Composition and Consciousness*

Andrew M. Bailey

June 2010

Abstract

Some strange cases have gripped philosophers of mind. They have been deployed against materialism about human persons, functionalism about mentality, the possibility of artificial intelligence, and more. In this paper, I cry 'foul'. It's not hard to think that there's something wrong with the cases. But what? My proposal: their proponents ignore questions about composition (questions about when some things make up another). And ignoring composition is a mistake. Indeed, materialists about human persons, functionalists about mentality, and believers in the possibility of artificial intelligence can plausibly deploy moderate theories of composition in defense of their views. And as it turns out, these strange cases constitute an interesting source of evidence for moderate theories of composition.

*Thanks to audiences at CU Boulder, Pacific University, and Wake Forest, and to Sean Choi, Jason Decker, Mike Rea, Jeff Speaks, and especially Leopold Stubenberg and the 'van Plantingwagena' reading platoon for helpful comments.

1. Introduction

A host of strange cases have gripped philosophers working on personal identity. John Locke told the story of a body-swapping prince and cobbler. Derek Parfit told the story of a double brain-transplant. Sydney Shoemaker told the story of a brain-*state* transfer.[1] And so on. The cases confused us, and induced inconsistent intuitions. Some cried 'foul', claiming that something had gone wrong.[2] The charge: debates about personal identity–when persons survive–had proceeded without reference to personal ontology–what persons *are*.[3] No wonder we got ourselves confused. Thinking about our *survival* without a thought as to our *nature* was a mistake.

In this paper, I too cry 'foul'. A host of strange cases ('the Cases') have gripped philosophers of mind. The Cases have been deployed against materialism about human persons, functionalism about mentality, the possibility of artificial intelligence, and more. We are given a scenario, encouraged to suppose that thought, consciousness, or understanding is not present in that scenario, and further encouraged to draw a conclusion with far-reaching implications.

It's easy to think that there's something wrong with this procedure. But what? My proposal: the procedure ignores questions about *composition* (that is, questions about when some things make up or are parts of another). And ignoring composition is a mistake. As it turns out, materialists about human persons, functionalists about mentality, and believers in the possibility of A.I. may quite plausibly deploy moderate theories about composition in defense of their views. The Cases don't pose much of a threat after all. Indeed, the Cases constitute an interesting source of evidence for moderate theories of composition.[4]

2. The Cases

I shall first say something about the Cases and the conclusions that have been drawn from them.

[1] *An Essay concerning Human Understanding*, book II, ch. 17, section 15; Parfit (1971); Shoemaker and Swinburne (1984): 108-110.

[2] See, e.g., Johnston (1987), Olson (1997), and van Inwagen (1997).

[3] It is not easy to explain the status and content of the question *what are we?*; but see Olson (2007): 3-22.

[4] The structure of my argument is inspired by Barnett (2008). Barnett argues that the best explanation of our intuitions about a variety of cases in the philosophy of mind is that only simple things can be conscious. If my arguments are sound, there is another explanation of these intuitions: the conditions under which composition occurs. As will soon be obvious, I also owe much to van Inwagen (1990): 119 and to Olson (2007): 228-232.

Person Pairs. David Barnett tells the story of a person-pair. It is comprised of two people; but it cannot be conscious. For even if one person is conscious and so is the other, the pair is not: 'you might pinch your arm and feel a pain. I might simultaneously pinch my arm and feel a qualitatively identical pain. But the pair we form would not feel a thing.' In light of this case, Barnett invites us to intuit: only simple things–things without proper parts–can be conscious.[5]

Chinese Brain. Ned Block tells the story of China, a nation realizing the functional role of an ordinary brain. Every Chinese citizen is equipped with a phone and a call-list. Whenever a citizen's phone rings, she phones other citizens on her call-list. She doesn't *say* anything on the phone; she just rings and hangs up. Others do the same, calling those on their call-list whenever called themselves. 'The call-lists would be constructed in such a way that the patterns of calls implemented the same patterns of activation that occur in someone's brain when that person is in a mental state–pain, for example. The phone calls play the same functional role as neurons causing one another to fire.'[6] Block invites us to intuit: the nation of China could not be conscious, whether or not the nation realizes the functional role of an ordinary brain.

