

Dave Beisecker

Zombies, Phenomenal Concepts, and the Paradox of Phenomenal Judgment

Abstract: *This paper explores the viability of rejecting a largely unchallenged third premise of the conceivability argument against materialism. Fittingly labeled 'type-Z' (for zombie), this reply essentially grants to the zombie lover, not just the possibility of zombies, but also their actuality. We turn out to be the very creatures Chalmers has taken such great pains to conceive and more conventional materialists have tried to wipe off the face of the planet. So consciousness (at least for us) is a wholly material affair. What is conceivable but non-actual are not zombies, but rather 'angelic' beings possessing an acquaintance with supermaterial phenomenal states. After showing how Chalmers' recent discussion of the phenomenal concepts strategy should incline those pursuing such a strategy toward a type-Z response, this paper relates type-Z materialism to similar replies that Chalmers has found 'hard to classify' and closes with a brief remark about how a type-Z materialist might reply to the knowledge argument.*

I. Introducing Type-Z Materialism

The conceivability (or zombie) argument against materialism starts from the apparent conceivability of a world that is a physical duplicate of ours but that altogether lacks qualitative consciousness. From there it proceeds in steps to the falsity of materialism. If, following

Correspondence:
D. Beisecker, Department of Philosophy, University of Nevada – Las Vegas,
4505 Maryland Pkwy., Box 455028, Las Vegas, NV 89154, USA.
Email: beiseckd@unlv.nevada.edu

Journal of Consciousness Studies, 17, No. 3–4, 2010, pp. ??–??

convention, we think of P as standing for the conjunction of all the physical or material facts that obtain in our own world and Q as some arbitrary phenomenal fact (such as 'I am conscious'), then the argument has the following canonical form:

- (P1) The conjunction of P & \sim Q is conceivable.
- (P2) If (P1), then the conjunction of P & \sim Q is possible.
- (P3) If P & \sim Q is possible, then materialism is false. So
- (C) Materialism is false.

A world in which P & \sim Q holds is typically thought of as a *zombie* world: one in which the distribution of material facts is just like ours, but in which some or all of its denizens altogether lack phenomenal consciousness. Thus the first premise is often glossed as asserting the conceivability of zombies, the second as inferring from this their possibility, and the third taking the possibility of zombie worlds to entail the falsity of materialism.

The conceivability argument has had no shortage of detractors, and on the face of it for good reason. According to the argument zombies are remarkably potent creatures, the bare conceivability of which is taken by proponents to spell the end of the world for a substantive metaphysical thesis. It wouldn't seem at all crazy to think that such powerful beings are utterly chimerical or (to use Dennett's turn of phrase) 'unimaginably preposterous.' It simply seems weird to hold that the contours of our imagination would be constrained by (or to constrain) the metaphysical limits of the way things actually are. Accordingly, most materialist opponents of the argument have seen fit to challenge either the first or the second premise. In his well-known taxonomy of responses to the conceivability argument, Chalmers dubs those who claim that despite appearances to the contrary, the conceivability of such powerful creatures is utterly fantastic 'type-A' materialists, while he calls those who balk at the second premise 'type-B' materialists. Though granting their conceivability in some complete but attenuated sense, type-B materialists deny that this conceivability is robust enough to license their bona-fide metaphysical possibility.

Curiously, the third premise of the argument has gone largely unchallenged, so much so that Chalmers hasn't even dignified its rejection with its own letter. Indeed, several presentations of the argument omit this premise altogether, as if its defense doesn't even merit discussion. The intuitive idea is that if we are compelled to accept the bona-fide metaphysical possibility of zombies, then the totality of physical facts (obtaining in our world) would seem to be compatible

with either the truth of some phenomenal claim or its negation. If so, then it would seem that the totality of physical or material facts could not settle or entail phenomenal facts, and a complete material characterization of the world would fail to catalog all that there is. Proponents of the argument take this to suggest that there must be an additional dimension of reality above and beyond the mundanely material — an immaterial or ‘super’-material¹ realm that eludes ordinary physical characterization or description. Thus, to deploy the Leibnizian metaphor Kripke made famous, it would seem to follow that once God determined the physical or material makeup of the world that he decided to actuate, he would then have had to have done something *in addition* to settle the phenomenal facts that evidently obtain in our world.

However, things cannot be as simple as all that, for the truth of materialism is actually quite compatible with the metaphysical *possibility* of supermaterialism. One can be a materialist without demanding that the totality of material facts *entail* the phenomenal facts. That is, materialism doesn’t have to carry in its wake strong modal implications. To see this, consider what we should say if, continuing the Kripke-Leibniz metaphor, God didn’t get around to settling all of the mundanely material facts of the world to be actualized until late Saturday, and knowing that Sunday was to be a day of rest, had simply stopped there *without adding anything in addition*. What we would find ourselves living in is, in effect, a zombie world: a thoroughly material world from the perspective of which the possibility of supermaterialism is robustly conceivable (and indeed deemed actual by zombie-Dave Chalmers!). So let me cheekily propose the letter ‘Z’ as a label for those who suspect the weak link in the zombie argument to be in the vicinity of the third premise (P3). I’ll explain later why I just said that the type-Z objection is lodged ‘in the vicinity’ of that third premise. The important bit is that the type-Z materialist is one who grants the conceivability *and* the possibility of zombies, yet still maintains that we live in a thoroughly material world (or what Stalnaker, 2002, calls a ‘z-world’). That is, the type-Z materialist will ultimately agree with the supermaterialist about the space of

