

Two Conceivability Arguments Compared*

Daniel Stoljar

Draft of October 6 2006

1. The first conceivability argument of my title is, of course, the famous one: the zombie argument against physicalism. In a version that is both familiar and relatively clear it goes like this:

Zombie Argument:

1. It is conceivable that I have a zombie-duplicate, i.e., someone identical to me in respect of all physical truths but different with respect to some phenomenal truth.
2. If it is conceivable that I have a zombie-duplicate, it is possible that I do.
3. Ergo, it is possible that I have a zombie-duplicate.

The conclusion of this argument (hereafter ZA) entails the falsity of physicalism because, technical details aside, physicalism is or entails the thesis that every psychological truth is entailed by some physical truth. If it is possible that I have a zombie duplicate however, then it is possible that the physical truths are as they are and some psychological truth is different. Hence 3 entails that physicalism is false.

The second conceivability argument is one that is almost as famous, though perhaps it is less famous for being a conceivability argument: the perfect actor argument against behaviorism (see, e.g., Putnam 1963, 1975). In a version that is both familiar and relatively clear, it goes like this:

Actor Argument:

4. It is conceivable that I have an actor-duplicate, i.e., someone identical to me in respect of all behavioral truths but different with respect to some phenomenal truth.

* [Acknowledgments]

5. If it is conceivable that I have an actor-duplicate, it is possible that I do.
6. Ergo, it is possible that I have an actor-duplicate.

The conclusion of this argument (hereafter AA) entails the falsity of behaviorism because, technical details aside, behaviorism is or entails the thesis that every psychological truth is entailed by some behavioral truth, i.e. some truth that concerns my actual and potential behavior. If it is possible that I have an actor duplicate, however, then it is possible that the behavioral truths are as they are and some psychological truths are different. Hence behaviorism is false.

I think a comparison of these arguments is the source of considerable insight about one of the central questions in philosophy of mind, viz., whether ZA is sound. In this paper I explain why.

2. I will begin by setting out two substantive assumptions about both arguments. It may be that you disagree with me about these assumptions. If so, the following discussion will have a conditional flavor. But even so, I think (hope!) the conditional is of interest. Be that as it may, my two assumptions are as follows:

Assumption #1: AA is sound.

Assumption #2: AA is of the same kind as ZA.

Let me start with some comments on the nature and plausibility of these assumptions.

First, to assume that AA is sound, and so that behaviorism is false, is not to deny that other arguments against behaviorism are also sound. Behaviorism has trouble explaining the sense in which one's actions are a causal expression of one's thoughts—I am not denying that this is a good reason for resisting behaviorism. And of course there are empirical reasons against behaviorism—I am not denying that. What I am insisting on is that *this* reason—i.e. AA—is a good reason for denying behaviorism. However, since there is no upper bound on the number of reasons there might be for denying behaviorism, in emphasizing one I am not at war with the others.

Second, to assume that AA and ZA are of the same kind is to assume at least that they are instances of at least one fairly general kind—viz., the kind *conceivability*

argument. But AA and ZA are also instances of kinds that are more specific than this. For one thing—this point will be important later—AA and ZA are a minimal pair. They are identical save for the fact that where AA has ‘behavioral’, ZA has ‘physical.’ Moreover, since a behavioral truth is itself a kind of physical truth, or at least so I will assume here, both arguments concern the relation of phenomenal truths to physical truths: ZA concerns the relation of phenomenal truths to *all* physical truths, AA concerns the relation of phenomenal truths to *some* physical truths, i.e., the behavioral ones. So not only are AA and ZA both instances of the kind *conceivability argument*, they are also instances of more specific kinds, e.g., *conceivability argument which concerns phenomenal truths*, and *conceivability argument which concerns the relation between phenomenal truths and some or all physical truths*. In the following, when I say that ZA and AA are of the same kind, I will let the context decide whether I have in mind the determine kind conceivability argument, or some relevant determinable of that kind—for most purposes it will not matter terribly which is intended.

