



Universiteit Utrecht

# THE ZOMBIE ARGUMENT

a discussion of the soundness of the zombie argument

Geke Pals – 4148606  
Bachelor Thesis Philosophy  
Faculty of Humanities  
Department of Philosophy and Religious Studies  
Utrecht University  
12.451 words

Supervisor: Menno Lievers  
Second Supervisor: Annemarie Kalis

July 5, 2016

## **ABSTRACT**

In this thesis I inquire whether the zombie-argument is sound, and thus whether it refutes physicalism of the mind. The zombie-argument states that (P1) zombies are conceivable; (P2) if zombies are conceivable, then zombies are metaphysically possible; (P3) if zombies are metaphysically possible, then physicalism is false; (C) thus, physicalism is false. The zombies referred to by the argument are philosophical zombies, which are creatures physically identical to human beings, but lacking phenomenal consciousness. I inquire whether the zombie-argument is sound by investigating the two controversial premises of the argument, namely (P1) zombies are conceivable, and (P2) if zombies are conceivable, then zombies are metaphysically possible.

The thesis consists of three parts. The first part consists of an introduction in the subject, including an overview of physicalism of the mind, the zombie-argument and an overview of modal logic and the two-dimensional semantic framework. In the second part I discuss the first premise of the zombie-argument. I present two arguments saying that zombies are not conceivable: one by Robert Kirk, one by Robert Stalnaker. I criticize both arguments by showing that they are flawed, and conclude that the first premise of the zombie-argument is sound. In the third part I discuss the second premise of the zombie-argument. I present David Chalmers's argument in favour of it, and I present Stephen Yablo's argument against Chalmers. I also provide my viewpoint on the second premise, in which I attack Chalmers. This brings me to the conclusion that the second premise of the zombie-argument isn't sound. Therefore, the zombie-argument isn't sound, and physicalism of the mind is not refuted by it. This answers the main question of this thesis.

# TABLE OF CONTENTS

INTRODUCTION.....	3
PART I – THE ZOMBIE ARGUMENT.....	5
Section 1 – Physicalism of the Mind .....	5
Section 2 – What are Zombies?.....	6
Section 3 – The Zombie-Argument in detail.....	7
Section 4 – Modal Logic and the Two-Dimensional Semantic Framework.....	8
PART II – THE CONCEIVABILITY OF ZOMBIES.....	11
Section 1 – About ‘Conceiving’ .....	11
Section 2 – Robert Kirk on the Conceivability of Zombies .....	13
Section 3 – Criticism of Kirk’s argument.....	15
Section 4 – Robert Stalnaker on the Conceivability of Zombies .....	16
Section 5 – Criticism of Stalnaker’s argument.....	17
Subconclusion.....	19
PART III – THE CONCEIVABILITY ARGUMENT .....	20
Section 1 – About ‘Possibility’ .....	20
Section 2 – David Chalmers, The Conceivability Argument and Zombies .....	21
Section 3 – Stephen Yablo on the Conceivability Argument.....	23
Section 4 – My thoughts on the Conceivability Argument.....	24
Subconclusion.....	29
CONCLUSION.....	30
BIBLIOGRAPHY .....	31

## **INTRODUCTION**

Philosophy of Mind is a discipline in philosophy, which is mostly concerned with the famous mind-body problem, brought forward by René Descartes in his *Meditations*. This problem is a fundamental one, as it is concerned with the question what we human beings consist of. Do we consist of two entities – mind and body – or are we just bodies? These are the two main positions in the debate: the former being dualism, of which Descartes is the founding father; the latter being materialism or physicalism. The difference between the two is in their answer to the ontological question what the mind is: is it physical or non-physical?

The mind possesses both mental states and consciousness, which itself can be split into properties like self-consciousness, sentience, and phenomenal consciousness. The latter is what this thesis is concerned with. Phenomenal consciousness consists of the what-it-is-like feeling that is familiar to all of us: there is something it is like for you to eat your favourite ice-cream, just like there is something it is like for you to feel the hot sun on your skin. Intuitively you might think that this what-it-is-like feeling can't be purely physical: how could physical properties provide for such a feeling? As I describe in this thesis, proponents of physicalism of the mind do have a coherent story that explains how phenomenal consciousness can be physical, and physicalism is a major position in the mind-body debate. However, there are also opponents. This thesis will be concerned with one of the arguments that opponents use in order to show that physicalism is false. This argument is concerned with philosophical zombies – it is called the zombie-argument. Zombies are creatures physically identical to human beings, the only difference being that they don't have the phenomenal consciousness that we have: there is no what-it-is-like to be a zombie. Proponents of the zombie-argument argue that if such zombies are possible, this would refute physicalism of the mind. Why? Because physicalism states that human beings are only physical and have phenomenal consciousness; our phenomenal consciousness is thus constituted from our physical properties, only. This would mean that creatures physically identical to us should also have phenomenal consciousness as provided by their physical properties. However, zombies are by definition physically identical to us but do not have phenomenal consciousness, which thus refutes physicalism.

The relevance of the discussion between dualism and physicalism works in two ways. First, an answer to the mind-body problem gives us an answer to the question what we are, and provides us with more knowledge about ourselves. Information like this is useful in social sciences like psychology, as this is concerned with an explanation of our behaviour, which is closely tied to what we consist of. Secondly, an answer to the mind-body problem can help us in the development of artificial minds, which is currently a hot topic with the rise in popularity of robots and other forms of artificial intelligence. Artificial models of the mind in these areas are based on human minds, and more knowledge of what human minds consist of will thus provide us with better insight into how these artificial minds should be constructed. The relevance of the mind-body debate is thus twofold: both in a better understanding of our own minds, and in a better construction of artificial minds.

The method used in this thesis is the common method used in philosophy. I will critically analyse various arguments and the way concepts are used, and on the basis of that I will form my own criticisms and a personal opinion on the zombie-argument.

The introduction just provided is only a short one. In the rest of this thesis I will provide a broader survey of the terms and positions involved in the zombie-argument. The main question of this thesis is whether the zombie-argument is sound, and thus whether physicalism is refuted by it. I answer this question by investigating the two controversial premises used in the zombie-argument.

This thesis consists of three parts. Part I consists of an introduction into the subject and the context of the zombie-argument. I give a description of the position of physicalism of the mind, of the definition of zombies, and of the zombie-argument in detail. In this introductory part I also give an overview of modal logic and the two-dimensional semantic framework.

Part II of this thesis is concerned with the first premise of the zombie-argument, which states that zombies are conceivable. First I present different definitions of ‘conceivable’, and next I consider two criticisms of this premise of the argument. In return, I analyse these criticisms.

Part III of this thesis is concerned with the second premise of the zombie-argument, which is called the Conceivability Argument. I first provide a brief summary of various definitions of ‘possibility’. Next, I outline an argument in favour of the Conceivability Argument, given by a prominent participant in the debate, David Chalmers, and I consider one criticism of his position. Lastly, I provide my own viewpoint on the matter.

I refute the second premise of the zombie-argument, which forms the conclusion that the zombie-argument is not sound.

## **PART I – THE ZOMBIE ARGUMENT**

This part provides an introduction to the zombie-argument. The zombie-argument is an argument used by opponents of physicalism of the mind, one of the main positions in the mind-body debate. Section 1 consists of a survey of physicalism, and how physicalism provides us with phenomenal consciousness. In Section 2, I explain what zombies are in detail, and why they form a problem for physicalism. In Section 3, I give a detailed account of the zombie-argument. I also indicate which premises of the argument are controversial, and in what way. These remarks form the starting points for the rest of my thesis, in which I discuss the premises in detail. However, an introduction into modal logic and the two-dimensional semantic framework is needed first, as I do in Section 4. Both are used throughout the thesis, so some knowledge about it is required.

### ***Section 1 – Physicalism of the Mind***

Briefly stated, physicalism is the metaphysical thesis that everything is physical, or that everything supervenes on the physical. Supervenience means that two sets of properties are related to each other in the following way: there cannot be a difference in one set without also a difference in the other set.<sup>1</sup> In the case of physicalism of the mind this means that consciousness supervenes on the physical, i.e. that there couldn't be a change in the physical properties without also a change in consciousness.

The supervenience of the mental on the physical has been explained in different ways by philosophers. Joseph Levine defines it as 'realizing'. In this sense, A, which is physical, realizes B, which doesn't have to be physical. This realization involves metaphysical necessitation, but only in one way, namely in the way that when A occurs, it also always realizes B. B on the other hand can be realized in multiple ways, so it doesn't have to be that A always occurs when B occurs. Levine then formulates the Materialist Thesis as follows:

“Only non-mental properties are instantiated in a basic way; all mental properties are instantiated by being realized by the instantiation of other, non-mental properties.”<sup>2</sup>

Robert Kirk explains supervenience by arguing that every physicalist is committed to the Strict Implication Thesis:

“P strictly implies Q.”<sup>3</sup>

In which P is the conjunction of all true statements in the vocabulary of idealized contemporary physics, and Q is the conjunction of all true statements in psychological language about the individuals whose existence physicalists assume to be provided for by P. Lastly, this 'strictly implying' means that 'P and not-Q' involves inconsistency or incoherence. This means that it is absolutely impossible that P is true and Q false.

An important thesis of physicalism is the causal closure of the physical. This means that every physical effect must have a physical cause – it can't be uncaused or have a non-physical cause.<sup>4</sup> There is thus no

---

<sup>1</sup> Brian McLaughlin and Karen Bennett, "Supervenience," in *The Stanford Encyclopedia of Philosophy (Spring 2014 Edition)*, ed. Edward N. Zalta, <http://plato.stanford.edu/entries/supervenience/> (consulted June 6, 2016).

<sup>2</sup> Josh Levine, *Purple Haze* (New York: Oxford University Press), 21.

<sup>3</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 10.

<sup>4</sup> Daniel Stoljar, "Physicalism," in *The Stanford Encyclopedia of Philosophy (Spring 2016 Edition)*, ed. Edward N. Zalta, <http://plato.stanford.edu/entries/physicalism/> (consulted May 24, 2016).

way that non-physical properties play any role in the explanation of physical effects. So far, all the evidence suggests that the whole physical world is closed under causation.<sup>5</sup> However, I will not discuss the causal closure of the physical any further in this thesis, but take it as an important thesis of physicalism.