Blockhead. Block tells another story; but this one takes place not in a nation but in a head. Some very tiny people all inhabit the head of an ordinary human person. The tiny people together realize all the the functional roles of an ordinary brain (each tiny person playing the role, say, of a bundle of neurons). Block invites us to intuit: some system of tiny people inside a head could not be conscious, whether or not the system realizes the functional role of an ordinary brain.[7]

Bees. Hilary Putnam tells a story of a swarm of bees that realize the same functional states as a human organism. Barnett elaborates: 'suppose that over the horizon we spot what appears to be a colossal human marching toward us, destroying everything in its path. As it nears, we see that it is in fact an enormous swarm of bees. In deciding whether to fire missiles at it, we calculate the projected suffering of each individual bee, but not of the swarm itself, for the

[5]Barnett (2010): 161. In his (2008), the conclusion is: unless we are justified in believing that only simple beings can be conscious, our intuitions about some stories in the philosophy of mind (Barnett considers each of the Cases and several others) lack the significance attributed to them.

[6]Cole (2004): section 2.3. For the original case, see Block (1978).

[7]Block (1978).

idea that the swarm itself might experience pain seems absurd.'[8]

Chinese Room. John Searle tells a story of a monolingual English speaker, sitting alone in a room. She has pieces of paper with Chinese writing on them, and a set of extensive instructions. Whenever an outsider slips a piece of paper with Chinese writing on it into the room, our subject consults her instructions (written in English) to see which of her pieces of paper to slip outside the room. To outsiders, it seems as though the insider is receiving questions in Chinese and answering them in Chinese. But the insider doesn't understand Chinese. Nor does the system comprised of the insider, her instructions, the pieces of paper, and the room. Searle invites us to intuit: while the system plays the functional role of someone who understands Chinese, there is no understanding.

Rarified Brain. Peter Unger and Arnold Zuboff tell stories of a brain whose hemispheres are scattered abroad, each connected to the other by radio signals. In Unger's tale, the hemispheres are themselves halved and then halved again until every neuron is separated from any other neuron by a vast distance. Unger invites us to intuit: as the brain becomes more and more scattered, its being conscious is less and less determinate. Once the brain is *scattered enough*, it is no longer conscious. But what counts as *scattered enough* is a vague matter; and so too, *being conscious* is vague.[9]

3. Composition

Thus, the Cases. In this section, I say something about composition and briefly sketch some of the main views.

There's something composed of ('a sum of', 'a fusion of') some objects, let us say, just in the case that there is something overlapped by anything overlapping those objects.[10] The Special Composition Question (Spec) asks: under what circumstances do some objects compose a thing? Put 'practically', what must one do to some objects to get there to be a thing composed of those objects?[11] Proposed answers to Spec include:

[8]Barnett (2008): 3. For the original case, see Putnam (1967).

[9]Zuboff (1981) and Unger (1990): 177ff. A similar case is discussed in Dennett (1981).

[10]x overlaps $y =_{df}$ some one thing is a part of both x and y. Read 'part' in this definition such that everything is a part of itself; x is a *proper* part of $y =_{df}$ x is a part of y and $x \neq y$.

[11]For more on Spec, see van Inwagen (1990): 21-32.

Organicism. 'Get the objects related to each other such that they're all caught up in the same life process; then they'll compose a thing.'[12] On organicism, there are trees but not apples. There are songbirds and dogs, but no computers. (Apples and computers are not organisms; trees, songbirds, and dogs are.)

Causalism. 'Get the objects related to each other such that together they have non-redundant causal powers that they wouldn't otherwise have; then they'll compose a thing.' According to its (only, so far as I know) defender, Trenton Merricks, causalism implies that while there are thinking agents like you and I, there are no computers, baseballs, stars, or undetached brains.[13]

Disjunctivism. 'Get the objects functionally unified or such that they and all their parts are rigidly bonded; then they'll compose a thing'.[14] On disjunctivism, there are organisms and some artifacts; but there is no solar system, no sum of your house and mine, and no fusion of all the coffee in the land.

Ordinary Brutalism. 'Since there is no informative (non-mereological) answer to the Special Composition Question, the best answer we can give is this: get the objects arranged such that they compose something. And happily, all and only the ways of so arranging objects correspond to the sortals we ordinarily deploy when talking about medium-sized dry goods.'[15] On ordinary brutalism, there are songbirds, dogs, computers, and statues. But there isn't anything composed of my keys and my favorite dog, nor are there sums of apples and trees.

Universalism. 'Nothing; for any bunch of objects there's something composed just of them.'[16] On universalism, there are songbirds and dogs and apples and trees, and sums of apples and trees and more besides; there is even a sum of my keys, my favorite dog, and all the coffee in the land.

[12]Defended in van Inwagen (1990).

[13]Merricks (2001).

[14]Hoffman and Rosenkrantz (1997): 80-90, 128-134. I follow Olson (2007): 227 in this characterization of the view. The parts of organisms are, for Hoffman and Rosenkrantz, paradigms of functional unity.

[15]Mereological brutalism (*sans* any mention of ordinary sortals) is introduced and defended in Markosian (1998). I have added the 'ordinary' component to Markosian's suggestion. A view much like ordinary brutalism is deployed in Gilmore (forthcoming): 7ff. Further views in the ballpark are defended in Elder (2005), Thomasson (2007), and Korman (2007).