Third, to assume that ZA and AA are both conceivability arguments, and so are of the same kind, is not to say that we have a terribly good grasp on the nature of the kind conceivability argument, or indeed on any of the relevant determinables of that kind. On the contrary, I take it to be an on-going question in philosophy what the precise nature of conceivability reasoning is. Some good questions that arise in this context, for example, are: ‘What precisely is the nature of conceiving?’ ‘What precisely is the nature of possibility?’ and ‘What precisely is the connection conceivability and possibility, i.e. why if at all is the first evidence for the second?’ I am not going to go into these questions in any detail here. The idea is rather to draw attention to AA and ZA in order to draw attention to certain typical examples of reasoning, and then to assert that the two examples are of the same kind. I take it that we are sometimes in a position to say that two things are of the same kind, without being in a position to say (or say very clearly) what the nature of that kind is. I think a situation of this sort obtains here.

Fourth, to assume that AA and ZA are of the same kind might be thought to be controversial in the following way. It is an empirical fact that good actors exist—e.g. Glenn Close exists and she is a good actor. This suggests that the possibility at issue in AA—the possibility of my having a behavioral duplicate who is not a psychological duplicate—is not a terribly remote one; indeed, it might even be

nomologically permissible, i.e., a possibility permitted by the empirical laws. On the other hand, the possibility that is at issue in ZA—the possibility of my having a physical duplicate who is not a psychological duplicate—is (it is often said) an extremely remote one, and one that is almost certainly nomologically ruled out. This apparent difference in the relative remoteness of their key examples might suggest in turn that AA is a *very* different kind of argument to ZA, contrary to my assumption #2. In particular, one might think that, in the case of AA, one is reasoning about a possibility on the basis of various actual facts, whereas in the case of ZA, one is reasoning about a possibility but not on the basis of various actual facts.

However, this objection fails to distinguish between merely good actors (like Glenn Close) and perfect actors. Good actors are (to put it somewhat roughly) people who know how to contort their voices and bodies in order to produce a desired affect in an audience (or in some cases are people good at taking instruction from directors on how to contort their voices and bodies in order to produce these effects). Perfect actors, by contrast, are people who duplicate actual people (like me or Glenn Close) in all actual and potential respects relevant to their behavior. But the possibility of there being perfect actors, as opposed to merely good actors, is in fact extremely remote, and moreover, is plausibly inconsistent with the actual laws of nature. So there is no failure of analogy between ZA and AA on this score.

Fifth and finally, to assume that AA and ZA are of the same kind is not to deny that the route from 3 to the denial of physicalism is less direct than the route from 6 to the denial of behaviorism. Behaviorism is often thought of as, if true, an a priori and necessary truth. If so the inference from 6 to the denial of behaviorism is straightforward: if 6 is true, a possibility exists which behaviorism says does not exist. By contrast, physicalism is often thought of as, if true, an a posteriori and contingent truth. If so, the inference from 3 to the denial of physicalism is not straightforward; in particular, if 3 is true, a possibility exists which physicalism *need not* say does not exist. How then is the inference from 3 to the falsity of physicalism supposed to go?

This topic has been discussed extensively in the relevant literature, so here I will be brief (see, e.g., Jackson 1998, Chalmers 2002, Stoljar 2006 ch.2). Physicalism is or entails the thesis that every psychological truth is entailed by some physical truth. This thesis is contingent because it is a contingent matter which truths are the psychological ones and which truths the physical. If the psychological truths concern operations in your soul and the physical truths concern extension, physicalism is false;

but if the psychological truths concern experiences as we ordinarily conceive of them, and the physical truths are as complicated and variegated as the ones that (we assume) obtain in this world, physicalism might be true.

On the other hand, while physicalism is contingent, there is a closely related thesis that is necessary. For consider a statement *S* which summarizes the physical truths, construed broadly to include topic neutral and totality truths; and now consider a different statement *S** which summarizes the psychological truths. If physicalism is true, the conditional formed from these statements—i.e., ‘If *S* then *S**’—is necessary. The reason is that if physicalism is true, *S* entails *S**, and yet it is not contingent which truths *S* summarizes; *mutatis mutandis* for *S**. On the other hand, if the conclusion of *ZA* is true, and it is possible that I have a zombie duplicate, the conditional ‘if *S* then *S**’ is contingent. But if this conditional is contingent, physicalism is false. Putting all this together: it is true that 3 is not inconsistent with physicalism alone, but 3 *is* inconsistent with physicalism combined with the empirical assumptions (which physicalists are committed to) that *S* is the statement that summarizes the physical truths, and *S** is the statement that summarizes the psychological truths.