This thesis is focussed on physicalism of the human mind, and specifically on phenomenal consciousness. This is the part of consciousness that is concerned with the what-it-is-like feeling, familiar to us all: there is something it is like for you to taste your coffee in the morning, and there is something it is like for you to smell the blossoming trees in the spring. The properties concerned with phenomenal consciousness are called 'qualia' in the philosophy of mind. These qualia are smells, feelings, pain, colours, etc.<sup>6</sup> I don't use the term 'qualia' in the rest of this thesis, because one can discuss what the term 'qualia' exactly entails. Therefore, when speaking of 'phenomenal consciousness', I will be concerned with the intuitive what-it-is-like feeling. Phenomenal consciousness is the biggest challenge that physicalism faces: many objections to physicalism rely on arguments concerning phenomenal consciousness – as does the zombie-argument.

### **Section 2 – What are Zombies?**

So what are these zombies that are supposed to refute physicalism? The zombies we talk about in this thesis are not the kind of zombies you would see in a horror movie, but are creatures just like us humans. They are usually called 'philosophical zombies', in order to distinguish them from the brain-eating creatures in horror scenarios. The general way to describe philosophical zombies is the following: imagine that there exists, in this world or in some other world, an exact physical twin of yourself. This zombie-twin looks like you, behaves like you, and matches you in every detail of your body, brain and behaviour. It is a particle-for-particle duplicate. Upon seeing your zombie-twin, nobody would notice any difference between you and your zombie-twin. Everyone would therefore also treat your zombie-twin as if it were conscious. However, this is the crux: by definition philosophical zombies have no phenomenal consciousness. So physically and behaviourally your zombie-twin is an exact duplicate of you, but the only difference is that it has no phenomenal consciousness –no what-it-is-like feeling.<sup>7</sup>

So why would these philosophical zombies form a problem for physicalism? As we have seen, physicalism entails that everything is physical or supervenes on the physical. This is thus also the case for phenomenal consciousness if physicalism is true in our world. Phenomenal consciousness therefore comes into being by means of human beings as physical entities. However, as philosophical zombies are physically exact duplicates of human beings but without phenomenal consciousness, this means that if they exist, the physical alone isn't enough to produce phenomenal consciousness. Opponents of physicalism conclude that something else – something non-physical – is needed to produce phenomenal consciousness, which is exactly what physicalists deny.

The central question in this discussion is not whether zombies can actually exist in our world – it is usually assumed that they cannot – but whether they, or a whole zombie world, are possible in a metaphysical sense, which would be a problem for physicalism.

In this section I have explained what philosophical zombies are and why they form a problem for physicalism. In the next section I describe the zombie-argument in detail.

---

<sup>5</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 8.

<sup>6</sup> Chris Buskes and Herman Simissen, ed., *Analytische Filosofie* (Nijmegen: Vantilt, 2014), 195.

<sup>7</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 3.

### **Section 3 – The Zombie-Argument in detail**

Now it is time to look at the zombie-argument in detail. This is how Tim Crane presents the argument:

- “(P1)           Zombies are conceivable.
- (P2)           If zombies are conceivable, then zombies are metaphysically possible.
- (P3)           If zombies are metaphysically possible, then physicalism is false.
- (C)            Therefore: physicalism is false.”<sup>8</sup>

The argument in this form is simple, and it can easily be concluded that it is valid. We thus have to consider whether the premises are justified, which would make the argument sound. Let’s take a closer look at the premises.

Premise (P3) is the most uncontroversial of the premises. As I described in the previous section, it is clear that the metaphysical possibility of zombies forms a problem for physicalism: because physicalism states that our phenomenal consciousness is constituted from our physical properties alone, the case of zombies – by definition creatures physically identical to us but without phenomenal consciousness –, refutes physicalism. I will therefore not pay further attention to (P3) in this thesis, and conclude that this premise stands – as is mostly consented.<sup>9</sup>

The justification of premises (P1) and (P2) is trickier to establish. Let’s take a look at premise (P1) first. When looking at this premise, an important question immediately arises: what does it mean to conceive of something? As I discuss in Part II, there are several varieties of conceiving. There are philosophers who refute premise (P1), and argue that zombies aren’t conceivable, because their conceivability is incoherent or already assumes that physicalism is false. In Part II I discuss these criticisms.

Next is premise (P2). This is an important premise for the argument, as it forms the bridge between conceivability and possibility, which creates the problem for physicalism. The premise can be broken down into smaller premises, which together have famously been called the ‘Conceivability Argument’. It runs as follows:

- “(P2.1)        Zombies are conceivable.
- (P2.2)        Whatever is conceivable is possible.
- (P2.C)        Therefore zombies are possible.”<sup>10</sup>

The difficulty with this argument is premise (P2.2). It is a widely discussed question in metaphysics whether conceivability entails possibility. In order to answer this question, the definition of ‘conceivability’ is important, just like the definition of ‘possible’. A strong proponent of the Conceivability Argument is David Chalmers. His main argument for showing that conceivability entails possibility is the use of his two-dimensional semantic framework, which I discuss in Section 5 of this part. In Part III, I apply this two-dimensional semantic framework directly to the case of zombies. There are also critiques, and I discuss one of them as well in Part III. Lastly, I provide my own viewpoint.

---

<sup>8</sup> Tim Crane, *Elements of Mind* (New York: Oxford University Press, 2001), 99.

<sup>9</sup> Tim Crane, *Elements of Mind* (New York: Oxford University Press, 2001), 100.

<sup>10</sup> Robert Kirk, “Zombies,” in *The Stanford Encyclopedia of Philosophy (Summer 2015 Edition)*, ed. Edward N. Zalta, <http://plato.stanford.edu/entries/zombies/> (consulted May 31, 2016).

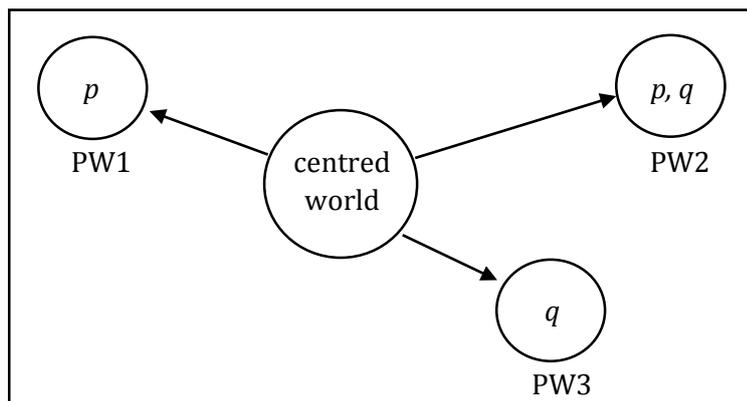
Before we continue to Part II and Part III of this thesis, in which I discuss premises (P1) and (P2), an introduction of modal logic and the two-dimensional semantic framework is needed. Both are used throughout this thesis, so some knowledge about it is required.

#### **Section 4 – Modal Logic and the Two-Dimensional Semantic Framework**

The two-dimensional semantic framework is an extension of the traditional one-dimensional modal logic. I therefore first provide a short introduction to modal logic.

Modal logic is the field of logics that talks about expressions like ‘it is necessary that’ and ‘it is possible that’. In order to determine whether a statement is necessary or possible, modal logic appeals to possible worlds: if a statement is necessary, this means that it is true in all possible worlds; if a statement is possible, this means that it is true in at least one possible world.<sup>11</sup> Examples of necessary statements are mathematical statements like ‘ $2+2=4$ ’; examples of possible statements, or contingent statements, are statements like ‘the table is white’. An intuitive way of talking about contingent statements is to say ‘it could have been different’; and an intuitive way of talking about necessary statements is to say ‘it could not have been different’.

An important feature of traditional one-dimensional modal logic is that it assigns the truth-value of statements according to just one parameter, which is possible worlds. This means that if you want to determine the truth, possibility or necessity of a statement, you take a centred world – usually the actual world, but not necessarily – as your starting point, and from that point you check the truth-value of the statement in all the possible worlds available. By doing this, you acquire a list of truth-values in the various possible worlds, from which you can define the truth, possibility or necessity of a statement. Figure 1 (next page) gives an example of such a model. In this model, we are inquiring whether statement  $p$  is possible and/or necessary. We start at the centred world, and then look at the possible worlds PW1, PW2 and PW3. As we can see,  $p$  is true in PW1 and PW2, but not in PW3 – in that world only statement  $q$  is true. We thus conclude that  $p$  is possible, but not necessary. This means that  $p$  is a contingent statement.



**Figure 1**

---

<sup>11</sup> James Garson, “Modal Logic,” in *The Stanford Encyclopedia of Philosophy (Spring 2016 Edition)*, ed. Edward N. Zalta, <http://plato.stanford.edu/entries/logic-modal/> (consulted May 27, 2016).

The two-dimensional semantic framework is developed in order to refine modal logic by extending the traditional framework with another parameter, making it two parameters. These parameters have to do with the dependence of the referent of a statement to the state of the world. This can be done in two ways. The first way is the way the referent of a statement is fixed in the *actual world*. This actual world can be any possible world, and thus also a world that is different from the world we live in. The reference of a statement thus depends on what that world is like: if it turns out one way, a concept or statement picks out one thing, but if it turns out another way, the concept or statement picks out something else. The second way is the way the referent of a statement is fixed in a *counterfactual world*. This thus takes as a premise the way the reference is already fixed in the world we live in; this is the starting point from which you look at other possible worlds.

Chalmers calls these two different ways of dependence the primary and secondary intension. Let's take a closer look at them:

“The *primary* intension of a concept is a function from worlds to extensions reflecting the way that actual-world reference is fixed. In a given world, it picks out what the referent of the concept would be if that world turned out to be actual.”<sup>12</sup>

As an example Chalmers takes the famous case of water. Chemistry has shown us that water consists of the molecules H<sub>2</sub>O, and therefore the identity statement ‘water is H<sub>2</sub>O’ is considered a truth in the world we live in. However, you still have the feeling that this could have been different. This is according to the primary intension of water: in different possible worlds, the concept of water could have different references. For example, in a possible world it could have turned out that XYZ was the fluid in the oceans, which would mean that in that world ‘water’ would refer to XYZ. The primary intension thus doesn't take the way the actual world is as a presupposition, but leaves it open to the way the possible world turns out. Let's say that in the case of water it picks out the *watery stuff* in a world, which can be H<sub>2</sub>O, XYZ, or something else – as long as it resembles the watery substance of our world.

The secondary intension is different. It takes the way the references of the world we live in are as a presupposition, and on that basis considers statements in counterfactual worlds. Thus, to take the case of water again, the reference of ‘water’ is fixed to H<sub>2</sub>O in the world we live in. Therefore, the secondary intension of ‘water’ picks out H<sub>2</sub>O in all worlds.