[16]Defenses of universalism include Lewis (1986): 212-213, Rea (1998), and Sider (2001): 120-132.

Nihilism. 'Nothing anyone could do could make multiple objects compose one thing, for nothing could have any proper parts.'[17] On nihilism, the world is populated by only mereological simples (things without proper parts). The nihilist's world is a sad one: there are no apples, no songbirds, and no dogs, only simples arranged applewise, songbird-wise or dogwise.

Organicism, causalism, disjunctivism, and ordinary brutalism are moderate views. Their answer to Spec–their verdict on when composition happens– is 'sometimes, and under certain circumstances'. Universalism and nihilism are extreme views. Their verdicts are, respectively, 'always' and 'never'.

4. Composition and the Cases

The answer to Spec bears on the Cases; in this section, I show how this is so and what follows.

I shall assume that thinking (or understanding or being conscious) isn't something that can be done collectively.[18] Jim and Jill can collectively *surround a building*; they might be surrounding a building even when it is false that *Jim is surrounding a building or Jill is surrounding a building.* But Jim and Jill cannot collectively *believe that p* or *feel a pain.* They can 'together' do this, as it were, only if there is something of which they're both parts that feels a pain. There must be something there, some one thing, to feel that pain. Similarly, the only way for my parts to together desire food is to compose something, some one thing (viz., me) that desires food.[19]

The Cases invite us to intuit that some things don't together think, or aren't together conscious, or don't together understand. In Person Pairs, Barnett says that a pair of persons don't together feel pain. In Chinese Brain, Block says

[17]Defenses of nihilism include Dorr (2005) and Sider (manuscript). A similar view is defended in Horgan and Potrc (2008).

[18]A stipulation about properties *collective* and *distributed*: where 'aa' is a plural term, a property F is collective iff $\diamond(Faa \wedge \exists y(y \text{ is among } aa \wedge \neg Fy))$. F is distributed iff F is not collective. Say I: mental properties are distributed. Thinking isn't something a plurality can do. For necessarily, it is true that *Jim and Jill feel pain* only if it is true that both *Jim feels pain* and that *Jill feels pain.*

Objection: some sentences apparently express truths and predicate collective mental properties of a plurality. Example: 'my parts are together thinking that p'. Reply: sentences like that express a truth only if they are loose-talk for theses like *my parts compose something that is thinking that p.*

[19]I owe this way of thinking about things to van Inwagen (1990): 115-123. But while van Inwagen applies this point only to Chinese Room (119), I shall apply it to all of the Cases and draw some rather general morals from the exercise.

that the the citizens of China are not together conscious. In Blockhead, Block says that the tiny people in a head are not together conscious. In Bees, Putnam and Barnett say that a swam of bees are not together conscious. In Chinese Room, Searle says that a room and instructions and monolingual English speaker do not together understand Chinese. And in Rarified Brain, Unger says that some scattered neurons are not together thinking or conscious. Suppose that Barnett, Block, Putnam, Searle, and Unger are right about all this. Does that tell against materialism about human persons, functionalism about mentality, or the possibility of A.I.?

It does not. The subjects of each Case can together think (be conscious, understand) only if they compose something that thinks (is conscious, understands). The defender of materialism about human persons, functionalism about mentality, or the possibility of A.I., then, has an interesting response to all of these cases. The response is this: there is no thinking (no consciousness, no understanding) in the Cases because there is *nothing there to think*. Pairs of people do not exist, and hence are not conscious. Chinese rooms composed of a subject, some inscribed rules, and some sheets of paper don't exist either; and so they do not understand. And so on.[20] This line has a certain dialectical virtue. The functionalist, say, needn't bite any bullets and insist that something is conscious in Blockhead. Instead, she may *grant* the central intuition of the case (that the tiny people are not together conscious). And she may explain this intuition without back-peddling on her functionalism. She may explain it by insisting that the tiny people do not together compose something, and hence do not together compose something that feels.

If any of the moderate answers to Spec are correct, this line is quite plausible. On organicism, the only composite objects are organisms. But a pair of people (something having persons as its immediate parts) isn't an organism, nor is the nation of China, something composed of tiny people in a head, a swarm of bees, a Chinese Room, or a rarified brain. On organicism, it's not just that these things don't think; it's that there aren't any of them.[21]

[20]Interpreting proponents of the Cases as intending to express a distributed mental property is implausible. For example, in the Chinese Room, it is clear that the subject does not understand Chinese, and neither does her stack of papers or her instruction booklet. Those banal observations aren't Searle's point. His claim is not just that the system doesn't have a *distributed* understanding. Rather, it's that the system doesn't have a *collective* understanding of Chinese either.

