1 Introduction

In this paper I develop a new argument against interactionistic substance dualism, the view that human beings consist of physical bodies and non-physical souls which interact with each other. Substance dualism is not popular nowadays, so arguing against it might seem like floggin a dead horse. Why then should philosophers be interested in yet another argument against substance dualism? Whereas the traditional arguments against substance dualism, the interaction problem\textsuperscript{1}, the pairing problem\textsuperscript{2}, the argument from the closure of the physical\textsuperscript{3}, and the problem of specifying criteria of singularity and identity for souls\textsuperscript{4} are metaphysical arguments based on metaphysical assumptions, my argument will be purely semantic. I will try to show that a reductio of substance dualism can be construed with premises that are purely analytic or conceptual truths. If such a purely semantic reductio of interactionistic substance dualism succeeds this has some interesting consequences: The denial of interactionistic substance dualism would then turn out to be a conceptual truth and hence not a substantive metaphysical position.

\textsuperscript{1}For an overview over various formulations of the interaction problem see Robinson 2012, section 3.1.
\textsuperscript{2}Kim 2005, p. 78 ff.
\textsuperscript{3}McLaughlin 1993.
\textsuperscript{4}Strawson 1974, p. 190 ff.
Roadmap: I first define interactionistic substance dualism. I then proceed in two steps to develop an argument against it. First I have a look at a soul with a specific causal profile and argue for the modest claim that the non-physicality of this specific soul is inconsistent with a certain metaphysical view about properties. Then I replace the metaphysical premise in that argument with a related semantic one and generalize the argument to all minds to obtain a general reductio of interactionistic substance dualism. In the last section I consider some of the consequences of the argument for the principle of the causal closure of the physical and deflationism in philosophy of mind. I will argue that closure of the physical is a conceptual truth and that neither physicalism nor substance dualism are substantive metaphysical positions.

2 The Target: Interactionistic Substance Dualism

At the core of (interactionistic and non-interactionistic) substance dualism lies the following thesis:\(^5\)

**Dualism:** Every human being is composed of a part that is purely mental and a part that is purely physical.

What does it mean to be a purely mental part? The standard view, I take it, entails the following condition:\(^6\)

**Nonphysicality:** For any object $x$, if $x$ is purely mental then $x$ is a non-physical object.

Nonphysicality expresses only a necessary condition for being purely mental. This leaves logical space for non-physical non-mental objects. This seems right, various philosophers, among them Spinoza, have argued that there are, or at least could be,

---

\(^5\)See e.g. Hawthorne 2007, p. 97.  
\(^6\)See e.g. Ravenscroft 2005, p. 9.
objects which are neither mental nor physical. Nonphysicality shows that “being purely mental” is not on par with “being physical”. Rather it depends partially on what “being physical” means. Dualism and Nonphysicality are together compatible with epiphenomenalism, the claim that mental states do not influence physical objects. We can therefore introduce the following thesis to arrive at interactionistic substance dualism:

**Interactionism:** The purely mental part of a human being causally interacts with the purely corporeal parts.

The conjunction of Dualism, Nonphysicality, and Interactionism is equivalent to generic interactionistic substance dualism, the target of my paper. Some forms of interactionistic substance dualism may make additional claims, such as specific claims about where and how the purely mental part of a human being interacts with the purely physical parts. But since all these variations will still be committed to these three theses, arguing against them amounts to a general attack on interactionistic substance dualism.

## 3 Divine Amputation

There is a (merely?) possible world in which God exists and is a centre of phenomenal consciousness without any physical properties. He is what Richard Swinburne calls a spirit, a person without a body or a non-embodied person. And he has amazing causal powers. For any event which has ever been caused by a micro- or macrophysical object in the universe, God could have caused that event directly. Since God is a centre of phenomenal consciousness he has the power to bring about causal relations between a centre of phenomenal consciousness, namely himself, and physical objects.