The biggest distinction between the two intensions should be clear: the primary intension is independent on empirical factors and therefore a priori, whereas the secondary intension is dependent on empirical facts, i.e. on the way the world we live in has turned out to be.

So why does Chalmers design this two-dimensional semantic framework? Why doesn't he just work with the traditional modal logic? The reason is to provide an answer to so-called ‘necessary a posteriori’ statements, as proved by Saul Kripke. Traditionally it was thought that only a priori statements could serve as a guide to necessity. However, Kripke famously proved that a posteriori statements could do so as well. The well-known example for this is ‘water is H<sub>2</sub>O’. Kripke showed, using his theory of rigid designators, that this statement is necessary: because it is an identity-statement with rigid designators on both sides, the statement is true in all possible worlds and thus necessary.<sup>13</sup> However, you still have

---

<sup>12</sup> David Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996), 57.

<sup>13</sup> Joseph LaPorte, “Rigid Designators,” in *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), ed. Edward N. Zalta, <http://plato.stanford.edu/entries/rigid-designators/> (consulted June 8, 2016).

the feeling that water could not have been H<sub>2</sub>O, that it could have been XYZ or something else. Traditional modal logic wasn't able to respond to these intuitive feelings, so the two-dimensional semantic framework came into play. As described above, this framework both admits the necessity of these a posteriori statements, and enables space for the intuitive feeling that it still could have been different. So, to summarize, the statement 'water is H<sub>2</sub>O' is necessarily true according to the secondary intension, and contingently true according to the primary intension.

For now, this short overview of the two-dimensional semantic framework will do. In both Part II and Part III of this thesis, its ideas return and are applied to the case of zombies.

## **PART II – THE CONCEIVABILITY OF ZOMBIES**

In this part I discuss two criticisms of the first premise of the zombie-argument, i.e. that zombies are conceivable. In order to do so, I first elaborate on the definition of ‘conceiving’ in Section 1. Next, I discuss two arguments for the inconceivability of zombies. In Section 2 I present Robert Kirk’s argument, which I criticize in Section 3. The second argument for the inconceivability of zombies is by Robert Stalnaker. I describe his argument in Section 4, and criticize it in Section 5. After that, I have shown that the arguments against the conceivability of zombies have failed, and that the first premise of the zombie-argument still stands.

### ***Section 1 – About ‘Conceiving’***

Before we start discussing whether zombies are conceivable or not, it is important to know what exactly is meant by ‘conceivable’. Authors usually use similar definitions, so I provide a survey of them that captures the general definition.

First, Chalmers offers three dimensions of differences between notions of conceivability. The first dimension of differences is between *prima facie* conceiving and ideal conceiving:

“S is *prima facie* conceivable for a subject when S is conceivable for that subject on first appearances. [...] S is ideally conceivable when S is conceivable on ideal rational reflection.”<sup>14</sup>

This distinction is pretty straightforward. It can happen that S is *prima facie* conceivable for someone, but after further reflection it turns out that S is actually not ideally conceivable. A good example for cases like this are mathematical statements which turn out to be true.

The second dimension of differences is between positive and negative conceivability:

“Negative notions of conceivability hold that S is conceivable when S is not *ruled out*. [...] The central sort of negative conceivability holds that S is negatively conceivable when S is not ruled out a priori, or when there is no (apparent) contradiction in S. [...] Positive notions of conceivability require that one can form some sort of positive conception of a situation in which S is the case. [...] Overall, we can say that S is positively conceivable when one can imagine that S: that is, when one can imagine a situation that verifies S.”<sup>15</sup>

Positive conceivability entails negative conceivability: if you can coherently imagine a situation in which S, this means that there is no contradiction in S as well. It doesn’t work the other way around: there can be cases where you cannot rule out S – thus you can negatively conceive S –, but where you can’t imagine a situation that would verify S – thus you can’t positively conceive S.

---

<sup>14</sup> David Chalmers, “Does Conceivability Entail Possibility?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 147.

<sup>15</sup> David Chalmers, “Does Conceivability Entail Possibility?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 149-150.

The third dimension of differences is between primary and secondary conceivability:

“We can say that *S* is *primarily conceivable* (or *epistemically conceivable*) when it is conceivable that *S* is *actually* the case. We can say that *S* is *secondarily conceivable* (or *subjunctively conceivable*) when *S* conceivably *might have been* the case.”<sup>16</sup>

This distinction is similar to the distinction between primary and secondary intensions, discussed in the previous part. We can take the example of ‘water is H<sub>2</sub>O’ to make it clear. Primary conceivability considers a possible world as actual, and doesn’t take the world we live in as a presupposition. Accordingly, it is primary conceivable that water is not H<sub>2</sub>O, but instead is XYZ. Primary conceivability is always an a priori matter: a statement is primary conceivable if and only if it is not ruled out a priori. Secondary conceivability considers a possible world as a counterfactual world, and takes the way the world we live in as a presupposition. Because the reference of water is fixed to H<sub>2</sub>O in the world we live in, it is not secondary conceivable that water is not H<sub>2</sub>O; it is only secondary conceivable that water is H<sub>2</sub>O. Secondary conceivability is thus dependent on empirical facts about the world we live in.

So what kind of conceivability do we need for the zombie-argument? It is important to note that the conceivability we need is supposed to lead us to possibility, as this is required by the argument. Chalmers is clear about the first dimension: prima facie conceivability is no good guide to possibility. We thus need ideal conceivability. Concerning the second dimension, Chalmers argues that positive conceivability is a better guide to possibility than negative conceivability. Finally, concerning the third dimension, Chalmers describes that primary conceivability functions as a guide to *primary possibility*, and that secondary conceivability functions as a guide to *secondary possibility*.<sup>17</sup> More on these two varieties of possibility in Part III of this thesis. For now, it is important to note that Chalmers considers ideal positive conceivability as the best guide to possibility.

Kirk also discusses the best way to understand ‘conceivability’. According to him, “a proposition or situation is conceivable only if no amount of a priori reflection on it would reveal contradiction or other incoherence.”<sup>18</sup> It thus seems that Kirk agrees with Chalmers on the aspect of ideal conceivability, but allows negative conceivability to be a good enough guide to possibility.

Chalmers thus uses a stronger notion of conceivability: he requires positive conceivability, while Kirk settles for negative conceivability. Even though it would be easier for opponents of the zombie-argument to attack positive conceivability – because positive conceivability requires more than negative conceivability –, they are satisfied with the harder-to-refute notion of negative conceivability. For proponents of the zombie-argument, the use of negative conceivability is strong enough, seen as its use only makes their argument harder to refute and therefore stronger. Therefore, in the rest of this thesis, I use the notion of (ideal) negative conceivability, as this is strong enough for both the proponents and the opponents of the zombie-argument.

---

<sup>16</sup> David Chalmers, “Does Conceivability Entail Possibility?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 157.

<sup>17</sup> David Chalmers, “Does Conceivability Entail Possibility?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 159-161.

<sup>18</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 28.

## Section 2 – Robert Kirk on the Conceivability of Zombies

Kirk argues that there is no coherent way to properly conceive of zombies. In this section I describe his argument. Kirk's concern is mainly focused on the causal closure of the physical. This means that, if zombies are conceivable, the conscious experiences of non-zombies must play no essential part in the causation or explanation of their behaviour. This is a consequence of the fact that zombies by definition behave in the same way as their human counterparts.

Kirk's argument is as follows:

- “(KP1) The e-quality story is not conceivable
- (KP2) If zombies were conceivable, the e-quality story would be conceivable
- (KC) Zombies are not conceivable”<sup>19</sup>

First, it is important to give an exposition of the so-called e-quality story and a definition of the related 'e-quality'. Kirk provides 5 theses:

- “(E1) The world is partly physical, and its whole physical component is closed under causation: every physical effect has a physical cause.
- (E2) Human beings stand in some relation to a special kind of non-physical properties, *e-quality*. E-quality make it the case that human beings are phenomenally conscious.
- (E3) E-quality are caused by physical processes, but have no physical effects: they could be stripped off without disturbing the physical world.
- (E4) Human beings consist of nothing but functioning bodies and their related e-quality.
- (E5) Human beings are able to notice, attend to, think about, and compare their e-quality.”<sup>20</sup>

Kirk's goal for (KP1) is to show a priori that the e-quality story is not coherent, because theses (E1)-(E4) are incompatible with thesis (E5).

Let us first consider what (E5) exactly means. The notion (E5) expresses what Kirk calls 'epistemic intimacy' or what Chalmers calls an 'intimate epistemic relation'.<sup>21</sup> This epistemic intimacy means that we can notice, attend to, think about, and compare the what-it-is-like of our experiences. Can these cognitive processes be *about* e-quality?

In order to answer the question, it is important to note that according to the e-quality story, e-quality are completely inert: they have no effects on the physical world or among themselves – this because of (E3). To show that we can't have epistemic intimacy with e-quality – which is what (E5) requires –, Kirk designs the example of sole-pictures. For this, we need a typical zombie-twin – *Zob* –, who lacks e-quality and therefore lacks phenomenal consciousness. According to the e-quality story, it is the case that if we add e-quality to *Zob*, he will suddenly become conscious (mainly according to theses (E2) and (E4)). The question is the following: how can *Zob*, after adding the e-quality, immediately have epistemic intimacy with his e-quality? An absurd alternative of the story is given by Kirk: let us imagine that the e-quality which are suddenly given to *Zob*, do not occur in his brain, but instead occur on the sole of his feet as constantly changing pictures. These 'sole e-quality' would still have the same role as 'brain e-quality', namely the role described to them in the e-quality story, and thus are compatible with theses (E1)-(E4). Now, is *Zob* in any way epistemically intimate with his 'sole e-quality'? It seems clear that he is not. This

---

<sup>19</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 39.

<sup>20</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 40.

<sup>21</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 42-43.

is so, because according to the story, Zob never even notices his sole e-qualia (I mean, who looks at the sole of their feet on a daily basis?), which means that they are in no way epistemically relevant to his cognitive processes: not in the thinking itself, and not in what he is thinking is about. Thus, there is no epistemic intimacy with them.

What this story intends to illustrate is that the fact that certain items are caused by cognitive processes ((E3)) is not sufficient to show that these cognitive processes are *about* the items that they cause, or that these items are in any way epistemically relevant to them. This means that the e-qualia as described by theses (E1)-(E4) cannot produce the epistemic intimacy required by (E5). We have thereby established (KP1): the e-qualia story is not conceivable.