[21]Objection: brains can think. On organicism, there aren't any brains, since brains aren't organisms. So the organicist must say that both rarified and non-rarified brains alike are unable to think. Reply: on organicism, there aren't any composite embodied brains. But possibly, some brains are organisms: a (disembodied) brain-in-vat might be just such an organism–an

On the plausible supposition that pairs of people, nations, things arranged tiny-people-wise, a swarm of bees, Chinese Rooms, things arranged rarified-brain-wise have no non-redundant causal powers, causalism has the same results.[22] Such items fail to compose anything at all.

The same goes for disjunctivism. None of the subjects of the Cases display rigid bonding or functional unity. (Recall that on disjunctivism, the functional paradigm of functional unity is an organism; so the merely *computational* functional unity on display by, say, the tiny people in Blockhead isn't sufficient for them to compose something). And so they do not compose anything at all.

Similarly, on ordinary brutalism, the only composite objects there are answer to ordinary sortals for medium-sized dry goods. And the *exotica* described in the Cases are far from ordinary. Since I haven't described ordinary brutalism in much detail, it is not obvious that it tells against the existence of *all* of the subjects of the Cases. But it certainly tells against things like rarified brains.

Even one of the extreme answers to Spec–nihilism–has the same consequence. On nihilism, nothing has any proper parts. So two persons do not compose a person-pair, the citizens of China do not compose a nation, and so on.[23]

I have thus far shown a connection between the Cases and answers to Spec. And I have suggested that materialists (or functionalists, or A.I. theorists) may deploy moderate or nihilistic theories of composition to defend their views against the Cases. On my proposal, the functionalist, say, needn't insist that swarms of bees can feel pain. The central intuition many feel upon considering the Bees is that there isn't any pain there (aside from any pain felt by each bee individually), and the functionalist needn't deny this. She can simply explain this intuition by appeal to a theory of composition. The reason there isn't any pain is that there isn't anything there–there is no swarm composed of all the bees–to feel pain. And

organism shrunk down to the size of a brain, as it were. So the organicist may grant that a non-rarified brain-in-vat might be, and hence might be conscious. And she may also insist that when sufficiently rarified, the brain ceases to be conscious, to be an organism, and indeed, to be.

[22]And for a book-length defend of this supposition, see Merricks (2001).

[23]Objection: this move is plausible only if restricted (non-universalist) theories of composition are plausible. But they aren't. Organicism, causalism, and nihilism entail that there aren't any computers or books or LEGO sets. And that's not a consequence we can accept. Reply: no theory of composition is plausible. Nihilism rules out tables, sure. But universalism entails that there is a fusion of me and my iMac. That too is an unacceptable consequence. What matters here isn't whether a theory of composition is plausible, but whether it is more or less plausible than its competition. And it is far from clear that restricted theories of composition fare poorly on this count. Further, ordinary brutalism and disjunctivism are immune to the charge (they allow the existence of computers and books); so even if the charge has force, it doesn't tell against the application of *those* restricted theories of composition to the Cases.

the same goes for each of the Cases.

But we can go further still. A debate has reached a dialectical stalemate just in the case that (roughly) there's nothing proponents of one position in that debate can do to persuade rational proponents of other positions in that debate to change their minds.[24] One way of achieving stalemate is for all sides to disagree about the relevant philosophical data (to have different *intuitions*, as they say). This phenomenon can take on varying degrees of sophistication. In its more sophisticated form, all participants are well-versed in their opponents' maneuvers (and prepared with answers to all of them). But they nonetheless disagree. Plausibly, debates about composition (about the answer to Spec, in particular) have reached such a dialectical stalemate. Indeed, that debates about composition are considered (by, e.g., Eli Hirsch) *merely verbal* is, I think, an artifact of their having reached a stalemate. By way of illustration, here's a (fictional) dialogue between two philosophers. Let us call them 'PvI' and 'DKL':

> PvI: New debate topic: when does composition occur?
>
> I'll go first. I propose that some things compose another just in the case that they're arranged organism-wise. Only organisms have proper parts.
>
> DKL: It is utterly obvious that there are tables and chairs, and these things have proper parts. And so far as we know, tables and chairs aren't organisms.
>
> P: Not so fast. I concede that tables and chairs aren't organisms. And I concede that it is utterly obvious that there are things arranged tablewise. But it's another matter whether those things compose tables. And I shall paraphrase away any sentences you throw my way that seem inconsistent with my view.
>
> D: Hmm, paraphrase, you say?
>
> P: You shouldn't mind; we all have to paraphrase away the sentences that offend. Indeed, I once read about how to do this in a rather nice paper on holes.
>
> D: Fair enough. Well, I'll go now. I propose that some things compose another... *always*.
>
> P: It is utterly obvious that there isn't something composed just of Art Garfunkel, all the coffee in the land, and the capital building.

