The Combinatorics of Conscious Experience

James R. Shaw [†]

Contrast building blocks with jigsaw puzzle pieces. Building blocks can be laid out contiguously in many two and three dimensional arrays. But the jigsaw puzzle is different: its pieces normally only allow for a single contiguous two-dimensional arrangement.

We can put the contrast between the building blocks and puzzle pieces in terms of what I will call their *combinatorics*: the study of the ways these elements can be arranged and rearranged. Building blocks, by design, have a boring, permissive combinatorics. Puzzle pieces, by design, have a more interesting, restrictive one.

More broadly, the combinatorics of physical things like blocks and puzzle pieces is given by physics. Physical laws tell us when physical things can coexist in close quarters, or be superimposed (where "can" expresses a physical modality). This paper instead explores the combinatorics of *conscious* things: things like sensations, emotions, and thoughts. Which of these things can coexist 'in close quarters', when, and how? Do conscious experiences have a boring, permissive combinatorics, like the building blocks, or an interesting, restrictive one, like the puzzle pieces?

I'm going to argue that we can make sense of these questions and give plausible, revealing constraints on their answers through introspection and reflection. These constraints are significant, because they could provide the materials for a surprising *a priori* case against dualism.

Draft of May 18, 2014; please don't cite without permission

✉: James.A.R.Shaw@gmail.com

[†] In writing this paper I've benefited from helpful discussions with Jacob Beck, Andrew Chignell, Esa Diaz-Leon, Ellen Fridland, David Miguel Gray, Chris Hill, Kristen Inglis, Ole Koksvik, Japa Pallikkathayil, Adam Pautz, Kieran Setiya, and participants in the *2013 Midsummer Philosophy Workshop* in Edinburgh.

1 Combinatorial Spaces

I'm going to assume that conscious experience generally has a discrete structure in this sense: experience can be broken up in to isolatable component types that can sometimes be tokened in other experiences. Two caveats: These components needn't be atomistic. For example, I may treat a patch of color as a discrete element of an experience, without claiming it is, or has as a constituent, a smallest discrete patch of that kind. Also, the language most suited to talking about these discrete parts involves reifying them. I don't mean the reification to be an assumption on which my arguments rest. I suspect my claims could be translated into more acceptable discourse by those who find it suspicious, but I won't explore how here.

To investigate possible rearrangements of conscious experience, we need to settle two further issues: What notion of possibility is at stake? And what is involved in 'rearrangement'? We can resolve the first question by stipulation: we're concerned with metaphysical possibility. The answer to the second is more complicated. Rearrangements of conscious experience may have to be understood differently than spatial rearrangements of physical things, for example. Color sensations may have something analogous to a spatial character. But what about emotions, or thoughts?

We can make headway here by accommodating different senses of recombination—several 'dimensions' along which it seems conscious experiences could be rearranged. Let me briefly characterize some salient distinctions.

Interpersonal v. Intrapersonal. I'll assume there are 'streams' of consciousness—temporally extended, but 'unified' series of experiences.[1] Roughly, a stream of consciousness constitutes the experiences of some entity, like a person. We can ask, appealing to multiple streams of consciousness, whether there are possible streams of consciousness in which two constituent experiences are 'swapped'. This is a question of interpersonal conscious combinatorics. We can also focus on one stream at a time, asking how to swap elements into it, out of it, or within it. These are questions of intrapersonal conscious combinatorics.

Diachronic v. Synchronic. I'll assume that a stream of consciousness has a subjective temporal order (intuitively, the succession in which the several experiences seem, by the

[1] Granting, of course, that what precisely this unity consists in is theoretically contested (see, e.g., Bayne & Chalmers (2003), Bayne (2010)).

experiencer, to be had). We can ask of one or more streams of consciousness whether there are possible distinct streams of consciousness which are the result of swapping future elements into the past, or vice-versa. These are questions of diachronic conscious combinatorics. If we instead ask how experiences can be swapped into, out of, or within a stream at a single time, we are asking questions about synchronic conscious combinatorics.

Intramodal Combinatorics, and Concurrence. I'll assume that within a single stream, conscious experiences can be grouped into types (e.g.: visual, auditory, cognitive, emotional) which I'll call *modalities*. These groupings can be selected so that some overlap, or subsume each other. But, if groups are chosen properly, each should have an associated structured 'space' in which experiences of that modality figure. For example, the modality of visual experiences is associated with something like a field or spatial array. Emotional experiences may well be associated with a simple unstructured space in which an emotion merely is, or is not, present.[2] Once we make these distinctions, we can ask how the characteristic experiential constituents of a modal space can be swapped within that space, or how many experiences of one kind can 'fit' into that space at once, or over time, and so on.

The dimensions I've briefly catalogued here aren't obviously exhaustive. For example, one might ask 'higher order' questions about the structure of a given modal space. (Can such a space be 'expanded'—could there be a doubling of one's experienced visual field, for example?) But my first rough set of distinctions will be enough to get us started.

2 Methodological Remarks

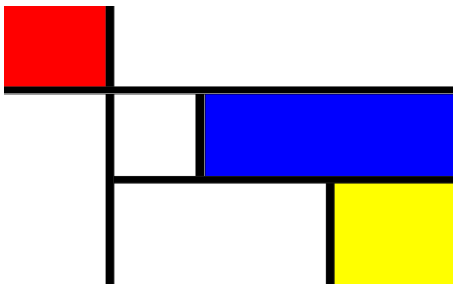
Having set up our questions, let me propose a methodology to address them: introspection and *a priori* reflection. This proposal should set off warning bells. There are well-known and well-founded doubts precisely about whether introspection and reflec-

[2] We should allow that a single modality can be associated with a grouping of spaces, especially if it is reducible to a bundle of more basic modalities. For example, emotions could turn out to be reducible to bundles of perceptual experiences from one or more perceptual modalities (cf. Prinz (2004)). Emotional space will then be constituted by the grouping of modal spaces associated with the forms of perception that contribute to emotional experience. We should also allow that modal spaces can 'overlap'. Perhaps there is a single space in which both a color is experienced to be, and a patch of roughness is felt.

tion can reliably give us modal information about conscious experience. I might seem to be adopting the contested methodology that led Descartes to conclude that mind and body are possible to separate, and leads some contemporary theorists to endorse the possibility of phenomenal zombies, or physical duplicates with inverted spectra.

We need special justification in appealing to introspection and reflection to resolve questions about what it is possible, or impossible, for conscious experience to be like. One way to supply it would be to draw on methodological defenses developed by dualists, like Chalmers (1996, 2009), who argues that we have evidence for a tight connection between *a priori* conceivability and metaphysical possibility. But this will not be my strategy. This is because the modal information we're aiming to acquire is importantly different from that the dualist seeks. The differences make the justification of our methodology easier than for dualists in one way, and harder in another. Before I explain why, it will help to discuss one relatively uncontroversial combinatorial possibility, and one relatively uncontroversial combinatorial impossibility, revealed by introspection and reflection.

Look at the Mondrian-inspired image below and as you attend to it, ask yourself whether it is metaphysically possible for there to be an experience just like the one you are having, but where the red square occurs to the right of the leftmost vertical black line, instead of to its left.