On to (KP2): would the conceivability of zombies entail the conceivability of the e-qualia story? Kirk argues for it by equating the conceivability of a world  $z^*$  – which he argues zombists are committed to – to the e-qualia story. Let's take a step back. Zombists claim that zombies are conceivable, and therefore also zombie worlds. Such a zombie world would be causally closed. We can now ask the question what would be needed to provide the inhabitants of zombie world  $z$  with consciousness: something has to be added, and this has to be non-physical. Let's call this item  $x$ . Zombists cannot deny that the conceivability of zombies would entail that it is also conceivable that a non-physical item  $x$  could be added to the zombie world  $z$  in order to transform it into a world  $z^*$ , in which the ex-zombie inhabitants would be conscious. They are also committed to the view that this  $x$  conceivably provides epistemic intimacy. This means that the conceivability of zombies would entail the conceivability of a world  $z^*$ , which satisfies the following conditions:

- “(Z1)  $z^*$  is partly physical, and its whole physical component is closed under causation: every physical effect in  $z^*$  has a physical cause.
- (Z2) The human-like organisms in  $z^*$  are related to a special kind of non-physical item  $x$ .  $x$  makes it the case that they are phenomenally conscious.
- (Z3)  $x$  is caused by physical processes but has no physical effects: it could be stripped off [the human-like inhabitants] without disturbing the physical component of  $z^*$ .
- (Z4) The human-like inhabitants of  $z^*$  consist of nothing but functioning bodies and their related  $x$ .
- (Z5) The human-like inhabitants of  $z^*$  are able to notice, attend to, think about, and compare the qualities of their experiences.”<sup>22</sup>

As we can see, (Z1)-(Z5) almost exactly mirrors (E1)-(E5), with just three differences: ' $z^*$ ' instead of 'the world', 'human-like inhabitants' instead of 'human beings'; and 'a special kind of non-physical item  $x$ ' instead of 'e-qualia'. However, these differences are not significant.<sup>23</sup> Therefore, we can equate theses (Z1)-(Z5) with the e-qualia story, which shows that if zombies are conceivable, then the e-qualia story would be conceivable as well. We thus have established (KP2).

So let's return to the original argument – remember:

- “(KP1) The e-qualia story is not conceivable
- (KP2) If zombies were conceivable, the e-qualia story would be conceivable

---

<sup>22</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 51.

<sup>23</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 52.

(KC) Zombies are not conceivable”<sup>24</sup>

I first showed that (KP1) is right, by showing that there is an incoherence in the e-quality story. After this, I showed that zombists are committed to the conceivability of a world  $z^*$ , which can be equated to the e-quality story. Thus, zombists are committed to the conceivability of the e-quality story, which proves (KP2). Together, they form the conclusion (KC) that zombies are not conceivable. Kirk has thus shown that the first premise in the zombie-argument is not right.

### Section 3 – Criticism of Kirk’s argument

In this section I argue that Kirk’s argument for the incoherence of the e-quality story isn’t valid. Because Kirk’s argument that zombies aren’t conceivable needs this incoherence of the e-quality story, I will in the end have argued that Kirk’s argument for the inconceivability of zombies isn’t sound.

Kirk’s argument for the incoherence of the e-quality story is based on the claim that theses (E1)-(E4) are incompatible with thesis (E5), which is the thesis that accounts for the epistemic intimacy with e-quality. He uses the example of the ‘sole e-quality’ to show this. So far, the argument seems sound to me. However, Kirk then makes the move that in fact the e-quality produced by theses (E1)-(E4) can’t possibly produce the epistemic intimacy required by (E5). To ensure I present Kirk’s view fairly, I quote the corresponding passage:

“If that is right, we can see a priori that the conception of consciousness envisaged in the first four clauses (E1)-(E4) of the e-quality story precludes conscious subjects from being able to think about, notice, attend to or compare the qualities of their experiences. The story allows individuals no more epistemic access to their e-quality than *Zob* has to his sole-pictures, or I have to the electric currents in my brain: that is, none. Thus the e-quality conception of consciousness is not just odd or empirically improbable but incoherent: clauses (E1)-(E4) could be true only if (E5) were false.”<sup>25</sup>

This is where the problem occurs. Kirk makes the step from insufficiency to impossibility without further explanation, and thus acts like insufficiency to produce is the same as *impossibility* to produce. A simple example can show that this is not true: if I want to make a delicious grilled ham and cheese sandwich, only bread, butter, cheese and a grill aren’t sufficient to do this. However, this doesn’t show that it is *impossible* for me to make a grilled ham and cheese sandwich: if I simply walk to the supermarket and buy some ham, I will be able to satisfy my hunger. Therefore, insufficiency isn’t the same as impossibility.

At this point you might say that in the grilled sandwich example it is right to simply add the ham in order to make the desired sandwich. However, you might say, in the case of an argument like the one by Kirk, you can’t just add stuff in order to make the argument compatible. Well actually, *you can*. This is the whole point of premises in an argument: to add enough ‘stuff’ – properties, requirements, and the like – in order to provide a compatible story or conclusion. It seems like Kirk treats (E5) as the conclusion that follows from (E1)-(E4). In that case, the insufficiency of (E1)-(E4) to provide for (E5) would indeed be a problem, because the validity of an argument depends on the sufficiency of the premises to provide for the conclusion. However, the e-quality story only consists of five *theses*, which are all equal and don’t have to provide for each other. It is therefore no problem that (E1)-(E4) do not provide for (E5), because

---

<sup>24</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 39.

<sup>25</sup> Robert Kirk, *Zombies and Consciousness* (New York: Oxford University Press, 1996), 47-48.

(E5) is a thesis just like the other theses, which stands on its own and doesn't need any other theses to provide for it. Put this way, it is clear that the move Kirk makes from insufficiency to impossibility isn't a valid move in this case. Therefore, theses (E1)-(E4) are compatible with thesis (E5), which makes the e-qualia story coherent, and therefore makes Kirk's argument against the conceivability of zombies unsound.

#### **Section 4 – Robert Stalnaker on the Conceivability of Zombies**

I now consider Stalnaker's argument for the inconceivability of zombies. He argues that there is too much noise at the key-concepts used in the zombie-argument, in the sense that it is not clear which definitions are used and what already has been presupposed. The noise occurs at the concepts of zombies and zombie-worlds. Stalnaker uses the strategy of rewriting these concepts in a theoretically neutral way, to show the problem in the argument: that the conceivability of zombies already presupposes that physicalism is false. Let's take a look at the rewritings of these concepts.

First, Stalnaker defines the '*z-world*', which is a possible world exactly like the actual world in all physical respects. This world contains nothing except what supervenes on the physical and is thus a physicalistic world. Anyone who believes that physicalism is true, will believe that the actual world is a z-world.<sup>26</sup> Take note that this definition of the '*z-world*' says nothing about the presence or absence of consciousness. This means that any physicalist who believes that we are conscious beings – which is every non-eliminative physicalist – will deny that the z-world is a zombie world, while any proponent of the zombie-argument will insist that the z-world is a zombie world.

Next, Stalnaker defines the '*a-world*'. This world is physically just like the actual world, and therefore also just like the z-world. However, in addition it has some properties not instantiated in the z-world – let's call them a-properties. A-properties are properties of conscious beings, and they are the properties we are referring to when we talk about consciousness, qualia, and the like.<sup>27</sup>

So now we have two worlds, of which only one can be the actual world. How do we decide what is the actual world? It all comes down to what consciousness is, as this property is the only difference between the two worlds. At least, if we assume that we human beings have consciousness – which is something that generally has been assumed. However, we do not know *what* consciousness is, seen as this is exactly what the argument is about. Stalnaker uses the following method: he considers both the z-world and the a-world as actual, in order to see what the consequences are.

So, first we consider the z-world as actual. If this is true, then anti-physicalists would have to admit that we actually are zombies in this world, because according to them our physical properties aren't sufficient to produce consciousness. This is a conclusion that seems counterintuitive: we clearly do have the intuition that we are conscious.

But can't we, inhabitants of the z-world, then still conceive of zombies? Shouldn't we trust this intuition? This is where Stalnaker uses Chalmers's two-dimensional semantic framework in order to show that this conceivability doesn't say anything about the possibility of zombies. He uses the example of water and H<sub>2</sub>O again. We can in some way conceive of water not being H<sub>2</sub>O, namely in the primary way of

---

<sup>26</sup> Robert Stalnaker, "What is it Like to be a Zombie?," in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 387.

<sup>27</sup> Robert Stalnaker, "What is it Like to be a Zombie?," in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 395.

conceiving. It is conceivable that the XYZ-theorist is right, and, if so, that we would refer to something other than H<sub>2</sub>O with the term 'water'. Let's take this counterfactual world as the actual world. So, in this world, the XYZ-theorist believes that 'water' refers to XYZ, and that this is necessary. She also grants that if the H<sub>2</sub>O-theorist is right, then we would all be referring to H<sub>2</sub>O with the term 'water', rather than to water in the XYZ-sense. Then, in this world, it would be conceivable that there be a possible world where there is no water, even though the lakes and oceans are filled with H<sub>2</sub>O. This is surely impossible, as this is exactly the way our actual world is: filled with H<sub>2</sub>O. However, it is still *conceivable* for the XYZ-scientist, just as it is conceivable for us to describe the XYZ-world as a possible world with XYZ in the lakes and oceans, but no water. This shows that if something is conceivable, it isn't therefore possible as well. So, the intuitions about the conceivability of zombies in the z-world don't say anything about the possibility of zombies, and therefore don't form any problem for physicalism.<sup>28</sup>

Next, we consider the a-world as actual. From this point we can consider a counterfactual world, namely the zombie-world. The reasoning is well known: if we didn't have any of the a-properties, but the world was physically just as it is, then we wouldn't be conscious. We would thus be in a zombie world. However, in order to consider this counterfactual, the assumption of the a-world as being the actual world has to be made. If we assume we live in a z-world, which is what physicalists claim, then zombie worlds wouldn't be conceivable, seen as it is in the definition of the z-world that it only contains the physical and that which supervenes on the physical, and thus also consciousness – assumed that we are conscious. A physically similar world to the z-world but without consciousness thus wouldn't be conceivable.

This means that in order for zombie worlds to be conceivable – which is needed for the zombie-argument – we already have to assume that physicalism is false and that we thus live in an a-world. This undermines the whole argument: an argument can't be valid if you first have to assume the thing you are arguing for.

