1984). Others say that it is the human body (Williams 1957; Wiggins 1967), or the human brain (Nagel 1986, 39–41), or whatever exact set of physical mechanisms continuously subserves basic experiential capacities (Unger 1990, 139–169). Some philosophers might even identify the subject with the world itself, understood as a single being (Śāṅkara 1994; Schopenhauer 1969). But what unites these views is the idea that the subject of experience is the thing which lies beneath particular experiences, which persists while particular experiences arise and cease, which remains the same even when very different experiences belong to it.³

I will make a sequence of claims about these two conceptions of subjecthood. First, both connect to something important in our intuitions about personal survival. Second, both connect to important epistemic interests in our thinking about the world. Third, both connect to important practical interests that the notion of a subject is bound up with. As a result, it would be inappropriate to dismiss either as entirely without merit or significance. And while there might be ways to partially combine or reconcile them, I will in this book remain neutral, talking about subjects according to both the psychological and metaphysical **(p.39)** conceptions without committing to either definitely being the thing we mean to pick out with personal pronouns, human names, and terms like “subject,” “self,” and “person.”

2.1.2. Intuitive Foundations of the Two Conceptions

The best way to bring out the resonance of both conceptions with our intuitions is to consider Williams’s 1970 paper “The Self and the Future.” In this paper, Williams describes the same sequence of events in two ways, generating opposite intuitive reactions. The first description is as follows: two people (“A” and “B”) are sent into a futuristic device, which exhaustively scans and records the wiring of their two brains, and then reconfigures each brain in accordance with the model provided by the other. When they come out, they are “switched”: there is a person with A’s old body, reporting B’s old memories, self-identifying with B’s name, and in every respect acting like B, and there is another person with B’s old body, reporting A’s old memories, self-identifying with A’s name, and in every respect acting like A. As these two people meet each other, the most natural thing for them to say is “Wow! You have my old body, and I’m in your old body. We have switched bodies—I, A (B), now inhabit B’s (A’s) body.” And it is very tempting, from a third-person point of view, to agree with them—to think that two people have switched bodies. That is just what the psychological conception predicts: all the interesting psychological relations now hold between the present person in one body and the past person in the other body. Yet since

no substance has passed between them, the metaphysical conception must disagree.

The second description of these same events, however, makes it seem as though the opposite has happened: two people have retained their bodies but had their psychologies radically altered. This description focuses on the forward-looking perspective of the people involved: before entering the device, each is told in detail about what will be done to them. First, they will have their memories wiped away; then, they will have a series of false memories implanted; then, they will have their temperament systematically altered, and then likewise with all their other personality traits, background knowledge, beliefs, and so on. While this prospect is certainly hugely significant and weighty, perhaps terrifying, each stage seems to be something that the individuals themselves can survive—losing memories, changing temperament, and so on. Indeed, there is no reason in principle why this could not all be done while the person remains awake, enjoying an uninterrupted stream of consciousness even while the contents of that stream change radically. If they are then told that the memories and personality they are to end up with are modeled on another person, what difference does that make? That is a fact about **(p.40)** where the instructions came from, but does not change what happens to them, there, in that device, having their brain scrambled.

It seems to me, when considering the case in this way, that the subjects would be perfectly reasonable in thinking of the device as imposing a series of major changes on a single continuous person.⁴ But over the course of these changes, all interesting psychological relations between their “new self” and their “old self” are broken, so the psychological conception of subjecthood must deny this. The metaphysical conception, however, can exactly capture what is going on here: a single underlying substance (whether we identify it as the body, the brain, the organism, or something else) is persisting through unusually rapid and radical changes in its properties.

At this point we might start to debate which set of intuitions is stronger, or more reasonable, or better-grounded, or easiest to explain away somehow. But I want to hold onto both, side by side, and think about what kind of function they serve. I think that both serve valid functions, both practical and epistemic. Let me begin with the epistemic functions: How do these conceptions help us to make sense of a perplexing, mutable, world?

2.1.3. Functions of the Two Conceptions

On the one hand, the psychological conception of subjects is epistemically useful because the definition of a psychological subject is inherently so rich in implications: knowing that a person before us is psychologically continuous with a previous person we knew tells us a lot about how that person will behave, far more than knowing any particular psychological trait of theirs. In deploying this

concept we can satisfy a general epistemic goal: to identify rich continuities, where much is preserved across time, and deploy concepts to track them.

On the other hand, the metaphysical conception of subjects is epistemically useful for the opposite reason: the definition of a metaphysical subject allows it to undergo extensive changes, and thus can be applied in a greater range of cases (e.g., identifying someone with their brain allows us to make sense of the influence on them of prenatal events, postdementia events, and any other events which affect the brain even when psychological continuity is broken). In deploying **(p.41)** this concept we can satisfy a different epistemic goal: to identify underlying fundamentals, things which are preserved through many changes, and deploy concepts to track them.

Next, what are the practical functions of “subject” (what Parfit 1971 refers to as “what matters” in survival)? Many of the answers to this question seem to track the psychological conception of subjects, but one, I will argue, better tracks the metaphysical conception.

Here are some of the things that are at stake in asking whether the person before us is identical to a past person P: Are they bound by P’s commitments (promises, contracts, debts, etc.)? Are they culpable for P’s past crimes, or laudable for P’s past heroism? Are they the right person to continue the projects that P started, and carry out the plans that P laid? Do they have the shared history with P’s friends and family that gave them their deep emotional relationship to P? Should those friends and family embrace P fondly, or introduce themselves? I think that if we suppose that the person before us is psychologically continuous with P, that will provide a very good reason for treating them “as P” in all of these respects. Since that is what the psychological conception says matters, the prominence of these questions seems to support that conception. Psychological continuity with P at least provides a better reason for treating them as P than would their possession of P’s body, brain, or even soul, if there was no psychological continuity. After all, continuity is what would let this person recognize and reciprocate the love of P’s friends and family, acknowledge remorsefully or proudly P’s past deeds, recall P’s commitments and why P took them on, pick up P’s plans right where P left off, and so on. Of course, if they did not have P’s body or brain, they might react with denial and treat all of this as basically illusory: “I seem to remember doing that awful thing, but it wasn’t me! I feel great affection for your familiar face, even though this is the first time we’ve met!” But for them to embrace the “practical identity” of P would at least feel more natural.

But now consider person P themselves, contemplating a possible future and wondering whether anyone in that future will be *them*. Part of what matters to them may be questions like “Will anyone honor my commitments and repay my debts?,” “Will anyone comfort my friends and family?,” “Will anyone remember

my life and experiences?," or "Will anyone complete the plans I've laid?" But this is surely not the only thing that matters to them; after all, these questions are not wholly different from questions someone can ask about the world after they themselves die, when there is no question of identity. Something else that matters to them is what they can appropriately anticipate—whether they can anticipate pleasant experiences, unpleasant experiences, or no experiences at all. If someone in the future completes my plans, repays my debts, or even remembers (p.42) my experiences, they are not me if it would be a mistake for me to look forward to their pleasures and dread their suffering. But what determines the appropriateness of anticipation? Presumably I can anticipate any experience that would be part of the same stream as my present experiences. And it seems that I could enjoy a continuous, unbroken stream of consciousness even while my memories, intentions, personality, and so on were radically changed (as Williams's story brings out). Thus it seems that the metaphysical conception is better-placed to capture the relationship between subjects and anticipation than the psychological conception.

None of the intuitive claims that I have made in the past few pages is uncontroversial, but I think that taken together they support the conclusion that we have at least two distinct, and independently attractive, ways of thinking about experiences as "belonging to something." Rather than try to adjudicate which of them best deserves the label "subject" (and the associated status of being referred to by "I"), I will keep both notions in play.

2.1.4. Combination for Substrates and for Personas

We could use the term "metaphysical subject" for the kind of thing that the metaphysical conception is about, if it is true, and the term "psychological subject" for the kind of thing that the psychological conception is about, if it is true. Then it would be an open question which of "metaphysical subjects" or "psychological subjects" exist. I will use the terms "substrate" and "persona" for the things that *would be* metaphysical and psychological subjects, respectively, if any existed. So, given some plausible assumptions about what underlies human experience, the substrate of my experiences is a human brain, and if the metaphysical conception of subjects is correct, then that brain, or the larger bodily or organismic structure that contains and supports it, is a subject, and in fact is *me*. But there is also what we may call "the Luke Roelofs identity," the persona that my experiences fit into (roughly, what would be destroyed if I suffered thoroughgoing retrograde amnesia and massive personality alteration); if the psychological conception of subjects is correct, then that persona is a subject, and in fact is *me*. By formulating things in terms of substrates and personas, we can proceed without deciding which one is really "the subject." We can say that in Williams's thought experiment, two substrates switch personas, and vice versa; we can say that in dissociative identity disorder, and in fictions

like that of Jekyll and Hyde, a single substrate may sequentially support two or more personas.⁵

(p.43) What is the relationship that holds between the persona that is a candidate referent for a given person's name and the substrate that is a candidate referent for that same name—the relationship between Luke's brain and the Luke Roelofs identity, for example? It is hard to say exactly, but roughly we can say that the substrate, by underlying the experiences which together fit into the persona, also underlies the persona. We might put this in terms of occupying or implementing roles (analogizing it to the relation between brain states and functional states), or in terms of material constitution (analogizing it to the relation between lumps of clay and statues made out of them), or even in terms of performance (analogizing it to the relation between an actor and a character they play, so that the Luke Roelofs identity is something that my brain *does*). However we spell it out, this does not seem to be a relationship between two entities of the same category, in the way that would be needed for either to be a part of the other. Subject-combination could be combination of substrates into substrates, or personas into personas, but not of substrates into personas, or personas into substrates.