[1] ‘Super-materialist’ is my preferred term for those who follow Chalmers and accept that the zombie argument (or the knowledge argument) demonstrates that there must be more under heaven and earth than is countenanced by (mundane) materialism, for the moniker nicely captures an ambivalence in how the position is understood. To their materialist opponents, supermaterialism advocates us to accept the existence of spooky, *supernatural*, *non-material* features of the world, while supermaterialists themselves think they are simply urging us to acknowledge an underlying, intrinsic facet of our material existence, which has heretofore eluded systematic scientific investigation.

conceivability and possibility, yet nevertheless deny that the argument carries with it any substantial metaphysical implications, for when all is said and done, there is no legitimate reason why we couldn't turn out to be the very kinds of creatures the possibility of which Chalmers has taken such great pains to establish (and that his more conventional materialist foes have taken to eradicate).²

II. Type-Z and The Paradox of Phenomenal Judgment

The type-Z response calls upon us to take seriously a long-recognized creepiness about zombies, which Chalmers has labeled 'the paradox of phenomenal judgement' (Chalmers, 1996, ch. 5). By supermaterialistic lights, nothing anyone can say will convince them that they are zombies, or that they lack the consciousness of 'regular' (that is, *actual*) folk. And there is nothing anyone can *say* in defence of the claim that we are *not* zombies that zombies themselves wouldn't find equally convincing. They harbor the same zombic hunches that we do, and they find the hard problem of consciousness every bit as hard as we do. By design, the apparent compellingness of zombie arguments will not distinguish us from them. Perhaps stranger still, it would even seem possible for *non-zombies* to learn how to master phenomenal concepts and come to talk effectively about their conscious experience *from zombies*. For a supermaterialist, then, either zombies are horribly mistaken about the nature of their own conscious experience, or their phenomenal concepts turn out — perhaps unbeknownst to them — to apply to something very different from that to which supermaterialists take our own phenomenal concepts to apply, if indeed they apply to anything at all.

This paradox of phenomenal judgment points to the fact that there are two distinct perspectives and ways of talking in play in the conceivability argument: that of the zombie and that of the non-zombie. The conclusion of the zombie argument is true for the latter, but not for the former. Thus, the whole question about the truth of

[2] To be sure, this option has occasionally been entertained, only to be discarded almost as soon as proposed. Witness Dennett's (1991) playful comment that perhaps the most satisfying response to the zombie argument is to conclude that zombies are 'not just possible. They're actual. We're all zombies.' (p. 406). However, Dennett declines to treat this option seriously and warns us in a footnote that 'it would be an act of desperate intellectual dishonesty to quote this assertion out of context!' Instead, his reasoned response to the zombie menace is a thoroughly conventional Type-A reply (and as such direct opposed to the one explored here); zombies must be stamped out as 'unimaginably preposterous.' See also Dennett (2008). In effect, Dennett is committed, not to the explanation of the explanatory gap (in accordance with the phenomenal concepts strategy), but rather to explaining it *away*.

materialism for us turns on whether our own *way of talking* accords with that of a zombie or that of a non-zombie with an appropriate contact or acquaintance with immaterial properties of consciousness. The trouble is that by design, we just can't *say* all that much about what it is with which we are supposedly acquainted, and which is supposed to inoculate us against zombification. More and more philosophers are now responding to this creepiness by exploring the zombie way of talking about their apparently subjective feelings, and at least a few of them have found this perspective a great deal more congenial than supermaterialists typically allow. Zombies have gotten a bad rap; being one wouldn't be such an awful fate after all (see, for instance, Moody, 1994; Balog, 1999; Skokowski, 2002; Stalnaker, 2002; Lynch, 2006; Frankish, 2007; Brown, CO1).

In short, then, the position the type-Z materialist tries to stake out as *actual* is just like the one that is envisioned and granted to be *possible* by proponents of the conceivability argument: that an explanatory gap between the phenomenal and the material exists and nevertheless there could be creatures who talk about consciousness much like we do yet turn out not to have any sort of contact with super-material properties (or to have 'genuine' consciousness *as supermaterialists understand it*). The type-Z materialist thus finds awkward and contrived any attempts by supermaterialists to close off this admittedly possible scenario as somehow impossible after all. For those attempts only play into the hands of their type A or B opponents who've suspected all along that there is something deeply wrong with the very idea of a zombie. Instead, the claim that we are *not* zombies has to rest on some contingent fact, the truth of which we have some independent justification for holding, justification which the type-Z materialist doubts will ever be forthcoming.³ The type-Z materialist claims, then, that whatever it is like to be me just might be equivalent to whatever it is like to be one of those things Chalmers and his ilk envision as a 'zombie'. Call them what you like; I just might be one of *them*. By

[3] This position is very close to Stalnaker (2002). However, there is a slight twist. Stalnaker agrees that the supermaterialist has stacked the deck by illicitly framing the zombie argument in terms that aren't appropriately topic-neutral. If one gives the notion of 'phenomenal consciousness' an appropriately topic-neutral reading, then Stalnaker concludes that the notion of a zombie is ultimately incoherent, a conclusion that makes Stalnaker look more like a type-A materialist. By contrast, my version provides topic-neutral readings of *both* the notions of a zombie *and* — granting the supermaterialist's understanding of the term — that of phenomenal consciousness, with the consequence that I arrive at the conclusion that we must be zombies. Although these positions might just be semantic variants of one another, I submit mine to be rhetorically more robust, since it grants more to the supermaterialist.

supermaterialists' lights, then, (though not the type-Z materialists'!), there is nothing it is like to be me.