Altogether, Stalnaker has argued that zombies are not conceivable, by arguing that you first have to assume that physicalism is false in order for zombies to be conceivable. It is an invalid move to first assume the thing that you are arguing for. Therefore, zombies aren't conceivable.

### ***Section 5 – Criticism of Stalnaker's argument***

In this section I argue that the argument Stalnaker uses to show that zombies aren't conceivable isn't valid. In his argument, Stalnaker uses the following strategy: he first assumes a possible conclusion of the zombie-argument – namely, either that we live in a physicalistic world (z-world) or in a non-physicalistic world (a-world) –, and from thereon he checks whether the argument still works. In particular, Stalnaker checks whether one of the premises still holds, namely that zombies are conceivable. If you are arguing along this strategy, it's not a surprise that the argument only works in the case that the conclusion fits with what the argument is arguing for. Let's look at a very simple example of a valid argument, in order to see what Stalnaker is doing:

- (SP1) Socrates is a philosopher
- (SP2) If Socrates is a philosopher, then Socrates is smart
- (SC) Therefore, Socrates is smart

---

<sup>28</sup> Robert Stalnaker, "What is it Like to be a Zombie?," in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 398.

We use the same strategy as Stalnaker: we assume first the truth and then the falsity of the conclusion, so first that Socrates is smart and then that Socrates is not smart. Let's start with the first: Socrates is smart. Everything is fine: it is clear that this assumption doesn't form any problem for the premises. Now we take the second possible assumption: Socrates is not smart. Now it is clear that one of the premises has to be false, because the premises don't support this conclusion and the argument is clearly valid. So either (SP1) or (SP2) is false if we assume that Socrates is not smart.

This clearly demonstrates the problem. You can't prove that an argument or a premise in an argument is wrong by assuming the truth or falsity of the conclusion: that is what the whole argument is for. This is a way of turning the whole argument upside down. To come back to Stalnaker's specific argument: it doesn't make sense to assume both the z-world and the a-world in order to check the soundness of the zombie-argument. You have to go into the argument without any assumptions: we don't know whether we live in a z-world or an a-world, as this is what the argument is supposed to tell us. From thereon we have to check whether it is conceivable and possible that there are zombies.

At this stage, it seems like Stalnaker's discussion of the conceivability of zombies from the z-world is relevant, as this explores the conceivability of zombies regardless of assumptions about which world we live in. However, Stalnaker mixes up the link between primary conceivability and secondary possibility when he uses Chalmers's two-dimensional framework. This part of my critique is inspired by Chalmers's answer to Stalnaker<sup>29</sup>. In the argument, Stalnaker asks us to consider the XYZ-world – where water is XYZ – as actual. This is a necessary identity-statement, just as is the case with water being H<sub>2</sub>O in our world. Considering this world, it is conceivable, Stalnaker says, that water is actually H<sub>2</sub>O and thus not XYZ. However, note that this can only be conceivable in the primary way, seen as in the XYZ-world the identity between water and XYZ is necessary. The next step is important. Let's see how Stalnaker describes it, so there can be no confusion about the right words to use:

*“But now consider whether it is conceivable that there should be no water in the world at all, even though H<sub>2</sub>O fills the lakes and streams, and falls from the sky on rainy days, just as it does in the world as we believe it to be? That is, try to envision a possible world physically and chemically exactly like our, but with no water. This, we say, is surely impossible, but the XYZ theorist disagrees. He thinks that the world we think is actual is a world that is correctly described in this way – as a world with H<sub>2</sub>O, but no water, in the lakes and streams.”<sup>30</sup>*

The impossibility Stalnaker now talks about is secondary impossibility, because the H<sub>2</sub>O-world is taken as a counterfactual to the XYZ-world, and in this sense is proven to be impossible. Because of the necessity of water being XYZ in the XYZ-world, there is no 'water' in this sense if we consider the H<sub>2</sub>O-world. However, if we would have considered whether the H<sub>2</sub>O-world was possible in the primary way, this wouldn't have caused any problems: because in that case, you would understand 'water' as 'watery stuff', which can apply both to XYZ and to H<sub>2</sub>O. Stalnaker thus makes a mistake in his argument. He argues that the link between conceivability and possibility doesn't work, but he only does this for primary conceivability and secondary possibility. This is a move that you're not allowed to make in the two-

---

<sup>29</sup> See: David Chalmers, “Does Conceivability Entail Possibility?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 198-199.

<sup>30</sup> Robert Stalnaker, “What is it Like to be a Zombie?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 398.

dimensional semantic framework, as Chalmers clearly describes.<sup>31</sup> Therefore, the argument that Stalnaker uses to show that conceivability doesn't entail possibility is not valid.

I have shown in this section that Stalnaker makes two mistakes in his argument. First, the argument that zombies aren't conceivable isn't valid, because it's not a valid move to first assume a possible conclusion to the argument in order to show that one of the premises in the argument is wrong. Second, Stalnaker's argument to show that conceivability doesn't entail possibility isn't valid, because he actually shows that primary conceivability doesn't entail secondary possibility, which is a move you're not allowed to make in the two-dimensional semantic framework.

### ***Subconclusion***

This part focussed on the first premise of the zombie-argument, i.e. the conceivability of zombies. First, I showed that the right notion of 'conceivability' is ideal negative conceivability, which I use in the rest of this thesis. In Section 2 I presented Kirk's argument for the inconceivability of zombies, for which he argued by showing that his e-quality story is incoherent. In Section 3 I showed that his argument was flawed: Kirk treated one of the premises of the e-quality story as a conclusion, which is why he thought the argument was incoherent. In Section 4 I described Stalnaker's argument for the inconceivability of zombies, for which he argued by showing that the conceivability of zombies presupposes the falsity of physicalism. In Section 5 I showed that his argument too was flawed, in two ways: he used an invalid form of reasoning by first presupposing the conclusion of the argument, and he mixed up the primary and secondary intensions in his argument. The conclusion of this part is that the first premise of the zombie-argument still stands: zombies are conceivable.

---

<sup>31</sup> David Chalmers, "Does Conceivability Entail Possibility?," in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 161-165.

## **PART III – THE CONCEIVABILITY ARGUMENT**

In this part I discuss the Conceivability Argument, which constitutes the second premise of the zombie-argument (see p.6). As a reminder, the argument is as follows:

- “(P2.1)           Zombies are conceivable.  
(P2.2)           Whatever is conceivable is possible.  
(P2.C)           Therefore zombies are possible.”<sup>32</sup>

The argument thus aims to form the bridge between conceivability and possibility, which is needed for the zombie-argument to work. First, I describe different definitions of ‘possibility’ and which definition we need for the zombie-argument to work. Next, I give an overview of Chalmers’s positive account of the Conceivability Argument. In reply to this I examine Stephen Yablo’s critique of Chalmers. Lastly, I provide my viewpoint on the Conceivability Argument, in which I show that Chalmers’s Conceivability Argument isn’t sound. Therefore, I refute the second premise of the zombie-argument

### **Section 1 – About ‘Possibility’**

In Part II, I discussed the definition of ‘conceivability’, and came to the conclusion that the notion of ideal negative conceivability is strong enough for both the proponents and the opponents of the zombie-argument. I now consider what kind of possibility is needed for the zombie-argument to work.

It is customary to distinguish between logical, metaphysical and physical possibility. Logical possibility is the most inclusive notion, and includes any proposition that logic allows or even necessitates – as is the case with tautologies. Metaphysical possibility is narrower, and excludes the logical possibilities that conflict with the nature of things that could have existed. The third kind of possibility is physical possibility, which excludes the logical and metaphysical possibilities that conflict with the physical laws of nature.<sup>33</sup> This type of possibility is thus the narrowest. I will not be concerned with physical possibility in this thesis.<sup>34</sup>

So which notion of possibility do we need for the zombie-argument to work? Most philosophers hold that metaphysical possibility is needed for the zombie-argument to be a true objection to physicalism.<sup>35</sup> The main proponent of the zombie-argument, Chalmers, also agrees that metaphysical possibility is at issue in the Conceivability Argument, and thus the zombie-argument as well.<sup>36</sup>

Therefore, in the rest of this thesis, I assume that metaphysical possibility is needed in the zombie-argument. In the next section I discuss Chalmers’s argument for the Conceivability Argument. On the basis of his two-dimensional semantic framework he defines two notions of possibility: primary

---

<sup>32</sup> Robert Kirk, “Zombies,” in *The Stanford Encyclopedia of Philosophy (Summer 2015 Edition)*, ed. Edward N. Zalta, <http://plato.stanford.edu/entries/zombies/> (consulted May 31, 2016).

<sup>33</sup> Anand Vaidya, “The Epistemology of Modality,” in *The Stanford Encyclopedia of Philosophy (Summer 2015 Edition)*, ed. Edward N. Zalta, <http://plato.stanford.edu/entries/modality-epistemology/> (consulted June 8, 2016).

<sup>34</sup> Chalmers also makes a distinction between primary and secondary possibility, but I discuss that in the next section.

<sup>35</sup> Robert Kirk, “Zombies,” in *The Stanford Encyclopedia of Philosophy (Summer 2015 Edition)*, ed. Edward N. Zalta, <http://plato.stanford.edu/entries/zombies/> (consulted May 31, 2016).

<sup>36</sup> David Chalmers, “Does Conceivability Entail Possibility?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 146.

possibility and secondary possibility. I describe how they are related to the logical and metaphysical possibility just discussed.

### **Section 2 – David Chalmers, The Conceivability Argument and Zombies**

In this section I discuss Chalmers's Conceivability Argument, using his two-dimensional semantic framework. As I described, this framework uses the primary and secondary ways of conceiving – primary in the world considered as actual, secondary in the world considered as counterfactual. Chalmers links these two ways of conceiving with two ways of possibility, which he calls 1-possibility and 2-possibility.<sup>37</sup> Primary conceivability (or 1-conceivability) implies 1-possibility, and secondary conceivability (or 2-conceivability) implies 2-possibility. You can't mix them up by saying that 1-conceivability implies 2-possibility.

Chalmers compares his two senses of 'possibility' with the distinction between logical and metaphysical possibility:

“A statement is “logically possible” in this sense if it is true in some world when evaluated according to primary intensions; a statement is “metaphysically possible” if it is true in some world when evaluated according to secondary intensions. The relevant space of worlds is the same in both cases.”<sup>38</sup>

Chalmers's main claim for the link between conceivability and possibility is that there seems to be no reason to deny it. The only worries there could be about the link apply at the level of statements, not worlds, namely:

“either we use a statement to misdescribe a conceived world [...], or we claim that a statement is conceivable without conceiving of a world at all [...].”<sup>39</sup>

So we take it from Chalmers that primary conceivability entails primary possibility and that secondary conceivability entails secondary possibility. How does this apply to the case of zombies? As we saw in the previous section, metaphysical possibility is needed for the zombie-argument to work. We thus need secondary possibility in Chalmers's sense, which can only be reached from secondary conceivability.