Surely this is possible, and we know it. Perhaps we know by imagination, or perhaps by reflection. But, most importantly, we do *not* know it from experience. We haven't had this experience before. Nor do we extrapolate from previous experiences to tell that this is a metaphysical possibility.

To appreciate the latter claim, call the rearranged experience "*R*" and grant you haven't experienced *R* yet. Past experiences could only provide evidence for the possibility of *R* if they formed patterns in which rearrangements *like R* were actually expe-

rienced. If so, then you could have inductive grounds for believing that the laws governing the formation of conscious experience are compatible with ***R***. Since nomological possibility entails metaphysical possibility, ***R*** should be metaphysically possible as well.

Suppose this is your basis for believing ***R*** metaphysically possible. Now suppose scientists discover that the neurophysiological state corresponding to the rearranged experience I've described is impossible—say, by discovering more about the physical laws governing the neural bases of conscious color experience. If so, you should conclude, or at least become more confident, that anyone's having the experience I described is nomologically impossible. But that was supposedly your basis for believing it metaphysically possible. Thus you should become less confident in this imagined scenario of ***R***'s metaphysical possibility, since your information about nomological possibility has been revealed as unreliable in this case. But surely you shouldn't be less confident of ***R***'s metaphysical possibility. It would be an oddity if metaphysically contingent laws happen to rule out ***R***'s nomological possibility, but ***R*** would still clearly be a metaphysically possible experience.

Accordingly, though we are quite confident that various rearrangements of color experiences are metaphysically possible, that confidence isn't justified empirically. Of course, experiences might be a prerequisite to gaining this knowledge. But if so, they are either acquainting a thinker with the relevant spatial and color concepts, or making the possibility easier to imagine—not supplying justification.

Introspection and reflection also acquaint us with a familiar combinatorial *im*possibility concerning color experience: two colors can't both be experienced at the same point in the visual field at the same time. Is there a possible experience, for example, just like that you have when looking at the previous image, but where the red square is experienced within the area where yellow is experienced *while* yellow continues to be experienced throughout that area?

If we don't misconstrue the request (so that it could be satisfied by blurred vision, quickly alternating experiences, experiencing a 'blend' of the two colors, etc.) it should be clear that the rearrangement of color experiences is impossible. Note that the impossibility here isn't obviously just a logical one—a fact famously of great consternation to the early Wittgenstein. In fact, the claim might even provide an interesting distinction between the combinatorics of visual and auditory phenomenology. Perhaps two distinct sounds can be experienced as if coming from the exact same location in 'auditory space' at the same time.

We are about as certain of color exclusion as we are of any metaphysical impossibility that isn't a claim of logic or mathematics. Again, how do we know this? Perhaps you find that you cannot imagine the rearrangement, or you have a more intellectualized way of detecting the incompatibility. But whatever the answer, it should be clear that introspection and reflection suffice. Here there isn't really much temptation to advert to empirical justification: it's not clear how it could possibly ground the certainty we have in color exclusion.

The two foregoing examples seem to show that at least *some* modal information about conscious combinatorics—both about metaphysical possibility and impossibility—must be reliably accessible by introspection and reflection. How could this be? And how does accepting it situate us in relation to familiar dualist appeals to conceivability to learn about metaphysical possibility?

Let me address the second question first, by noting an important difference between our appeal to introspective/reflective methodology and that of dualists. Dualists want introspection and reflection to provide reliable modal information about the *relationship* between conscious experience and the physical world. We might say: they want combinatorial information about the physical and mental domains taken together. It is precisely the alleged access to these relationships that generates the most resistance to the methodology. What if we have two faculties of imagination or conceiving, or two methods of imagining, one of which is more properly connected with the physical or third-personal setting, and the other of which is connected with the mental or first-personal setting? If physicalism were true, the interaction of these two faculties might generate an illusion of more combinatorial possibilities than there really are. Or what if we have two fundamentally distinct sets of concepts—one phenomenal, one physical—of the same entity, which again makes our imagination a poor guide to which mental/physical combinations are metaphysically possible?[3]

All of these ways of undermining the reliability of introspection and reflection are irrelevant to the questions we're addressing. We want to consider conscious experience *with indifference to* the physical. It is implausible that we have two faculties, both devoted exclusively to first-personal representation, or two fundamentally distinct, non-overlapping sets of concepts that apply only to conscious experiences. Accordingly, standard methods of resisting the dualist's use of introspection and reflection may well succeed. But if so, their success would still leave intact applications of that methodology

[3] See, e.g., Nagel (1974) fn.11, Hill (1997), McLaughlin (2000), Stoljar (2005).

to conscious combinatorics in isolation.

So in one way, our appeal to introspective/reflective methodology is significantly easier to justify than the dualist's. But in another way, our appeal requires saying a little more. The dualist is mostly concerned with establishing the existence of certain metaphysical possibilities: these alone would suffice to refute physicalism. But we, on the other hand, will also be interested in whether we can establish metaphysical *im*possibilities. We want to know not only which combinatorial possibilities conscious experience affords, but also whether there are any combinatorial rearrangements it precludes.

This concern starts to push our earlier, first question: How could introspection and reflection furnish us with knowledge of both combinatorial possibilities and impossibilities? This is an important and challenging question, and I confess I don't have a general answer to it. But the fact that we know *some* combinatorial possibilities and impossibilities for conscious experiences merely by introspection and reflection suggests that there must be an answer that facilitates our project, at least to some degree. So let me venture a very rough proposal, consisting of two claims, that I think has some plausibility as a starting point.

> *Essential Combinatorial Bases*. The metaphysical combinatorial behavior of some set of entities is settled largely or exclusively by their essential properties.
>
> *Experiential Acquaintance*. Merely having several experiences acquaints us with enough of their essential properties to settle much of their combinatorial behavior vis-à-vis each other.

Essential Combinatorial Bases has the virtue of explaining how we come by knowledge of metaphysical combinatorial possibilities for physical objects in our environment. Lois wonders: "is it metaphysically possible for Clark Kent to exist without Superman?" Or, while pointing at Uranium-235: "is it metaphysically possible for there to be an enduring mile-high cube of *this*?" Assuming that identity and material constitution are essential, to answer these questions Lois may first have to learn whether or not Clark Kent actually is Superman, or to learn the physical properties of Uranium-235. But it seems that once she knows all this, any further work in settling the answers to her combinatorial questions can be settled by reflection alone.

The intuitive reason for this, and the intuitive justification of *Essential Combinatorial Bases* itself, should be relatively clear. One way to ascertain a combinatorial pos-

sibility is to witness it, actually. But ascertaining non-actual combinatorial possibilities for some objects requires an act of imagination, or reflection, on the basis of some information about those objects. What information? The information needed to properly survey the interactive behavior of the objects in various metaphysical possibilities. Essential properties are just those that must be retained in all metaphysical possibilities, and so are poised to do the work of constraining or permitting the objects to interact in various ways across those possibilities. Accordingly the crucial information needed, before abstract reflection comes into play, is adequate knowledge of essence.