Moreover, the distinction between substrates and personas connects with the distinction, drawn in chapter 1, between aggregates and structure-specific wholes. Personas are defined by a certain psychological structure, and so a composite persona must be thought of as a structure-specific whole, not as an aggregate. But a substrate of experience might be either; for example, the brain is a prime candidate for being the substrate of human experiences and could plausibly be regarded as a structure-specific whole, existing only as long as its physical parts remain in certain complex biological relationships. On the other hand, an aggregate of substrates might also be a substrate of experience; indeed, aggregates are in some ways better candidates for substratehood because of their relatively more fundamental, "lower-level" status relative to structure-specific wholes. The existence of a structure-specific whole is grounded in the aggregate of its parts instantiating a certain structure, and in that sense the aggregate underlies the structure-specific whole. If we want to identify the substrate of an experience with what ultimately underlies experience, it might make sense to look "under" the structure-specific whole that underlies an experience, to the aggregate that underlies it in turn. At least, that is one attractive way to think about subjects, and my aim is to have something to say for all the attractive ways to think about subjects.

(p.44) So for a systematic defense of combinationism I will have to consider three sorts of entity as candidate subjects: aggregates of substrates (see chapter 3), structure-specific wholes composed of substrates (see chapter 5), and composite personas (see chapter 7).

2.2. What Is the Unity of Consciousness?

If combinationists can address the subject-summing and privacy arguments, perhaps they can show that parts of me having experiences could explain my having those experiences too. And so if many parts of me have various experiences, that might explain why I myself have many experiences. But just as no set of bricks automatically forms a house, no set of experiences automatically forms the kind of unified total experience, the kind of “phenomenal field” of consciousness, that human beings have. After all, these experiences are experiences of many distinct subjects, so shouldn’t they be in some sense “isolated” from each other, forming separate fields of experience just as different human beings seem to have separate fields?

We can illuminate this point by considering what it might be like to be a three-headed monster, with three separate brains each having partial control over its body (or alternatively, what it would be like to be a human being who somehow grew fully functional brains in their feet). Surely there is a sense in which, rather than there being any single “something it is like” to be such a being, there are in fact three things it is like. This creature would have its head-one experiences, its head-two experiences, and its head-three experiences, but it would have them separately, not together. So combinationism should be able to explain why, despite our own consciousness arising from many distinct parts of us, we do not find ourselves in the situation of this imagined monster. What “unifies” all the experiences that our brains give rise to?

But maybe I’ve put the cart before the horse? The above way of putting the problem of unity assumes that we can talk about the three-headed monster as being itself a subject of experience (in addition to its three, or more, smaller conscious parts). We might, however, think that the whole monster is not a subject precisely *because* it lacks unity. That is, we might think that subjects essentially have unified experiences—that unity is part of what it means to be a subject of experience. This would make sense especially if we were thinking of composite subjects as structure-specific wholes rather than as mere aggregates: unity of consciousness is a very natural sort of structure to use for individuating subjects. We would then say: there is a subject of experience for every set of unified experiences, and when not all the experiences associated with a system **(p.45)** are unified (as with our three-headed monster), that system is not a subject, though it may contain subjects (the three heads). In chapters 5–8 I explore this idea of subjects individuated by unity, but of course this increases, rather than diminishes, the importance of providing a combinationist account of how experiences get unified.

The problem is, it is far from obvious what “unity” means here. Although both “the unity of consciousness” and “the subject of experience” are philosophical jargon, the latter at least can be anchored to very familiar structural features of language, namely the way that we always ascribe experiences *to* someone and

speak of things feeling some way “to” or “for” this someone. It is not as common for us to talk about unity, and so it is not as clear that we should look for one thing that this term “unity” picks out. Whereas different accounts of “the subject” seem to *disagree*, providing rival accounts of the same thing, it is more natural to see different accounts of “conscious unity” as just picking out different things. So in this section I review five things that could be called “conscious unity” and how they relate to unity-based challenges to combinationism.

2.2.1. Five Sorts of Unity

So, first, we might talk about someone’s mind or consciousness being “unified” when it is internally consistent, and conversely speak of it being “split” or “divided” when they harbor inconsistent beliefs, attitudes, or judgments. Schechter (2013) uses the term “coherence unity” for something like this consistency of contents (cf. Korsgaard 1989, 176), and it does seem to be one major part of our normal sense of a subject as having a unified mind. A single subject should not think inconsistent things at the same time, and the occasional exceptions are cause for criticism. For distinct subjects, or even for the disunified experiences in two heads of a three-headed creature, inconsistency would simply be normal, unsurprising disagreement. Call this relation among experiences “global consistency”: a set of experiences are “globally consistent” if none of them contradict any of the others.⁶

Second, we might say that two experiences are unified if the present state of each is in part a causal result of the other. For example, suppose I am distracted from my friend’s words by a traffic light turning green: my perception of my friend’s voice would be different (namely, attended-to) were it not for my more salient perception of the light. And my perception of that green circle would be **(p.46)** different were it not for the contrasting black around it, for then I would not see it as a circle specifically. This kind of mutual influence is pervasive in normal human consciousness: you cannot change one experience without its making a difference to many others.⁷ Call this relation “causal interdependence.”

Causal interdependence seems to be connected to global consistency, in that it is because our many conscious thoughts and perceptions can freely influence each other that they are kept consistent. That is, one form of causal interdependence holds when two experiences, which would otherwise have had conflicting contents, interact to produce a sense of tension or to amend one another, so that we end up with only consistent experiences. Without causal interdependence—with experiences that were “cut off” from affecting each other—global consistency would be unlikely to arise or persist.

One particularly important pattern of causal interconnection is what is called “access-unity” (Bayne and Chalmers 2003, 31). A single state is access-conscious if the subject can use its content to guide their various cognitive activities; for instance, a perception of someone shouting “Fire!” is access-conscious if I can report that I heard it, can remember hearing it later, can form plans and inferences based on the fact that someone shouted “Fire!,” and so on. If I cannot do these things, the perception is not access-conscious even if it is like something for me to hear it, and even if it makes a difference to my mood or dispositions. Access-consciousness, thus defined, is a capacity for a particular sort of causal interdependence between the access-conscious state and other mental states. Access-unity is then defined thus: two experiences are access-unified if the subject can use both of them, and the conjunction of their contents together, to guide the same cognitive activities. If, like a split-brain patient, I can respond intelligently to seeing “key” on the left and to seeing “ring” on the right, but not to seeing “keyring,” then my visual perceptions are access-disunified.

Access-unity is a particular pattern of causal relations, a particular way that two experiences can interact: we can define any number of other specific such (p. 47) patterns, and some might be good candidates for being thought of as forms of conscious unity. For example, my various desires interact with my beliefs in such a way that generally, if I desire X and believe that Y will lead to X, I will start to desire Y: this is one sense in which my beliefs and desires are unified. My various feelings of pleasure interact with my perceptions and desires in such a way that if I feel pleasure while perceiving Z, I will generally come to desire Z. And so on. We can refer to these sorts of relations as forms of “functional unity,” with access-unity being just one of them.

A fourth way that our experiences are connected is through their representational contents: they present themselves as being about different aspects of the same thing. In perception this primarily takes the form of what is called “feature binding,” the relation between experiences of (say) redness and squareness that obtains when we see a red square. But it also involves tying experiences together into, for instance, a single spatially connected visual scene, or an overall sense of a single body with sensations all across its surface. And there are also the relations that knit together elements of thoughts (like the concepts “snorkeling” and “bad idea” in the thought that snorkeling is a bad idea). Call both these relations, and any others that serve to form complex conscious representations out of simpler ones, “representational unity.”

Representational unity seems to hold only within the minds of individual human beings. But it does not seem to be pervasive within those minds: sometimes I have unrelated experiences which form no complex representation but which are still had together. So while this relation may be sufficient for two experiences to be “unified” in the sense we are looking for, but it cannot be necessary.⁸ Thus some philosophers (e.g., Dainton 2000; Bayne and Chalmers 2003; Bayne 2010)

have suggested that there is a further, fifth, sort of unity which is pervasive in ordinary human consciousness: phenomenal unity. The thought is that for two experiences to be representationally unified in any way, they first need to be in some more basic sense experienced together, “there beside each other” ready to be connected. Unless they share a phenomenal field in this basic way, establishing any richer sort of unity will be impossible, like drawing a line to connect items on two different maps.