Chalmers states his Conceivability Argument in the case of zombies as follows:

- “(1)  $P \wedge \neg Q$  is ideally primarily positively (negatively) conceivable.
- (2) If  $P \wedge \neg Q$  is ideally primarily positively (negatively) conceivable, then  $P \wedge \neg Q$  is primarily possible.
- (3) If  $P \wedge \neg Q$  is primarily possible, materialism is false.
- (C) Materialism is false.”<sup>40</sup>

In which P is the conjunction of physical truths about the world, and Q a phenomenal truth. However, as we just saw, we do not need primary possibility of zombies in order to refute physicalism, but we need

---

<sup>37</sup> David Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996), 67.

<sup>38</sup> David Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996), 68.

<sup>39</sup> David Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996), 68.

<sup>40</sup> David Chalmers, “Does Conceivability Entail Possibility?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 196.

secondary possibility. This means that we have to build a bridge in the argument from the primary conceivability to the secondary possibility. There is one way to do this: we have to show that the primary and secondary intensions of both P and Q coincide. That way it would be possible to go from the primary conceivability of them to secondary possibility.

The claim that the primary and secondary intension of Q coincide is easy to make according to Chalmers. He argues for it in the following way:

“What it takes for a state to be a conscious experience in the actual world is for it to have a phenomenal feel, and what it takes for something to be a conscious experience in a counterfactual world is for it to have a phenomenal feel. The difference between the primary and secondary intensions for the concept of water reflects the fact that there could be something that looks and feels like water in some counterfactual world that in fact is not water, but merely watery stuff. But if something feels like a conscious experience, even in some counterfactual world, it is a conscious experience. All it means to be a conscious experience, in any possible world, is to have a certain feel.”<sup>41</sup>

He thus concludes that the primary and secondary intensions coincide in the case of Q. Now we need to do the same for P. This is a harder claim to make: P could easily be similar to the case of water being H<sub>2</sub>O, in which case the primary and secondary intensions would be different. This would mean that physical concepts have their reference fixed, just like the reference of ‘water’ is fixed to H<sub>2</sub>O. However, Chalmers asks us to consider the consequences this would have. Let us consider a conceived world W, in which the primary intension of P $\wedge$ ¬Q holds:

“Because the primary intension of P holds, this world must be structurally-dispositionally isomorphic to the actual world, with the same patterns of microphysical causal roles being played. If P’s secondary intensions fails, it can only be because these microphysical causal roles have different categorical bases in W (or, just possibly, no categorical bases at all).”<sup>42</sup>

What does this mean? Let’s keep up with the analogy of ‘water’. The fact that it is primary conceivable that ‘water’ is XYZ, doesn’t mean that this XYZ can just be anything. It still has to be ‘watery stuff’, which means amongst other things that oceans and lakes have to contain it, and that it has to be wet. This is what Chalmers means with the claim that the world W must be ‘structurally-dispositionally isomorphic’ to the actual world. This leads us to the following conclusion:

“What results is a view on which the existence of consciousness is not necessitated by the structural or dispositional aspects of the microphysics of our world, but is necessitated by the categorical aspects of microphysics (the underlying categorical basis of microphysical dispositions), perhaps in combination with structural-dispositional aspects.”<sup>43</sup>

Chalmers calls this view *panprotopsychism* or *Russellian monism*, as it has been put forward by Russell. It consists of the thesis that consciousness is constituted by intrinsic properties, which serve as the categorical bases of microphysical dispositions. These properties are either phenomenal or

---

<sup>41</sup> David Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996), 133.

<sup>42</sup> David Chalmers, “Does Conceivability Entail Possibility?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 197.

<sup>43</sup> David Chalmers, “Does Conceivability Entail Possibility?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 198.

protophenomenal: they don't have to be phenomenal properties, but they stand in a constitutive relation to phenomenal properties.

Chalmers questions whether this view of panprotopsychism qualifies as a version of physicalism, because it is not sure whether the underlying protophenomenal properties are physical properties. At least the following can be said:

“We need only note that if it is a sort of physicalism, it is a quite unusual sort, and one that many physicalists do not accept.”<sup>44</sup>

Chalmers thus gives us a choice: either P's primary and secondary intensions coincide, or panprotopsychism is true. Seen as panprotopsychism isn't generally accepted by physicalists, Chalmers draws the conclusion that P's primary and secondary intensions must coincide.

This is what Chalmers needed in order for his Conceivability Argument to be true: because we have established that both Q's and P's primary and secondary intensions coincide, we can overcome the gap between primary conceivability and secondary possibility in this case. Thus, Chalmers's Conceivability Argument in the case of zombies is sound, and thus, physicalism is false.

In the next section I consider a critique of Chalmers's Conceivability Argument.

### ***Section 3 – Stephen Yablo on the Conceivability Argument***

Yablo argues that the second step in the Conceivability Argument is invalid. That is, he argues that you can't make the step from logical possibility to metaphysical possibility. In Chalmers's terminology: you can't make the step from primary possibility to secondary possibility. Yablo thus directly attacks Chalmers.

Chalmers states that the difference between logical and metaphysical possibility comes down to statements, not to worlds:

“It follows from all this that the oft-cited distinction between “logical” and “metaphysical” possibility stemming from the Kripkean cases – on which it is held to be logically possible but not metaphysically possible that water is XYZ – is not a distinction at the level of *worlds*, but at most a distinction at the level of *statements*.”<sup>45</sup>

On the basis of this, Yablo reconstructs Chalmers's Conceivability Argument for zombies as follows<sup>46</sup>:

- “(a) it is conceptually possible that there be zombies, so
- (b) zombie worlds are conceptually possible, so
- (c) zombie worlds are metaphysically possible.”<sup>47</sup>

---

<sup>44</sup> David Chalmers, “Does Conceivability Entail Possibility?,” in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 198.

<sup>45</sup> David Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996), 67-68.

<sup>46</sup> Note: Yablo uses ‘conceptual possibility’ where Chalmers uses ‘logical possibility’. They mean the same and the terms can be used interchangeably.

<sup>47</sup> Stephen Yablo, “Concepts and Consciousness,” in *Philosophy and Phenomenological Research* 59 (1999), 459.

However, these premises don't actually support each other. What (a) supports is:

(b') it is conceptually possible that there be zombie worlds.

However, in order to get to (c), you need:

(b'') there are conceptually possible zombie worlds.

You thus need to bring the conceptual possibility in actual existence, or in other words, you need to create a link between the *de dicto* possibility of zombie worlds, as asserted by (b'), to the *de re* possibility of zombie worlds, as asserted by (b''). It thus seems that Chalmers implicitly assumes CP:

CP If it is conceptually possible that E, then E's primary intension contains at least one world.<sup>48</sup>

This premise allows Chalmers to make the move from *de dicto* possibility to *de re* possibility, because he assumes the existence of at least one world in conceptual possibility.

But why should we believe CP? Proponents of it might say something like the following: in order to know whether E is conceptually possible, you first have to understand it; and understanding is knowing its truth-conditions, i.e. how the truth-value depends on worldly context. Now, if you say that E is conceptually possible, which is what CP does, you thereby say that E is true in some such worldly context. Thus, E is true in some possible world considered as actual. And this means that E's primary intension contains at least one world. However, Yablo argues that the step from truth-condition to the truth of E in some possible world is not valid. All that the understanding of a statement and thus the knowing of its truth-conditions demands, is to know what a world *has to be like* to verify a statement. It doesn't need the *existence* of such a world – that is a completely different case. The justification for CP thus doesn't work, and CP can't be assumed.

Yablo has thus shown that CP, the implicit assumption on which Chalmers's Conceivability Argument rests, is not true. This is a direct attack on Chalmers, and it means that Chalmers's Conceivability Argument is not valid.

#### ***Section 4 – My thoughts on the Conceivability Argument***

Yablo makes a strong argument that creates difficulties for Chalmers. I can't find any flaws in the argument, and Chalmers's initial reply to Yablo seems weak to me.<sup>49</sup> Chalmers needs a better refutation of the argument, if he wants to hold his position.

In this section I provide my viewpoint on the Conceivability Argument. I examine Chalmers's two-dimensional semantic framework and show that it creates two problems: the framework in itself creates problems because of its presuppositions, and the way Chalmers uses the framework in the zombie-case creates problems in the coinciding of the primary and secondary intensions.

---

<sup>48</sup> Stephen Yablo, "Concepts and Consciousness," in *Philosophy and Phenomenological Research* 59 (1999), 460.

<sup>49</sup> See: David Chalmers, "Materialism and the Metaphysics of Modality," In *Philosophy and Phenomenological Research* 59 (1999), 480-483.

### *Survey*

We need to take a small step back in order to regain a perspective on what is actually happening here. We are discussing the Conceivability Argument, which is an argument that aims to provide a link between conceivability and possibility. ‘Conceivability’ in the sense that I use it here means ‘ideal negative conceivability’, which means that S is conceivable, if on ideal rational reflection S is not ruled out a priori. There are two kinds of this conceivability – primary and secondary conceivability –, which leads to two kinds of possibility: primary (logical) possibility and secondary (metaphysical) possibility. In order for the zombie-argument to work, we need metaphysical possibility, thus secondary possibility.

What we are concerned with is how to go from conceivability to possibility, i.e. how to go from primary conceivability to primary (logical) possibility, and how to go from secondary conceivability to secondary (metaphysical) possibility. Let’s start with the latter first.

### *Secondary Intensions*

I argue that the use of secondary intensions – thus both secondary conceivability and secondary possibility, as well as the link between them – relies on claims about essence. Whenever you claim that two properties can be separated in a counterfactual world, you claim that the two properties aren’t connected to each other in a necessary way. For example, I can see that the white colour of the table my laptop rests on could have been separated from the table, i.e. could have been different. This is because the colour white and the table are not necessarily connected in any way. A different case of this is the famous statement “Water is H<sub>2</sub>O”. As shown by Kripke, this is a necessary identity statement and it is therefore impossible for water and H<sub>2</sub>O to come apart – it is even secondarily inconceivable for them to be apart if you have the right amount of information. Necessary identity claims appeal to the essence of a thing: water has the essence that it is H<sub>2</sub>O. Essences are by definition necessary, which means that they are true in every possible world. The essences in the world we live in are relevant for the secondary intension, because you consider a possible world as counterfactual; you thus take the way our world is as a presupposition, upon which you consider a counterfactual world.