3. The two assumptions about *AA* and *ZA* that I have just set out do not entail that *ZA* is sound. It might be true both that *AA* is sound and that *ZA* is of the same kind as *AA* without its being true that *ZA* is sound; an argument may be of the same kind as a sound argument without itself being sound. Nevertheless, as I said earlier, these assumptions *do* tell us a considerable amount about various proposals that have been made about *ZA*, and in particular about where it might go wrong. It is to these proposals that I now turn.

We may begin by considering an idea which, while perhaps not quite as common now as it once was, is nevertheless surprisingly common: the skeptical position about conceivability reasoning. The skeptic about conceivability says, in effect, that *ZA* is unsound because, in general, arguments of this style are an utter fantasy. “Facts about conceivability are at the end of the day facts about psychology” the skeptic says, “but facts about possibility are facts about something that obtains independently of psychology. What then does the first have to do with the other? Trying to figure out what the possibilities are on the basis of conceiving is like mathematics on the basis of toothache”.

There is much to say against skepticism of this sort, but for us the important point is that it is in violation of our two assumptions. The skeptic about conceivability arguments says that *ZA* is unsound *because* (a) it is a conceivability argument, and (b) all such arguments are unsound. However, if *ZA* fails because it a conceivability argument, then, by assumption #2, *AA* fails also because it is a conceivability argument. But by assumption #1, *AA* does not fail; *a fortiori*, it does not fail because it is a conceivability argument. Hence skepticism is mistaken.

Might a skeptic respond by denying assumption #1? Of course—in fact, a skeptic is *committed* to denying assumption #1. And I will not here attempt to dislodge such a skeptic from their position in any serious way. (In calling the assumptions I have set out ‘assumptions’ I mean in part that I am not going to defend them.) Let me say only that, for my part, denying assumption #1 is not only desperate but involves a considerable amount of *mauvais fois*. When you teach a philosophy of mind course, you routinely begin with dualism and explain to the students why it is implausible. Next you turn to behaviorism and explain to the students why it is implausible. After that, you move on to other things. But wait! How did you convince the students that behaviorism is implausible? The usual answer is that you used a conceivability argument such as *AA* to do this. (Of course, you might have used a slightly different conceivability argument, e.g. the super-spartan argument (Putnam 1963) or the blockhead argument (Block 1981) or something similar; the point remains the same.) On being presented with these arguments most people judge that they are convincing. The best explanation of this is that they *are* convincing. Hence, in denying that *AA* is sound, skeptics about conceivability are at odds, first, with the practice of philosophy of mind—in which these arguments are routinely advanced—and, second, with what that practice is indicative of, i.e., the soundness of at least some conceivability arguments.

Philosophers who appear skeptical about conceivability arguments in their writing or teaching may of course be using such skepticism to make a point about such arguments rather than to reject them. For example, they might be making the point that the basic nature of such arguments is ill understood. However, if this is the use to which skepticism is put I have no objection. As I have said, I *agree* that arguments of this style are not very well understood. The point is not to deny this but to insist that, whatever is the ultimate philosophical story about conceivability

reasoning, that story will have to somehow accommodate the data that AA is sound and that AA and ZA are of the same kind.

4. Skepticism about conceivability is no good if it is intended as a proposal about where the mistake in ZA if any lies. But if we limited ourselves to skepticism our comparison between ZA and AA would be not so interesting; as noted earlier skepticism of this kind is not that much in vogue today. However, while contemporary proposals about where ZA goes wrong (if it does) are not skeptical, they may be criticized in precisely the same way as skepticism.