If we want to know the combinatorial possibilities for objects in our environment, introspection and reflection may not suffice. Where conscious experience seems different is that, unlike with physical objects of our environment, we are able to access essential features of conscious experiences merely by attending to them as they are experienced. Indeed, the qualitative features of an experience of some type seem to be both necessary and sufficient for it to be of that type. Having a particular blue experience, for example, could not have been the same type-experience while seeming to be more yellow. If the experience were to seem different, it would *ipso facto* be a different experience. Similarly if some other experience were to feel exactly like it, then it would have to be an experience of blue.

If we trust our judgments about color rearrangement, I conjecture, it is because we presuppose or endorse *Experiential Acquaintance*: merely having experiences acquaints us with enough of their essential properties to ascertain their combinatorial behavior vis-à-vis each other simply through reflection.

Note that *Essential Combinatorial Bases* and *Experiential Acquaintance* don't entail that we are capable of knowing every fact about conscious combinatorics, even of experiences we have actually had. First, experiences may have essential properties relevant to their combinatorial behavior that experience does not acquaint us with.[4] We should mind this possibility. But note that this poses special problems for assessing combinatorial possibilities and not combinatorial *im*possibilities. Any impossibility soundly assessed in reflection on the basis of a limited range of essential properties will equally well be detected under an expansion of essential properties. This point is important to bear in mind because, in what follows, the most philosophically significant

[4] The idea that introspection falls short of capturing the qualitative natures of phenomenal properties will be entailed by many kinds of reductionist treatments of qualitative experiences. I'll set aside the separate, more radical, contention that introspection represents phenomenal properties as having qualitative natures they do not have—see Pereboom (2011) chs.1-4.

combinatorial claims are claims of impossible rearrangement.

There are other pitfalls to the introspective/reflective methodology I'm prosing. Introspection is fallible since we may be unable to recognize possible rearrangements because of the cognitive costs in envisaging a complex rearrangement. And introspection itself may at times be an unreliable way of getting a cognitive grip on particular experiences, especially if they go by quickly, or are indistinct.[5] If so, we may not be in a position to assess combinatorial facts about such experiences.

Neither are our reflective judgments infallible. We may incorrectly conclude a rearrangement is possible because we fail to adequately attend to certain experiential details, or because we fail to consider a conscious experience 'all at once' that seems possible when considered piecemeal, or because we simply misclassify a possible conscious experience that we are conceiving.

We are prone to error in all these ways, and perhaps others. But this doesn't mean that we must reject our introspective access to the combinatorics of conscious experience wholesale. Rather, we need to be careful to rule out the obvious possible pitfalls as we seemingly succeed, or fail, in conceiving a combinatorial possibility.

In other words, introspection and reflection are fallible, but generally reliable guides to the combinatorics of conscious experience, especially combinatorial impossibility. When introspection delivers strong verdicts on combinatorial possibility or impossibility, and we have taken suitable care to rule out obvious abuses or confusions in our methodology, we have sufficient reason to take those verdicts at face value. Our next question is: what verdicts do introspection and reflection seem to deliver?

3 Cognitive Phenomenology

I leave it to the reader to try a little experimentation on her own. My suspicion is that most will concur that even a cursory effort provides quite a bit of evidence for the claim that, on the whole, the conscious combinatorics of most modalities seems to be *boring*, like that of my earlier building blocks. It seems that all sorts of intramodal permutations, diachronic and synchronic, interpersonal and intrapersonal, are all clearly conceivable, and hence (barring special explanations) possible. Color experiences seem like they can be arbitrarily permuted within a single visual field, and across visual fields

[5] Perhaps even careful, reflective acts of introspection are prone to error, as has been argued by Schwitzgebel (2008).

(subject to color exclusion), at once and over time. The same goes for the sensory modalities of gustatory, auditory, and tactile phenomenology (of course with restrictions furnished by the modal spaces associated with each modality). Olfactory modal space as experienced by humans seems simple, so that intramodal rearrangement doesn't become a question.[6]

There are a few tricky, and hence potentially interesting, cases here and there. Investigations of 'saturation' of concurrent phenomenology seem challenging for those modalities which, unlike visual experience, seem to allow multiple experiences at once at the same 'point' in modal space. How many distinct sounds is it possible to experience at a single location in auditory space, without their 'blending' into a new sound? How many emotions is it possible to experience at once? How many flavors is it possible to experience at what feels to be the tip of one's tongue? It seems like there has to be some limit—still any limit is hard to ascertain. It also seems quite plausible that any difficulties in assessing the existence, and extent, of any such limits is hindered by cognitive limitations in our faculties of imagination. There may be interesting combinatorial constraints here, and we may never know them.

These obstacles seem to be the exception rather than the rule. In general, introspection seems to reveal very few positive constraints on metaphysically possible combinatorial permutations of conscious experiences. If all there were to the combinatorics of conscious experience were such pervasive, permissive permutability, it would unlikely have any independent interest. But I think the combinatorics of conscious experience is not quite so simple. There is a particular phenomenal modality which seemingly gives rise to a very interesting, constrained combinatorics: cognitive phenomenology.

I'll make this case by considering a series of permutations of cognitive and non-cognitive phenomenology that I think strongly resists imaginability very much like in cases of color exclusion. But for these cases to have importance for conscious experience, we need to take a final detour to discuss cognitive phenomenology itself. The nature of cognitive phenomenology, and even its existence, has been contested by philosophers and psychologists. I need to clarify what theses about cognitive phenomenology my subsequent cases will presuppose.

Cognitive phenomenology consists in conscious experience that characteristically accompanies personal-level cognitive activities such as thinking, wondering, supposing,

[6] But perhaps olfactory phenomenology seems to be located in physical space—roughly where one's nose is felt to be? Perhaps the question of whether there are possible experiences of smells near one's fingers or toes then makes sense.

and intending. Contrast three positions on the existence of such phenomenology and its relation to broadly sensory experience.[7]

> *Strong conservatism*: there are no conscious experiences that characteristically accompany personal-level cognitive activities.
>
> *Weak conservatism*: though there are conscious experiences that characteristically accompany personal-level cognitive activities, they are constituted by sensory or affective phenomenology (e.g., pictorial/symbolic imagery).
>
> *Liberalism*: there are *sui generis* conscious experiences that characteristically accompany personal-level cognition.

Suppose you look out across rolling green pastures on a tour of Scotland. As you do, the thought occurs to you that this terrain could be used for golf. At this moment, you have a certain set of perceptual experiences of lush green against a grey backdrop, and the sound of wind in your ears. Perhaps you also have some affective reaction—a sense of calm and relaxation. But is there also some distinctive 'way it is like' to have your thought about golf?

According to strong conservatism, the answer is 'no'. You are thinking about golf, and presumably you know it, but your having this thought need have no effect on the character of your conscious experience. According to weak conservatism, by contrast, having the thought characteristically makes some difference to your experience—but it is one that can be spelled out in terms of changes to your perceptual and affective phenomenology. Perhaps you experience a faint image of a person golfing on the greens, or an image of the word "golf". According to the liberal, the difference may go beyond any such phenomenology—a *sui generis* experience of thinking, thinking about golf, or thinking the very thought about golf that you do.

This brings us to a first presupposition of my arguments. To accept the importance of my ensuing thought experiments, we at least need questions about combinatorics of cognitive phenomenology to make sense as questions *about phenomenology*. For these purposes strong conservatism must fail. Otherwise, when we ask "was it possible to think p instead of thinking q?", we aren't necessarily asking questions about conscious experience at all.