[24] I borrow 'dialectical stalemate' from Fischer (1994).

D: Not so fast. I concede that there isn't (quantifier restricted to *things we have reason to ordinarily talk about*) anything composed just of Art Garfunkel, all the coffee in the land, and the capital building. But there is (quantifier unrestricted) such a thing. And I shall explain away any sentences you throw my way that seem inconsistent with my view by appeal to such restricted quantification.

P: Hmm, restricted quantification, you say?

D: You shouldn't mind; we all have to interpret some sentences as employing restricted quantifiers. When I interrupt this little exchange to tell you 'there's no beer', it is obvious that my domain of quantification is restricted to the fridge.

P: Fair enough. But I hope your beer sentence was just an example; I'm thirsty.

DKL and PvI have both advanced views with significant weaknesses, views that are subject to obvious objections. They're both prepared to defend their views against those objections. And they're each prepared to concede that the defenses mounted by the other (e.g., paraphrase and appeal to restricted quantification) have something going for them. The characters in our dialogue have reached something of an impasse. What's to say? Here is a modest proposal: reflection on the Cases is helpful; and it is especially helpful for philosophers that are dialectically positioned as DKL and PvI are.[25]

Metaphysical theories that explain phenomena (including our intuitions about particular cases) are, *ceteris paribus*, better than those that do not. And metaphysical theories that give *unified* explanations are better still. Some answers to SPEC can provide a unified explanation to each of the Cases; others cannot. On universalism, there are–or could be–person-pairs, nations, swarms of bees, and rarified brains. So the universalist does not, *qua universalist*, have any unified explanation of the Cases. Organicism, causalism, disjunctivism, ordinary brutalism (maybe), and nihilism fare better. For they can explain why we're convinced that in each of the Cases there is no thinking (no consciousness, no understanding); it is because there is nothing there to think. This explanation is not, of course, conclusive evidence for a restricted theory of composition. But it does show one dimension on which these theories fare better than universalism.

[25]I shall not in this paper argue that David K. Lewis and Peter van Inwagen are dialectically positioned just as DKL and PvI are. But that suggestion is plausible, I think; for similar stalemates appear in many of the debates relevant to SPEC (e.g., the possibility of gunk and the bearing of vagueness on composition).

Here's another way to put the point. Debates about composition are tricky. The various participants are entrenched, unlikely to change their minds, and perhaps rationally so. If anything might resolve these debates, it is the introduction of a new kind of evidence. If I'm right about the Cases, they constitute a new kind of evidence in debates about Spec. I come, then, as the bearer of good news. For news of *any new evidence* is good news.

I have sketched two moves. First, I have suggested that materialists about human persons, functionalists about mentality, and believers in the possibility of A.I. have a new defensive move to make. They may marshall moderate theories of composition to resist arguments turning on the Cases. Second, defenders of moderate theories of composition may deploy the Cases on behalf of their theories. It would perhaps be dialectically dubious for one and the same philosopher to make both of these moves. But that point is consistent with my claim that each move is individually interesting and plausible.

5. Objections and Replies

5.1. Composition Doesn't Explain Enough

Objection: you have given an explanation of the Cases ('Composition'). Your explanation is that there's no consciousness in the Cases because there's nothing there to be conscious. But there are other thought experiments in the philosophy of mind that Composition cannot explain:

Zombies. Possibly, I have a physical duplicate that is not conscious; there is nothing it is like to be that zombie. So the mental does not supervene on the physical.[26]

Mill. Consider a human organism–or perhaps just a brain–as a mill, enlarged so that one might fit inside it. Step into the mill, into the brain; examine its inner workings, its cogs and systems. Look carefully at the neurons as you would the pulleys and planes of a mill. The cogs will turn and the neurons will fire, but that's all that is happening. Try as you might, you will not find consciousness anywhere. So human brains or organisms are not conscious.[27]

[26]See, e.g., Chalmers (1996).

[27]See, e.g., Leibniz (1979): 536. I have suggested that Leibniz' thought experiment involves an *inference* from facts about the parts (cogs, neurons) to facts about the whole (mill, brain); whether this suggestion is charitable is a matter I shall not address.

What's more (my objector continues), there are theses that give a unified explanation of the Cases, Zombies *and* Mill. For example: that consciousness is impossible. Or: that only mereologically simple beings can be conscious. Or: that nihilism is true. Or: that consciousness is inscrutable, something about which we aren't cognitively equipped to think. This tells against your project; for there are explanations superior to Composition.