The essence is the nature of a thing; it is that without which it could not be what it is.<sup>50</sup> The essence of a thing always has to be present in the thing considered, seen as it is that which makes the thing what it is. This is how we’ve determined that the essence of water is H<sub>2</sub>O: every time you see water, you can see that it is composed of the molecules H<sub>2</sub>O. ‘Essence’ is a metaphysical notion: a thing has an essence regardless of whether we know it or not. Before we knew what molecules were, the essence of water was already H<sub>2</sub>O, and the statement ‘water is H<sub>2</sub>O’ was already necessarily true. This is what makes the use of essences problematic: there could be essences of things that we don’t know of (yet). In a hypothetical situation, there could be another property essential to water that we are not able to discover with the (scientific) instruments that we have at this moment. Just like we thought, before the discovery of molecules, that all there was to water was ‘watery stuff’ – even though H<sub>2</sub>O was already essential to water –, it is possible that the thought that all there is to water is H<sub>2</sub>O is wrong. This links the discovery of essences inextricably with the scientific possibilities and discoveries of the moment. The essences itself stand apart from this, but the *discovery* of the essences are dependent on science. I call this the science-dependence problem of the discovery of essences.

---

<sup>50</sup> Simon Blackburn, *The Oxford Dictionary of Philosophy (3 ed.)* (Oxford University Press, 2016), <http://www.oxfordreference.com/view/10.1093/acref/9780198735304.001.0001/acref-9780198735304> (consulted June 22, 2016).

There are two strategies you can take to object to the zombie-argument, using the reliance of essences in the secondary intension. You can argue that the use of essences is unreliable because of the science-dependence of the discovery of essences, and that therefore the use of the secondary intension is problematic. The other strategy is to argue that physical processes are essential to phenomenal consciousness, just like H<sub>2</sub>O is essential to water. This way, the secondary intension would be useless for proponents of the zombie-argument. I combine both strategies, to make my argument as strong as possible.

If you claim that zombies are secondarily conceivable, you claim that our physical properties are not identical with our phenomenal consciousness. You thus claim that it is not entailed in the essence of phenomenal consciousness that it is physical or supervenient on the physical. However, isn't phenomenal consciousness always realised by a physical process? There are two ways to show that this is so. I will take the case of pain as an example to make the argument a bit less abstract. Rewritten, then, the claim I object to is that it is not essential to pain that it is physical or supervenient on the physical. I show that a physical process is essential to pain, using two arguments.

First, science has proved that whenever we experience pain, a physical process is happening in our body: a signal is transmitted from the nerves where the pain is happening (e.g. your toe when you stub it) to your brain, where the famously coined process of 'the firing of C-fibers' is happening. This happens every time you stub your toe, so couldn't we say that this physical process is essential to pain, just like H<sub>2</sub>O is essential to water? It would then not be secondarily conceivable for pain and its physical process to be apart: thus, zombies wouldn't be secondarily conceivable.

Opponents might object to this argument with the counterexample of phantom pains. Phantom pains are pains people feel of body parts that have been amputated. Opponents argue that phantom pains form a counterexample for the claim that pain is always realised by a physical process, because in the case of phantom pain the stubbed toe of the example *doesn't exist anymore*, because it has been amputated. The pain I feel is then 'only in the head', and not based on any physical process. However, this isn't right. The opponent is right that the pain can't be traced back to the body part that the person claims to be in pain in is – as this body part has been amputated –, but this doesn't mean that no physical process is happening and that the pain is 'only in the head'. The general consensus in medical science is that phantom pain is caused by nerve endings that still send pain signals to the brain, or that mirror neurons in the brain generate the experienced pain.<sup>51</sup> There is thus still a physical process happening! Therefore, phantom pains aren't a good counterexample to the claim that a physical process is essential to pain. The conclusion that zombies aren't secondarily conceivable still stands.

The second argument to show that a physical process is essential to pain is inspired by Maurice Merleau-Ponty. He argues that the mind is inseparable from our physical body.<sup>52</sup> We can translate this to the case of pain: when we experience pain, we need our body and its processes to be able to do this. If we didn't have a body, we wouldn't be able to experience pain. This means that pain couldn't be what it is without our physical body. As we saw earlier, this is what it means to be the essence of a thing: that without which it could not be what it is. Therefore, our physical body and the processes happening in it are

---

<sup>51</sup> Bishnu Subedi and George T. Grossberg, "Phantom Limb Pain: Mechanisms and Treatment Approaches," in *Pain Research and Treatment* (2011), consulted June 22, 2016, doi: 10.1155/2011/864605.

<sup>52</sup> Jack Reynolds, "Maurice Merleau-Ponty (1908-1961)," in *Internet Encyclopedia of Philosophy*, <http://www.iep.utm.edu/merleau/> (consulted June 25, 2016).

essential to pain. Again, this means that it is not secondarily conceivable for pain and its physical processes to be apart; thus, zombies aren't secondarily conceivable and possible.

I have argued that the use of the secondary intension is problematic because of its reliance on essences. There are two problems with this: the first is the science-dependence of the discovery of essences; the second problem is that physical processes are essential to phenomenal consciousness. I argued for both problems, and thereby showed that the secondary intension is useless for proponents of the zombie-argument. Next, I consider primary intensions.

### *Primary Intensions*

I now consider the link between primary conceivability and primary (logical) possibility. Primary conceivability isn't concerned with any essential claims made in our world, seen as it considers a statement in a world considered as actual. Any identity claim made in the world we live in is thus irrelevant. However, this doesn't mean that we are entirely free in making claims. For example, we could consider a world physically exactly like ours, where 'water has tails' would be true: that is, if we consider in that same world that people refer to horses when they speak of 'water'. If we allow this, everything is allowed and we seem to lose grip on our reality entirely. We thus need some rules in order to stay seated. Chalmers describes it himself:

“Because the primary intension of P holds, this world must *be structurally-dispositionally isomorphic to the actual world*, with *the same patterns* of microphysical causal roles being played.”<sup>53</sup> [italics added]

This means that in the case of 'water', the primary intension of it must still resemble our notion of 'water' in some way, which is what Chalmers calls 'the watery stuff' that is characteristic of water.

It gets tricky when we consider the case of phenomenal consciousness. We don't need the fixed reference of phenomenal consciousness in our world, but we still need something that resembles our notion of 'phenomenal consciousness'. What is this and how do we determine this? The most likely answer would be that phenomenal consciousness in its primary intension is the *feeling* of phenomenal consciousness, the what-it-is-like feeling discussed previously. If we consider phenomenal consciousness as this what-it-is-like feeling, it seems reasonable that we can conceive it both with and without physical processes, as we are not committed to anything more than just the feeling. This would mean that zombies are primary conceivable and thus primary (logical) possible in this sense.

But what does this logical possibility give us? 'Logical possibility' only means that logic permits that it is possible, i.e. that the proposition considered does not cause a contradiction. So applied to this case, this means that it is not logically contradictory that phenomenal consciousness occurs without a physical process realising it, as long as we consider phenomenal consciousness as just the what-it-is-like feeling. Here the reasoning seems to be circular. First, we establish that the primary intension of phenomenal consciousness is the what-it-is-like feeling that belongs to phenomenal consciousness. On the basis of this we conclude that this feeling can be both with and without physical processes realising it, and that both of these are thus (logically) possible. This is no surprise, as we first established that phenomenal consciousness doesn't need more than the what-it-is-like feeling. We thus base our conclusion on the assumption made in the establishing of the primary intension.

---

<sup>53</sup> David Chalmers, "Does Conceivability Entail Possibility?," in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne (New York: Oxford University Press, 2002), 197.

This points to the problem of primary intensions in general, and specifically the link between primary conceivability and primary (logical) possibility. The use of primary intensions gives you the freedom to decide what the primary intension of the thing considered is, just as long as it resembles the original thing. We can question whether a resemblance like ‘watery stuff’ to ‘water’ is the right resemblance, and there might be cases where the resemblance is even harder to determine.

In the specific case of phenomenal consciousness, the primary intension brings us three problems. First, the claim that the resemblance of phenomenal consciousness is the what-it-is-like feeling. As just mentioned, this is problematic because of the circularity this brings in the argument. Second, the uncertainty whether the what-it-is-like feeling is the right primary intension of ‘phenomenal consciousness’. How do we know for certain that this is the intension that we should use? Third, the question arises what the use of this logical possibility is: we only establish that something isn’t contradictory. This doesn’t seem to add much to possibility-debates, as the general use of ‘possibility’ is metaphysical. This is also the case in the zombie-argument, as I showed in Section 1 of this part. Therefore, even if we can overcome the first two problems the primary intension brings us, we still run into the problem what this logical possibility brings us. In any case, it doesn’t bring proponents of the zombie-argument any further. The strategy of using the primary intension in the zombie-case is thus also problematic for zombie-proponents.

#### *From Primary Intensions to Secondary Intensions*

In the case of phenomenal consciousness, Chalmers argues that the primary and secondary intensions coincide. He needs this, because in his argument he establishes that zombies are primary conceivable, and from that he needs to move to the secondary possibility of zombies. For this, he needs a link from primary conceivability to secondary possibility, which he makes by coinciding the primary and secondary intensions of ‘phenomenal consciousness’ (see Section 2 of this part for an extensive discussion).

His argument for the coinciding of the primary and secondary intension is that in both the actual world and a counterfactual world (thus both according to the primary and the secondary intension), all there is to phenomenal consciousness is to have a phenomenal feel: “All it means to be a conscious experience, in any possible world, is to have a certain feel.”<sup>54</sup>

As I just argued, this is accepted for the primary intension, all of its flaws aside. But Chalmers ascribes ‘phenomenal feel’ too easily to the secondary intension as well. He just states that all there is to phenomenal consciousness is to have a phenomenal feel, in both intensions. However, I have just argued that the use of secondary intensions is problematic because of its reliance on essences. Namely, two problems arise: the science-dependence of the discovery of essences; and I argued that physical processes are essential to phenomenal consciousness. This is exactly opposed to Chalmers’s claim that all there is to phenomenal consciousness in the secondary intension is to have a phenomenal feel.

I can’t find any extensive argument in which Chalmers argues for one of the following things: (a) that zombies are secondarily conceivable; and (b) that all there is to phenomenal consciousness in both intensions is to have a phenomenal feel, other than just stating this. I find it odd that Chalmers never discusses whether zombies are secondarily conceivable, as this would provide a much easier way to secondary possibility than the detour he takes now.