---

7 Shoemaker 2003, p. 143.
10 Or if you think causal relations obtain between events: He has the power to bring about causal relations between events involving himself and events involving physical objects. In what follows I will assume that causal interactions are relations between objects. Anything I say about causal relations between objects can be translated into event-terminology with the following translation scheme: An object...
Now suppose God one day decides to do the following to you: He destroys your right foot in a single instant. But you won’t notice any of this because God henceforth plays the causal role that your foot would have played. Every single causal interaction your foot or any of its parts at any level of decomposition would have had with the world is mimicked by God. This thought can be made more precise. Let “$f_1$”, “$f_2$”, etc. be constants which refer to (proper or improper) parts of your right foot, and “$o_1$”, “$o_2$”, etc. constants which refer to objects which are not part of your foot. $T[C_1 f_1 o_1, C_2 f_2 o_2, ..., C_n f_n o_n]$ is a theory describing all the causal interactions between parts of your right foot and other objects, where the dyadic predicates $C_n$ express causal relation tokens between foot-parts and other objects during a certain time span. Every single causal interaction any part of your foot ever had with objects other than its own parts during that time will be captured in $T$. It is a complete causal history or profile of the parts constituting your right foot for a certain time span. The third assumption we made about God was that for any event event which has ever been caused by a micro- or macrophysical object in the universe, God could have caused that event directly. Translated into talk about causal relations between entities this means that for any causal relation between two objects $x$ and $y$, God could have been one of the relata of that causal relation. This means God could also have caused all the things that in the actual world are caused by your right foot. The theory that this is so can easily be obtained by modifying theory $T$. We simply have to substitute all the proper names for parts of your right foot for $g$, a proper name for God, to get the following theory: $T^*[C_1 g o_1, C_2 g o_2, ..., C_n g o_n]$.

Let a $T$-world be one in which $T$ is true and a $T^*$-world one in which $T^*$ is true. The way $T^*$ is obtained from $T$ guarantees that nobody could notice the slightest dif-

---

4

---

\[=\text{translation}\]

An event involving object $x$ exemplifying a property $P_m$ at $t_m$ stands in a causal relation $R$ with an event involving object $y$ exemplifying a property $P_n$ at $t_n$. This scheme is based on a variation of Kim’s theory of events, see Kim 1993, p. 8 ff.

\[11\]

I make the simplifying assumption that all objects are either part of your right foot or not and do not change that status.
ference between a T-world and a T*-world, not even you. If your foot in the T-world causes certain visual, tactile, auditory, and olfactory perceptions, then God causes those perceptions in the T*-world. Or to put it more bluntly, if your right foot looks, feels, sounds, and smells a certain way in a T-world, then in a T*-world God looks, feels, sounds, and smells the same way.

4 The Modest Claim

Is the divine amputation story consistent? Can God have the specific causal profile specified in T* and still be a centre of phenomenal consciousness without any physical properties? To decide that we need some sort of criterion of physicality.

4.1 The Direct Approach: Physicality of Properties

When is an object physical? Daniel Stoljar considers the following rough and ready criterion:

x is a physical object if and only if x has (or has enough of) the following properties: it has size, shape, extension in space, the capacity to move and be moved, the capacity to undergo various processes such as bending, breaking, and burning, and perhaps most importantly it has solidity or bulk – that is, it is intrinsically such that it resists or would resist pressure from other physical objects, for example, pressure from human bodies.12

This criterion has some obvious problems. It is circular because it mentiones other physical objects. It seems suitable only for macroscopic physical objects and not for physical elementary particles. It is vague due to the “enough of” clause. I suggest that we use instead the following criterion:

12Stoljar 2010, p. 52.
**Physicality of Objects:** For any object x, x is physical iff x has at least one physical property.

This criterion gives us the right results on an intuitive understanding of physical properties: All elementary particles of physics, my chair, my body, and my computer are physical objects, whereas God before the amputation is not a physical object. The burden of the criterion lies on the understanding of physical properties and in controversial cases such as God after the amputation an intuitive understanding will not be enough. We need a criterion for the physicality of properties. Here is a standard criterion:

**Physicality of Properties:** For any property x, if x is expressed by a predicate in the ideal theory of physics, or x is necessitated by such a property, then x is a physical property.

Unfortunately there is a well known objection to this formulation which goes by the name of Hempel’s dilemma. In what follows I do not want to rely on Physicality of Properties to develop my argument, for there is a way to decide whether God has physical properties which avoids problems related to Hempel’s dilemma.

### 4.2 The Indirect Approach: Sameness of Properties

Instead of asking whether God has physical properties in the T*-world we could ask whether God has the same properties in the T*-world your foot has in the T-world. If the answer is yes, then we know that God has physical properties in a T*-world, for your right foot surely has many physical properties in a T-world.