III. Type-Z and the Phenomenal Concept Strategy

I shall have more to say about how to flesh out the Type-Z response further. However, my initial aspirations will be more modest. For starters, I'd like to suggest that Chalmers' recent discussion of the phenomenal concepts strategy should actually steer stalwart materialists toward a type-Z position.⁴ Roughly, *the phenomenal concepts strategy* is that of giving a materialistic explanation, not of consciousness itself, but rather of what Chalmers calls 'the next best thing,' the explanatory gap. Those who adopt such a strategy are likely to endorse those intuitions about zombies and colour-deprived Mary that are so suggestive of an explanatory gap between the mental or phenomenal and the more mundanely material. However, unlike those who take these considerations to suggest that there is an additional supermaterial realm that eludes ordinary physical description, the phenomenal concept strategist holds that the epistemic gaps revealed by the zombie and Mary intuitions need not ramify into an ontological one. Rather, the modal slippage or flexibility we observe between phenomenal and physical descriptions is due to peculiarities in the *concepts* with which we *talk* or *think* about phenomenality.

The phenomenal concepts strategy is usually associated with materialisms of the type-B sort. Since type-B materialists deny that the admitted conceivability of zombies entails their genuine possibility, their supermaterialist opponents have typically tried to expose an inherent instability in the type-B position by questioning whether one can genuinely conceive of something which turns out not to be possible after all. In reply, type-B materialists naturally draw upon those portions of Kripke's work suggesting that not everything that seems conceivable will turn out to be a genuine possibility. The identity between water and H₂O is necessary, despite the fact that one can conceive of worlds in which those two concepts come apart. In turn, supermaterialists insist that the analogy between consciousness and water ultimately breaks down, because it is apparent that we can conceive of the possibility that consciousness is not a material state in a more direct fashion than we can conceive of the possibility that water is not H₂O. When we conceive of a world in which water is not H₂O, what we are *really* imagining strictly speaking is a world in which

[4] See Chalmers (2007), which shall hereafter be abbreviated as PCEG.

either the concept of water or that of H₂O applies to something other than water. In the case of phenomenal concepts, however, no such room is open; we cannot help but conceive of a pain as such. That is, phenomenal concepts cannot be pried apart from their referents in the same fashion as material ones.⁵

The thing we find so peculiar about the phenomenal realm is the special epistemic position or privilege we have (or at least sometimes seem to have) with respect to our own phenomenal states, a position that Chalmers frequently glosses as *acquaintance*. Unlike thoughts about the rest of reality, one's sincere thoughts or avowals about how things appear are especially resistant to third-person challenge. In the phenomenal realm, no epistemic gap opens up between appearance and reality. While you might be mistaken about the *actual* colour, taste or smell of the fruit in front of you, typically you cannot question what colour, taste or smell it *appears* to you to have. In short, the qualitative character of one's own experiences seems somehow self-intimating, suggesting perhaps that those aspects are somehow constitutive of the concepts bound up in our thoughts about conscious experience.⁶

Chalmers doesn't object to the idea that phenomenal concepts need to given a different sort of analysis from 'ordinary' concepts. Rather, he suspects that any such analysis will not be available to a thorough-going type-B materialist — at least not to one who takes the materialism part seriously. For merely pointing to a difference between phenomenal and material concepts does not let a materialist off the hook. Any materialist pursuing the phenomenal concept strategy further needs to tell a materially acceptable story about how we could ever so much as come to *grasp* phenomenal concepts so that we could be capable of entertaining genuine, self-intimating phenomenal thought.⁷ This, I take it, is the launching point for the master argument Chalmers has recently raised against the strategy. While I think Chalmers' considerations do indeed expose an instability in conventional type-B positions, it ultimately shows that the phenomenal concept strategist should instead 'go zombie (Z)' on Chalmers.

Chalmers begins in his typical fashion with a conceivability exercise intended to ascertain whether a candidate phenomenal concept strategist thinks that our grasp of phenomenal concepts is itself

[5] Kripke himself took this tack, and we have Chalmers to thank for unpacking and resurrecting Kripke's dense comments on the subject.

[6] It's this very epistemic peculiarity of the phenomenal that ultimately accounts for Chalmers' inability to drive a wedge between the primary and secondary intensions of our phenomenal concepts.

[7] Levine (2007) calls this the 'materialist constraint'.

materially explicable: is it conceivable for a being to be physically just like us, and nevertheless lack the psychological characteristics that allow us to grasp phenomenal concepts? Depending upon the answer provided, Chalmers constructs the following dilemma (2007, p. 174):

- (1) If it is conceivable for a being to be physically just like us and yet lack the psychology to grasp phenomenal concepts, then this ability itself would not seem to be materially explicable. The phenomenal concept strategist would then be forced into conceding that the material facts cannot account for something that is evidently psychologically significant about us: our possession of phenomenal concepts.
- (2) But on the other hand, if the phenomenal concept strategist avows that it is inconceivable for a being to be physically just like us and to lack the psychological ability to entertain genuine phenomenal thought, then it would seem that zombies would also be capable of harboring phenomenal thought. But that would put us and zombies on the same epistemic plane, and for Chalmers, that's just plain crazy talk.