The most common contemporary proposal about where ZA goes wrong is (as I have called it elsewhere; see Stoljar 2005) ‘the phenomenal concept strategy.’ The general idea of the phenomenal concept strategy is to say that, instead of thinking about experiences, we should think instead about how experiences are represented in speech and thought. As such the phenomenal concept strategy is an instance of a common manoeuvre in philosophy whose paradigm case is emotivism in ethics. Just as emotivists say that *moral* talk and thought have special features and that when we attend to these features, metaphysical questions about morality will disappear or at least emerge as more tractable, so too the phenomenal concept strategist says that *phenomenal* thought and talk have special features and that when we attend to these features, metaphysical questions about experience (e.g. of the kind raised by ZA) will disappear or emerge as more tractable. More particularly, the phenomenal concept strategy begins with the point that any attempt to expound ZA will involve the employment and possession of phenomenal concepts, such as the concepts of pain, itch and so forth. But these concepts, it is alleged, have some special features, and that when these special features are sufficiently attended to, it will become apparent that the argument is a failure.

Now, different exponents of the phenomenal concept strategy have different features or alleged features of phenomenal concepts in mind. Some say they have special epistemological features. Others say they are a special kind of indexical. Some say they have a special possession condition, e.g., they constitutively involve having the relevant experiences. As such the phenomenal concept strategy raises a number of different concerns (see, e.g., Perry 2001, Sturgeon 2000 and Tye 1999 among many others; for further criticism and references see Stoljar 2005). However, the point for us is quite independent of these different versions of the phenomenal

concept strategy, and will apply whatever version is in play. For the fact is that *any* version of the phenomenal concept strategy is in violation of our two assumptions. If ZA fails because of the special features of phenomenal concepts, then, by assumption #2, AA fails also because of the special features of phenomenal concepts. But by assumption #1, AA does not fail; *a fortiori* it does not fail because of the special features of phenomenal concepts. Hence the phenomenal concept strategy is mistaken.

It might be thought that to possess a phenomenal concept presumably involves the sort of cognitive achievement that would be impossible if behaviorism is true; hence if behaviorism were true, we could not have phenomenal concepts in the relevant sense. But this is no objection to my objection. The point is not that behaviorism is true or that proponents of the phenomenal concept strategy are committed to it; it isn't and they aren't. The point is that we are entitled, or take ourselves to be entitled, to use the perfect actor argument to explain—for example, to explain to a student—what is wrong with behaviorism, but if the phenomenal concept strategy is correct we are not so entitled. Hence the phenomenal concept strategy is false.

5. The phenomenal concept strategy is the dominant proposal about ZA in the recent literature, but other proposals face the same objection. To take another example, consider the Footnote Eleven View (as I will call it) due originally to Nagel (see Nagel 1974, fn.11) and developed by Hill 1997, Hill and McLaughlin 1999 and noted sympathetically by Block and Stalnaker 1999. The starting point of the Footnote Eleven View is that there are (at least) two quite distinct kinds of imagining or conceiving. *Sympathetic* imagination occurs when one imagines oneself undergoing a certain experience—being in pain, say. In such cases, Nagel says in footnote 11, one puts oneself in a position that “resembles” the pain itself. *Perceptual* imagination occurs when one imagines oneself perceiving (or seeming to perceive) some event in the world. In such cases, Nagel says, one puts oneself in a position that “resembles” seeing the event in question. With this distinction established, the proponent of the Footnote Eleven View goes on to say that what is wrong with ZA is that it illegitimately blends these two kinds of imagination together.

In fact, there are a number of ways to develop this basic idea (see Doggett and Stoljar, forthcoming), but on one rather straightforward version, the Footnote Eleven View denies the first premise of ZA, viz. 3:

3. It is conceivable that I have a zombie-duplicate, i.e., someone identical to me in respect of all physical truths but different with respect to some phenomenal truth.

If we assume that ‘conceivability’ and ‘imaginability’ are in these contexts rough synonyms, 3 is equivalent to 7:

7. It is imaginable that there is someone identical to me in respect of all physical truths but different with respect to some phenomenal truth.

According to the Footnote Eleven view, however, what we are justified in asserting is not 7 so much as 8:

8. It is *perceptually* imaginable that there is someone identical to me in respect of all physical truths and it is *sympathetically* imaginable that there is someone different with respect to some phenomenal truth.