What we now need is not a criterion for the physicality of properties, but for the sameness or identity of properties. *Structuralism* is one metaphysical view of properties

---

13Stoljar 2010, p. 93 ff.
14By the same properties I mean that if your foot has a property, then God has that property too. I do not mean that if God has a property, then your right foot has it too.
which offers such a criterion. Examples of structuralist theories of properties are Whittle's causal nominalism\textsuperscript{15}, Shoemaker’s non-reductive causal theory\textsuperscript{16}, and Hawthorne’s causal structuralism\textsuperscript{17}. At the core of structuralist theories lies the following claim about property individuation:

**Structuralism:** Properties are individuated by their nomological roles/causal features.\textsuperscript{18}

Does God have the same properties in a T*-world as your foot in a T-world according to Structuralism? If properties are individuated by their causal features and God has the same causal features in a T*-world as your right foot in a T-world, then God has the same physical properties your right foot has in a T-world. Therefore God has at least one physical property and is thus, according to Physicality of Objects, a physical object. The concrete consequences of the positive answer vary depending on the specific form of structuralism in question. According to Whittle’s causal nominalism God’s phenomenal consciousness is not a physical property because it does not belong to the right ontological category. A property simply is an object’s satisfying a certain functional role and God’s phenomenal consciousness is, according to the first assumption about God, not an object’s satisfying a certain functional role. So on this view God has both physical properties (mass, weight, colour, surface structure, etc.) and mental properties (phenomenal consciousness). Interactionistic substance dualism collapses into property dualism.

According to Shoemaker’s non-reductive causal theory there are, in contrast to causal nominalism, quiddities, intrinsic natures of physical properties, but the necessary and sufficient condition for being a certain quiddity can be given in terms of its causal relations.\textsuperscript{19} The only thing which makes a contribution to God’s causal powers is his

\textsuperscript{15}Whittle 2009.
\textsuperscript{16}Shoemaker 1980.
\textsuperscript{17}Hawthorne 2001.
\textsuperscript{18}Whittle 2009, p. 249; Locke 2011, p. 348 f.
\textsuperscript{19}This short summary of Shoemaker’s theory is due to Hawthorne, see Hawthorne 2001, p. 376.
phenomenal consciousness, for God simply is a centre of phenomenal consciousness. So if quiddities are individuated by their causal powers, then the phenomenal consciousness that is God is the quiddity of his physical properties. Hence God has only physical properties. Interactionistic substance dualism collapses into property physicalism. There is room here for a variety of views here about which elements of God’s phenomenal consciousness are the quiddities of which physical properties. And according to some of these views there may be a stock of elements of phenomenal consciousness which end up not being a property’s quiddity. Then we get property dualism again instead of property physicalism.

We’ve started with the assumption that God is a purely mental entity without any physical properties and a certain causal profile. We’ve shown that this is inconsistent with a certain view about the metaphysics of properties: God after the amputation is (by stipulation) non-physical and, if structuralism is true, also physical. Contradiction! This is a fairly modest result and does not yet put a lot of pressure on the substance dualist. He has two options: he can either give up structuralism, or he can accept that minds with certain kind of causal profile are physical, but deny that ordinary souls have that kind of causal profile. In the next section I want to generalize the argument from God in the divine amputation case to all minds and replace Structuralism with a purely semantic premise.

5 The Sweeping Claim

To arrive at the conclusion that interactionistic substance dualism is conceptually incoherent we need to replace all premises in the argument with analytic truths or conceptual truths in some wider sense. Moreover, we need to generalize the argument from a purely mental object mimicking the causal profile of a foot, God in the divine amputation story,

---

20 Interactionistic substance dualism collapses into a specific kind of physicalism, namely non-reductive physicalism. Other names are type-F monism, realistic monism, or Russelian monism. See e.g. Strawson 2006, p. 4; Chalmers 2010, p. 133.
to all minds with all sorts of causal profiles.

5.1 Premise 1: Semantic Structuralism

Consider the following claim about physical property predicates (PPPs):

**Semantic Structuralism:** The meaning of PPPs is such that whether an object falls under their extension is determined by its causal profile alone.

Semantic Structuralism is a claim about the meaning of predicates expressing physical properties such as “being spherical”, “being red”, “having a mass of 1kg”, etc, and as such it is an analytic truth if it is true at all. It is not a metaphysical position rejecting quiddities. There might well be quiddities according to Semantic Structuralism, they simply play no role in determining the extension of PPPs. If a theory of physics is true in a world, then, necessarily, it is also true in any other world which is causally isomorphic with it. Or to put it in Lewis’ terms: the ideal theory of physics has one unique actual realization but many possible realizations.\(^{21}\) Russell had something like Semantic Structuralism in mind when he made his famous observation that physics does not tell us anything about the intrinsic character of the entities and properties that it discusses, only about their structural relations.\(^{22}\)

Is Semantic Structuralism true? In what follows I want to briefly mention three arguments in its favour. The first one is simply that it follows from the Ramsey-Carnap-Lewis account of theoretical terms\(^{23}\), the most plausible account of theoretical terms we have, and many PPPs are theoretical properties.