Now I'm not convinced that a proper-thinking materialist needs to be particularly concerned about being gored by either of these horns, for the whole dilemma is premised on the idea that we could not turn out to be zombies. The first horn is troublesome only if we insist that our own phenomenal concepts are inflated beyond those of a zombie, while the second horn is a problem only under the assumption that our epistemic situation is richer than a zombie's. Either way, Chalmers' dilemma is going to impale only those materialists who are wed to the idea that a zombie's conceptual or epistemic position must differ from our own. And that is precisely what a materialist of the type-Z stripe is prepared to deny. Such a materialist maintains that we ourselves happen to be in the same epistemic and conceptual position as Chalmers envisions zombies to be in. We have learned how to embrace explanatory gap intuitions, yet without the 'benefits' of, or any contact with, full-blooded super-material conscious experience. Consciousness (at least for beings like us) is a wholly material affair. Instead, what is conceivable, but most likely non-actual, are not *zombies*, but rather 'angelic' beings possessing an acquaintance with *super*-material phenomenal properties.

IV. The Type-Z Diagnosis of the Conceivability Argument

The suggestion that we just might turn out to be the subjects of Chalmers' thought experiments is apt to be met with incredulous, boggle-eyed stares. Doesn't the proper conception of a zombie rule out this possibility? Here, in defence of the second horn of the dilemma, is how Chalmers tends to express his reaction to the type-Z suggestion that ours just might be (or be equivalent to) a zombie world:

Here, the natural response is that this scenario is simply not what we are conceiving when we conceive of a zombie. Perhaps it is possible to conceive of a being with another sort of state — call it 'schmonsciousness' — to which it stands in the same sort of epistemic relation we stand in to consciousness. Schmonsciousness would not be consciousness, but it would be epistemically just as good. It is by no means obvious that a state such as schmonsciousness is conceivable, but it is also not obviously inconceivable. However, when we ordinarily conceive of zombies, we are not conceiving of beings with something analogous to consciousness that is epistemically just as good. Rather we are conceiving of beings with nothing epistemically analogous to consciousness at all.

Put differently, when we conceive of zombies, we are not conceiving of beings whose inner life is as rich as ours, but different in character. We are conceiving of beings whose inner life is dramatically poorer than our own (Chalmers, 2007, p. 186; see also p. 177).

In short, then, Chalmers tells us that a being that is epistemically equivalent to us just doesn't match our intuitive conception of what a zombie is. We cannot be zombies, because by definition zombies are just like us, *except* that they must *lack our conscious experience*, and this consigns them to a dramatically less-accurate self-conception than genuinely conscious beings. While I believe (indeed, *know*) that I am conscious and have states with remarkable qualitative character available to introspection, the corresponding thoughts of a zombie wouldn't really be true.⁸ Thus the intuitive resolutions of debates about consciousness *ought* to go differently in a zombie world than they do in ours (Chalmers, 2007, p. 178). By way of illustration, Chalmers invokes a hypothetical debate between zombie eliminativists and realists, but we can just as well consider a genuine Zombie-Dave and the belief he holds to the effect that he is not a zombie. While the 'real' Dave Chalmers (in whatever possible angelic world that might be) knows this, poor, unfortunate Zombie-Dave cannot, simply because it's not true. Hence, even if we are kind enough to allow some

[8] Chalmers (2007), p. 179. Note that I've avoided characterizing these zombic beliefs as *false*, for they might instead turn out to be empty.

of a zombie's phenomenal belief analogues to be true, zombies are bound to be in a different epistemic position from 'real' folk.

To be sure, Chalmers treads a fine line here. If zombies really are as introspectively impoverished as he suggests, then that threatens to make the explanation of their beliefs that correspond to our beliefs about consciousness utterly mysterious. It also makes their epistemic practices appear remarkably off the mark. Why wouldn't their lack of justification for their apparently introspective beliefs show up for them, thereby rendering them disinclined ever to engage in introspective reports? But if they do turn out to be remarkably less inclined to introspection than 'fully-conscious' beings, then that only plays into the hands of more conventional type-A materialists, who have argued all along that on serious consideration, a zombie's remarkable incapacity for self-reflection will ultimately ramify into serious behavioral or linguistic deficits. Chalmers seems perfectly prepared to bite this bullet, telling us that all sorts of strange things can happen in possible worlds.⁹ However, that remark will be satisfying only if we can rest assured that this strangeness has indeed been sealed off or quarantined from the actual world. And the type-Z materialist suspects that any attempt to hold the zombies at bay from this world will only prove to beg the question in favour of supermaterialism.

The type-Z complaint, then, is that the supermaterialist has illicitly stacked the deck from the very outset by deploying a conception of a zombie that smuggles in an illicit comparison to us, and consequently begs the question in favor of supermaterialism. Consider, for instance, Chalmers' more recent and careful presentation of the conceivability argument (Chalmers, 2009, pp. 314–15). There he recognizes that in order to ensure that the conceived world will not turn out to be the actual world, he needs to add another premise, (we can think of it as P0), to the effect that P & Q is true of this the actual world. So the world which the conceivability argument asserts to be conceivable in the first premise (namely P & ~Q) cannot be the actual one. But that only does part of the trick, for there is a subtlety in how Q needs to be characterized. If it is characterized in the way that is applicable to a zombie's 'conscious' (or perhaps 'schmonscious') states, then ~Q might well be true of an angelic, *supermaterial* world, and the conceivability argument would basically amount to a 'reverse-zombie' argument. In order to block this possibility, Chalmers further stipulates that Q must be given a 'positive' characterization: that is, it

[9] See Chalmers (1996), p. 180, and also section 4 (5) of the expanded (online) version of Chalmers (2009).

must hold in all worlds ontologically equivalent to or richer than our own. Thus any world in which $\sim Q$ is true must be one which *lacks* something contained in ours.