(p.48) Phenomenal unity is sometimes said to admit of no further analysis (e.g., Dainton 2000). But other writers (Bayne and Chalmers 2003; Bayne 2010) point out that it does appear to be tied to “phenomenal subsumption”: two experiences are phenomenally unified when they are contained as parts within a larger experience, so that there is not only “something it is like” to undergo the first and “something it is like” to undergo the second, but also “something it is like” to undergo both. That is, phenomenal unity is a matter of the composition of experiences into wholes which are also experiences. On this view, the phenomenal unity of everyday human consciousness consists in the fact that all of a given human’s experiences are subsumed by their single total experience, but no such total experience subsumes their experiences together with someone else’s. And this seems to be the fundamental sense in which a three-headed monster has disunified experiences: there is no *single* thing it is like to be that monster, because there is no single experience that contains as parts all the experiences arising in its three heads.

So “the unity of consciousness” can be resolved into at least five sorts of unity: global consistency, causal interdependence, access-unity and other forms of functional unity, representational unity, and phenomenal unity. When I speak of “unity” in what follows, I will usually mean this cluster of relations: the experiences of one human are “unified” in that they are generally connected to each other in many of these ways.

2.2.2. Is Conscious Unity Transitive?

One debated question about the formal features of unity relations, which will prove important later on, is whether they are *transitive*: Does it follow from one experience being unified with a second, and the second with a third, that the first and third must also be unified? If unity is transitive, then there could not be a system with a “Y-shaped” unity pattern, where two subjects share some but not all of their experiences, and the experiences of each are fully unified with one another, but each one’s unshared experiences are not unified with the other’s. That is, I could not have two unified experiences, of a sound and a smell, while you share my experience of the smell but not the sound, and also experience a sight, with the sight and smell unified as much as the sound and smell are unified.

The plausibility of transitivity depends on the particular kind of unity being considered. Representational unity seems not to be transitive; to pick a particular variety, predication, there can be a conscious thought that, say, “the archbishop’s happiness was fleeting,” where fleetingness is predicated of happiness, and happiness of the archbishop, but fleetingness is not predicated of the archbishop, **(p.49)** showing that whatever relations are binding together the constituents of that thought are not always transitive. Global consistency is not transitive: consider the sets P, not-Q; not-Q, R; not-P, R. The first two are consistent, the second two are consistent, but the first and third contradict one another. Access-unity and also other forms of functional unity seem not to be transitive: Why couldn’t there be a system whose cognitive mechanisms were able to synthesize and employ contents A and B, or contents B and C, but which for whatever engineering reason were not set up to do so with A and C together?

On the other hand, causal interdependence is plausibly transitive, if understood in the weakest possible way: from the fact that A can influence B, and that B can influence C, it does seem to follow that A is capable of having some sort of impact on C, however slight and indirect. But usually we are interested in interactions that reach some minimum of strength, and that sort of interdependence is not transitive: A might have a big impact on B, and B a big impact on C, but A only a small impact on C.

It is for phenomenal unity that the question of transitivity is most controversial. There is no obvious contradiction in the idea of A and B being subsumed within one phenomenal field, and B and C being subsumed within another, so that both pairs are unified but A and C are not. But some writers have felt that this sort of situation is incoherent in some more subtle sense. I am not sure whether phenomenal unity is transitive: when transitivity would matter, I will simply note the relevance of the question. But even if phenomenal unity is not transitive, the other relations discussed above are plausibly not, and consequently “unity” as a whole will be nontransitive even if it has transitive elements.

2.2.3. Which Sort of Unity Is Most Fundamental?

I have distinguished five different sorts of “unity relations” that experiences can stand in. One important question is how these different relations are related to one another, and in particular how phenomenal unity relates to the other four. Here I think there are broadly two different attitudes that can be taken, which I will call “reductionism” and “primitivism” about phenomenal unity.

“Reductionism” is the idea that phenomenal unity is not really anything over and above the other four sorts of unity (e.g., Shoemaker 2003): all that it is for two experiences to be phenomenally unified is for them to be causally interdependent enough, consistent enough, functionally unified enough, perhaps representationally unified, perhaps access-unified. Since things like causal interdependence or access-unity come in different forms or different degrees,

this amounts to thinking that phenomenal unity is a somewhat messily defined thing, **(p.50)** a cluster-concept, something which holds when we have “enough” of the other sorts of unity, even though it is very hard, perhaps impossible to say precisely how much is “enough.”

By contrast, “primitivism” about phenomenal unity (e.g., Dainton 2000; Bayne 2010) is the view that two experiences being phenomenally unified goes beyond any other form of unity: not only are they interacting, consistent, jointly accessible, and so on, but they are *also* subsumed by a composite experience, and this is not simply a way of summarizing the former facts.

These two views have different implications for the separability of different sorts of unity. If reductionism is true, then any experiences which are phenomenally unified will also be unified in other ways, since that is all phenomenal unity is. But if primitivism is true, then it is at least conceptually possible that two experiences might be phenomenally unified despite not being causally interdependent, jointly accessible, consistent, or representationally unified. Indeed, if primitivism is true, it is not strictly impossible that my experiences and yours are phenomenally unified, despite the lack of any other sorts of unity among them.

The choice among these two views is likely to track views about the fundamentality of consciousness itself. Phenomenal unity involves the subsumption of experiences by a composite experience; to put it another way, it involves a pair of conscious states being itself a conscious state. If consciousness is a fundamental property, so that whether some event or state is conscious is a fundamental fact, then it would be natural to extend that to whether a pair of conscious states is conscious, and conclude that whether two experiences are phenomenally unified is likewise a fundamental fact. That is, primitivism about consciousness and about phenomenal unity go hand in hand. By contrast, reductionism about phenomenal unity goes hand in hand with reductionism about consciousness, since the former implies that whether a pair of experiences is itself an experience depends on whether they are “sufficiently” unified in other ways, which, as noted above, is likely to allow for some vagueness, some borderline cases where there is no good basis for saying that the unity relations involved are sufficient or not. That in turn implies that it might sometimes be indeterminate whether something is an experience, which fits awkwardly with taking consciousness to be a fundamental property.

2.2.4. Can Conscious Unity Hold between Distinct Subjects?

The combinationist needs to explain how a composite mind, whose experiences come to it from its many individual parts, can come to be knitted together in all these various ways. More specifically, they need to explain how to understand **(p.51)** unity as simultaneously both a within-subject relation, among the experiences of a single whole, and also a between-subjects relation, among the

experiences of many parts. We are familiar with the former, but much less so with the latter: How could the relations that tie together all of my experiences hold also between one subject and another?

For some forms of unity, this question is actually fairly easy to answer. In the case of global consistency, for instance, we have a perfectly good grasp of what it is for two people's experiences to not contradict each other, and there is thus no special problem of explaining the consistency of a composite subject's total experience, beyond explaining why particular component experiences tend to be consistent with each other.

Causal interdependence is likewise easily accounted for by purely causal unifying relations: all we need are some causal laws according to which the various parts of the world interact with and affect each other, and we can readily explain how it is that particular experiences come to be causally interdependent.

Moreover, we need not worry that two subjects' experiences being unified in these ways would deprive those subjects of their boundedness, merge them so fully as to no longer be really *two* subjects. Indeed, it might seem that global consistency and causal interdependence are both things that routinely hold between the experiences of different human beings: when two people are talking, their auditory experiences are causally dependent on the other's vocal intentions, which in turn are influenced by their previous auditory experiences. And they might well end up believing very similar things, forming a largely consistent shared worldview.

Of course, it is natural at this point to object by saying that the interactions between two conversational partners' experiences are "indirect," while those among one person's experiences are "direct" (cf. Schechter 2009, 153, 2015, 505). But this distinction, though intuitive, is hard to define exactly (though see Schechter 2018, 95–105); in both cases there is some crucial causal intermediary, whether it is sound waves traveling through the air or nerve impulses traveling down axons. Certainly the causal interactions within one brain are *faster* and harder to interfere with than the causal interactions between two brains, but so what? I think what really lies behind this desire to distinguish direct from indirect interactions is a concern with *phenomenal* unity. The thought is that causal interdependence in the absence of phenomenal unity is just not the same as causal interdependence with phenomenal unity, because phenomenal unity, by linking experiences, putting them "in contact" with each other, allows for a special, "direct" sort of causal interaction. Consequently, the intuition goes, just establishing causal interdependence is not enough to unify two minds, without phenomenal unity.