Reply: that Composition cannot explain the results of *all* thought experiments in the philosophy of mind doesn't conclusively tell against my argument. For it still neatly explains *some* cases. Furthermore, Composition has, I think, more prior plausibility than the theses that consciousness is impossible or that nihilism is true or that only simple beings may be conscious.[28] Anyone with a prior commitment to our own consciousness and our own materiality, at least, has reason to prefer Composition over these other explanations. Finally, it's not clear that these explanations are inconsistent with Composition. So citing them doesn't immediately tell against my argument.

5.2. Composition Doesn't Explain At All

Objection: Composition isn't an explanation, it's a cheat. Consider the problems with analyzing 'S knows that p' as 'S truly and justifiably believes that p'. Once upon a time, Gettier and his followers proposed counterexamples to that analysis. One goes like this: a person sees what is apparently a sheep (but what is in fact a wolf in sheep's clothing) on the hill and believes that there is a sheep on the hill. There is, in fact, a sheep on the hill (beyond the person's sight, say); so the belief is true. But it is not known. So some justified, true beliefs are not known. Something is missing in the analysis: but what could it be? What is the fourth condition? What necessary condition on knowing goes unsatisfied in the case such that its going unsatisfied explains why the person doesn't know? This is the Gettier problem.[29]

It would be an odd philosopher who replied to Gettier with this speech: 'I am an eliminativist about mental subjects. There are not (and could not be) such things as persons. This is what explains why the person does not know that there's a sheep on the hill. The person does not know, because she does not exist. And things that don't exist can't know. Problem solved!' This is hardly an explanation. But, my objector continues, Composition is an 'explanation' much along these lines. It 'explains' the Cases only by depriving them of a subject matter. And that is entirely unsatisfying.

[28] See my (manuscript) for more on this.

[29] Thanks to [omitted] for discussion on this point and to [omitted] for the Gettier analogy.

Reply: I'll offer two analogies of my own to show that *explanation by elimination* is a reasonable philosophical tool. Reflection shows, I think, that the moves I have advocated are more similar to the following cases than they are to the reply to Gettier envisioned by my objector.

Example 1: there are puzzles associated with the Christian doctrine of the Trinity. It is difficult to explain how it could be that there is one God in three persons. There are data (the various theses that comprise the doctrine) that apparently imply a contradiction. What's to say? The atheist has an easy answer. The puzzles of the Trinity are to be resolved by elimination. There are no gods, and hence no one God in three persons.

Example 2: there are puzzles associated with vagueness. On the one hand, it seems that there could be no fact about the matter about whether some atom is part of, say, a statue. On the other, it's odd that a thesis like *the atom is part of the statue* could fail to be metaphysically settled. What's to say? Those who deny the existence of composite objects–mereological nihilists–have an answer. They can explain–or explain away–this puzzle. Says the nihilist: there aren't–and couldn't be–any composite statues. All theses presupposing otherwise are false. Nihilism thus solves all such puzzles of vagueness, and *this counts in its favor*.

There are puzzles associated with the Cases. For example: if functionalism about consciousness is true, then it certainly *seems* that there should be consciousness in cases like Blockhead. What's to say? Proponents of the cases suggest that the functionalist must abandon her functionalism. But I suggest another way out. She may instead deny something presupposed by the proponent of the case. She may deny that there are–or could be–such things as Chinese rooms. There is, I grant, something thin about explanation by elimination. But it is a new move for the functionalist to make; and as I've argued above, it's a move that has a lot going for it.

5.3. Composition is a Shallow Explanation

Objection: perhaps Composition explains the Cases (in some thin sense of 'explains'). But there is an important datum in the neighborhood that Composition can never explain: if there were (*per impossible*!) such things as person-pairs, Chinese rooms, and the like, they would not be conscious. Thus, Composition is a shallow explanation at best. Materialists, functionalists, A.I. theorists, and the like had best look elsewhere if they wish to give a plausible reply to the Cases.[30]

[30]Thanks to [omitted].

Reply: I appeal to my previous analogies. Consider again the puzzles of vagueness and the mereological nihilist's solution to them. Suppose we were to reply to the nihilist: 'yes, but if there *were* statues, then there *would* be puzzles of vagueness, so your nihilism has explained nothing!' This reply is unpromising. Nihilism in fact solves these (though perhaps not all) problems of vagueness. If nihilism were false, there would be puzzles in need of resolution. But that is just to state of the nihilism's primary advantages: it avoids such puzzles. Similar if the Triune God of the Christian were to exist, there *would* be puzzles in need of resolution. But this doesn't somehow tell against atheism as a resolution of those puzzles. It is just to state one advantage of atheism.