---

[7] For various sides of the debate on cognitive phenomenology see, e.g., the readings in Bayne & Montague (2011).

Strong conservatism has many adherents, but it is in some ways an extreme position. Many individuals report noticeable changes in the character of their experiences as they take up different cognitive relations to different contents. The strong conservative may have to posit a costly introspective error theory to cope with these cases. Also, if our attitudes have introspectable phenomenology, this may contribute to a theory of our special access to knowledge of when, and what, we are thinking.[8] The issues here are thorny, and I won't have time to argue against strong conservatism here. Instead, I will argue for my desired conclusions conditional on the existence of some form of cognitive phenomenology.

There are further questions about cognitive phenomenology I haven't yet discussed: Just how 'fine grained' is such phenomenology? For example, does every change in attitude, or the content to which it is borne, require a change in cognitive phenomenology? And in what sense is the phenomenology of cognition 'characteristic' of it? Answers to some of these further questions will be important for my arguments. But we should defer the questions for now. It will be much easier to appreciate what added assumptions about cognitive phenomenology are needed after we consider some key cases.

4 Combinatorics of Cognitive Phenomenology

If cognitive phenomenology exists, we are in a position to explore its combinatorial behavior in part by ascertaining which cognitive states are possible on their own, and in conjunction with other experiences. I claim that when we do this, we find cognitive phenomenology constrained alongside certain constraints on rational cognition. I'll start with a simple case which begins to press the basic idea before building to more complex, much more problematic cases.

Consider the conscious experiences that would typically accompany the following events.

> *Cheerios.* You see a bowl of Cheerios in front of you, and intentionally bring a spoonful of cereal to your mouth, and chew. Your mouth fills with sensations of a mealy texture and vaguely sweet taste. You intentionally focus your attention on the texture and taste, wondering how to characterize

[8] For a careful explanation of what is 'special' about such access, and why it needs accounting for, see Byrne (2005, 2011).

> it, and think: this tastes sweet and pleasant. You calmly, but enthusiastically (intentionally) bring more spoonfuls to your mouth.

Here is my question: is it possible to hold all experiences in *Cheerios*, fixed, including all experiences associated with personal-level cognition, while swapping the experiences of mealy texture and mild sweetness for some experiences of intense pain. For example, can we swap them for the sensations one would experience when chewing on razor-sharp shards of glass that shred and pierce your gums. This is what it would have to be like:

> *Not so Cheerios.* You see a bowl of Cheerios in front of you, and intentionally bring a spoonful of cereal to your mouth, and chew. Your mouth fills with an agonizing sharp pain with each chew—as if you were chewing on razor-sharp shards of glass. You intentionally focus your attention on this experience, wondering how to characterize it, and think: this tastes sweet and pleasant. You calmly, but enthusiastically (intentionally) bring more spoonfuls to your mouth, each time creating a new explosion of pain.

I claim that we (I at least) cannot coherently imagine what it is like to be the character described in *Not so Cheerios*. The problem is this: as soon as I imagine or conceive of what it is like to experience the pain of chewing on razor sharp shards of glass, I imaginatively feel my thoughts (and hence my cognitive phenomenology if there is any) pushed around, or constrained. I can't imagine what it is like to attend directly to this explosive, sharp pain over an extended period and misidentify it as a sweet pleasant taste. I strain but fail to imagine, through misidentification or otherwise, eagerly intending (with associated cognitive phenomenology if intention has any) bringing this pain about to myself, with no other thought in my mind.

I think the resistance to conceivability or imaginability here is incredibly strong. Moreover, I think that if there is such resistance, it indicates the presence of a metaphysical impossibility. To help make out why, let me address some worries for both theses.

First, one might claim that this could well be an instance where cognitive costs to our imaginative capacities are interfering with our ability to imagine the desired scenario. But this can't be right. The case is too simple for this to be the problem. In fact, it actually becomes *easier* to imagine a scenario like *Not so Cheerios* the more features of

the conscious experience we are permitted to alter. This shows that our obstacles can't stem from overtaxing imaginative resources.

Second, one might make a semantic objection: *Not so Cheerios* is *meaningful*. It says something intelligible. So surely what it describes is both conceivable and possible. The objection misfires: Perhaps the story is meaningful. But the story is not uncontroversially meaningful in any sense that entails possibility (or a suitable sense of conceivability). A story in which a mathematician succeeds in squaring the circle is certainly meaningful and intelligible in some sense. But what it describes is known to be metaphysically impossible. And it is probably not imaginable or conceivable in its details. It is not enough to engage my task to read idly over the description of the story and 'understand'. It requires a conscious, focused imaginative act. It is this active, imaginative engagement with the story that matters, and it is this that I claim is not possible.

Third, one might think that *Not so Cheerios* is possible as a kind of dissociation. It would involve a kind of bewildering experience to feel this intense pain, the failure of recognition and the movements of one's arms, and as if they were out of one's control. If you are thinking this, you are not undertaking the assigned task which is to arrive at *Not so Cheerios* via *Cheerios* by altering only one aspect of your conscious experience: sweet taste swapped for crippling sharp pain. If you imagine shoveling painful cereal into your mouth against your will (as if you were 'stuck' in someone else's consciousness), you may imagine something possible, but you are not imagining my scenario.

A fourth, related, worry is that one might think that this case is conceivable in light of masochists—those who experience pains as pleasurable. The same problem arises for this objection. Masochists arguably either feel pleasure accompanying pain, or experience things that are typically painful as in fact pleasurable (at least to some degree). Neither helps in our case since we are stipulating both the existence of a painful sensation and that it is unaccompanied by pleasure.

A fifth worry accepts that masochists aren't the right candidates to help conceivability along. Instead we should imagine a character hardened over time to pain—someone who, through years of training, is so indifferent to the experience of pain that she can mistake it for light sweetness. I'm not sure that such an individual is coherently conceivable either. But even if she was, an example of an extended stream of consciousness will suit my purposes just as well. I submit that if you imagine temporally extended streams in which *Cheerios* and *Not so Cheerios* can be embedded, with no experience, or memory, of training to cope with pain, etc., the pull to inconceivability will be more

strongly felt, and ultimately unavoidable.

A sixth worry is that I am succumbing to one of the pitfalls of introspective methodology. Perhaps introspection on the nature of pain poorly reveals its characteristic motivational effects, or its identifiability on scrutiny. With regards to the former worry, there are even empirical results from pain asymbolics that may bear on the question. Patients with pain asymbolia, when subjected to characteristically painful stimuli, exhibit no propensity to aversive behavior but maintain (sometimes with surprise) that they continue to have a sensation that is recognizable as pain. Doesn't this show that we shouldn't trust our judgments about *Not so Cheerios*?