---

<sup>54</sup> David Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996), 133.

Another thing that is remarkable about Chalmers's strategy is that he takes quite some space and time to build his two-dimensional semantic framework, to finally conclude that in the case of phenomenal consciousness the primary and secondary intensions coincide. He even argues that this must be the case for all the physical truths in the world as well. It looks like a magic trick, where the magician focusses the viewer's attention on some complicated model, while the real trick happens at a completely different point: in seemingly simple statements like "All it means to be a conscious experience, in any possible world, is to have a certain feel."<sup>55</sup>

This magician's strategy is the one that Chalmers uses. He creates the two-dimensional semantic framework, which turns out to be problematic for both the primary and secondary intension, as I just showed. Then, he doesn't use the two intensions when considering the zombie-argument and phenomenal consciousness, by stating that the intensions coincide in this case. He doesn't provide a real argument for the latter, other than just stating it. First, to state that some things just are the case isn't a good argument. Second, the claim that the primary and secondary intensions coincide is exactly opposed to what I argued for at the beginning of this part, namely that physical processes are essential to phenomenal consciousness.

Combining all of this, I have shown that there are mistakes in Chalmers's Conceivability Argument in the case of zombies. Therefore, Chalmers's argument is not sound. Seen as the Conceivability Argument is needed in order for the zombie-argument to work, and the Conceivability Argument is not sound, the zombie-argument is not sound as well. I have thus refuted the zombie-argument by showing that (P2) is not right.

### ***Subconclusion***

This part focussed on the second premise of the zombie-argument, i.e. on the Conceivability Argument. I first showed that we need metaphysical possibility in order for the zombie-argument to work. In Section 2 I gave an overview of Chalmers's Conceivability Argument in the case of zombies. He argued that zombies are primary conceivable, and that because of the coinciding of the primary and secondary intensions, zombies are secondary (metaphysical) possible. In Section 3 I described Yablo's critique of Chalmers. He showed that Chalmers's argument relies on the silent assumption CP, which can't be justified. This creates a problem for Chalmers. Lastly, in Section 4, I defended my viewpoint on the Conceivability Argument and showed its two problems: the notions of the primary and secondary intension are problematic because of their presuppositions, and the way Chalmers uses the two-dimensional semantic framework in the zombie-case is problematic because it is opposed to the conclusions I drew from the presuppositions of the primary and secondary intension. The conclusion of this part is that the second premise of the zombie-argument has been refuted. This means that the zombie-argument isn't sound and thus has been refuted.

---

<sup>55</sup> David Chalmers, *The Conscious Mind* (New York: Oxford University Press, 1996), 133.

## **CONCLUSION**

In this thesis I inquired whether the zombie-argument is sound, and thus whether physicalism of the mind is refuted by it. I did so by investigating the two controversial premises of the argument, namely (P1) zombies are conceivable, and (P2) if zombies are conceivable, then zombies are metaphysically possible.

The thesis consisted of three parts. The first part consisted of an introduction in the subject, including an overview of physicalism of the mind, a detailed version of the zombie-argument and an overview of modal logic and the two-dimensional semantic framework.

In the second part, I discussed the first premise of the zombie-argument, which is that zombies are conceivable. I presented two arguments saying that zombies are not conceivable: the first one by Kirk, the second one by Stalnaker. I immediately showed that both arguments were flawed, and therefore form no problem for the first premise. I concluded this part by saying that the first premise of the zombie-argument still stands, and thus that zombies are conceivable.

In the third part, I discussed the second premise of the zombie-argument, which is the Conceivability Argument. I presented Chalmers's argument in favour of it, and I then presented an argument by Yablo that directly attacks Chalmers. Lastly, I provided my viewpoint on the Conceivability Argument, in which I concluded that Chalmers's Conceivability Argument isn't sound because of several mistakes in the argument. Combining Yablo's critique and my critique, I concluded that the second premise of the zombie-argument has been refuted.

The main question of this thesis was whether the zombie-argument is sound, and thus whether physicalism of the mind is refuted by it. In Part II, I showed that the first premise of the argument is sound; in Part III, I showed that the second premise of the argument has been refuted. Therefore, I conclude that the zombie-argument isn't sound, and that physicalism of the mind has not been refuted by it.

This doesn't mean that I have proven that physicalism of the mind is true. I have only proven that one of the main arguments against it, the zombie-argument, isn't sound, and therefore doesn't form a problem for physicalism of the mind.

This thesis is of course not the end of the zombie-argument and the discussion of physicalism of the mind. First of all, I have only provided several viewpoints on the matter: there is much more literature that can be read and discussed. I considered Chalmers and his two-dimensional semantic framework as the main proponent of the zombie-argument, but other philosophers could have been mentioned as well. The time and space of this thesis restricted me in not considering all of them. Secondly, different branches of philosophy could have been used to consider the zombie-argument. I mainly used arguments from logics and metaphysics in this thesis, but I could have used arguments from philosophy of language as well. These can be considered as flaws of this thesis, but at the same time they offer future points of research. Future research can be done both of the zombie-argument itself, at the points just mentioned, and of other arguments pro or con physicalism, like the Knowledge Argument. This way, the discussion whether physicalism is the right position in the mind-body debate can come closer to an agreement, one step at a time.

## **BIBLIOGRAPHY**

- Almog, Joseph. *What Am I?* New York: Oxford University Press, 2002.
- Buskes, Chris and Simissen, Herman, editors. *Analytische Filosofie*. Nijmegen: Vantilt, 2014.
- Brueckner, Anthony. "Chalmers's Conceivability Argument for Dualism." In *Analysis* 61 (2001), 187-193.
- Blackburn, Simon. *The Oxford Dictionary of Philosophy (3 edition)*. Oxford University Press, 2016.  
<http://www.oxfordreference.com/view/10.1093/acref/9780198735304.001.0001/acref-9780198735304> (consulted June 22, 2016).
- Chalmers, David. "Does Conceivability Entail Possibility?" In *Conceivability and Possibility*, edited by Tamar Szabó Gendler and John Hawthorne, 145-200. New York: Oxford University Press, 2002.
- Chalmers, David. "Materialism and the Metaphysics of Modality." In *Philosophy and Phenomenological Research* 59 (1999), 473-496.
- Chalmers, David. *The Character of Consciousness*. New York: Oxford University Press, 2010.
- Chalmers, David. *The Conscious Mind*. New York: Oxford University Press, 1996.
- Crane, Tim. *Elements of Mind*. New York: Oxford University Press, 2001.
- Della Rocca, Michael. "Essentialism versus Essentialism." In *Conceivability and Possibility*, edited by Tamar Szabó Gendler and John Hawthorne, 223-252. New York: Oxford University Press, 2002.
- Dennett, Daniel. "The Zombic Hunch: Extinction of an Intuition?" In *Royal Institute of Philosophy Supplement* 48 (2001), 27-43.
- Frankish, Keith. "The Anti-Zombie Argument." In *The Philosophical Quarterly* 57 (2007), 650-666.
- Garson, James. "Modal Logic." In *The Stanford Encyclopedia of Philosophy (Spring 2016 Edition)*. Edited by Edward N. Zalta. , <http://plato.stanford.edu/entries/logic-modal/> (consulted May 27, 2016).
- Heil, John. *Philosophy of Mind*. New York: Routledge, 2013.
- Kirk, Robert. "Zombies." In *The Stanford Encyclopedia of Philosophy (Summer 2015 Edition)*. Edited by Edward N. Zalta. <http://plato.stanford.edu/entries/zombies/> (consulted May 31, 2016).
- Kirk, Robert. *Zombies and Consciousness*. New York: Oxford University Press, 1996.
- Kripke, Saul. *Naming and Necessity*. Oxford: Basil Blackwell, 1980.

- LaPorte, Joseph. "Rigid Designators." In *The Stanford Encyclopedia of Philosophy (Spring 2016 Edition)*. Edited by Edward N. Zalta. , <http://plato.stanford.edu/entries/rigid-designators/> (consulted June 8, 2016).
- Levine, Joseph. *Purple Haze*. New York: Oxford University Press, 2001.
- Loar, Brian. "David Chalmers's The Conscious Mind." In *Philosophy and Phenomenological Research* 59 (1999), 465-472.
- Lycan, William G. "Stalnaker on Zombies." In *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 133 (2007), 473-479.
- Marcus, Eric. "Why Zombies are Inconceivable." In *Australasian Journal of Philosophy* 82 (2004), 477-490.
- McLaughlin, Brian and Bennett, Karen. "Supervenience." In *The Stanford Encyclopedia of Philosophy (Spring 2015 Edition)*. Edited by Edward N. Zalta. , <http://plato.stanford.edu/entries/supervenience/> (consulted June 6, 2016).
- Reynolds, Jack. "Maurice Merleau-Ponty (1908-1961)." In *Internet Encyclopedia of Philosophy*. <http://www.iep.utm.edu/merleau/> (consulted June 25, 2016).
- Stalnaker, Robert. "What is it Like to be a Zombie?" In *Conceivability and Possibility*, edited by Tamar Szabó Gendler and John Hawthorne, 385-400. New York: Oxford University Press, 2002.
- Stoljar, Daniel. "Physicalism." In *The Stanford Encyclopedia of Philosophy (Spring 2016 Edition)*. Edited by Edward N. Zalta. <http://plato.stanford.edu/entries/physicalism/> (consulted May 24, 2016).
- Subedi, Bishnu and Grossberg, George T. "Phantom Limb Pain: Mechanisms and Treatment Approaches." In *Pain Research and Treatment* (2011). Consulted June 22, 2016. doi: 10.1155/2011/864605.
- Vaidya, Anand. "The Epistemology of Modality." In *The Stanford Encyclopedia of Philosophy (Summer 2015 Edition)*. Edited by Edward N. Zalta. <http://plato.stanford.edu/entries/modality-epistemology/> (consulted June 8, 2016).
- Yablo, Stephen. "Concepts and Consciousness." In *Philosophy and Phenomenological Research* 59 (1999), 455-463.
- Yablo, Stephen. "Textbook Kripkeanism & The Open Texture of Concepts." In *Pacific Philosophical Quarterly* 81 (2000), 98-122.
- Yablo, Stephen. "Coulda, Woulda, Shoulda." In *Conceivability and Possibility*, edited by Tamar Szabó Gendler and John Hawthorne, 441-492. New York: Oxford University Press, 2002.