On the other hand, 8 will not yield any claim that threatens physicalism; in particular, 8 will not yield 7 courtesy of the perceptual/sympathetic distinction.

Once again there is much to say about this response to ZA. But for our purposes the problem with the Footnote Eleven View is, again, that it is inconsistent with our assumptions. If ZA fails because of the perceptual/sympathetic distinction in imagination, then, by assumption #2, AA also fails because of the perceptual/sympathetic distinction. But by assumption #1, AA does not fail; *a fortiori*, it does not fail because of the perceptual/sympathetic distinction. Hence the Footnote Eleven View is mistaken.

6. I have argued that some of the most influential proposals in the literature today about where ZA goes wrong (assuming it does) face a simple but devastating problem. The problem is that if they were sound, they would also tell us that AA goes wrong. But since AA does not go wrong—so, anyway, I have assumed—there must

be something wrong with these proposals. Of course, I have not gone into detail about what precisely *has* gone wrong with these proposals. So one way in which the discussion might now be pursued would be follow up on what is wrong with them. In the next part of my discussion, however, I want instead to consider the issues I have been discussing from a different angle.

Many philosophers in the world today are interested in the question of whether ZA is sound, and many—for example, those I have just been discussing—take a negative view. But some take a positive view, arguing that ZA is sound (see, e.g. Chalmers 1996, 2002, and, in a different way, Kripke 1980). What does the comparison between ZA and AA tell us about the plausibility of their position?

Well, as I noted earlier, it clearly does not *follow* from the assumptions that AA is sound, and that AA is the same kind as ZA, that ZA is sound. Nevertheless, it might be thought AA and ZA are in some sense deeply analogous, and thus, while it does not follow from our assumptions that ZA is sound, this is nevertheless strongly suggested by them. One way to develop this point is as follows: behavioral truths are, not only a sub-set of the physical truths (as I noted earlier), but a *representative* sub-set (or representative sample) of the physical truths. If behavioral truths are a representative sample of physical truths, then presumably what goes for behavioral truths goes for physical truths more generally. If so, from the soundness of AA, the soundness of ZA follows.

Now, in order to assess this suggestion we would need to spell out the sense in which behavioral truths are a representative sample of physical truths. I take this to mean that behavioral truths have some property, F, such that F projects in an obvious way to all of the physical truths. But what property could F be? It is clearly no good to say that F is the property of being behavioral. Behavioral truths certainly have that property, but since physical truths do not, being behavioral is not a feature in virtue of which behavioral truths are a representative sample of the physical truths. Nor is it any good to say that F is the property of being physical. It is true that behavioral and physical truths both have that property, but this will not advance the polemical purposes of the philosopher who is trying to convince the not-already-convinced that ZA is sound on the ground that AA is.

7. Even if F is neither the property of being behavioral or the property of being physical, there are other possibilities. One proposal is that F is, to put things at first a

bit vaguely—being perceptually available. As partial motivation for this view, we might remind ourselves that it is at least initially plausible to suppose that any truth which is perceptually available at least in a relatively direct or unaided way does not entail any phenomenal truth. (We can take this either as a claim about the nature of phenomenal truths, or about the limits of perception, or both). Moreover, remember the original rationale for behaviorism. Behaviorists were motivated largely by a desire to solve the skeptical problem about other minds, a problem that starts from the assumption that perceptually available truths (again: at least of a relatively direct or unaided sort) do not by themselves entail any phenomenal truths. The beauty of behaviorism, at least in the eyes of behaviorists, is that it denies this assumption and so denies skeptics about other minds their starting point.

So one suggestion about the property F in virtue of which behavioral truths are a representative sample of the physical truths is that it is the property of being perceptually available. But the problem with this is that being perceptually available is not a general feature of physical truths. It might be that behavioral truths are perceptually available—in fact, they are defined so as they are. But this is clearly not a feature of physical truths in general. Physical truths are not in general perceptually available even in principle. So it can't be that F is the property of being perceptually available.