(p.52) What about representational unity? We might think that some sorts of social interaction establish a kind of unity of representations. Sometimes I

categorize things using concepts that I don't fully understand, relying on other members of my linguistic community to fix their meaning; e.g., when I believe that there are elm trees and beech trees but have no idea what differentiates the two, the meaning of my belief depends not on anything in my mind but on what botanists associate with the two words I use (Putnam 1975; Burge 1979). And sometimes it seems that groups can know things that no individual member knows, through a sort of "distributed knowledge" (Hutchins 1995). It would be nice if we could see this as a sort of primitive representational unity among different subjects (and in chapter 5 I will argue that we can), but at first glance it does not seem to be enough. The connections between two members of a linguistic community or social group seem not to touch their consciousness: even if there is something like a complex thought spread across multiple minds, it is not a complex *conscious* thought. These connections seem to be "indirect" in exactly the same way as I noted above for causal interdependence between subjects: they need phenomenal unity if they are to be a genuine sort of unified consciousness.

Finally, what about functional unity? Considering access-unity as an example, we can observe that access is in large part a matter of causation: if there are two experiences, and they have the power to jointly cause some activity to be performed in a way that reflects the conjunction of their contents, then they are to that extent access-unified; if they have such powers for all the relevant activities that a subject is capable of, they are access-unified for that subject. So can the experiences of two humans be access-unified? There is of course a sense in which we can, and often do, "access" the contents of each other's experiences—by asking each other questions, for instance. And if we count one person's experience as "access-conscious" for the other person when the former tells the latter about it, making it accessible to both, then there is no problem saying that such a state is then access-unified with the other experiences of the second subject. (If I see "key" on the left and you tell me you see "ring" on the right, I can access the joint content "keyring.")

But this is still intuitively different from a single person's access-unity, again because the access here is so "indirect": I can access your experience only because I can access another state that carries information about it (my auditory experience of your words), and you may or may not provide me with that access. The same indirectness can obtain within a single person—we might, for instance, be able to recognize our repressed anger only by reflecting on our memories of how we have acted over the past few days, and this should not make the anger **(p.53)** count as access-conscious. The difference in this within-subject case seems to be largely about speed, reliability, automaticity, and similar factors: Can the anger be accessed quickly, reliably, and fully, and does it spontaneously suggest itself when I consider a question to which it is relevant? We can easily imagine making the information-sharing between two people extremely fast, reliable, automatic, etc. (perhaps with futuristic brain technology, as imagined in

my chapter 8 thought experiment). Insofar as it still seems as though the between-subjects access is too “indirect” to establish access-unity, I suspect that what really matters is, once again, phenomenal unity.

So can conscious unity hold between subjects? Partly. Most types of unity can sort of hold between subjects, but the lack of one type (phenomenal unity) fundamentally changes their character, making them “indirect,” ersatz versions of the relations that hold within a single subject. The combinationist’s challenge, then, concerns principally phenomenal unity: How can this crucial relation hold between experiences of different subjects?

This challenge will look very different to primitivists and reductionists about phenomenal unity. Reductionists must think the problem is in a certain sense merely apparent, since phenomenal unity, the key missing ingredient, is really nothing more than the right degree and mixture of the other forms of unity. Our intuition that without phenomenal unity, the other forms are missing a crucial sort of “directness” may be correct in a sense, but only in the sense that without *enough* of the other forms, any particular instance of them will be missing something. The causal interdependence of my experiences and yours when we converse seems like a very different thing to the causal interdependence among my own experiences; the former seems “indirect” in contrast to the latter’s “directness.” But if reductionism about phenomenal unity is true, then all that it would take to remove this indirectness is to add “more of the same”: more causal relations, more mutual access, more transfers of information, between my experiences and yours. By contrast, if primitivism is true, then no amount of strengthening, enriching, or complicating the merely causal relations among two subjects’ experiences will be enough to establish phenomenal unity. Phenomenal unity, on such a view, is its own thing, and which experiences it connects are not determined just by how experiences causally and informationally interact.

Because the challenge posed by conscious unity for combinationism is so different for reductionists and primitivists, I will develop two distinct solutions: in chapter 3 I present a primitivist account of phenomenal unity, while from chapter 5 onward I present a reductionist account.

(p.54) 2.3. Five Arguments against Combinationism

Let us return the five arguments that mental combination is impossible. I have drawn them from the existing literature on the combination problem, in particular the work of James, Goff, Chalmers, and Coleman. They do not all appear in quite the forms I have given them; indeed different writers on the combination problem typically differ both in which arguments they regard as the most important and in how they formulate those arguments. Rather than laying out and comparing all these authors’ different versions, I have simply presented my own formulations, which I have tried to make clear, concise, and rigorous, and have tailored to bring out the commonalities among the different

arguments.⁹ In this section I present the arguments and briefly note the options for responding to them; in the next section I consider those options more closely.

2.3.1. The Subject-Summing Argument

The simplest argument against combinationism is that “subjects don’t sum”: knowing what it’s like to be each of some collection of subjects seems impotent to tell us anything about what it’s like to be *someone else*. And since the whole they compose is not any one of them, it seems to be an example of “someone else.” A popular way to dramatize this point is inspired by the “philosophical zombie,” a being physically and functionally identical to a human being but wholly lacking in consciousness, whose apparent conceivability threatens physicalism. Analogously, the critic of combinationism imagines a “microexperiential zombie” (cf. Carruthers and Schechter 2006, 38; Goff 2009b; Chalmers 2017, 187–188), a physical duplicate of an actual person, each particle of which is conscious but which is itself completely lacking in consciousness. If this creature is conceivable, then premise **3a** is supported and an explanatory gap between subjects is established. A similar point is made by James (1890, 160) and others, who imagine the component subjects as a group of people, and point out that whatever their individual experiences, and whatever their relations to each other, it still seems a further question whether the group they form is itself a conscious subject.

(p.55)

Premise A1: For there to be composite subjectivity, facts about the intrinsic properties of, and real relations between, a set of subjects must ground and explain the fact that the whole they form is itself a conscious subject.

Premise A2: If one set of facts grounds and explains another, it should be inconceivable that the former obtain but the latter do not.

Premise A3 (Independence): For any set of intrinsic properties of, and real relations between, a set of subjects, it is conceivable that those subjects have those properties and relations, but no other subject exists.

A4 (from A2 and A3): No set of facts about the intrinsic properties of, and real relations between, a set of subjects grounds and explains the fact that the whole they form is itself a conscious subject.

A5 (from A1 and A4): Composite subjectivity is impossible.

Since premise **A1** draws directly from the definition of composite subjectivity, I will set aside the option of denying it. Denying premise **A2** involves insisting that there can be genuine combination even if our efforts at conceptual analysis and understanding will never succeed in grasping the connection—that even though the whole’s consciousness is explained by facts about its parts, conceiving of those facts about the parts does not force us also to conceive of the whole as

conscious. The remaining option (which I am going to defend) is to deny premise **A3** (the principle I am calling “Independence”): to affirm that there is some set of conditions that component subjects can meet, such that it will become inconceivable that the whole they form is not itself a conscious subject.

2.3.2. The Unity Argument

The idea behind our second argument is that the experiences of a single composite subject must be unified in a way that contrasts with the apparent separateness of experiences had by different subjects. And yet if the whole gets one experience from one part, and another from another part, surely their experiences will be spread out across subjects rather than unified. And on the face of it this will be true whatever the character of those experiences, and however those parts are related. We know conscious unity as a within-subject relation, something connecting the many experiences within one subject’s phenomenal field. We do not know how to see it as, or even relate it to, a between-subjects relation, holding between two or more different subjects. Hence we cannot see what relation among component subjects and their experiences could suffice to unify them. Critics of panpsychism could thus imagine another kind of “zombie,” a “fragmented zombie” (cf. Goff 2009b, 299–300): a being physically and functionally identical to a human being, made out of conscious particles, and undergoing all the experiences of its **(p.56)** parts, but with no unity among them. If this being is conceivable, then we seem to have a second explanatory gap facing experiential combination.

Premise B1: For there to be composite subjectivity, facts about the individual experiences that a composite subject derives from its parts, and the real relations among them, must ground and explain the fact that those experiences are unified.

Premise B2: If one set of facts grounds and explains and entails another, it should be inconceivable that the former obtain but the latter do not.

Premise B3 (Separateness): For any set of experiences deriving from distinct subjects, standing in any real relations, it is conceivable that the whole formed by those subjects not have those experiences in a unified way.

B4 (from B2 and B3): No set of facts about the individual experiences that a composite subject derives from its parts, and the real relations among them, can ground and explain the fact that those experiences are unified.

B5 (from B1 and B4): Composite subjectivity is impossible.

The combinationist’s options here are similar to those for the subject-summing argument: denying premise **B1** means abandoning combinationism, and denying premise **B2** means rejecting an attractive model of what “explanation” means; the most interesting option is to deny premise **B3** and try to make sense of

conscious unity as something explained by the way that experiences spread among separate subjects are related.