Whether pain asymbolia has any relevance to *Not so Cheerios* will of course depend on what pain asymbolia amounts to. One account takes pain to be a hybrid phenomenal state, with a broadly sensory and a broadly affective component that can come apart. Patients with pain asymbolia have the characteristic sensory experience (and so identify it as pain) but lack the affective experience (thereby lacking the motivation to exhibit aversive behavior).[9] This account involves an interesting, and plausible, suggestion: that painful experiences are *experienced differently* (in the phenomenal sense) by patients with pain asymbolia. In other words, one is in a position tell whether one has pain asymbolia from the character of the painful experiences one has, not merely by an inference from one's own reactions to felt pain. If such a claim is true, then pain asymbolia is irrelevant to my examples: we can merely stipulate that the case involves pains as *we* experience them, with attendant affective phenomenology. There are other accounts of pain asymbolia. For example, pain asymbolia may be a form of dissociative disorder, in which patients lack a capacity to care about their own bodily integrity.[10] Even in this case, there is an open question about whether having the relevant dissociative disorder results in differences in affective phenomenology while experiencing pain sensations.

Whether pain asymbolia is connected with differences in phenomenology is worth thinking more about. But regardless, the issues raised by pain asymbolia don't get to the heart of *Not so Cheerios*, which involves not only motivational responses, but more straightforward cognitive/recognitional responses, which are in part supposed to be the occasion of the motivational responses. Our experiencer, while attending to pain, actively fails to classify it as pain. Instead, she classifies it as a completely unrelated sensation, then forms intentions to continue eating because of that misidentification. Pa-

[9] Grahek (2007).
[10] Klein (ms.).

tients with pain asymbolia don't seem to make this kind of situation any more intelligible.

I recognize that not everyone may sense the resistance to imagination I'm claiming occurs for *Not so Cheerios* as strongly as I. But note that this case is merely a simplest instance of resistance to combinatorial rearrangement that can be pressed in increasingly extreme ways. Consider another case, which intensifies the recognitional abnormalities of *Not so Cheerios*.

> *Office.* You have a long, ten hour day doing office work of all kinds: paperwork in which you recognize, intentionally write on, manipulate, and file paperwork; work conversations in which you see and recognize coworkers, recall what relationship they have to you and your work, and discuss work related matters; presentations; computer work; and so on.

Is it possible to hold all experiences in *Office*, including all experiences associated with personal-level cognition, fixed while, every other hour, all sensory and affective phenomenology drop away. Let's call this situation *Office Blackout*. In *Office Blackout*, for five interspersed hours of the day, one has no visual, auditory, tactile, olfactory, gustatory, or affective sensations. Nonetheless, the interruption induces no reaction: no 'wondering what just happened', no surprise or concern at suddenly lacking all sensations, nor even the thought that this change was somehow expected. The lack of phenomenology induces no change in one's broader cognitive reactions either: one continues, for example, to think there is a stapler in front of one, that one sees it (in just the way one always has), one intends to grab the stapler, one thinks that one hears one's colleague's muffled speech through an office wall, one intends to shout to be quiet, after shouting one thinks one has heard one's own shout, and on this basis concludes one might have shouted a bit too loud, etc.

My sense is that such a set of systematic misidentifications and reactions based on those misidentifications, as if no change in phenomenology had occurred, is too absurd. Merely imagining all sensation being cut off without noticing or caring (even to recognize it as expected) seems troublesome. But somehow to imagine one's cognitive life going on systematically confused about one's own sensations and what they deliver (when what is delivered is simply nothing) seems challenging for a few seconds, let alone an hour, let alone several hours in alternation.[11]

[11] Note the possibility of this character is significantly more problematic than the already contested

The earlier caveats apply: make sure that you are not imagining yourself 'trapped' in another consciousness, with your own separate reactions. Don't suppose that someone trained or prepared for these events is at issue: we can build in that there is no experience, or recollection, of training, no expectation of aberration, and so on. If we are careful in these ways, I believe we experience a very strong resistance to conceiving of the suggested conscious stream—one that prevents me, at least, from being able to do it at all.

To imagine this isn't even to imagine the limiting cases of recombination that press my worries. Take the streams of consciousness constituted by two, full human lives of equal duration. Then ask yourself whether there are possible streams of experiences in which we hold the cognitive phenomenology of each fixed, while swapping all other phenomenology. As in the case of Office Blackout, there is going to be such systematic, egregious error in judgment and confusion in intention that it weighs against any way of taking the sensory and affective phenomenology of one stream and binding it into some familiar 'unity' with the cognitive phenomenology of the other stream. And we can continue by adding lifelong intense pains, subtracting experiences, adding odd ones, redistributing the phenomenology of attentional shifts, cutting and pasting the non-cognitive phenomenology of different lives into a life's worth of cognitive phenomenology similarly cobbled together from different agents, and so on. I claim that somewhere along this chain of increasingly unintelligible pairings of non-cognitive and cognitive phenomenology, we arrive at cases that are simply inconceivable (for myself, and I suspect many others, this occurs right at the outset).

What is happening in these cases? Part of what is happening is that we are taking more or less familiar theses about conscious experience in its relation to cognition, and showing how these theses produce constraints on the combinatorics of the associated phenomenology. We start, for example, by reflecting on characteristic motivating force of affective phenomenology or pain, or the introspective accessibility of the broad features of some sensory phenomenology. We note that this characteristic motivation, or accessibility, places constraints on intentions and beliefs in the face of the relevant phenomenology. We then extrapolate from those constraints on intentions and beliefs,

'super-blindsighter' of Block (1995)—the blind-sighter for whom "[v]isual information of a certain limited sort (excluding color and complicated shapes) from his blind field simply pops into his thoughts in the way that solutions to problems we've been worrying about pop into our thoughts." This character is stipulated to know that he is not having visual experiences—the opposite judgments from those that, in our case, are supposed to generate problems. Moreover, in having 'thoughts pop into his head', the super-blindsighter may have an added bit of surrogate phenomenology that is supposed to make his intentions and pronouncements intelligible, which our character must lack as well.

through the phenomenology that would accompany them, to constraints on the space of conscious experience.

Another, perhaps more fitting, way of putting the point is in terms of rationality: if we push ourselves to consider highly extreme forms of irrationality—irrationality in judgments about our experiences and irrationality in intentional action based on them—then we eventually come across irrationality too gross, too systematic, to be possible. But if irrationality is a property of personal-level cognition, and personal-level cognition is accompanied by characteristic phenomenology, then certain kinds of phenomenology will be impossible to have as well.

If so, the combinatorics of cognitive phenomenology, and hence the combinatorics of conscious experience more generally, is interesting: it has, what we might call a *rationally restrictive combinatorics*—a combinatorics constrained to preclude the possibility of egregious violations of rationality.[12]

Earlier, I deferred the question of what added theses about cognitive phenomenology would be needed to make my test cases interesting. Now we're in a position to state an additional assumption my arguments require, which is a constraint both on the 'fineness of grain' of cognitive phenomenology, and a constraint on the sense in which cognitive phenomenology is 'characteristic' of certain cognitive states.

> *cp-sufficiency*: For each attitude of *V*ing that *p*, which cannot coexist with total non-cognitive phenomenology *NCP*, there is at least one bit of cognitive phenomenology *cp*, such that it is metaphysically necessary that if one instantiates *cp*, one *V*s that *p*.