Of course this is not to say that in the past, philosophers might have *thought* that physical truths were perceptually available, at least in principle. In her classic confrontation with Arthur Eddington, L. Susan Stebbing (1958, p.56) speculates that Eddington was an “inveterate visualizer” in that he assumed that elementary particles and related phenomena are in principle immediately perceptually available somewhat in the way that rocks are. Stebbing may or may not have been right about Eddington, but she is surely right that we have a tendency to think about the physical in visual terms. To the extent that we succumb to that tendency we will be inclined to think that behavioral and physical truths *are* similar in that both are perceptually available. Moreover, to the extent that we succumb to that tendency, we will be inclined to believe that ZA is sound if AA is. So at this point we have a potential diagnosis of the way in which ZA and related arguments have such a grip on our imagination, i.e., our inveterate visualizing. What we do *not* have, however, is a reason to believe that ZA is sound. For whether or not Stebbing is right that Eddington or anybody else is an inveterate visualizer, she is surely right that inveterate visualizing is deeply mistaken.

8. A different proposal about F—i.e., the property in virtue of which behavioral truths are a representative sample of the physical truths—is that both sorts of truths concern states of affairs that conform to a certain metaphysical category, while phenomenal truths concern states of affairs that conform to a different, and contrastive, metaphysical category. The rationale for this idea lies in the very common thought that phenomenal truths tell us about *intrinsic* features of mental states, whereas behavioral or physical truths only concern extrinsic (for example, relational or dispositional) features. Wilfred Sellars (1963) is one philosopher who talks this way for example; another is David Chalmers (1996).

In fact there are different ways to develop this point, depending on how exactly the intrinsic/extrinsic distinction is to be understood (see, e.g., Humberstone 1996). Moreover, there are related proposals in the vicinity that concentrate, not on the intrinsic/extrinsic distinction, but on the categorical/dispositional distinction or—related to this—the occurrent/dispositional distinction. For our purposes however, it will not matter precisely what metaphysical contrast is in play, and I will present the issue in terms of the intrinsic/extrinsic distinction. Focusing on that distinction, we now confront a different suggestion about the property in virtue of which behavioral truths are a representative sample of physical truths, viz., both are extrinsic, i.e. concern extrinsic states of affairs. In particular, if AA is sound, and if AA is sound because behavioral truths are extrinsic while phenomenal truths aren't, then the same is true for ZA, and it is sound too.

However, there are two problems with this suggestion. First, the intrinsic/extrinsic distinction cross cuts the phenomenal/physical distinction. There are certainly intrinsic phenomenal features of mental states, but there are phenomenal features that are non-intrinsic too, e.g., relations of unity and difference between various aspects of a course of experience (see, e.g., Bayne and Chalmers 2003) and between different experiences. Similarly, there are certainly extrinsic physical features, but there are physical features that are intrinsic too. Newton thought that atoms are spherical—surely *that* is a feature of an atom that is both physical and intrinsic.

The second and more important problem is that, while there is a philosophical problem about whether physical truths entail phenomenal truths, there is no comparable problem whether extrinsic truths entail intrinsic ones. It is presumably an

extrinsic fact about George that he is married to Laura, but is an intrinsic fact about the pair of them (i.e., both George and Laura) that they are married. So here we have an extrinsic fact entailing an intrinsic one. Or, to take a different example, it is an extrinsic fact about you that you are exactly similar to someone whose intrinsic features are thus and so. But this extrinsic feature entails precisely what your intrinsic features are. So the mere fact that a truth is extrinsic does not place any restriction on its ability to entail an intrinsic truth.

In view of the fact that the metaphysical distinction between intrinsic and extrinsic states of affairs is apparently quite independent of the issues surrounding ZA, it is something of a mystery why so many reach for this distinction when they want to bring out the persuasiveness of KA and arguments like it. The impression one has is that this sort of talk is often intended as a metaphysical *diagnosis* of why arguments like AA and ZA work. In my view, however, this gets matters back to front. For one thing, most of the talk of intrinsic properties in philosophy of mind seems to me at least to be just talk about essential properties—when the critic of physicalism says that physicalism leaves out something *intrinsic* to mental states, all that seems to be meant is that physicalism leaves out something *essential* to mental states, e.g., their phenomenal character. Moreover, if talking of intrinsicness is just a roundabout way of saying that various physicalist proposals leave something out, then in turn it is just a roundabout way of repeating the basic point behind the ZA—it therefore does not in any way explain or diagnose the basic point behind ZA.