2.3.3. The Privacy Argument

Philosophers writing about the combination problem often assume, implicitly or explicitly, that for a composite subject's experiences to be fully grounded in those of its parts, there must be particular experiences that are simultaneously undergone by both the whole and some part of it. This is compatible with the whole having some experiences that no single part has, if they are built up out of experiences of different parts. And it is compatible with the parts having some experiences that the whole lacks, which, for whatever reason, are never "incorporated" into the whole. But something, it is often assumed, must be shared. The problem is that a central fact about consciousness, which distinguishes it from physical properties, is that it is "subjective" rather than "objective." Of course these terms have many meanings, but the idea here is that my consciousness is specifically mine, and directly knowable exclusively by me, in a way that contrasts with "public" facts about material things. To know the shape of a table just requires (p.57) gaining certain evidence, and anyone is in principle equally capable of doing that. But to know my feelings means, in the primary instance, to feel them, and only I can do that: others must hope that those feelings correlate with some observable physical manifestation. And this seems to rule out the kind of "experience sharing" needed for experiential combination.

Premise C1: For there to be composite subjectivity, at least some experiences belonging to a composite subject must also belong to one or more of their component subjects.

Premise C2: Experiences are essentially subjective, tied to a particular perspective.

Premise C3 (Privacy): If experiences are essentially subjective, an experience of one subject cannot belong to any other subject.

C4 (from C2 and C3): Experiences cannot belong to more than one subject.

C5 (from C1 and C4): Composite subjectivity is impossible.

There are three ways for a combinationist to respond to this argument. One is to deny premise **C1**, as Basile (2010, 111) and Goff (2017a, 194ff.) do. Goff puts this by saying that while the experiences of the parts may "constitute" the experiences of the whole, they do not "characterise" it (194): the whole does not itself undergo those experiences. Another option is to deny premise **C2**, and deny that there is any such thing as the special "subjectivity" of experiences. As far as I know, this view is likely to be found only among hard-core physicalists, and moreover it is essentially just a more extreme form of the final option (my preferred one), namely to deny premise **C3** (the principle I am calling "Privacy")

and try to make the sharing of particular experiences consistent with their being in some significant way “subjective.”

2.3.4. The Boundary Argument and the Incompatible Contexts Argument

I will present the last two arguments together, because I think they are best viewed in terms of a dilemma arising from a set of shared premises: each argument tries to show that one or the other horn of the dilemma is unacceptable. They both begin by assuming the starting points of the previous two arguments: that experiential combination must involve the sharing of particular experiences and that it must yield conscious unity in the whole.

The arguments then pose a dilemma: consider some pair of experiences belonging to a composite subject, each of which is shared with a different component subject, **(p.58)** and which the whole experiences as unified. Must they also be unified “for the component subjects”? That is, is conscious unity subject-invariant, something that holds between two experiences in themselves, and so remains constant whichever subjects we consider them relative to? Or is it subject-relative, something that connects experiences only relative to a particular subject, so that the experiences can be unified relative to one subject, but not unified relative to others?¹⁰ Either supposition is then taken to be unacceptable.

Suppose, for instance, you are now experiencing two things together (e.g., seeing someone’s mouth move, while also hearing them sing), and each is also being experienced by one part of you. Clearly these two experiences will color one another: each movement of the mouth will be seen as the source of a particular note of the song, and the aesthetic quality of each will be influenced by the presence of the other; the experiences are to some extent interdependent. But what about the part of you that is experiencing just, say, the visual image of the mouth? Do they experience it as a musical, noise-producing mouth, or not? If they do, it seems to follow that they are also aware of the sound, and so our initial supposition that they shared only part of your experience collapses. This is the boundary argument: the component subject disappears into the whole. The key assumption (what I will call “Boundedness”) is that a subject’s consciousness must be “bounded” in the sense that their experiences can be unified only with other experiences *of theirs*.

But suppose my part does not experience the visual image as a noise-producing mouth, but simply as a mouth; then they will be experiencing the visual image differently from me, and surely if they experience it differently, it is a different experience. So our initial supposition that the very same experience is being shared by me and my part collapses. This is the incompatible contexts argument: a single experience is forced to have two incompatible characters, reflecting its unification with two different sets of experiences.

For an even simpler example, consider the experience of seeing blue side by side with red; the red here appears as limited by and contrasting with the blue (and vice versa). If some part of me is experiencing the red alone, does it experience it the same way I do (as limited by and contrasting with another color), or not? If the former, surely that experience implies also experiencing the blue; if the latter, surely that is a different experience of red from the one I am having. Either way, **(p.59)** we do not get substantive experiential combination into a composite subject with unified consciousness.

As with the unity argument, the difficulty of seeing within-subject and between-subjects relations as two sides of the same coin is at the heart of the problem. If we try to think of two distinct component subjects as having unified consciousness, we run into the boundary argument; if we try to keep conscious unity as a within-subject relation by relativizing it to subjects, we run into the incompatible contexts argument. Something has to give.

The Boundary Argument

Premise D1: For there to be composite subjectivity, at least some experiences belonging to a composite subject must also belong to one or more of their component subjects, which do not also share all the whole's other experiences.

Premise D2: For there to be composite subjectivity, the composite subject must have unified experiences.

Supposition S1 (Invariant Unity): If the experiences which a composite subject shares with its component subjects are unified with each other for the composite subject, they are also unified with each other for those component subjects.

Premise D3 (Boundedness): For any experience e_1 belonging to a subject s , if another experience e_2 is unified with e_1 , then e_2 must also be had by s .

D4 (from D3 and S1): All component subjects in a unified composite subject will have the same total set of experiences; i.e., there is no genuine combination of experiences going on.

D5 (from D1 and D4): Composite subjectivity is impossible.

The Incompatible Contexts Argument

Premise E1: For there to be composite subjectivity, at least some experiences belonging to a composite subject must also belong to one or more of their component subjects, which do not also share all the whole's other experiences.

Premise E2: For there to be composite subjectivity, the composite subject must have unified experiences.

Supposition S2 (Relative Unity): If the experiences which a composite subject shares with its component subjects are unified with each other, they may nevertheless not be unified with each other for those component subjects.

(p.60) Premise E3 (Interdependence): The phenomenal character of an experience often depends partly on its phenomenal context, i.e., on the set of other experiences it is unified with.

Premise E4: The phenomenal character of an experience is essential to it.

E5 (from E3 and S2): A single experience, shared by part and whole, could have a different phenomenal character relative to the part and to the whole.

E6 (from E4 and E5): If a single experience were shared by part and whole, it could differ from itself in its essential properties, and thus be numerically distinct from itself.

E7 (from E6, by reductio): Composite subjectivity is impossible.

The combinationist has three options here. First, they might try to dodge both arguments, by denying the shared presuppositions that get them started; in particular, they might deny that particular experiences are shared between parts and whole (which was also one option with the privacy argument). Second, they might accept supposition **S1**, that unity is subject-invariant, not subject-relative, and then rebut the boundary argument by denying premise **D3**, “Boundedness.” This would mean denying that unity between two subjects’ experiences confers all of those experiences on both subjects: even if my experiences are unified with yours, that would not give me your experiences. Third, they might accept supposition **S2**, that unity is subject-relative, and rebut the incompatible contexts argument by denying either premise **E4** or premise **E3**; since premise **E4** seems almost tautological, the more likely option is to deny **E3**, the principle of Interdependence.

2.4. Responses to the Five Arguments

Coleman (2017, 259) aptly summarizes the gist of the above five arguments when he claims, “The self-contained nature of these units [subjects], their phenomenal unity and boundedness, make them singularly unsuitable to the constitution of any further entity.” These arguments present internal problems, not bridging problems: they could be made however science revealed the world to be organized, as could the responses I will now evaluate. In the coming chapters I will engage more directly with the empirical facts about the universe, human brains, and human psychology, so as to connect the schematic defense of combinationism given here with concrete cases. But the five arguments of this chapter point to some basic questions all combinationists must answer, and in examining the possible answers, we see the abstract skeleton of any fleshed-out combinationist theory.

(p.61) 2.4.1. Subject-Summing and Privacy

I think the best responses to the subject-summing argument and the privacy argument are to deny premise **A3** (“Independence”) and premise **C3** (“Privacy”). Let me say briefly why I prefer these to the other options.

One other option for responding to the subject-summing argument is to deny premise **A2**, which says that for x to fully explain y , it must be possible for a sufficiently ideal reasoner to see, from the conception of x , that y follows necessarily—i.e., that it is impossible to consistently and fully conceive of x in the absence of y .