[12] My examples here have focused on incompatibilities between certain cognitive states and non-cognitive phenomenal states. Another promising route to restrictions on the combinatorics of conscious experience, as helpfully pointed out to me by Kieran Seitya and Will Davies, would be to consider impossible combinations of cognitive states alone. One could maintain it is impossible to hold certain egregiously irrational sets of occurrent beliefs, for example, either on the basis of something like the kind of imaginative resistance I've explored here, or perhaps as a consequence of a more general theory of intentional states such as the 'constitutive ideal of rationality' advocated in Davidson (1980). Such a position may be stronger than, weaker than, or simply supplement my argument for a rationally restrictive combinatorics, depending on what assumptions are required to support the impossibility claim regarding combinations of cognitive states taken in isolation. I won't explore the question of which weakest assumptions could do this work here. However, in n.17, after arguing that a rationally restrictive combinatorics could present challenges for the dualist, I'll give some reasons why cases appealing to cognition in isolation may not have the same dialectical force for these purposes as combinations of cognition and non-cognitive phenomenology.

According to cp-sufficiency, there needn't be a unique way it is like to be in the relevant cognitive state of *V*ing that p. In fact, cp-sufficiency even allows that one could *V* that p without having any attendant cognitive phenomenology at all. But cp-sufficiency does require that there be *some* phenomenology 'characteristic' of *V*ing that p in the sense that is impossible to instantiate cp without *V*ing that p as well. If so, we can show cp is incompatible with the totality of non-cognitive phenomenology NCP by *reductio*: if it were possible for cp to co-occur with NCP, then since having cp is sufficient for *V*ing that p, it would be possible to *V* that p in the face of NCP, contrary to our assumption.[13]

Cp-sufficiency presupposes the existence of some cognitive phenomenology, and so entails the denial of strong conservatism. But neither liberalism nor weak conservatism entails cp-sufficiency. For example, a weak conservative may believe that the characteristic phenomenology that accompanies the thought that snow is white is simply having an auditory or visual image of a sentence like "snow is white". Had one learned a different language, one could perhaps have had the same auditory or visual image while thinking a different thought (or thinking no thoughts at all). Or one could be a liberal who takes cognitive phenomenology to be individuated very coarsely, so that all thoughts regardless of content come with the same *sui generis* phenomenology.

So cp-sufficiency constitutes a stronger assumption than the denial of strong conservatism. Even so, many of those who take cognitive phenomenology seriously for the kinds of reasons I mentioned before should be willing to accept it. If cp-sufficiency fails, we cannot obviously appeal to cognitive phenomenology to account for introspec-

[13] The general connections here between cognitive phenomenology, especially on the assumption of cp-sufficiently, and something like combinatorial principles has also been noted by Pautz (2013). Pautz explores a use of these connections to push for an inverted conclusion from my own. I have assumed, largely unargued, that cognitive phenomenology exists and cp-sufficiency holds, and will presently explore its unusual consequences. Pautz, by contrast, wants to use these unusual consequences in a *reductio*: to show that the relevant forms of cognitive phenomenology don't exist. Despite these differences in methodology, Pautz and I ultimately argue for similar claims of what follows *conditional* on the existence of the relevant cognitive phenomenology. So the examples I have chosen, and those that Pautz treats, can be seen as standing in a relationship of mutual support, at least insofar as they are in service of these conditional claims. Pautz is largely concerned to argue against advocates of cognitive phenomenology who are motivated to resolve skeptical problems for intentional properties—authors who think phenomenology *grounds* intentional properties. I have not cited these among my motivations for endorsing the existence of cognitive phenomenology, and I am inclined to agree with Pautz that cognitive phenomenology is likely unable to serve the foundational role of resolving, say, Kripkean meaning skepticism. Accordingly, I won't have more to say about his use of our shared conditional claim here.

tive judgments about cognitive phenomenology. Those judgments, such as they are, tend to motivate relatively fine-grains of cognitive phenomenology. Nor, if we deny cp-sufficiency, will appeals to cognitive phenomenology obviously be of any help in explaining our special access to our own cognitive states and their contents. For example, consider the character I gave above as an example of a kind of weak conservative who denies cp-sufficiency. If we adopt her views, it should be clear that the kind of auditory or visual phenomenology characteristic of thought won't be sufficient to explain how we know which of two homonymous expressions we are thinking.

In any event, as stated before, these claims about the existence and nature of cognitive phenomenology go beyond the scope of this paper. So let me continue to explore what conclusions we can draw conditional on the truth of cp-sufficiency.

5 Physical Combinatorics and Physicalism

As with many introspective investigations, the heavy lifting in my argument will be done through the reader's own reflective work. So I won't belabor the points raised in the previous section. Now I want to turn to a separate issue: if conscious experience does in fact have some form of rationally restrictive combinatorics, why care?

A restrictive combinatorics of conscious experience would be extremely significant, because it would generate an interesting asymmetry between the combinatorics of the phenomenal and the physical. The asymmetry threatens to generate unprecedented, and unexplained, aberrations in the structure of logical space. This is because the metaphysical combinatorial structure of the physical seems, up to the involvement of laws, boring, simple, permissive.

The basic idea here is that if we consider the non-phenomenal in isolation, we expect a kind of 'plenitude' of possibilities familiarly discussed by Lewis (1986). As Lewis put it, somewhat metaphorically: "There are no gaps in logical space; no vacancies where a world might have been, but isn't."[14] To express this plenitude in his framework, Lewis adopts a 'principle of recombination' according to which parts of possible worlds can be patched together to form new worlds. As Lewis put it: "Roughly speaking, the principle is that anything can coexist with anything else, at least provided they occupy distinct spatiotemporal positions."[15] The recombinations that afford such coexistences must, of course, also exhibit a kind of proper spatiotemporal 'fit'.

[14] *ibid.* p.86.
[15] *ibid.* p.88.

Lewis's particular formulation of his principle of recombination is arguably facilitated by two of his more controversial metaphysical theses: Humean Supervenience about natural laws and a counterpart theoretic treatment of metaphysical modality *de re*. But the motivations for Lewis's plenitude principle are independent of such theses, and the principle can and should be preserved by weakening it slightly if the two theses are dropped.

We can, for example, guard against concerns from *de re* modality simply by stipulation—that is, by asking about the combinatorics of the physical when worlds are characterized in *purely qualitative* terms. This allows us to sidestep the controversial question of whether Kripke himself could have had different parents than he actually did, and instead ask only the much less controversial question of whether there could have been someone qualitatively like Kripke with parents qualitatively different from those Kripke actually had. Of course *this* is metaphysically possible. More generally, once we are liberated from concerns about *de re* modality we tend to find a free, permissive combinatorics of worlds qualitatively characterized that marks a noteworthy property of plenitude in logical space.