It should be clear that all the work is being done by this stipulated characterization of Q . From the very get-go, this characterization has the effect of closing off the very possibility the type- Z materialist insists might be actual. Insofar as zombies are defined as *both* different from us *and* as lacking any contact or acquaintance with supermaterial qualia, the supermaterialist presupposes that our conscious experience must be something above and beyond the mundanely material. However, as the ontological argument for the existence of God shows, seemingly innocuous definitions can be far from such. In order to prevent their issuing suspiciously ampliative inference tickets, we need to be reassured that they are metaphysically conservative. And from the type- Z perspective, this one evidently fails that test, because it illicitly centers the actual world off of a materialistic one. If we are to be appropriately open-minded about the nature of consciousness, it should take more than mere stipulation to rule out the admittedly-conceivable hypothesis that ours is a zombie world. The issue of interest is not whether we differ from zombies, but whether our form of consciousness is of an angelic or a zombic sort (or equivalently, whether our *concepts* of consciousness *ought* to be appropriate to angelic or zombic worlds — whether or not they actually are).

From the type- Z vantage point, then, the major problem with the conceivability argument does not really lie with either the space of conceivability or of possibility (as more conventional type- A or B materialists insist). Rather, the trouble from the type- Z perspective is that supermaterialists insist upon conducting the argument in a manner of speaking that is appropriate only if we do indeed reside in an angelic, supermaterial world. As Chalmers fully acknowledges in the quoted passage above, there are two different ways of talking in play: that involving ‘consciousness’, which accords with an angelic way of speaking, and that involving ‘schmonsciousness’, which accords with a zombie’s way of talking. The conceivability argument gets the mileage that it does only by forcing materialists to engage the argument in a way that presupposes that the ‘zombic’ way of talking could not be our own.

It is instructive to compare this type- Z response to other closely related replies to the conceivability argument that Chalmers (2009) finds ‘hard to classify’. Consider, first, the ‘conditional analysis’ objection advanced by Hawthorne (2002) and Braddon-Mitchell

(2003). Like type-Z materialism, the conditional analysis approach also aims to allow the applicability of our own concepts of consciousness in a thoroughly material, zombish world. However, unlike the type-Z strategy, the conditional analysis approach doesn't fully acknowledge two *distinct* ways of talking, but rather attempts to subsume both zombic and angelic manners of speaking under a single overarching system of concepts structured in part by hidden conditionals. For reasons given by Alter (2007), such an attempt to unify angelic and zombic ways of speaking can appear clumsy and ham-handed. Moreover, this tacit commitment to a unified way of speaking makes it difficult for proponents of conditional analysis to pinpoint where exactly the original conceivability argument goes wrong. Braddon-Mitchell, for instance, denies (P1), while Hawthorne remains more circumspect, appearing to lean more toward the rejection of (P2) or possibly (P3).

The overall lesson for materialists is that they must be careful *not* to engage the argument in the supermaterialist's own preferred parlance, for when they do, a smart opponent like Chalmers (or Alter) is able to paint them into uncomfortable corners. Consider, for instance, my earlier suggestion that the problem with the conceivability argument lies in the vicinity of (P3). That is true, *but only if we are prepared to adopt a zombie's perspective or way of talking*. Things look otherwise if we use terms in the way that Chalmers would prefer. For according to his manner of speaking, when type-Z materialists suggest that our own epistemic and conceptual perspective is no different from that of zombies, it looks instead like they are 'really' committed to the idea that $P \ \& \ \sim Q$ is true of our own world. Thus it would seem that type-Z materialism must really be committed to the denial of (P0), and so must endorse an implausible ('Churchlandish') type-E *eliminativism* about consciousness. But that of course isn't quite right, for zombies (and their type-Z advocates), as opposed to eliminativists, do not and need not eschew everyday, ordinary talk about consciousness. Their project instead is to reconceive it in an appropriately materialist fashion.

Something similar is going on with various 'reverse-zombie' conceivability arguments (see Brown, CO1; Frankish, 2007; Stalnaker, 2002; Balog, 1999). Such reverse-zombie exercises strongly suggest the unsoundness of the conceivability argument, without specifying exactly where the original argument goes wrong. In that respect, they resemble Gaunilo's rejection of the ontological argument. Following Stalnaker, Brown suggests that the weakness in the conceivability argument lies with (P1); zombies only *seem* to be conceivable. Balog suggests instead that the culprit is (P2), while

Frankish doesn't really say. Reverse-zombie considerations are presented as if they come from a vantage point other than that from which the original conceivability argument is presented. Like the original conceivability argument, they get their parallel mileage by adopting a manner of speaking that is appropriate only from a purely materialistic, zombic perspective. It's little question, then, why Chalmers finds them so unconvincing, and why we seem to be left with what Balog calls (CO1) a 'standoff'.