The second one is that rejecting Semantic Structuralism leads to a whole host of implausible consequences. If Semantic Structuralism is false and PPPs are as fine-grained as quiddities and not just as causal profiles, then theories of physics can be false.

\(^{21}\)Lewis 2009, p. 207.
\(^{22}\)Russell 1927.
\(^{23}\)Lewis develops this claim in Lewis 2009.
in possible worlds in which they correctly predict every single observations one might make. This would be the case if we consider the ideal theory of physics for the actual world in a twin-physics-world, a world which is causally perfectly isomorphic to ours but in which the role that certain quiddities play in our world are played by different quiddities. But that seems wrong, surely a theory of physics is true in every world in which it is empirically adequate. Suppose an actual world physicist in possession of the ideal theory of physics travels to a twin-physics-world. What would he have to do to make the theory true in that world? Reintroduce the theory? How does one do that, by proclaiming loudly that one reintroduces it?

The third argument is that the content of concepts is usually only as fine-grained as necessary for the purposes of the concept. The expression “Jonathan’s copy of CrunchBang Linux” is fairly fine-grained, it picks out a piece of software that is installed on my laptop. But it is not extremely fine-grained: If I restart my laptop and the software gets allocated slightly different memory it still falls under the extension of that expression. And the reason for this is, I think, that I have introduced the expression with a certain purpose, namely to talk with friends about certain modifications I had to make to my copy of CrunchBang in order to run certain programmes. And for that purpose it is completely unnecessary to distinguish different copies of CrunchBang based on their memory allocation. So with what purpose have PPPs been introduced? I think they have been introduced to describe the causal structure of the world. And therefore they are not more fine-grained than necessary to accomplish this.

We can use Semantic Structuralism to obtain the same result in the divine amputation case as we did with Structuralism. This means that we’ve successfully replaced a metaphysical premise with a purely semantic one.
5.2 Premise 2: Soul Zombies

What is left now is to replace the divine amputation story with something appropriate. This *something* has to be more general in order to cover minds with all kinds of causal profiles and it has to be an analytic or conceptual truth.

**Soul Zombies**: For every being with a soul, there is a possible world in which there is an exact duplicate of that being, but the events which are caused by its soul in the actual world are caused by a peculiar physical particle. This particle changes its properties always just in such a way that it brings about the events in the merely possible world which the soul brings about in the actual world.

The modal claim in soul zombies seems plausible. Granted, such a particle might have to violate the laws of our universe due to its rapid and random change of properties. And perhaps the particle would even need to have physical properties which are not instantiated in this world. But this does not seem to be a problem: Anyone who thinks that there are possible worlds where only a lawless physical universe exists thinks that there could be physical objects which rapidly and randomly change their physical properties. One could even give an argument for this conclusion, an argument which the dualist will accept: It is conceivable that there is a physical particle which randomly and rapidly changes its properties just in such a way that it mimics the causal profile of the soul of a specific person in the actual world. What is conceivable is possible. Therefore it is possible that there is such a particle. If Soul Zombies is true then we can simply apply Semantic Structuralism again to show that the souls in the actual world, which according to interactionistic substance dualism have no physical properties, have the same physical properties as their physical counterparts in the soul-zombie world. So we have again a contradiction.
But is Soul Zombies, a modal statement, a conceptual or analytic truth? This is more controversial, but I think it is. Here is an argument: In Meaning and Necessity Rudolf Carnap tried to identify necessity with analyticity and apriority.\textsuperscript{24} According to Carnap “necessarily p” is true if and only if “p” is L-true\textsuperscript{25}, where L-truth is truth in virtue of the semantic rules of a language\textsuperscript{26}. Suppose Carnap’s project succeeded. Would that turn Soul Zombies into an analytic truth? The modal logic S5 contains the following plausible iteration axiom: $\Diamond A \rightarrow \Box \Diamond A$. So if the modal claim Soul Zombies is true and S5 gets things right, then Soul Zombies is necessarily true. And if Carnap’s claim is right, then this means that Soul Zombies is an analytic truth.