9. I have taken issue both with some reasons for thinking that ZA fails and with some reasons for thinking it successful. But I have not so far attended directly to the question that lends our discussion most of its interest, viz., is ZA sound or not? This question is too large for the present paper. Nevertheless, I will close by suggesting that a comparison of AA and ZA provides considerable guidance here too.

As we noted earlier, AA and ZA constitute a minimal pair: they are identical save for the fact that AA has ‘behavioral’ where ZA has ‘physical.’ On the other hand, if AA is sound and ZA is not, there must be a relevant difference between them. Now, where could this relevant difference be? Well, if AA and ZA constitute a minimal pair, the difference must be located (somewhere!) in the difference between behavioral truths and physical truths. That is the only place it could be. Indeed, it is this fact that to a large extent explains the failure in the proposals considered earlier,

the phenomenal concept strategy and the Footnote Eleven View. These views focus on elements *shared* by the ZA and AA—hence they could not in principle tell us what is wrong with ZA if there is nothing wrong with AA.

Once it is appreciated that any response to ZA must start from the *difference* between ZA and AA, however, a number of possibilities naturally suggest themselves. One takes up the point mentioned earlier that, while there is a sense in which behavioral truths are perceptually available there is no comparable sense in which physical truths are. This suggests that, as regards ZA, we are not entitled to its first premise, viz.,

- 1 It is conceivable that I have a zombie-duplicate, i.e., someone identical to me in respect of all physical truths but different with respect to some phenomenal truth.

What we are entitled to instead is only a weaker premise, viz., 9:

9. It is conceivable that I have a *perceptually-available-zombie-duplicate*, i.e. someone identical to me in respect of all *perceptually available* physical truths but different with respect to some phenomenal truth.

Factoring this premise into the reasoning involved in ZA we will arrive at the view, not that phenomenal truths are contingently connected to all physical truths, but only that they are contingently connected to the perceptually available ones. But this conclusion a physicalist might well be able to live with.

This proposal has some of the features of the proposal that I think is ultimately correct, but it is nevertheless implausible. For it is not at all obvious that we *are* only entitled to 9. For example, functionalists of varying sorts postulate facts that are not in any obvious sense perceptually available, but nevertheless functionalism is usually thought of as lying in the target range of ZA.

There is, however, a slightly different proposal, which in effect generalizes the point about perception. Behavioral truths are not only perceptually available in a way in which physical truths are not, they are also *epistemically available* in a way in

which physical truths as a class are not. This suggests that, as regards ZA, what we are entitled to is not its first premise but rather 10:

10. It is conceivable that I have an epistemically-available-zombie-duplicate, i.e. someone identical to me in respect of all *epistemically available* physical truths but different with respect to some phenomenal truth.

Factoring this premise into the reasoning involved in ZA we will arrive at the view, not that phenomenal truths are contingently connected to all physical truths, but only that they are contingently connected to the epistemically available ones. But this conclusion, because it shows only the incompleteness in our knowledge, a physicalist might well be able to live with.

10. The response to ZA that replaces 1 with 10 is what I have called in previous work the epistemic view (see Stoljar 2006), though the route to that view that I have just offered is different to any proposed before. The epistemic view clearly raises a series of questions beyond the scope of this paper, and I will therefore leave those questions unanswered here. The important point for present purposes is rather to demonstrate how much may be wrung from the apparently simple contrast between ZA and AA. In particular, I have argued on the basis of the comparison (a) that the most influential current proposals about where ZA fails (assuming it does) are wrong; (b) that two very natural proposals about why ZA does not fail (assuming it does not) are likewise wrong; and finally (c) that one important respect in which AA and ZA differ is that behavioral truths are epistemically accessible in a way that physical truths are not. This latter fact, I also suggested, may be exploited to show that ZA is unsound.