By way of reply, Chalmers maintains that reverse-zombie arguments are not parallel to the original conceivability argument after all. They present us with scenarios that, according to Chalmers, are not directly conceivable but rather conceivable only 'at arms length' or in some attenuated 'meta' sense.¹⁰ However, this reply has a vague whiff of the paradoxical to it, for the reverse perspective from which things are only conceivable in this attenuated 'meta' sense is at the same time the very materialistic scenario that the original zombie argument so stridently insists is conceivable in a much stronger sense. Once again, we come face-to-face with the remarkably ambivalent attitude proponents of the original conceivability argument adopt towards zombies. While they are conceivable, they aren't conceivably *actual*.¹¹ One might well take the whole point of reverse-zombie considerations to be that of showing that the notions of direct and meta-conceivability are themselves up for grabs. For whether or not one takes some situation to be directly conceivable or conceivable only 'at arms length' depends upon one's presuppositions about the nature of the actual world. The reverse-zombie exercises reverse Chalmers' presuppositions; they exploit the idea that *for all we know*, ours is a world in which the reverse-zombie scenarios are (directly) conceivable. Once again, Chalmers' insistence on the mere meta-conceivability of reverse-zombies seems yet another expression of his insistence upon a way of talking that begs the question against type-z materialism by illicitly centering the actual world off of a materialistic one. Chalmers simply finds the truth of supermaterialism to be more natural (conceivably true or *actual*) than the actual truth of materialism. In any event, we can see that the issue at hand is not really about what is conceivably possible, but rather conceivably actual (or epistemically conceivable). And the weird interplay in these arguments is not between

[10] See Chalmers (2009), p. 332. It is interesting to note that in his reply to reverse-zombie arguments, Chalmers presupposes that those who subscribe to the objection commit themselves to the necessity of materialism. In so doing, he completely overlooks the possibility that the challenge might instead be 'in the vicinity' of (P3).

[11] Reverse-zombists adopt the same attitude towards angels.

conceivability and possibility as the type-B materialist suspects, but rather between conceivability and actuality — just as the type-Z materialist would have us believe.

V. Type-Z and the Knowledge Argument

My focus so far has been on the alleged *epistemic* shortcomings of zombies relative to us (corresponding to the second horn of Chalmers' dilemma). Let's turn now to the first horn, and their alleged *conceptual* shortcomings. The intuition guiding the supermaterialist here is not that zombies are horribly misinformed about consciousness, but rather that their situation renders them unable to grasp some of our own 'pure' phenomenal concepts, and so they are not capable of our range of phenomenal thought. Once again, Chalmers must be very careful not to let this alleged deficiency manifest itself in our comparative verbal behavior. Nevertheless, the upshot is the same as before: what with their allegedly limited conceptual repertoire, zombies would have to be in an impoverished epistemic situation relative to the one he takes us to be in.

So what ultimately drives the idea that zombies are conceptually inferior to us? Here Chalmers calls upon additional considerations surrounding the knowledge argument (suggesting, incidentally, that the knowledge argument might be deeper). Specifically, he bids us to consider a zombie in the place of colour-deprived Mary (Chalmers, 2007, p. 178). Once again, he'd have our intuitions informing us that a 'zombie-Mary' would not acquire precisely what Mary acquires when she leaves 'Kansas' (her black and white environs). While zombie-Mary might acquire some new knowledge concerning her internal states, she nevertheless wouldn't acquire the full-fledged, 'cognitively significant' first-person knowledge of her own introspective, experiential states that we have (and that Mary comes to have). Though zombie-Mary might acquire some new ways of representing her discriminatory states in thought and talk, she wouldn't acquire all that we think Mary acquires when we say she has finally learned 'what it's like' to see red.

Notice that at least for the sake of argument Chalmers grants that zombies might have a great number of concepts that are in the vicinity of what he terms 'pure' phenomenal concepts: various concepts that relate inner discriminatory states to external, publically available

states of affairs.¹² The story here appears to be a familiar, broadly Sellarsian account of how creatures already capable of reasoning about an external world can extend their vocabulary to encompass concepts of inner discriminatory states, including those of experience, with respect to which they can come to have a privileged epistemic position.¹³ By working ‘from the outside in’, such an account is tailor-made for zombies. It shows how they could begin to make observations about, and so come to be distinctly aware of, those internal discriminatory states responsible for their being disposed to make reports about how things appear to them.¹⁴

The question, then, is whether those comparative phenomenal concepts that Chalmers is willing to grant to zombies are sufficient to exhaust all of our own. Chalmers purports to have an argument establishing that we do indeed possess ‘pure’, non-comparative phenomenal concepts.¹⁵ Roughly, he tells us that Mary’s concept of what it’s like to see red cannot be identified with any comparative, indexical, or demonstrative phenomenal concept (that is, one that Zombie-Mary can also have), because Mary can always have a ‘cognitively significant’ (non-trivial) thought to the effect that having a red-experience picked out in whatever comparative, indexical, or demonstrative fashion you might like is like *this* (where *this* allegedly picks out the experience of red in a distinctive, purely phenomenal fashion). At this point, one might complain that Chalmers refuses to give Mary’s

[12] Or at least he does so for the sake of argument. See Chalmers (2003), pp. 223ff and (2007), pp 178–9.