5.4. Hive Minds are Possible

Objection: your argument assumes that mentality is distributive. But it isn't. Hive minds, after all, seem possible. Haven't you watched a Borg episode of Star Trek? The assumption that thought requires a thinker is antiquated and Cartesian, hardly fit for we post-Hume philosophers familiar with the Borg and their ilk.

Reply: I have no argument, only a confession. If it really seems that hive minds are possible, then adopting Composition for the reasons I give has a price attached. But for what it's worth, I pay the price willingly, for *I* can't see that hive minds are possible. I'm not even sure I can conceive of them. I can conceive of some things that together think, but the only way I can conceive of them doing this is by composing something that thinks. I cannot shake myself free of the conviction that every thought has something thinking it, some one thing to whom the thought belongs. If you disagree with me on this point, I invite you to read me as arguing for the conditional: if mentality is distributive, then Composition best explains the Cases.

Objection: the hive mind objection can be sharpened. The problematic assumption isn't so much that thoughts need thinkers. Rather, it's that thinkers are *prior to* thoughts. This suggests that there's a theater in which thoughts dance, all the while being watched by a 'self'. That is the way of the Cartesian, and it is a mistake.[31]

Reply: in assuming that every thought has a thinker, I take on no commitments about priority relations between the two. Perhaps thoughts are, somehow, more fundamental than or prior to thinkers (whatever that may mean). Perhaps not. Nor do I take on any commitments about Cartesian theaters. All that matters for my argument is that thoughts need thinkers. Some things can together

[31]See Dennett (1994): 540 for more against such 'Cartesian Materialism'.

think only by composing something that thinks.

5.5. Intuitions about Composition and Consciousness do not Covary

Objection: if Composition is a good explanation of the Cases, then our intuitions about, say, consciousness, and about composition should covary. But they don't. Here's one way to draw this out. If universalism is true, then composition happens in every case. Suppose so. So in Chinese Room, there *is* a system, something composed of the subject, the papers, the instructions, etc. But even on the supposition that universalism is true and that *there is a system there to understand*, it seems still that there is no understanding. It seems then, that it really is intuitions about mentality that drive our reactions to the Cases, and not just intuitions about what there is.

Reply: I'm not sure what to say. But here are two replies. First, an interesting covariance *can* be shown. Take an ordinary human organism. Such a thing *might* be conscious. Only those committed to the falsity of materialism about human persons should disagree with me on this point. Now consider a series of changes whereby the organism is rarified, its parts slowly drifting from each other (this is, essentially, the Rarified Brain case discussed above). Somewhere along the series we will be inclined to say that the thing couldn't be conscious. Where in the series? It's a series, so it's hard to say. But this suggestion strikes me as good as any: when the thing's parts cease to display the functional unity and cooperation typically displayed by an organism's parts–when it ceases to be an organism, in other words. If that's right and if Organicism is true, then consciousness stops when composition does.[32]

Second, I'm not confident that we take seriously the gerrymandered objects in which universalists believe. We may nod along when asked to suppose that universalism is true and hence that there could be, say, fusions of cabbages and kings–or person-pairs. But can we *really* countenance such things? Are we equipped to clearly think of them? I confess that I, at least, have difficulty so doing. I *think* I understand what such objects would be and can list their parts and some of their properties. But something about them remains obscure. This kind of obscurity is perhaps common in metaphysics. Questions like 'what would the world be like if universalism were true?' are obscure in just this sense (unlike, say, questions like 'what would the world be like if I were able to fly?').

When my objector notes that even on universalism, things like person-pairs

[32]A similar moral could be drawn from a similar story, that of an organism's parts slowly being replaced with synthetic duplicates (the story begins with a conscious organism, proceeds through a cyborg stage, and concludes with an unconscious android).

couldn't be conscious, I'm inclined to agree. But it's not clear that my agreement issues from convictions about mentality (consciousness, understanding) rather than ontology. Perhaps evolution has not gifted us with the cognitive equipment to think about the gerrymandered objects of the universalist. So it should come as no surprise that try as we might, we are still convinced that such objects couldn't be conscious or whatnot.[33]

6. Conclusion

If there's something wrong with the Cases, it is this: composition has been ignored. Philosophers of mind could learn from metaphysicians, and perhaps even deploy theories of composition in defense of their theories about mentality. But metaphysicians, too, have something to learn. For the Cases give us data, data that can help adjudicate disputes over when composition occurs. What an irenic suggestion!

[33]There's a lot more to say on this point. Another notion worth pursuing: *natural boundaries*. Cameron (2007): 276 explains: 'If a white sheet of paper has a red circle on it, both the paper and the circle have a natural boundary. In contrast, the red circle whose centre is the same as that of the main circle, but whose radius is, say, half as great, does not have a natural boundary: this inner circle does not contrast with its surroundings in the appropriate way.' Here's a suggestion: even if there were such things as systems or person-pairs or rarified brains, their borders wouldn't be natural. Combine this move with the thesis that only naturally bounded objects can bear mental properties, and we have an explanation of the Cases that's not that far from Composition.