Unfortunately Carnap’s project is generally considered to be a failure due to criticism by Quine\textsuperscript{27}, Kripke\textsuperscript{28}, Putnam\textsuperscript{29}, and others. But there is a very promising successor to Carnap’s project, a philosophical package called epistemic two-dimensional semantics without brute necessities which has been developed by David Chalmers.\textsuperscript{30} According to this package all necessity is either due to the primary intension of an expression (the Fregean sense) or due to the secondary intension (the Kripkean intension). And both options are in some broad sense semantic or linguistic.\textsuperscript{31} Brute necessities would be necessities which are neither explicable in terms of the Fregean sense of an expression nor the rigidity of an expression.\textsuperscript{32} They would indeed be worthy of the name “metaphysical necessities”. And it seems plausible that there aren’t any of those. As Cian Dorr puts it:\textsuperscript{33}

\textsuperscript{24}Carnap 1947.
\textsuperscript{25}Carnap 1947, p. 174.
\textsuperscript{26}Carnap 1947, p. 7ff.
\textsuperscript{27}Quine 1951.
\textsuperscript{28}Kripke 1972.
\textsuperscript{29}Kripke 1972.
\textsuperscript{31}Saying that it is metaphysically necessary that the Morning star is the Evening star is a bit of a misnomer: there is nothing metaphysical about this necessity. This kind of necessity is due to the rigidity of the expressions involved, and rigidity is certainly a semantic feature.
\textsuperscript{32}For a more precise characterization of brute necessities see Whittle 2010.
\textsuperscript{33}Dorr 2008, p. 53. Dorr does not use the terminology of two-dimensional semantics, but what he means by real definitions is broad enough to capture the two kinds of necessity which exist according to two-dimensional semantics.
[T]he idea of a metaphysically necessary truth whose necessity does not flow from real definitions plus logic really should seem quite strange. A notion of necessity that allowed for such necessary truths would seem uncomfortably like nothing more than an extra-strong variety of nomological necessity. But when something strikes us as impossible - say, the hypothesis that some duplicate of an electron is not itself an electron, we don’t just think of it as ruled out by a “law of metaphysics”: we feel that in some important sense, the idea just makes no sense at all.

If there are indeed no brute necessities and all necessities are in a broad sense semantic or linguistic, then we can reiterate the argument above and conclude that Soul Zombies is in a broad sense a semantic or linguistic truth. This means that we now have all we need. For every possible soul we can use Soul Zombies to arrive at a possible world in which a physical particle has the same causal profile as that soul, and then we apply Semantic Structuralism to show that the soul has the same physical properties as that particle and is hence not a purely mental thing in the sense required for interactionistic substance dualism. This is a reductio of interactionistic substance dualism based on conceptual premises alone.

6 Consequences

In what follows I will have a brief look at some of the consequences of the argument I’ve developed in this paper.

6.1 Non-Interactionistic Substance Dualism?

Does the argument work against all forms of substance dualism or only against interactionistic substance dualism? This depends on whether Soul Zombies also covers cases where a soul has an empty causal profile, i.e. where it doesn’t interact with any physical
object. Are there possible worlds with physical particles that do not interact with any
other physical objects under any circumstances? I’m not sure about this and will leave
this question open.

6.2 Closure of the Physical

The principle of the closure of the physical is a principle which often plays a role in
arguments against dualism. The generic formulation is:

**Closure of the Physical:** Every physical effect has physical sufficient
causes.

There are various views about the nature of the principle, some see it as a methodo-
logical guiding principle, others as a consequence of principles of physics such as the
law of conservation of energy. If the argument I have developed here is sound, then
Closure of the Physical is in a broad sense a conceptual truth. Even if a purely mental
mind would interact with a physical object it would thereby acquire physical properties
and therefore such a interaction would not violate the principle. And the same hold for
any other non-physical entity which might interact with physical objects. Therefore it
is conceptually impossible to violate the principle of the closure of the physical.

6.3 Conceptual Physicalism

A similar argument as the one developed here can be applied to interactionistic idealism,
the view that only non-physical minds exist and interact with each other. Therefore
both interactionistic substance dualism and interactionistic idealism are conceptually
incoherent. This means that the negation of both of those views, the view that there
are only physical things, is itself a conceptual truth. And if physicalism in this sense
is a conceptual truth then it is not a substantive metaphysical thesis. Exactly which

---

34 See e.g. Sturgeon 2008, p. 413f. Vicente 2006, p. 150.
35 This was Thomas Hobbes' view, see Hobbes 1889, p. 55, 1651, p. 24f.
form of physicalism is true is a substantive dispute: is it property dualism, reductive physicalism, or non-reductive physicalism?

So the argument here leads to a limited kind of deflationism in philosophy of mind, for it demotes physicalism to a somewhat uninteresting conceptual truth. But unlike similar deflationary positions in the philosophy of mind such as that of Carnap or Wittgenstein the view suggested in this paper does not rely on a general criterion of meaning to show that there is something wrong with substance dualism. It is based on specific semantic claims about a certain domain of language and therefore restricted to claims formulated in that language.

References


At least as far as the argument in this paper is concerned. Perhaps there are other arguments which show that there is something wrong with the distinctions between those views.