[13] This is roughly the second episode of Sellars’ ‘Myth of Jones’ (parts XIV–XVI of Sellars, 1956/97). The key is to understand a sense impression of some property *x* as the discriminatory state one is in when one is disposed (under normal conditions) to judge that *x* is present. It follows that the discriminatory state one is in when one reports that one has a sense impression of *x* is the same state that would (under normal conditions) incline one to judge that *x* is present. Hence, an impression of *x* will, as it were, be equivalent to an impression of an impression of *x*, and no gap opens up between one’s actually having an impression of *x* and one’s being disposed to say so. Please notice, by the way, that the Sellarsian notion of a sense impression, so construed, is routed through a definite description. As a result, it’s easy to see why the identity between phenomenal states (construed in terms of sense impressions) and material states specified in a more-or-less scientific fashion would seem so persistently contingent, thereby creating the appearance of an unbridgeable explanatory gap. This connection of phenomenal concepts with descriptions also explains how one can coherently deny (P3) and thus block the inference from the mere possibility of zombies to the falsity of materialism. There’s more to be said about unpacking phenomenal concepts in terms of Sellarsian sense impressions (concerning, e.g., transparency, ‘what it’s like,’ and the explanatory gap), but I’ll spare you the details. The curious can look at Beisecker (2005).

[14] Or if you prefer, a story about how zombies of one sort (namely Dennett’s (1998) ‘zimboes,’ who cannot report on their inner discriminatory states) can bootstrap themselves into becoming something greater.

[15] See especially Chalmers (2003), p. 257 as well as (2007), p. 178ff.

allegedly purely phenomenal conception of *this* any fuller specification. However, perhaps the deeper way to lodge this worry is to raise what we might call ‘the zombie’s revenge’ — that, as in the original paradox of phenomenal judgment, any such argument or line of reasoning should, by hypothesis, be equally available and compelling to zombies, and so is undermined. Zombie-Mary would claim to have such thoughts as well, though by Chalmers’ lights, they’d be empty, because there is nothing (super-material) in her world to answer to her allegedly pure phenomenal concepts. The question facing us, then, is whether our own thoughts are similarly vacuous, a worry made all the more vivid by Chalmers’ failure to give pure phenomenal concepts anything but a wholly negative characterization. As externalists are fond of reminding us, certain aspects of our own concepts may depend upon features of the world well outside of our cognitive ken, meaning that should the world turn out to be other than how we commonly suppose it to be, our very own concepts may be quite unlike how we commonly suppose them to be.

So once again, it would appear we are left with a standoff. There remain two ways of speaking in play here, and the type-Z materialist stubbornly insists that Chalmers has failed to establish that our own way of talking accords more with an angelic world than a zombic one. Much of what Chalmers subsequently has to say in this connection is directed against specific ways on offer to adopt the phenomenal concept strategy — particularly those self-styled type-B proposals that claim that pre-release Mary lacks the requisite experience to possess certain concepts, the phenomenal ones, which refer to their referents in some distinctive fashion that’s routed through experience.¹⁶ I happen to agree with Chalmers that these accounts don’t capture all that is bound up in attributions of knowledge of what it’s like, although for different reasons. I remain unconvinced that Mary’s failure to know what it’s like to see red can be traced to a gap in her conceptual repertoire, for her lack of experience with red things doesn’t seem to constrain her ability to entertain thoughts about what it would be like. Even without any contact with a coloured reality, she can still wonder what it would be like for her to see red, know that her earlier experiences are not what it’s like to see red, and surmise (probably correctly) that her similar physiological makeup suggests that what it would be

[16] These proposals try to unpack the possession of phenomenal concepts somehow in terms of demonstrative or quotational modes of presentation that are unavailable to subjects until they find themselves in the appropriate kind of experiential state. The governing idea is that upon exposure to phenomenal red, we become capable of entertaining whole new forms of thought. Stoljar (2005) has called this ‘the experience thesis’.

like for her to see red is pretty much like what it would be like for the rest of us to see red. She could even have mistaken thoughts about what it's like — say, by spying a green tomato rogue epistemologists have told her is red.¹⁷

There is, however, another option available, which doesn't explicate post-release Mary's new knowledge with a sudden expansion of her conceptual repertoire. Instead, this strategy points out that she lacks the 'experience' (here glossed as a history of visual contact with coloured things) to entertain relevant experiential thoughts *responsibly*. That is, post-release Mary acquires (eventually) not the capacity to entertain new thoughts, but rather a *justificatory* status which allows her to claim correctly that she now *knows* what it's like to see red. If we are guided by the idea that all our knowledge, including that involving our concepts of inner experience, is a product of training and attunement to the world,¹⁸ we might well maintain that our ability to identify or classify a given sensation as a red sensation — or to know what it's like to see red — depends upon an understanding that those are the sensations generally indicative of red things in an external world. Accordingly, we might hold that our capacity to know what it's like to see red depends upon an *awareness of our own reliability* at being able to detect red things. Merely being a reliable detector of redness is not sufficient for such knowledge; the awareness one must have of one's reliability can come about only through 'experience,' understood once again as a history of contact with red things. So even though — as if by magic — pre-release Mary might happen to be a perfectly reliable reporter of red things, she nevertheless lacks the experience to have the adequate awareness of this reliability. *That* is actually the reason we can perfectly well suspect that she lacks genuine knowledge of what it's like. But then notice how this should be an especially amenable option for a type-Z materialist, since it would seem that zombie-Mary could also acquire a post-release history of contact with a red reality as well, which would allow her to gain a similarly elevated epistemic status with respect to her observation reports.

[17] Consider Dennett's (2007) example of the blue banana.

[18] Guided, that is, by a rejection of any 'myth of the given' (Sellars, 1956/97). The subsequent sketch is also broadly Sellarsian, especially the requirement not just of mere reliability, but also *an awareness* of reliability (see section VIII of the Sellars piece).