References

[1] Author. (manuscript).

[2] Barnett, David (2008). 'The Simplicity Intuition and Its Hidden Influence on the Philosophy of Mind' *Noûs* 42: 308-355.

[3] _____ (2010). 'You Are Simple' in *The Waning of Materialism* (George Bealer and Robert Koons, eds). Oxford: Oxford University Press.

[4] Block, Ned (1978). 'Troubles with Functionalism', in C.W. Savage (ed.) *Perception and Cognition*. Minneapolis: University of Minnesota Press.

[5] Cameron, Ross P. (2007). 'Subtractability and Concreteness'. *The Philosophical Quarterly* 57: 273-279.

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

[7] Cole, David (2004). 'The Chinese Room Argument', *The Stanford Encyclopedia of Philosophy*, URL = http://plato.stanford.edu/archives/fall2008/entries/chinese-room/.

[8] Dennett, Daniel (1981). 'Where am I?' in his *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge: MIT Press.

[9] _____ (1994). 'Get Real: Reply to My Critics', *Philosophical Topics* 22: 505-568.

[10] Dorr, Cian (2005). 'What We Disagree About When We Disagree About Ontology', in *Fictionalism in Metaphysics*, Mark Kalderon (ed). Oxford: Oxford University Press.

[11] Elder, Crawford (2005). *Real Natures and Familiar Objects*. Cambridge: The MIT Press.

[12] Fischer, John Martin (1994). *The Metaphysics of Free Will*. Oxford: Blackwell.

[13] Gilmore, Cody (forthcoming). 'Sider, the inheritance of intrinsicality, and theories of composition', *Philosophical Studies*.

[14] Hoffman, Joshua and Gary Rosenkrantz (1997). *Substance: Its Nature and Existence*. London: Routledge.

[15] Horgan, Terrance and Potrc, Matjaz (2008). *Austere Realism: Contextual Semantics Meets Minimal Ontology*. Cambridge: MIT Press.

[16] Johnston, Mark (1987). 'Human Beings', *Journal of Philosophy* 84: 59-83.

[17] Leibniz, Gottfried (1979). *Leibniz Selections* (Philip P. Weiner, ed.). New York: Charles Scriber's Sons.

[18] Lewis, David K. (1986). *On the Plurality of Worlds*. Oxford: Blackwell.

[19] Korman, Daniel Z. (2007). 'The Naive Conception of Material Objects: A Defense'. Dissertation, University of Texas at Austin.

[20] Markosian, Ned (1998). 'Brutal Composition', *Philosophical Studies* 92: 211-249.

[21] Merricks, Trenton (2001). *Objects and Persons*. Oxford: Oxford University Press.

[22] Olson, Eric T. (1997). *The Human Animal: Personal Identity Without Psychology*. Oxford; Oxford University Press.

[23] _____ (2007). *What Are We?*. Oxford: Oxford University Press.

[24] Parfit, Derek. (1971). 'Personal Identity', *Philosophical Review* 80: 3-27.

[25] Putnam, Hilary (1967). 'The Nature of Mental States', in N. Block (ed.), *Readings in Philosophy of Psychology*. Cambridge: Harvard University Press.

[26] Rea, Michael C. (1998). 'In Defense of Mereological Universalism', *Philosophy and Phenomenological Research* 58: 347-360.

[27] Searle, John (1980). 'Minds, Brains, and Programs', *The Behavioral and Brain Sciences* 3: 417-424.

[28] Shoemaker, Sydney and Richard Swinburne (1984). *Personal Identity*. Oxford: Blackwell.

[29] Sider, Theodore (2001). *Four-Dimensionalism*. Oxford: Oxford University Press.

[30] _____ (manuscript). 'Against Parthood', presented at the Pacific APA, 2009, URL = http://tedsider.org/papers/nihilism.pdf

[31] Thomasson, Amie (2008). *Ordinary Objects*. Oxford: Oxford University Press.

[32] Unger, Peter (1990). *Identity, Consciousness, and Value*. Oxford: Oxford University Press.

[33] van Inwagen, Peter (1990). *Material Beings*. Ithaca: Cornell University Press.

[34] _____ (1997). 'Materialism and the Psychological-Continuity Account of Personal Identity'. *Philosophical Perspectives* 11: 305-319.

[35] Zoboff, Arnold (1981). 'The Story of a Brain', in D. Dennett and D. Hofstadter (eds.), *The Mind's I*. New York: Basic Books.