At least, we find such plenitude 'up to the involvement of laws', as I stated before. For those who endorse Humean Supervenience (roughly: the claim that there can be no difference in laws without a difference in distributions of local qualitative character), the involvement of laws is moot, since laws arise from patterns that manifest themselves at worlds where rearrangements yield suitable, law-like patterns in distributions of qualitative character. But denying Humean Supervenience doesn't obviously interfere with the existence of the free combinatorics we find on Lewis's view. If laws are superadditions to worlds that 'govern' objects or properties, we should still expect to have lawless worlds that fill modal space in the way Lewis envisages. We will simply also have qualitative copies of the better behaved subsets of these worlds that are compatible with sets of laws. This is the noteworthy sense in which the combinatorics of the physical is free 'up to the involvement of laws'.[16]

[16] This permissive combinatorics arguably persists even if we are necessitarians about lawhood, like Shoemaker (1980, 1998), though such views at least call for a slightly more delicate treatment. For example, if properties are constitutively linked to laws, perhaps by being individuated with respect to causal powers, we will obviously not find those particular properties re-instantiated in violation of their associated laws. But we can plausibly find 'similar' but distinct properties that will do the relevant work. (For example, a property—schmarge—that is not identical with, but is very much like like, charge, where attraction between opposite schmarges is slightly stronger, or weaker, than would be permitted for charge proper, and so on). If so, the kinds of combinatorial permutations Lewis en-

In sum, we have good reason to endorse something like the following weakened version of Lewis's principle of recombination.

> *Modal Plenitude Thesis for Qualitative Physical Combinatorics*: Qualitative recombinations of the physical, spatio-temporal parts of a metaphysically possible world are, up to the involvement of laws and spatiotemporal fit, metaphysically possible.

But although we have good motivations for this limited kind of permissive physical combinatorics, we've just seen that cp-sufficiency would ensure there is no corresponding permissive conscious combinatorics. That is, the following thesis would fail.

> *Modal Plenitude Thesis for Qualitative Conscious Combinatorics*: Qualitative recombinations of the phenomenal components of a metaphysically possible experience are, up to the involvement of laws and fit in the associated modal spaces, metaphysically possible.

If so, there is an interesting asymmetry between the combinatorics of physical and conscious entities. This raises the question: who is best positioned to accommodate and explain this asymmetry between the mental and the physical?

One might initially think it is the dualist. After all, both substance and property dualists are in a position to make a fundamental distinction between the physical and the mental that could account for their different combinatorial structure.

But any advantage here is illusory. While the dualist is positioned to *posit* an asymmetry, she is very poorly positioned to *explain* it. This is because the qualitative character of cognitive and non-cognitive phenomenology, as revealed to us in introspection, doesn't seem like it possesses the relevant structural properties that make it clear why the incompatibilities we've uncovered arise at all, let alone why they are metaphysically necessary incompatibilities.

Consider, for example, that the only combinatorial impossibilities we've allowed for qualitative physical combinatorics are essentially those that arise due to a certain failure

---

visaged would come with changes in property distributions in the strict sense, but there will still be a mapping between the Lewisian possibilities and possibilities compatible with necessitarianism that reveals an interesting, very weak kind of combinatorial freedom for the latter view. Even this highly qualified kind of permutability will not be mirrored in conscious combinatorics if the arguments of §4 go through.

of 'fit'. Any such preclusions will be readily explicable. It is clear why rearrangements of objects populating incongruent spaces may not yield new metaphysical possibilities: essentially due to the relevant incongruence. And it should also be clear why even rearrangements of objects populating congruent spaces may not be subjected to overlap: essentially due to constraints on how the relevant spaces are permissibly 'filled'. When it comes to the combinatorics of the phenomenal, we have some parallel cases of combinatorial restriction that can seemingly be explained in the same way. For example, in cases of color exclusion, the dualist is in a position to make a plausible analogous move to that above: qualitative features of the relevant bits of incompatible phenomenology, and the way they occupy their respective modal space, reveal why the incompatibility arises.

But unlike with color exclusion, the qualitative features of the incompatible cognitive and non-cognitive phenomenology don't seem to make evident why the incompatibility arises in the same way—it is not clear what structural features of the phenomenology could do the relevant work. As such, the incompatibility simply has to be posited as brute and *sui generis*—unlike any other kind of combinatorial fact concerning either conscious or non-conscious combinatorics.

By contrast, certain kinds of physicalist are much better situated to account for the non-trivial combinatorics of conscious experience. This is because they can give reductions of conscious experiences to physical entities with a structure that would directly explain the relevant conscious combinatorial behavior. Consider in this regard forms of reductive functionalism. According to such positions, sensory or affective experiences are reduced to physical (or other non-phenomenal) states individuated on the basis of their functional roles, comprising characteristic causal powers. It is a standard, characteristic causal power associated with the functional role of a sensory or affective experience that, *ceteris paribus*, it engenders the corresponding belief that one is having that sensory or affective experience. It is a standard, characteristic causal power associated with the functional role of affective experiences or pains that, *ceteris paribus*, they engender certain sorts of intentional action.

The question of which particular causal roles to pick here is much less important than the availability of a particular *form* of explanation for the reductive physicalist. The reductive physicalist is positioned to allow the combinatorial structure of the mental to 'piggy back' on that of the physical in a way that leaves no explanatory anomalies. This advantage of the physicalist is perhaps best brought out by contrasting the dualist and

reductive physicalist pictures of logical space.

On the dualist picture, logical space is populated by a hybrid of two combinatorial structures associated with physical and phenomenal entities respectively (constrained, perhaps at times, by contingent law-like relationships between them). Since conscious experiences have a non-trivial combinatorics, there will arise *sui generis* gaps in logical space—places in the combinatorially conceived space of possibilities which are 'vacant' owing to constraints on the existence of cognitive phenomenology. These vacancies are empty places in logical space where worlds would have been generated by a permissive qualitative combinatorics. The vacancies or gaps exist for only very select types of conscious experiences, with no explanation for why the gaps occur where they do, in precisely the way they do.

By contrast, for the reductive physicalist, logical space is uniform, and gapless. All metaphysical possibility is dictated independently of conscious experience. And considered independently of conscious experience, the combinatorics of entities has a natural, uniform, permissive combinatorial structure. Any combinatorial incompatibilities arise naturally, essentially from structural features of objects and the spaces they occupy. Once this space of metaphysical possibility is established, we can check which entities within any given metaphysical possibility have the appropriate roles to constitute conscious experience. When we do this, we will never find the offending combinations of cognitive and non-cognitive phenomenology, simply because of the conditions that must hold for such phenomenology to be in place. Proper reductive analyses of conscious states will ensure that no set of metaphysically possible entities could ever constitute it. That is, we will have a uniform logical space in which we can see, from the 'bottom up', why conscious experience has a rationally restrictive combinatorics. This will follow directly from what conscious experiences are. So it is in fact the physicalist—really, a reductivist—who is surprisingly best positioned to give a satisfactory explanation of the combinatorics of conscious experience.

Note that dualists can't capture this virtue of a physicalist position simply by elaborating their position with the introduction of laws governing the relevant phenomenology. Contingent laws couldn't even in principle do the relevant work, because the preclusions at issue are metaphysically necessary. But positing metaphysically necessary laws connecting the physical and the phenomenal will be anathema to the dualist. And positing metaphysically necessary laws connecting the phenomenal to the phenomenal would not really provide a deeper explanation of the incompatibility at all, and cer-

tainly not one which made it intelligible why the gaps in logical space arise precisely where they do. Instead, positing such laws would be tantamount to a concession that no interestingly deeper explanation can be given.

Note also that, for related reasons, the dualist can't try to ameliorate her situation by adopting and applying a functionalist framework to phenomenology itself, unaided by reduction. The functional reduction of the physicalist is explanatorily fruitful because of the ties it creates between phenomenology and structural properties of physical bases. But this is not how functionalism can be wielded by the dualist. Positing casual relationships between bits of phenomenology, whether they end up placing phenomenology in certain functional roles or not, simply won't help the dualist as we have just seen.