References

- Bayne, T. and Chalmers, D. 2003. 'What is the Unity of Consciousness?' in A. Cleeremans, ed., *The Unity of Consciousness: Binding, Integration and Dissociation* (Oxford: Oxford University Press), 23-58.
- Block, N. 1981. 'Psychologism and Behaviorism', *Philosophical Review*, 90: 5-43.

- Block, Ned and Stalnaker, Robert. 1999. "Conceptual Analysis, Dualism, and the Explanatory Gap", *Philosophical Review*, 108, 1-46.
- Chalmers, David. 1996. *The Conscious Mind*. New York: Oxford University Press
- Chalmers, David. 1999 "Materialism and the Metaphysics of Modality", *Philosophy and Phenomenological Research* 59:473-93.
- Chalmers, David. 2002 'Consciousness and its Place in Nature', in D. Chalmers, ed., *Philosophy of Mind: Classical and Contemporary Readings* (New York: Oxford University Press), 247-272.
- Doggett, Tyler and Stoljar, Daniel. Forthcoming. 'Imagination and the Mind-Body Problem'.
- Hill, Christopher S. 1997. "Imaginability, Conceivability, Possibility and the Mind-Body Problem," *Philosophical Studies* 87: 61-85.
- Hill, Christopher S. and McLaughlin, Brian P. 1999. "There are Fewer Things in Reality Than are Dreamt of in Chalmers's Philosophy," *Philosophy and Phenomenological Research* 59: 445-454.
- Humberstone, L. 1996. 'Intrinsic/Extrinsic', *Synthese*, 108(2): 205-67.
- Jackson, F. 1998. *From Metaphysics to Ethics*. Oxford: Oxford University Press.
- Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Nagel, Thomas. 1976. "What Is It Like To Be A Bat?," reprinted in Nagel (1979).
- Nagel, Thomas. 1979. *Mortal Questions*. Cambridge: Cambridge University Press.
- Perry, J. 2001. *Knowledge, Possibility and Consciousness* (Cambridge, MA: MIT Press).
- Putnam, H. 1963. 'Brains and Behaviour', in R.J. Butler ed. *Analytical Philosophy, Vol.2*, (Oxford: Basil Blackwell & Mott), 1-19.
- Putnam, Hilary. 1975. *Mind, Language and Reality: Philosophical Papers, Vol.2*. (Cambridge: Cambridge University Press).
- Stalnaker, Robert. 2002. "What Is It Like to Be a Zombie?" in T. S. Gendler and J. Hawthorne, eds., *Conceivability and Possibility* (New York: Oxford University Press), 385-400, repr. in his *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays* (New York: Oxford University Press, 2003).
References are to the reprinted version.

- Sellars, W. 1962. 'Philosophy and the Scientific Image of Man' in R. Colodny, ed., *Frontiers of Science and Philosophy* (Pittsburgh: University of Pittsburgh Press, 1962): 35-78, repr. in his *Science, Perception and Reality* (London: Routledge and Kegan Paul, 1963), 1-40.
- Stebbing, L.Susan. 1958. *Philosophy and the Physicists* (New York: Dover) (Original Date: 1937).
- Stoljar, Daniel. 2005 "Physicalism and Phenomenal Concepts", *Mind and Language*, Vol.20, No.5, 469-494
- Stoljar, Daniel. 2006 *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness* New York: Oxford University Press.
- Stoljar, D. (forthcoming) 'Actors and Zombies', In A. Byrne and J. J. Thomson (eds.) *Content and Modality: Themes from the Philosophy of Robert Stalnaker* (Oxford: Oxford University Press).
- Sturgeon, S, 2000. *Matters of Mind: Consciousness, Reason and Nature* (New York: Routledge).
- Tye, M. 1999. 'Phenomenal Consciousness: The Explanatory Gap as a Cognitive Illusion', *Mind*, 108: 705-25.

Daniel Stoljar,
Senior Fellow, Philosophy Program
Research School of Social Sciences
Australian National University
Canberra ACT 0200 Australia
+61-2-6125-2076 (t)
+61-2-6125-3294 (f)
<http://philrsss.anu.edu.au/~dstoljar/>