Denying this principle (that explanation requires inconceivability) amounts to relaxing the standards for counting as a good explanation and saying that mental combination can meet those relaxed standards. The problem is that inconceivability does seem like a good standard for explanation and seems to be satisfied in canonical examples of physical combination. For someone who knows the microphysical facts about how water molecules interact at room temperature, their properties and their circumstances and the laws governing them, and who understands what it means to be “liquid,” the further conclusion that water is liquid seems to follow entirely *a priori*, simply by deducing that a substance whose parts interact in that sort of way will satisfy the definition of “liquid.” The idea of a substance whose molecular parts behave in just that way, but which is nevertheless not liquid, seems self-contradictory and thus inconceivable. Since this seems to be what makes physical combination so explanatory, it is a reasonable standard to hope experiential combination can meet. Consequently in this book I will largely set this form of response aside.¹¹

One option for responding to the privacy argument is to deny premise **C2**, the claim that experiences have a special epistemic status called “subjectivity.” Relying on this amounts to saying that mental combination is possible only if, as hardcore physicalists sometimes suggest, our most basic intuitions about the nature of consciousness are misleading and unreliable—that even the idea that each of us knows our own experiences in a distinctive, “from the inside” way, which is more direct than how others must know them, is wrong. This will be an unattractive result for many combinationists, especially those attracted to panpsychist views on which our acquaintance with consciousness is taken to provide (at least a glimmer of) insight into the deep nature of reality.

Finally, what about denying premise **C1** and trying to make combinationism compatible with privacy? If a form of combinationism that respects privacy can be **(p.62)** shown to work, I will have no complaints—but there are several reasons I consider this approach unpromising. Most fundamental is that it radically separates mental combination from physical combination, since the latter has the sharing of property instances as a central feature. Surely physical wholes routinely share physical property instances with their parts: a red

surface seems to share, with each of its red subsections, their particular instances of redness; a car may be dented when its roof is dented, and this seems to involve only one instance of the property “being dented.” In cases like these it seems natural to say that there is a single instance of the property in question, which can be truly ascribed both the whole and the part. These cases are clearly different from cases where there are two separate and independent instances of the same property (two red things side by side, two distinct dents, etc.).

Second, combinationism that respects privacy will have trouble avoiding causal competition between composite and component subjects. After all, experiences have causal powers, and if there really are two sets of experiences—those of the parts and those of the whole—then both should have separate causal influence. This forces combinationists to either deny any causal efficacy to the whole’s experience, or else allow them to intervene in the efficacy of the parts; it would be more appealing to be able to say that the whole has causal efficacy in virtue of its parts having causal efficacy, but this is hard to make out if experiences cannot be shared.

A third worry will come out more clearly in the next chapter, concerning phenomenal unity. To prefigure: given the popular view of phenomenal unity as subsumption by a composite experience, unity between the experiences of distinct subjects turns out to require experience sharing. Privacy thus makes it harder for combinationists to account for phenomenal unity.

Finally, accepting privacy has methodological costs. If experience sharing is possible, then when a composite subject (like, perhaps, one of us) introspects upon their own experiences, they may thereby also be introspecting upon (some of) the experiences of their parts. This means they may be able to discern directly certain things about how component subjects relate, since those relations will be present as structure in the composite’s own experience. Privacy rules this out, ensuring that we will only ever have introspective access to one level of experiential reality (our own), never to that which lies below. This inability to introspect onto component subjects is a major reason why Goff (2016, 294; cf. Miller 2018) argues that something crucial to explaining how minds combine is forever hidden from us: since we cannot detect the experiences of our parts introspectively, we have no way to understand how they combine. Insofar as we would like to *understand* combination, if possible, this is another reason to prefer views which deny privacy.

(p.63) In light of the above considerations, I think combinationists ought to respond to the subject-summing and privacy arguments by denying both independence and privacy:

Premise A3 (Independence): For any set of intrinsic properties of, and real relations between, a set of subjects, it is conceivable that those subjects have those properties and relations, but no other subject exists.

Premise C3 (Privacy): If experiences are essentially subjective, an experience of one subject cannot belong to any other subject.

I do, however, think these premises contain a kernel of truth: each claim would be true if read in terms of discreteness (nonoverlap) instead of in terms of distinctness (nonidentity). That is, they would be true if we read the mention of an “other” subject in either principle as meaning a subject “sharing no parts with the first-mentioned subject,” rather than as meaning “not the very same thing as the first-mentioned subject.” This would allow for entailments and experience sharing among overlapping subjects, both those who contain each other as parts and those who share a part but do not fully contain each other. Call the claims about nonidentity, which I reject, “Strong Independence” and “Strong Privacy”; call the claims about non-overlap, which I accept, “Weak Independence” and “Weak Privacy.”

Strong Independence: For any set of intrinsic properties of, and real relations between, a set of subjects, it is conceivable that those subjects have those properties and relations, but no other subject not identical to a member of that set exists.

Weak Independence: For any set of intrinsic properties of, and real relations between, a set of subjects, it is conceivable that those subjects have those properties and relations, but no other subject not overlapping with a member of that set exists.

Strong Privacy: If experiences are essentially subjective, an experience of one subject cannot belong to any other subject not identical with the first.

Weak Privacy: If experiences are essentially subjective, an experience of one subject cannot belong to any other subject not overlapping with the first.

Let me say a bit to explain the motivation for accepting the weak forms of these theses. The reason to accept weak privacy is that denying privacy across the board, we might lose our grip on a seemingly crucial difference between experiential and physical properties. I can perceive or infer other people’s shape and clothes and **(p.64)** height and weight in just the same way that they can (such properties are “public” or “objective”), but what they feel and perceive is immediately present to them alone, with a directness that sharply contrasts with the way others must guess about it (such properties are “private” or “subjective”). While it may be debated how exactly to spell out this difference,

intuitively we should avoid views which remove the contrast altogether. Fortunately, denying strong privacy need not remove this contrast, for weak privacy still ensures a major contrast between physical and experiential properties. If weak privacy is true, then I have privileged access to someone's experiences if and only if I either am them or overlap with them (sharing a part, or being part of them, or containing them as a part); thus my identity is crucial to my ability to know their experiences. Nothing comparable is true of their physical properties, which are entirely public.

Weak privacy then motivates weak independence, for the easiest way to make sense of explanatory relations among subjects is for those subjects to share experiences, and weak privacy restricts such sharing to overlapping subjects. To put it another way, given weak privacy, all that weak independence rules out is cases where facts about some subjects and their experiences and relations conceptually entail not just that there exists another subject, but also that there is a whole new set of experiences for the other subject to have, entirely distinct from those of the first subjects. And ruling that out seems quite reasonable, since even physical combination generally relies upon property instances being shared between parts and wholes. But strong independence rules out something much more modest and plausible: facts about some subjects and their experiences and relations conceptually entailing that the whole they form is also a subject of (some of) those same experiences. I do not think we should rule this out; I think such entailments can in fact be shown, and I will attempt to do so in the coming chapters. I think failing to distinguish strong from weak versions of independence has allowed the implausibility of explanatory relations between discrete subjects—subjects related like two separate human beings, like you and I—to infect the much more plausible idea of explanatory relations between distinct but overlapping subjects.

I thus deny strong independence and strong privacy, while still accepting weak independence and weak privacy. Subjects can share experiences with other subjects who overlap with them, and when they do thus share experiences, the experiential facts about one can be conceptually entailed by facts about the other.

How exactly this works, however, will depend heavily on what kind of things we take subjects to be, which is why I devoted section 2.1 to disentangling different conceptions of subjecthood. After all, if subjects are personas, analyzable into patterns among experiences, then strong independence does not seem as compelling: personas are nothing over and above properly related experiences, so if we (p.65) can get the experiences of a group of subjects to stand in the right relations, that provides the right sort of basis for a new subject. (The privacy argument, by contrast, is still pressing for personas.) Strong independence is more compelling on the metaphysical conception of subjects, according to which experiences do not constitute subjects but rather presuppose

the prior existence of a subject. When focused on the intuitions which support the metaphysical conception, it is easy to feel that no amount of “shuffling experiences around,” contriving the right relations and organization for them, could be enough to generate a new subject: doing psychology is missing the point when we should be doing metaphysics. But then if we shift back to focus on the intuitions supporting the psychological conception, we may feel that any position in basic metaphysics, shorn of detailed psychological structure, is itself missing the point, ignoring what is actually key to subjecthood.

I think part of what has made the subject-summing argument seem so strong to philosophers is that we have these conflicting intuitions about what subjects are, and no consistent theory can do justice to all of them. This is why I think it is necessary to separate different conceptions of subjecthood and develop theories of mental combination tailored to each. In this chapter I have laid out what all these theories will have in common: in the coming chapters I will look at how they will have to differ.