The problem here is endemic to the dualist metaphysics. The physicalist is poised to give conscious experience more structure than revealed in experience, via reduction, which will place important constraints on its combinatorial behavior. The dualist cannot safely acquiesce in the existence of such hidden structure unrevealed in experience, lest she undermine the reliability of the only intuitions which lend dualism argumentative support: intuitions concerning psycho-physical combinatorics that vindicate the conceivability of phenomenal zombies, duplicates with inverted spectra, and the like.

So a rationally restrictive combinatorics of conscious experience would constitute a very interesting phenomenon which cries out for an explanation that only the physicalist seems able to supply in a principled way. Not only can the physicalist give such an explanation, but some physicalist theories, like the reductive functionalist theories just canvassed, would have *predicted* these striking features of conscious combinatorics.

In order to arrive at this virtue of a physicalist position, of course, we had to consider the combinatorics of physical entities. In §1, I gave an argument that the methodology I would employ in this paper was less controversial than that of the dualist, because the dualist needed introspective and reflective access to combinatorial facts about both conscious experience and physical entities. So hasn't consideration of physical combinatorics now forced us into this methodology of the dualist, or a methodology just as controversial?

The answer is "no". Our methodology continues to be safer, and less controversial, though seeing why involves a subtle distinction. Recall that the dualist not only needs introspection and reflection to deliver combinatorial information about conscious and physical entities, but information about such entities *taken together*. This is precisely what generates the bulk of the worries noted in §1. Even though we have also appealed

to facts about conscious experience and the physical world, we sought each batch of combinatorial information *separately*, and then asked questions about the logical relations between the results. So even if we have two faculties of imagination, or two sets of fundamentally different concepts, for the mental and the physical, none of this will impugn the kinds of investigations we have undertaken, despite raising worries for classic dualist arguments.

In conclusion, let me take stock by reviewing the three main components of my argument so far. First, my arguments have proceeded on the assumption of cp-sufficiency, an undefended thesis about the existence, grain, and characteristic presence of cognitive phenomenology. Second, I've defended the reliability of a highly constrained class of *a priori* judgments: those (i) that are purely about phenomenal combinatorics, (ii) that concern impossibilities, and (iii) that are safely made in certain ways (e.g., are not about combinatorial rearrangements that are too complex, etc.). Finally, I've claimed that special thought experiments evoke *a priori* judgments, in the constrained class given by (i)–(iii), to the effect that certain kinds of highly irrational cognition are impossible in the face of certain phenomenally conscious states.

All three of these components are required to defend my intermediate conclusion that consciousness has a rationally restrictive combinatorics. But it is important to stress that only the first and second components are required to raise problems for dualism. This is because of a fact alluded to recently: while the physicalist is in a position to try to resist the reliability of the class of *a priori* judgments I've singled out and used in my thought experiments, the dualist is not. As argued in §2, the class of judgments I've appealed to seems like it should be at least as reliable, if not more reliable, than the class of judgments that dualists need to appeal to in order to argue against physicalism. As such, the dualist cannot easily reject my methodology without undermining her own. This opens up a special argumentative strategy for the physicalist suspicious of the dualist's methodology: grant it to her for the sake of argument, and show how considerations involving the combinatorics of conscious experience in isolation, via that methodology, undermine the dualist position 'from within'.[17] My arguments thus open up a new

[17] In n.12, I mentioned that we could try to argue for my intermediate claim about the rationally restrictive combinatorics of conscious experience in other ways than using introspective methodology—for example, by appealing to something like Davidson's constitutive ideal of rationality. Note that if we get to my intermediate conclusion in this way, though, we cannot discharge our methodological assumptions when trying to wield the restrictive combinatorics against the dualist, precisely because the dualist needn't be committed to the reliability of the relevant non-introspective methodology.

kind of dialectical opportunity for the physicalist, in which the methodology that dualists have employed for centuries is turned against them. Granting the verdicts about my thought experiments, the principal resource needed to effect this dialectical maneuver is a simple, but controversial claim about the phenomenology of thought. If this is right, the stakes in the debate over the existence and nature of cognitive phenomenology are substantially raised.

References

Bayne, T. & M. Montague. 2011. *Cognitive Phenomenology*. Oxford University Press, USA. [11]

Bayne, Tim. 2010. *The Unity of Consciousness*. Oxford University Press, Oxford. [2]

Bayne, Timothy & David Chalmers. 2003. "What is the Unity of Consciousness?" In *The Unity of Consciousness*. Oxford University Press, Oxford. [2]

Block, Ned. 1995. "On a confusion about a function of consciousness." *Behavioral and Brain Sciences*, vol. 18 (2): 227–87. [17]

Byrne, Alex. 2005. "Introspection." *Philosophical Topics*, vol. 33 (1): 79–104. [12]

—. 2011. "Knowing That I Am Thinking." In *Self-Knowledge*, Anthony Hatzimoysis, editor, 105–124. Oxford University Press. [12]

Chalmers, David. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press. [4]

—. 2009. "The Two-Dimensional Argument Against Materialism." In *Oxford Handbook to the Philosophy of Mind*. Oxford University Press. [4]

Davidson, Donald. 1980. "Mental Events." In *Essays on Actions and Events*, 207–225. Oxford University Press. [18]

Grahek, N. 2007. *Feeling Pain and Being in Pain*. MIT Press, Cambridge. [15]

Hill, Christopher. 1997. "Imaginability, Conceivability, Possibility, and the Mind-Body Problem." *Philosophical Studies*, vol. 87 (1): 61–85. [6]

Klein, Colin. ms. "What Pain Asymbolia Really Shows." [15]

Lewis, David. 1986. *On the Plurality of Worlds*. Blackwell Publishers. [20]

McLaughlin, Brian. 2000. "Self-Knowledge, Externalism, and Skepticism." *Proceedings of the Aristotelian Society*, vol. 74 (74): 93–118. [6]

Nagel, Thomas. 1974. "What is It Like to Be a Bat?" *Philosophical Review*, vol. 83 (October): 435–50. [6]

Pautz, Adam. 2013. "Does Phenomenology Ground Mental Content?" In *Phenomenal Intentionality*, 194–234. Oxford. [19]

Pereboom, Derk. 2011. *Consciousness and the Prospects of Physicalism*. Oxford University Press. [8]

Prinz, Jesse. 2004. *Gut Reactions: A Perceptual Theory of the Emotions*. Oxford University Press. [3]

Schwitzgebel, Eric. 2008. "The Unreliability of Naive Introspection." *Philosophical Review*, vol. 117 (2): 245–273. [9]

Shoemaker, Sidney. 1980. "Causality and Properties." In *Time and Cause*, Peter van Inwagen, editor. Dordrecht: D. Reidel Publishing. [21]

—. 1998. "Causal and Metaphysical Necessity." *Pacific Philosophical Quarterly*, vol. 79: 59–77. [21]

Stoljar, Daniel. 2005. "Physicalism and Phenomenal Concepts." *Mind and Language*, vol. 20 (5): 469–94. [6]