VI. Giving up the Ghost

The zombie wars in philosophy are far from over, and the type-Z response explored here suggests that these battles might have a fittingly cinematic resolution. As in the finer films of the zombie genre, the distinction between the living and the undead turns out to be much harder to draw than one might have first supposed. From the type-Z perspective, the whole question about the truth of materialism turns on whether our own *way of talking* accords with that of a zombie or that of a supermaterial being. The trouble is that by the very rules of the philosophical game being played here, supermaterialists cannot have much to *say* about supermaterial reality that is supposed to inoculate us against potential zombification. Their admission of the possibility of zombies shows that any alleged contact we might have with supermaterial reality isn't necessitated by the way in which we seem to talk about our so-called 'conscious' experience. We are thus forced to question whether our phenomenal judgments are indeed partly constituted by some super-material facet of our being. Accordingly, several philosophers have recently launched their own investigations into how zombies might come to talk in their funny, allegedly 'paradoxical' fashion. While such explorations might appear perverse from the supermaterialist perspective, the paradox might really lie in the supermaterialist's ambivalent attitude toward their conceptions: though robustly conceivable, supermaterialists nevertheless insist that zombies are not conceivably actual.

The precise nature of our conscious experience remains stubbornly opaque. It should come as no surprise then that when type-Z materialists gaze deeply into their own minds and souls, they find it to be every bit as dark as that of a zombie. To the extent that we can make sense of the inner lives of zombies, the type-Z materialist suspects that we'll actually wind up fleshing out ours. So they opt, quite literally, to 'give up the ghost'. The type-Z materialist thus bets that the trajectory in philosophy will turn out to be the same as it is in the finer zombie flicks. Ultimately, there's nothing we can do or say to keep our world safe from zombies; in the end we are likely nothing but meat.¹⁹

[19] I'd like to thank Sam Coleman, David Chalmers, Michael Wolf, Robert Howell, Tim Cleveland, Josh Weisberg, Pete Mandik, and especially Tim Connolly for help at various stages of this project. I'm inclined to attribute any remaining conceptual blindspots to my zombish utter lack of full-fledged phenomenal consciousness.

References

- Alter, Torin and Sven Walter (ed. 2007), *Phenomenal Concepts and Phenomenal Knowledge* (Oxford: Oxford University Press).
- Alter, Torin (2007), 'On the conditional analysis of phenomenal concepts', *Philosophical Studies*, **134**, pp. 235–53.
- Balog, Katalin (1999), 'Conceivability, possibility, and the mind-body problem', *The Philosophical Review*, **108** (4), pp 497–528.
- Beisecker, David (2005), 'Phenomenal consciousness, sense impressions, and the logic of "what it's like"', *Consciousness and Emotion* (special issue on agency, conscious choice and selective perception), ed. Ralph Ellis and Natika Newton, pp. 137–53.
- Braddon-Mitchell, David (2003), 'Qualia and analytic conditionals', *Journal of Philosophy*, **100**, pp. 111–35.
- Brown, Richard (CO1), 'The reverse zombie argument against dualism', *Consciousness Online* 2009. **[Citation will need to be updated]**
- Chalmers, David (2003), 'The content and epistemology of phenomenal belief', in *Consciousness: New Philosophical Perspectives*, ed. Quentin Smith and Aleksandar Jokic (Oxford University Press), pp. 220–72.
- Chalmers, David (2009), 'The two-dimensional argument against materialism', in *The Oxford Handbook of Philosophy of Mind*, ed. Brian McLaughlin (Oxford: Clarendon Press), pp. 313–35.
- Chalmers, David (2007), 'Phenomenal concepts and the explanatory gap,' in Alter and Walter (2007), pp. 167–94.
- Dennett, Daniel (1991), *Consciousness Explained* (Boston, MA: Little, Brown and Company).
- Dennett, Daniel (1998), 'The unimagined preposterousness of zombies: Commentary on Moody Flanagan, and Polger', Chapter 10 of *Brainchildren* (Cambridge, MA: MIT Press).
- Dennett, Daniel (2007), 'What Robomary knows,' in Alter and Walter (2007), pp. 15–31.
- Frankish, Keith (2007), 'The anti-zombie argument', *The Philosophical Quarterly*, **57** (229), pp. 650–66.
- Hawthorne, John (2002), 'Advice for physicalists', *Philosophical Studies*, **108**, pp. 17–52.
- Levine, Joseph (2007), 'Phenomenal concepts and the materialist constraint', in Alter and Walter (2007), pp. 145–66.
- Lynch, Michael (2006), 'Zombies and the phenomenal pickpocket', *Synthese*, **149**, pp. 37–58.
- Moody, Todd (1994), 'Conversations with zombies', *Journal of Consciousness Studies*, **1** (2), pp. 196–200.
- Sellars, Wilfrid (1956/97), *Empiricism and the Philosophy of Mind* (Cambridge, MA: Harvard University Press).
- Stalnaker, Robert (2002), 'What is it like to be a zombie?', in *Conceivability and Possibility*, ed. Tamar Szabo Gendler (Oxford: Clarendon Press), pp. 385–400.
- Skokowski, Paul (2002), 'I, Zombie', *Consciousness and Cognition*, **11**, pp. 1–9.
- Stoljar, Daniel (2005), 'Physicalism and phenomenal concepts', *Mind and Language*, **20** (5), pp. 469–94.