2.4.2. Unity, Boundedness, and Incompatible Contexts

What about the unity argument, the incompatible contexts argument, and the boundary argument? The approach outlined so far already implies certain commitments with regard to these arguments. First, since I accepted premise **A2** of the subject-summing argument, I am bound to similarly accept premise **B2** of the unity argument, which expresses the same idea (that full explanation of y by x requires the inconceivability of x without y). Thus a combinationist who accepts weak independence but denies strong independence will naturally be led to say that the problem with the unity argument is that the following premise is false, and that its falsity can be shown:

Premise B3 (Separateness): For any set of experiences deriving from distinct subjects, standing in any real relations, it is conceivable that the whole formed by those subjects not have those experiences in a unified way.

For the other two arguments, accepting weak privacy in place of strong privacy stops me from taking the first response I noted above, of dodging both arguments by denying that experiences are ever shared by distinct subjects. That leaves two (**p.66**) options: accept supposition **S1**, invariant unity, or else accept supposition **S2**, relative unity. Let us consider what these options come to.

The first option is to treat unity as subject-invariant, which means denying boundedness:

Premise D3 (Boundedness): For any experience e_1 belonging to a subject s , if another experience e_2 is unified with e_1 , then e_2 must also be had by s .

This means facing down the boundary argument head-on, affirming that a subject's phenomenal field can be "unbounded" in the sense that some or all of its experiences may be unified with other experiences outside its field, which that subject does not itself undergo.

The second option is to relativize unity to a subject, so that even if two experiences are unified for me, they may belong to two distinct parts of me, and be *disunified* for those parts. It follows from this that we must deny interdependence:

Premise E3 (Interdependence): The phenomenal character of an experience often depends partly on its phenomenal context, i.e., on the set of other experiences it is unified with.

If unity is relative to a subject, then the same experience could be unified with different sets of other experiences relative to different subjects having it; if interdependence were true, that experience would then have two phenomenal characters, reflecting these two sets, and would thus very plausibly be (as premise **E4** says) two different experiences.

We do not have to choose one of these responses to apply to all forms of conscious unity. For different forms of unity, interdependence, boundedness, invariant unity, and relative unity may be more or less plausible, so combinationists should consider "mixing and matching."

In fact, considering some forms of unity immediately suggests that they are better thought of in either invariant or relative terms. Causal interdependence, for example, seems to be a subject-invariant sort of relationship: if A has a certain causal effect on B, it does so whether we look at it with a larger or a smaller context in mind, whether we zoom out or zoom in. But functional unity (including access-unity) and global consistency are just the opposite: what role a certain state plays depends on what system we look at it in. For example, for two states to be access-unified is for them, and the conjunction of them, to be accessible for the guidance of things like action, report, memory, and **(p.67)** reasoning. But what are the relevant actions, reports, etc.? If we look at these states relative to a single human being, then we should judge their accessibility according to whether they can guide reports with that person's mouth, actions with that person's hands, reasoning in that person's head, and so on. But if we look at them relative to, say, a large social group, we should judge their accessibility by whether they can guide group-level pronouncements, collective actions, and so on. States might be jointly accessible for my individual uses but have little impact on, or be actively concealed from, the rest of some group I belong to; then we should say that they are accessible, and access-unified, relative to me, but not relative to the group. In light of this, it seems natural to prefer a subject-invariant approach for causal interdependence, which denies boundedness while accepting phenomenal interdependence, while also

preferring a subject-relative approach for functional unity and global consistency, perhaps accepting boundedness but denying phenomenal interdependence. With representational and phenomenal unity it is much less clear what to say; different combinationists might treat them either as relative or as invariant. For now the main thing to note is the implications of the choice: if a given form of unity is made division-relative, it must be denied that interdependence holds for that form; if it is made division-invariant, it must be denied that boundedness holds for that form.

In the coming chapters I outline three combinationist theories that develop different approaches to conscious unity. The first, panpsychist combinationism, focuses on phenomenal unity, understood as a primitive and subject-invariant relation; it thus accepts interdependence but denies boundedness for phenomenal unity. The second, functionalist combinationism, focuses on functional unity, which is subject-relative, and treats phenomenal unity as reducible to sufficiently rich forms of functional unity (and thus as also relative to subjects); thus it denies interdependence for both but can accept boundedness. Functionalist combinationism also develops an account of representational unity, treating it as subject-invariant and thus accepting interdependence while rejecting boundedness. Finally, psychological combinationism focuses not on accounting for unity relations themselves, but on the challenges posed when we try to individuate subjects purely according to the pattern of unity relations among experiences. In each case the challenge is to make sense of the various relations from two angles at once: as internal relations structuring a single subject's phenomenal field, and as external relations between the experiences of distinct subjects.

(p.68) 2.4.3. *Two Technical Objections*

I think the above arguments all raise deep and interesting objections to combinationism; before proceeding it is necessary to address two objections which are more abstract and technical in character.

The first worry is about individuation. I have defined “experiences” as instances of experiential properties. It's often assumed that instances of properties are individuated by the property instantiated, the object that instantiates it, and the time at which it is instantiated.¹² If so, it will be contradictory for one experience to belong to two distinct subjects, for the experiences of any two subjects must be accounted different experiences just in virtue of their different subjects. But this sort of argument would prove too much—it would show that no property instances, physical or mental, could ever be shared between a whole and a part of that whole, which does not appear to be true. Physical wholes routinely seem to share features with their parts—if my desk is scratched or splotched, and its top is scratched or splotched, it is wrong to say that there are two distinct instantiations of scratchedness or splotchedness, let alone two distinct scratches or splotches. Certainly, this argument could not show that conscious subjects

specifically are uncombinable, for it has nothing to do with consciousness. Thus I would advise combinationists simply to reject this way of individuating property instances, doing so in some other way, such as by sets of objects that have them, by their causes and effects, or by supervenience bases.¹³

The second technical objection to combinationism is a sort of “semantic competition” among subjects which comes from the “maximality condition” built into the way we speak. Many of our concepts are such that “parts of an F are not themselves Fs, [and are] disqualified as being Fs because they are . . . parts of Fs” (Sider 2001, 357). For instance, we would not call the left half of a rock “a rock,” even though it would be a rock were it separated from the right half—or at least, we do not count it in our answers to questions like “How many rocks are there?” or consider it a candidate referent for phrases like “that rock.”

It has been claimed that concepts like “subject,” “thinker,” and “mind” are maximal concepts (e.g., Burke 1994, 136; Sider 2001, 357). If they are, then we might be barred, just by definition, from describing any entity as a subject if it is part of another subject. However, even if our everyday concept of a subject incorporates a **(p.69)** maximality constraint, this is not a genuine objection to composition of subjects. Not all facts about the concepts we use reflect facts about the reality we use them to think about: we have to ask *why* our concept of a subject requires maximality. One possibility is that it is simply a practical simplification: nonmaximal entities are not salient to us in the way that maximal ones are, and we have no need to speak or think about them as such because speaking or thinking about the maximal entities that contain them will suffice for all our practical needs. This seems to be why “rock” is maximal: there is no difference in the intrinsic nature of maximal rocks and “large continuous parts of rocks” we simply get by more efficiently using a word that applies only to the former. Indeed the convenience of such habits of speech is such that in the remainder of this book I will use the terms “person” and “human being” as maximal. But if this is the only reason for the maximality constraint on “subject” and similar terms, then subjects-strictly-so-called, and things that would be subjects were they not contained in larger subjects, are equally interesting from a philosophical perspective.

2.5. Conclusions

Over the next six chapters I will develop three theories of mental combination: panpsychist combinationism, functionalist combinationism, and psychological combinationism. These theories will start from different conceptions of subjects, and of conscious unity, but they will also have certain things in common. All three will deny strong independence and strong privacy while accepting weak independence and weak privacy; that is, all will involve component subjects who, when suitably connected, conceptually necessitate composite subjects because they share their experiences with those composite subjects. And all three will deny separateness and, for each form of conscious unity, either boundedness or

interdependence: they will all treat conscious unity as explainable by relations that can hold both within and between distinct subjects.

My main reason for preferring combinationist theories which deny these principles (strong independence, strong privacy, separateness, and boundedness or interdependence) over those that accept them is that I think theories which deny these principles are more ambitious and more interesting, and I think philosophers ought, where possible, to defend the boldest, most ambitious, most interesting claims. Even if those claims turn out to be false, we usually learn more from their failure than from the defense of more timid views. The most interesting question for mental combination, fundamentally, is whether it could work like physical combination, whether it could be intelligible and explanatory in the same way that the combination of physical things is. Most of the alternative **(p.70)** options I discussed and set aside in section 2.4 were set aside because they involve retreating from this ambition.

Nevertheless, denying strong independence, strong privacy, separateness, and either boundedness or interdependence is compatible with sharply opposing views on various questions. In particular, it leaves open the following choices:

- Are subjects personas, or substrates of experience?
- Are composite subjects aggregates, or structure-specific wholes?
- Is phenomenal unity reducible to other forms of unity, or primitive?
- Is consciousness fundamental, or explained by some purely physical structure?

The three combinationist theories I will lay out in the coming chapters give different answers to these questions.

Panpsychist combinationism, focused on the simplest and most basic sort of consciousness but treating it as something metaphysically fundamental, will assume that subjects are substrates, not personas, that composites are aggregates, not structure-specific wholes, that phenomenal unity is primitive, and that consciousness is fundamental. It will answer the subject-summing and privacy arguments by taking consciousness to be, like the fundamental properties of physics, division-invariant: the consciousness present in a whole is just the same consciousness as is present in its parts, and it makes no difference how you slice it. And it will address the unity and boundary arguments by taking phenomenal unity to be, like the fundamental relations of physics, pervasive in its distribution, connecting different experiences regardless of how richly unified they are in other respects.

Functionalist combinationism, focused on building up complex conscious systems, also takes subjects as substrates, but treats them as structure-specific

wholes rather than aggregates, with the brain as the prototypical structure-specific substrate of experience. Brains are, after all, not metaphysically fundamental (like particles) nor straightforwardly identifiable with sets of fundamental things (like aggregates of particles), but equally they are not mere patterns among experiences (like personas).

Functionalist combinationism addresses the subject-summing and privacy arguments by analyzing subjecthood in terms of functional and informational structure: to be a subject is to be a system that processes information in the right way, so when component subjects process information among themselves in the right way, it necessarily follows that the system they form is also a subject. It likewise addresses the unity and boundary arguments by analyzing conscious unity (**p.71**) in information-processing terms: for experiences to be unified, they just need to interact in the right ways.

Functionalist combinationism can be taken in two ways. One, which I will call “pure functionalist combinationism,” denies the starting point of panpsychist combinationism by regarding consciousness as entirely a matter of functional structure, and thus not a fundamental property. The other, which I will call “functionalist-panpsychist combinationism” is intended to be compatible with panpsychist combinationism, adding functionalist analyses not to account for consciousness per se but only for the special sort of consciousness that humans and other brainy creatures have. On this second reading, functionalist combinationism is a theory of “complex structured human-like consciousness,” which is not the only sort of consciousness but rather just a particularly interesting sort that comes about when information is processed in the right way. (Compare the ecumenical, multilayered version of panpsychism-cum-functionalism offered by Rosenberg [2017].)

Finally, psychological combinationism treats subjects as personas (and therefore as structure-specific wholes). It is largely neutral on questions of metaphysics, including the fundamentality or otherwise of consciousness and of phenomenal unity, since personas are by their nature less metaphysically fundamental than experiences. Thus it largely ignores the subject-summing and unity arguments, presupposing experiences and unity relations rather than trying to explain them. It does, however, face distinctive and interesting issues with privacy and boundedness.

Of these three theories, different readers may find different ones more or less plausible and more or less relevant, and may reject one while accepting another. But my own view is that the theories are not rivals but different views of the same reality—a whole universe of composite subjectivity. I think aggregates exist as well as structure-specific wholes, and substrates of consciousness exist as well as personas, and it can be reasonable in different contexts to refer to different entities as “wholes” or “subjects.” The universe is a multilayered thing,

which looks very different when viewed through different conceptual lenses. My ambition with these three theories of combination is to show how mental combination runs through the universe in all of these many layers. **(p.72)**

Notes:

(1) I take “underlying” to be the sort of grounding that holds between an object and an event or process that “goes on in it”; e.g., a particular organ underlies certain chemical processes if facts about those processes are fully grounded in facts about the organ.

(2) Tye’s view (2003, 140–154) is an interesting blend of the two approaches: his notion of the “psychological frameworks” that individuate subjects is clearly an instance of the psychological approach, but the subjects themselves are not the frameworks but whatever physical organ implements them, in line with the metaphysical approach.

(3) Two partial criteria for these conceptions deserve noting. First, if someone thinks that the identities and intrinsic properties of subjects *supervene on* facts about experiences, they almost certainly accept the psychological conception. (But if they deny this, that suggests but does not imply acceptance of the metaphysical conception.) Second, if someone thinks that “subject of experience” is a *phase sortal*, i.e., a term which applies to an entity at some points in its existence but might not apply at other points (like “student” or “adult”; contrast essential kinds like “cat”), then they almost certainly accept the metaphysical conception. (But if they deny this, that suggests but does not imply acceptance of the psychological conception: both Lockean and Cartesian think that a subject is essentially a subject.)

(4) This intuition is sometimes harder to elicit than the previous one; that at least has been my experience talking about the paper with undergraduates. What may help to elicit it is Williams’s remarks at the end of the paper about easy adjustments that could be made to the procedure, like copying A’s psychology into B but leaving A unchanged, or giving A the psychology of B but copying A’s psychology into three other people, B, C, and D. No simple switching story can be told about these cases, but the idea that the people remain in their bodies, while their psychology is oddly scrambled, continues to make perfect sense of what is happening.

(5) Fictional stories of body swaps and so forth are often hard to interpret, because they do not specify the metaphysical mechanisms involved. If the characters in, for example, *Freaky Friday* are wholly material entities, and their body switch involves the rewiring of their two brains, then personas have moved while substrates stayed in place; but if they exchange immaterial *souls*, then personas have in fact remained with the same substrate, and what has happened

is that one (mental) substance has come apart from another (material) substance. It is understandable but aggravating how often this is left ambiguous.

(6) Since my focus is on the unity of *consciousness*, global consistency here is defined primarily among conscious thoughts and perceptions, not among background beliefs that one is not conscious of on a given occasion. But to the extent that background beliefs heavily inflect the way we perceive and think about things, global consistency in conscious states will tend to require, and ensure, global consistency in standing beliefs.

(7) Causal interdependence is hard to characterize if we think of experiences as momentary—as what I am experiencing at a given *instant*, so that two simultaneous experiences are more likely to be influenced by what I states, for then it seems like simultaneous experiences are rarely, if ever, causally interdependent: rather, each will owe its character to certain experiences that happened a moment earlier, but have now passed. To account for this we might say that while conscious unity is usually thought of as a synchronic relation, this particular type is actually diachronic, but holds (stipulatively) only over very short intervals (less than a second, say). Alternatively we might insist that causal interdependence is synchronic, but think of experiences as enduring rather than momentary, with successive, causally connected, qualitatively identical (or very similar) momentary experiences being counted as stages of a single enduring experience. Or, third, we might think of causal interdependence in “common cause” terms, as holding between two experiences which owe their character to (some or all of) the *same* earlier moments of experience. The particular way that we conceptualize the causal web that knits our experiences together is less important than having a label for referring to it.

(8) A possible exception is “spatial unity,” the type of representational unity that holds between two experiences when their objects are presented as occupying regions of the same space. Spatial unity seems pervasive in our perceptual experiences under normal conditions, but does not obviously extend to our nonperceptual experiences. Moreover, it seems that a subject with multiple distant sense organs would not experience spatial unity, but might nevertheless experience conscious unity if these organs fed information to a single brain (as argued by Tye 2003, 76–78; Dainton 2004, 9–10; Bayne 2010, 262–266; cf. Roelofs 2014b, 91–93).

(9) The first two arguments, which try to show an “explanatory gap,” appear in roughly the below form in Goff 2009a, 2009b; Chalmers 2017, 185–189, and Mendelovici 2018, 4–8; the former also appears less formally in Carruthers and Schechter 2006; and arguably in James 1890, 150–160. The next three are extracted from arguments and suggestions in Goff 2006; Coleman 2012, 2014, 2017; Rosenberg 1998; Dainton 2010; Basile 2010; Tononi 2012; Mørch 2014; Chalmers 2017; Miller 2017; Mendelovici 2018. Rosenberg (1998) and Miller

(2017) are especially concerned with the boundary argument, while the incompatible contexts argument appears, in slightly different forms, in Basile 2010, 108ff.; Coleman 2014, 34ff.; and Mørch 2014, 172–175. The privacy argument is presented explicitly in Robinson 2016, 130, and its first premise, in particular, is asserted or assumed in James 1909, 181; Goff 2006, 58–59; Basile 2010, 108; Coleman 2017, 256; Miller 2017, 12).

(10) The subject-relative/subject-invariant distinction is obviously connected to the more general distinction between division-relative and division-invariant properties. Given that different divisions of reality can yield different subjects (which combinationists must accept), what is subject-relative will also be division-relative (and vice versa), and what is subject-invariant will also be division-invariant (and vice versa).

(11) Moreover, the appeal to inconceivability is a central part of influential arguments against physicalism, and so for a constitutive panpsychist, who motivates their view by identifying shortcomings in physicalism, it would be dialectically useful to accept premise **A2**.

(12) For examples of this way of thinking of experiences, and property instances more generally, see Bayne 2010, 24–29; Ehring 1996, 462.

(13) For a defense of the causes-and-effects approach to individuating experiences, see Schechter 2010. For a defense of individuating mental property-instances by their supervenience base, tailored explicitly to allowing overlapping subjects to share particular mental states, see Sutton 2014.

Access brought to you by: