

# *De Se* Modal Illusions

Clas Weber

University of Western Australia & Eberhard Karls University of Tübingen

Research on personal identity often relies on imaginary cases and tends to theorise about our nature from the first-person perspective. In this paper I argue that a problem arises when we combine the two methods and assess imaginary cases from the first-person perspective. The problem is that the link between *de se* imagination and modality is broken. *De se* imagination regularly gives rise to *de se* modal illusions. *De se* modal illusions come in two varieties: there are *de se* illusions of possibility and *de se* illusions of impossibility. I present a systematic account of the two types of illusion. I further make the case that the proposed account is compatible with maintaining a general link between imaginability and possibility. I apply the account to two concrete cases and use it to undercut the motivation behind a central theory of personal ontology, Dualism, and a central account of personal persistence, Lockeanism.

## 1 Introduction

How do we know what is possible? The received view is that we can gain access to possibility through imagining or conceiving. How do we know what is possible *for us*? Here we usually seek the answer by engaging in imagination *from the first-person perspective*. Modal judgments based on such *de se* imagination are a crucial source of evidence in debates about our metaphysical nature and our persistence conditions.<sup>1</sup>

Consider Descartes' central argument for Dualism (Descartes, 1641). At the core of the argument is a *de se* modal premise, i.e. the claim that it is possible for us to exist as something immaterial. To assess the plausibility of this premise we typically engage in the following intellectual exercise: we try to imagine being an immaterial object. It seems that we can do this. We can picture a situation where we have a conscious perspective on the world, but where there is no material object associated with this perspective. We then take our ability to imagine such a scenario as evidence for the claim that *we could be immaterial*.

Consider next Lockeanism about personal persistence. Crucial support for psychological accounts of persistence comes from the *de se* modal claim that we could switch from one body to another (Locke, 1694; Shoemaker, 1984). How do we know that this is the case? We assess the modal judgment by trying to imagine a corresponding scenario from the inside. We try picturing

---

<sup>1</sup> A note on terminology: I will use "first-person imagination", "*de se* imagination" and "imagination from the inside" interchangeably.

our conscious perspective being first connected to our own body, and then to another body. This, too, we can do. We take this as evidence for the claim that *we could switch from one body to another*.

In this paper, I argue that we should not rely on *de se* imagination to support a particular view of the self. The problem is that *de se* imagination is not a reliable guide to *de se* modality. This skepticism about *de se* imagination is not motivated by a more global skepticism about the connection between imagination and possibility. The account developed here maintains that there is an intimate connection between imaginability and metaphysical possibility. Rather, my contention is that it is the specific features of *de se* imagination which can lead us astray.

Here is the plan for the paper. In §2, I will present some of the pre-theoretic data and explore the range of *de se* imagination. It will become apparent that *de se* imagination is almost entirely unconstrained. But we will also see that it is fenced in by its experiential character. In §3, I present an account of *de se* imagination that explicitly articulates some of the principles that underlie the observed data. I will then critically discuss a veridical approach to *de se* imagination and make the case that this approach is not promising in light of the data. Instead, I propose to adopt a non-veridical view of *de se* imagination. Adopting this view does not amount to severing the connection between *de se* imagination and modality altogether. In §5, I show how to integrate a non-veridical approach to *de se* imagination within a Kripkean account of modal illusions. I then present a systematic account of the modal illusions connected to *de se* imagination. I first explain in §6 which features of *de se* imagination generate *de se* illusion of possibility, and apply the analysis to two concrete cases: Dualism and Lockeanism. In §7, I explain how *de se* imagination can produce *de se* illusions of impossibility. §8 offers replies to objections, and §9 concludes.

## 2 The Scope of *De Se* Imaginability

*De se* or inside imagination contrasts with third-person or outside imagination (see Williams, 1973; Vendler, 1979; Shoemaker, 1994; Recanati, 2007; Ninan, 2008, 2009; Torre, 2021; Weber, 2023). We can illustrate the difference with an example from (Ninan, 2008). I can imagine from the outside that there is someone, perhaps myself, skiing down a slope, where I picture the scene from a detached third-person perspective. Alternatively, I can imagine being the skiing individual myself and experiencing the situation from the first-person perspective. Here I put myself into the shoes of the imaginary subject and simulate having her experiences and picture engaging in the relevant actions. Our central question is whether the second, *de se* form of imagination gives us reliable insight into what is possible for us.

To answer this question, we should first explore the scope of *de se* imagination. The range of scenarios which are accessible from the first-person perspective is extremely wide. *De se* imagination appears almost entirely unconstrained. We can imagine being a conscious physical being in a materialistic universe. We can imagine having no physical body at all and being an immaterial Cartesian soul. We can imagine from the inside being a different person such as Napoleon (Williams, 1973). We can imagine being a different species, as in a Kafkaesque scenario where we wake up as a beetle. We can imagine being an avatar with an inner life in a computer simulation. We can even imagine being a conscious artefact, such as an animated teapot, in a fairy tale scenario. And so on.

Our persistence, as imagined from the inside, seems equally free-ranging. From the first-person perspective, we can imagine undergoing radical transformations of our bodies, even leaving our bodies behind altogether. We can imagine surviving wild jumps in spatio-temporal location, e.g. as a time traveler in a malfunctioning time machine. We can imagine experiencing a complete change in our psychology. We can imagine surviving fission as Lefty or as Righty. We can imagine successively living the lives of different people such as first being Julius Caesar, then being Marie Curie, and finally Angela Merkel. One may think that this supports a view according to which we are featureless Cartesian souls. But we can, as Locke (1694) has pointed out, even imagine surviving the replacement of our immaterial substrate. It seems that Nagel (1986, p. 33) is right when he writes: “When I consider my own individual life from inside, it seems that my existence in the future or the past [...] depends on nothing but itself.”

Interestingly, we can equally imagine *failing* to survive at any moment throughout a hypothetical persistence scenario. I can imagine failing to survive fission, or ceasing to exist when my body gets destroyed, or when my psychology is altered in certain ways. I can imagine from the inside being Julius Caesar, then being reborn as Angela Merkel, but coming to an end during Merkel’s second term as Chancellor (picturing that someone else is then taking over her body).

It may appear as if *de se* imagination is entirely without bounds. But this is not so. All *de se* imaginable scenarios require the existence of a *conscious perspective* for us to occupy. In other words, it is not possible to imagine from the inside having no experiences whatsoever. We cannot *de se* imagine being an unconscious rock, or a philosophical zombie. And *pace* Blackburn, we cannot imagine our own funeral from the inside, and picture lying in our coffin “being stiff and experiencing nothing” (Blackburn, 1997, p. 198). Further, as Nagel (1974) has argued, the relevant conscious perspective cannot be radically unlike our own, e.g. by containing experiences associated with sensory organs that we are lacking, such as echolocation and electroreception. We cannot imagine from the inside being a bat, since we lack a clear conception of what this would be like. *De se* imagination requires a conscious perspective that is *accessible* for simulation.<sup>2</sup> These constraints on *de se* imagination are not imposed by our concept of the self. There is nothing conceptually incoherent in the assumption that we are at times unconscious or that we could have had bat-type experiences.

Summing up. *De se* imagination is extremely permissive, both with regards to the type of entity we can imagine being and in terms of which scenarios we can imagine surviving or failing to survive. It is not fenced in by qualitative facts—be they physical, psychological, or immaterial—

---

<sup>2</sup> Admittedly, the corresponding notion of *accessibility* is somewhat vague. The underlying idea is that there are certain experiential states which we can simulate and others which we cannot. Accessible experiential perspectives are not necessarily confined to experiential states we have already encountered. But we should be able to simulate what the experiences are like giving our existing phenomenal resources. For instance, we can imagine *observing a group of wombats having a party*, even though we have not actually experienced such a scene, by recombining familiar experiences in the right way. Further, we can plausibly also extrapolate from our given stock of experiences and imagine what it would be like to see Hume’s missing shade of blue. Even still, certain experiences seem out of reach to us through such processes of recombination and extrapolation. There is an intuitive division between accessible and inaccessible experiential perspectives. The following arguments do not require a decision of where exactly the dividing line between the two lies.

or by particular facts about our identity. The main requirement on *de se* imagination is the presence of a suitable experiential perspective.

### 3 *De Se* Imagination

The contrast between outside and inside imagination is an instance of a more general contrast between third-person and first-person attitudes. I will here use David Lewis (1979)'s influential theory of first-person attitudes to explain *de se* imagination, following proposals by Recanati (2007) and Ninan (2008, 2009).<sup>3</sup> Many of the following points are independent of the specifics of Lewis's theory and could be reformulated in other frameworks. Ninan (2009) and Recanati (2007) are sympathetic to a veridical account of *de se* imagination, according to which *de se* imagination is a reliable guide to *de se* modality. I will argue in §4 that there is good reason to reject this approach.

In the Lewisian framework, first-person attitudes are modelled as binary relations between a subject and a *centered* content. Centered contents are sets of centered possible worlds, i.e. triples of the form  $\langle \text{individual, time, world} \rangle$ . According to the traditional theory, attitude contents are in contrast taken to be sets of uncentered possible worlds or Russellian structures which determine such sets of uncentered worlds. Traditional contents are absolute, i.e. invariable in truth value across times or individuals. Centered contents on the other hand can vary in truth value across individuals and times. Consider e.g. my *de se* belief that *I am happy*. Its centered content is the set of centered worlds with happy center-individuals:  $\{ \langle i, t, w \rangle : i \text{ is happy at } t \text{ in } w \}$ . We can model *de se* imagination in the same way. A *de se* imagining of *being happy* is a binary relation between the imaginer and the set of centered worlds with happy center-individuals. Figuratively speaking, when one imagines from the inside being happy one puts oneself into the shoes of one of the happy individuals and engages in an offline simulation of their experiences. The difference between *de se* belief and *de se* imagination does not lie at the level of content, but rather in the nature of the different attitudinal relations. Lewis (1979) calls the belief-relation 'self-ascription' (of a centered content). Call the relation that holds between a subject and the centered content of her *de se* imagining *imaginative self-ascription*.

A notable consequence of the Lewisian theory is that the content of a typical first-person attitude does not explicitly represent the attitude holder.<sup>4</sup> For instance, my *de se* belief that I am happy has the same content as your *de se* belief that you are happy. Both relate us to the set of centered worlds with happy center-individuals. And the same holds for my and your *de se* imaginings of being happy. It may be surprising that the contents of *de se* attitudes in particular do not explicitly represent their subject. But the account is theoretically well-motivated. First, it manages to accommodate the special cognitive significance of first-person attitudes, unlike

---

<sup>3</sup>In recent work, Ninan has abandoned the Lewisian approach to first-person attitudes (at least for belief) and endorsed instead the traditional theory of attitudes (Caie and Ninan, MS).

<sup>4</sup>What does it mean for a content to represent an individual? There are two options. First, a content may encode an identifying property which uniquely singles out that individual. Second, it may be a singular content directly about that individual (in the Russellian paradigm, it literally contains the individual). The centered content of a typical *de se* attitude does not represent its subject in either way. There are exceptions, e.g. when I explicitly judge that *I am NN*.

the traditional theory. Further, it explains the fact that one can believe or imagine that one is David Hume without having a trivially unsatisfiable attitude, and that one can have first-person attitudes while being ignorant about who one is. Lastly, the account manages to capture the behavioral and cognitive similarities of subjects who share the same *de se* attitudes.

If the attitude holder is not represented in the content of a *de se* attitude, what explains that my *de se* attitudes concern me and that your *de se* attitudes concern you? What makes my *de se* belief that I am happy a belief about *myself*? The answer is that my *de se* attitudes depend for truth/satisfaction on *me*, while yours depend on how *you* are. In Lewis's framework, the subject's actual centered world serves as the *point of evaluation* at which a *de se* attitude is assessed for truth/satisfaction (Recanati, 2007); the subject's actual centered world is the triple consisting of <subject, subject's current time, subject's world>. A *de se* belief B is true iff the believer's actual centered world is an element of B's content. S's *de se* belief *I am happy* is true iff S is one of the happy individuals (at the time and world where she has the belief). Equally, we can say that a *de se* imagining I is *satisfied* iff the imaginer's actual centered world is part of I's content. When Ani *de se* imagines to be happy, then for things to be as Ani imagines them to be, *Ani* has to be happy. Analogously, Aren's imagining that he is happy is satisfied iff *Aren* is happy. This subject-relative notion of truth/satisfaction explains why our *de se* attitudes concern ourselves.<sup>5</sup>

It may appear somewhat artificial to evaluate an imagining for satisfaction. But there is an intuitive sense in which we can compare our actual situation to the situation we imagine, and assess whether things actually are as we imagine them to be. The perceived oddity is rather grounded in the fact that imagination and belief play different functional roles. For instance, I will not modify or retract my imagining upon learning that its satisfaction conditions are unfulfilled. And my imaginings, unlike my beliefs, do not typically interact directly with my desires in the production of behaviour. In spite of the differences in functional roles, *de se* belief and *de se* imagination have important structural similarities. First, their content does not involve the subject. Second, the connection to the subject is established outside of content through the respective attitudinal relation.

We can now formulate several principles about the content and structure of *de se* imagination that accommodate the data from §2. We have just noted that, just like in the case of other *de se* attitudes, the imaginer is not typically represented in the content of her *de se* imaginings.

(Subjectless Content) The content of S's *de se* imagining does not (automatically) represent S herself.

We have furthermore observed in §2 that *de se* imagination is restricted by its experiential character in two ways. First, the centre-individuals which are eligible candidates for modeling the content of *de se* imagination need to be conscious. Second, their experiential perspective has to be accessible to us.

---

<sup>5</sup> In this regard, my account differs fundamentally from that of Recanati, who postulates cases of inside imagination, 'quasi-*de se* imaginings', whose centered content is not ascribed to the imaginer (Recanati, 2007, pp. 206–7). Recanati proposes that when one imagines being Napoleon one ascribes the centered content of one's imagining not to oneself but to Napoleon instead.

(Consciousness) The content of *de se* imagination comprises only centered worlds whose center-individual is conscious.

(Accessibility) The content of *de se* imagination comprises only centered worlds whose center-individual has an accessible experiential perspective.

The *structural* feature of *de se* imagination that will be relevant for us concerns temporally extended *de se* imaginings. A single centered content represents how a subject imagines herself being for a time. Many of the relevant cases concern instead lasting persistence scenarios. We can model temporally extended *de se* imaginings using *sequences* of sets of centered worlds. For instance, when S *de se* imagines having a sharp pain which slowly recedes, we can model her imagining as a sequence of centered worlds, such that the respective centre-individuals have milder and milder pains:  $\langle \langle i_{\text{pain}_{10}}, t_1, w \rangle; \langle i_{\text{pain}_9}, t_2, w \rangle; \langle i_{\text{pain}_8}, t_3, w \rangle; \langle i_{\text{pain}_7}, t_4, w \rangle; \dots \rangle$ .<sup>6</sup> S successively imaginatively self-ascribes each element in the sequence, simulating an imaginary stream of consciousness. We can summarize this point as follows:

(Sequential self-ascription) The content of *de se* imagining a temporally extended episode is a sequence of sets of centered worlds which the subject successively imaginatively self-ascribes.

In the following, I will give an account of how these features of *de se* imagination give rise to the two different types of *de se* modal illusions.

#### 4 Against the Veridical Approach to *De Se* Imagination

What is the connection between *de se* imagination and *de se* modality? One view is that *de se* imagination is a reliable guide to *de se* modality. Call this the veridical approach to *de se* imagination. In contrast, I will argue that *de se* imagination is not a reliable guide to *de se* modality, at least in most philosophically relevant cases.<sup>7</sup> Here, *de se* imagination often leads to *de se* modal illusions. Call this the non-veridical approach to *de se* imagination.

Let us consider the veridical approach first. For instance, Ninan (2009) is strongly sympathetic to the veridical approach. According to his view, *de se* imagination is a guide to the truth of *de se* modal statements such as *I could be F*. Ninan endorses the following principle:

##### Centered Guide

Imagining from the inside is a guide to centered possibility. If I can imagine a centered content *p*, that is evidence that there is a centered world  $\langle w, x \rangle$  accessible from  $\langle \text{actual world}, \text{me} \rangle$  such that *p* is true at  $\langle w, x \rangle$ . (Ninan, 2009, p. 446), .

Importantly, according to Ninan (2009), the existence of an accessible centered possible world renders true a corresponding *de se* modal statement. S's statement *I could have been F* is true iff there is a centered world accessible from S's context whose centre-individual is F (Ninan, 2009, p. 447). Putting the two pieces together gives us the following principle:

<sup>6</sup> For simplicity, I will here work with the idealization that the subject's imagination is maximally specific.

<sup>7</sup> This observation has also been made by Nichols (2008). Nichols does, however, not offer a detailed account of *de se* imagination or of the connection between *de se* imagination and possibility.

### Veridical De Se Imagination

Imagining from the inside is a guide to *de se* modal truth. If I can imagine *being F*, that is evidence that the statement *I could be F* is true when uttered by me.

The veridical approach maintains that *de se* imagination provides reliable access to *de se* modal truth. Is this plausible? To assess this question consider again the data from §2. Combined with the data, the approach predicts that an extremely wide range of *de se* modal statements are true of us. One group of statements concerns what kind of entity we could be:

*I could be a purely physical being.*

*I could be a Cartesian soul.*

*I could be a beetle.*

*I could be a conscious teapot.*

*I could be Napoleon.*

...

Another group of *de se* modal statements concerns our persistence:

*I could survive as Lefty.*

*I could survive as Righty.*

*I could outlive my body.*

*I could survive a radical modification of my psychology.*

*I could survive the replacement of my immaterial substrate.*

*I could be Julius Caesar, then Marie Curie and finally Angela Merkel.*

...

And we have seen that it is equally imaginable that we cease to exist throughout these scenarios. So we also have the predicted truth of the following statements:

*I could fail to survive as Lefty.*

*I could fail to survive as Righty.*

*I could fail to outlive my body.*

*I could fail to survive a radical modification of my psychology.*

*I could fail to survive the replacement of my immaterial substrate.*

*I could be Julius Caesar, then Marie Curie and finally Angela Merkel, but cease to exist during her second term as Chancellor.*

...

There are several worries with these predictions. First, which *de se* modal statements are true of us is arguably constrained by facts about our nature. What we could be depends on what we are. For instance, if we are indeed material, we could arguably not be immaterial. Familiar material objects such as chairs and mountains are plausibly *essentially* material. Conversely, assuming that we are in fact immaterial suggests that we could not be material. Similarly, the assumption that we are biological organisms rules out the option that we could be artefacts. And so on. Importantly, the relevant facts about our nature do not lead to corresponding constraints on which scenarios are accessible in *de se* imagination. Even when I explicitly assume that I am material, I can still imagine being an immaterial soul. And the assumption that I am a human mammal

leaves the *de se* imaginability of my being a (conscious) invertebrate untouched. So even without taking a stance on what our underlying nature actually is, it is clear that there will be a disconnect between *de se* modal truths and the range of scenarios which are imaginable from the inside. After all, we are either material or immaterial, either mammals or not.

In current modal epistemology, this issue is known as the *problem of modal epistemic friction* (Vaidya and Wallner, 2021). To be reliable, our source of modal knowledge, here imagination, needs to be reigned in. We need an assurance that the scenarios which our imagination presents as possible do not violate the nature of the objects involved (Vaidya and Wallner, 2021, p. 51914). The above suggests that *de se* imagination, just like ordinary third-person imagination, needs to be fenced in by information about our essential properties (for the corresponding claim about third-person imagination see Roca-Royes, 2011; Vaidya and Wallner, 2021).<sup>8</sup> This raises the question of how substantive the required knowledge of essential properties is. Roca-Royes (2011) and Vaidya and Wallner (2021) make the case that knowledge of essence is substantive and that imagination can therefore not be the ultimate source of modal knowledge. My own sympathies lie instead with the deflationary approach of Chalmers (2002, 2010), according to which the essential knowledge required to solve the problem of modal epistemic friction is ultimately non-substantive and based on conceptual understanding together with information about ordinary *a posteriori* identities, such as the fact that water is H<sub>2</sub>O.<sup>9</sup>

Second, the veridical approach leads to a clash between first-person and third-person modal statements. On this view, the truth of S's first-person utterance 'I could be F' depends on the existence of a centered possible world whose center-individual is F. In contrast, the truth of her third-person utterance 'S could be F' depends on the existence of an uncentered possible world at which S is F (assuming an orthodox semantics for third-person modal statements).<sup>10</sup> Since I can e.g. imagine being Napoleon, the statement *I could be Napoleon* comes out true when uttered by me on the veridical approach. On the other hand, the statement *CW could be Napoleon* is false, since there is no possibility according to which CW is Napoleon (assuming that distinctness is necessary). But I am CW So could I, CW, be Napoleon? There seems to be no neutral vantage point from which to answer this question.

Third, the account seems to yield an unattractive primitivist picture of personal persistence. If it were true that one could both survive and fail to survive two qualitatively identical scenarios, then one's persistence would seem to be independent of qualitative facts. One would appear to be

---

<sup>8</sup> In addition, van Inwagen (1998) has raised the problem of relevant depth, i.e. the problem of how detailed our imaginings need to be to qualify as reliable. On the analysis of *de se* modal illusions proposed in this paper, increasing the depth of imagination is unlikely to change the situation, since the ultimate source of the unreliability of *de se* imagination consists in the fact that the imaginer is not part of the imagined content to begin with. Adding further details by specifying the imagined subject's exact mental state or by filling in facts about her physical environment is therefore unlikely to affect the reliability of the imagining or our confidence in whether or not the scenario is possible.

<sup>9</sup> Chalmers' approach is criticized in (Vaidya, 2008) and (Roca-Royes, 2011); for a (sketchy) response see (Chalmers, 2010, fn 3). Resolving this conflict is beyond the scope of this paper. I am grateful to an anonymous referee for encouraging me to include a discussion of the problem of modal epistemic friction in the paper.

<sup>10</sup> (Ninan, 2008, Appendix) offers a centered semantics for third-person modal statements, but this proposal seems appropriate only for certain epistemic uses of such statements.

radically different from other ordinary objects whose persistence is arguably grounded in certain qualitative continuity relations. *De se* imagination further suggests that our persistence is also free-floating from particular facts, since it appears possible from this first-person perspective to live the lives of several different people in succession, or to only live part of someone's life. Our persistence would seem to be brute metaphysical bed rock. This seems highly implausible. What, on this picture, could explain the obvious metaphysical and epistemological connections between facts about our survival and facts regarding the persistence of our parts?

Fourth, the approach faces counterexamples. A relatively uncontroversial example involves imagining being someone else. Since I am not in fact Napoleon and assuming that distinctness is necessary, it is not the case that I could be Napoleon. Still, I can imagine from the inside being him. In addition, there are certain *de se* modal truths which are out of reach of the first-person perspective. For instance, it is evident that I can be unconscious. After all, I often am unconscious when sleeping dreamlessly. However, we have seen in §2 that such scenarios are inaccessible in *de se* imagination. Cases like this are counterexamples to a strengthened veridical view on which the inaccessibility of a centered world according to which *p* is true indicates that *p* is impossible.

A proponent of the veridical approach might try to remedy the situation by restricting the relevant accessibility relation, i.e. by imposing constraints on which centred worlds are accessible in *de se* imagination (Ninan, 2009, §4.4). However, this move seems *ad hoc*. The notion of *de se* imaginability is meant to have independent psychological reality—we cannot simply decide by fiat which scenarios are *de se* imaginable and which are not. Furthermore, as I will show in §5, even explicitly taking into account the relevant *a posteriori* information about our essence fails to generate the necessary modal epistemic friction and does not seem to affect the imaginability of the corresponding *de se* scenarios. Lastly, the strategy could only work for illusions of possibility to begin with, as these are due to the overly liberal character of *de se* imagination. It would be inapplicable to illusions of impossibility, since the problem here is that the range of accessible scenarios is already too narrow. Imposing further restrictions will not help. Together, these points provide sufficient motivation to explore a non-veridical approach to *de se* imagination instead.<sup>11</sup>

## 5 *De Se* Cases and a Kripkean Account of Modal Illusions

One might worry that a non-veridical approach to *de se* imagination would completely detach it from possibility and thereby sever the link between imaginability and possibility altogether (Ni-

---

<sup>11</sup> The problems for the veridical approach raised in this section are based on a metaphysical interpretation of the relevant *de se* modal statements. Ninan (2009, p. 450) stops short of fully committing himself to such an interpretation, and mentions the option of using his framework as an analysis of conceptual/epistemic possibility instead. The above points can then be understood as making the case that the framework should be interpreted in this way. Ninan furthermore intends his notion of *de se* possibility to be distinct from orthodox metaphysical possibility (Ninan, 2009, p. 447). However, the presented objections are based only on the approach's predictions about the truth values of *de se* modal statements, together with the fact that these statements have a salient metaphysical reading, as witnessed by the fact that they are at the core of our theorizing about the metaphysics of the self.

nan, 2016). This would be an overreaction. Skepticism about the reliability of *de se* imagination is compatible with optimism about the general connection between imaginability and possibility. Accordingly, I will situate my account of *de se* modal illusions within a broadly Kripkean framework, which allows for modal error but preserves a systematic connection between conceivability and possibility (Ch. 3 Kripke, 1972). The account proposed here can be considered a complementary supplement to the Kripkean picture for the *de se* case.<sup>12</sup>

Negations of necessary *a posteriori* statements, such as *water is not H<sub>2</sub>O*, are paradigm cases of conceivability-possibility failures. It is conceivable that water is not H<sub>2</sub>O, but this not possible. Why is it impossible? The reason is, Kripke argued, that the original statement *water is H<sub>2</sub>O* is a true identity statement involving rigid designators. Rigid designators have constant intensions, i.e. their actual extension is also their extension in every possible world. If there is no possible situation in which water is not H<sub>2</sub>O, in which sense is this conceivable? That water is H<sub>2</sub>O was a genuine empirical discovery. In making this discovery various alternative outcomes needed to be ruled out. For instance, had it turned out that the substance in our rivers and lakes is XYZ, as on twin-earth, water would have turned out to be XYZ. In a certain sense, then, it is conceivable that water is not H<sub>2</sub>O. There is an epistemic possibility in which water is not H<sub>2</sub>O, but no corresponding metaphysical possibility. Does that mean that there is no connection between epistemic and metaphysical modality? Are there simply two different modal realms? Not necessarily. From the remarks in (Ch. 3, Kripke, 1972), one can construct a systematic account of modal illusions that preserves an intimate link between conceivability and possibility.

Which possibility are we conceiving of when we are considering the possibility that water may have turned out to be something other than H<sub>2</sub>O? Kripke (1972, Ch. 3) suggested that we are tracking a genuine possibility, but are misdescribing it. We are in fact conceiving of an *epistemic counterpart* of the described possibility, i.e. a situation in which we are interacting with a liquid which resembles water superficially, but which has a different chemical structure. It is tempting to describe this situation as one in which *water* is not H<sub>2</sub>O, but since the term “water” is rigid, this is a mistake. It is not a situation where water, i.e. H<sub>2</sub>O, is not H<sub>2</sub>O—there is no such situation. Instead, it should be described as a situation in which the clear drinkable liquid that falls from the sky and is in the rivers and oceans, i.e. *the watery stuff*, is not H<sub>2</sub>O. The imagined possibility does not verify the statement “water is not H<sub>2</sub>O”; rather, it verifies the related statement, “the watery stuff is not H<sub>2</sub>O” which captures the original statement’s qualitative content (Kripke, 1972, Ch. 3).<sup>13</sup>

---

<sup>12</sup> I do not want to take a stance on whether the account sketched here is completely faithful to Kripke’s actual proposal; see e.g. (Byrne, 2007) for thinking that it is not. In calling it ‘Kripkean’ I merely mean to indicate that it is inspired by Kripke’s remarks in (Ch. 3, Kripke, 1972). A systematic account of the connection between epistemic and metaphysical modality has been given by Modal Monists, such as Chalmers (1996, 2002) and Jackson (1998), with the help of *two-dimensional semantics*. In contrast, Modal Pluralists, such as Soames (2005, 2006), have argued that Kripke’s findings show that there are indeed separate modal realms.

<sup>13</sup> This approach has been systematized by Chalmers (2002, 2010) within a two-dimensional framework. According to Chalmers’ approach, *water is not H<sub>2</sub>O* is primarily conceivable and also primarily possible. There is a (metaphysically) possible world which verifies the statement’s primary intension, i.e. the dimension of meaning that captures its qualitative content. The statement is neither secondarily possible nor secondary conceivable. There is no (metaphysically)

The leading charge against the Kripkean account of modal illusions is that it is psychologically unrealistic, since it claims that we are often mistaken about what we are imagining. For instance, Hill writes: “[the model] is fundamentally misguided; for [...] in non-pathological circumstances introspection gives us pretty accurate access to the contents of our own states of imagination.” (Hill, 1997, p. 83). This objection underestimates both the ambition and the potential of the proposal. The Kripkean model is not simply postulating any old confusion about what we are imagining, but an extremely subtle one. Further, the postulated confusion is such that we may well recognize it upon reflection as an accurate description of our actual imagining. As Yablo (2006, p. 335) puts it: “what makes Kripke’s approach so convincing is that [psychological adequacy] is the standard he tries to meet, and mostly *does* meet”.<sup>14</sup>

Intuitively, the judgment that water may have turned out to be something other than H<sub>2</sub>O is based on the thought that we might have been in a situation where things appear to us just as they do, but where the watery liquid around us isn’t H<sub>2</sub>O but some other substance such as XYZ. In order to meet the standard of psychological adequacy, epistemic counterpart possibilities should then fulfil two requirements (Yablo, 2006). First, a counterpart situation needs to include a *facsimilie* of water, i.e. something that produces the same superficial appearance in the counterpart situation as H<sub>2</sub>O does in actuality. Second, *we* need to be present in the counterpart situation: “The counterfactual thing has to look the same, not to the counterfactual folks, but to us. [...] what seems possible [...] is not just that [low molecular energy] could have paraded itself in front of someone or other who felt it as hot, but that *I* could [...] have found it to be hot.” (Yablo, 2006, pp. 337–339) Similarly, the situation I am considering is one in which XYZ

---

possible world which verifies the statement’s secondary intension, i.e. the dimension of meaning which captures its subjunctive content.

<sup>14</sup> Some have objected that the Kripkean model renders our modal judgments implausibly indirect (e.g. Soames, 2006). In his analysis of the famous case of a wooden table that may have turned out to be made of ice, Kripke suggests that the relevant epistemic possibility involves a *different* table (Kripke, 1972, pp. 141–142). Soames (2006, p. 303) protests: “[...] surely that can’t be right. If it were, it would be hard to see how anyone could ever conceive of anything about a specific individual, or how anyone could ever have *de re* attitudes.” He proposes that Kripke instead extend his stipulative identification of individuals across metaphysical possibilities also to the case of epistemic possibilities. In response, it is important to note that proponents of the Kripkean approach can agree that our judgment *this table could have been made from ice* concerns the very table in front of us (the same holds for other *de re* thoughts). But they also endorse the additional claim that we arrived at this judgment by misdescribing a related possibility involving a different table. Further, it is a contentious theoretical assumption and not a pre-theoretic datum, as Soames seems to assume, that *de re* attitudes involve rigidity across *epistemic* possibilities. That this assumption is problematic, even for demonstrative attitudes, can be seen on independent grounds. Consider the following scenario from Perry (1977), here slightly modified. I am at the harbour facing a large ship. Since the mid-section of the ship is obscured by large building, I am uncertain whether I am looking at one ship or two. Looking first to the left side then to the right of the building, I judge: *this ship might not be the same as that ship*. The possibility I am entertaining is one where I am facing two ships, x and y. Which of these two ships is identical to the original ship? Since choosing either x or y would be arbitrary, the most plausible answer is that neither of them is. See (Chalmers, 2011, §8) for a related case involving names. To conclude, there seems little reason to think that the Kripkean model puts us cognitively out of touch with the world, and there is independent theoretical support for a model along Kripkean lines.

Another class of alleged counterexamples to the Kripkean model are *a posteriori* impossibilities involving the first person (Wright, 2002; Soames, 2006). The account of *de se* modal illusions presented in this paper is a way of incorporating such cases into the Kripkean framework.

appears to *me* the way H<sub>2</sub>O in fact does, rather than to some arbitrary subject with a potentially alien sensory apparatus.<sup>15</sup>

To summarize: On the Kripkean model, when we are subject to a modal illusion related to a necessary *a posteriori* statement S, the corresponding modal seeming is evidence for the existence of a related possibility, an epistemic counterpart, in which the negation of a related statement  $\neg S^*$  is true. Further, the epistemic counterpart situation contains a facsimile of the relevant object/stuff and the subject herself.

Yablo (1993) has presented a schema for how *a posteriori* necessities can block the reliability of imagination. Consider a certain *a posteriori* necessity  $p$ ;  $p$ 's *a posteriori* status suggests that its negation  $\neg p$  is imaginable. (Remember, however, that on the just presented Kripkean account of modal illusions, the seeming imaginability of  $\neg p$  should more accurately be described as the imaginability of its qualitative counterpart  $\neg p^*$ .) But since  $p$  is necessary,  $\neg p$  is impossible. So in the case of *a posteriori* necessities, (seeming) imaginability is not always a reliable guide to possibility.

We can apply Yablo's schema to the first-person cases. Consider, for instance, Materialism about personal ontology. On this view, each of us is a material entity. If we are in fact material, we plausibly could not be immaterial. Still, most Materialists will admit that Dualism could have turned out to be true; there is an epistemic possibility in which we are immaterial. Such Materialists should therefore count the statement *I am a material object* as an *a posteriori* necessity, just like *water is H<sub>2</sub>O*. According to Yablo's model, the *a posteriori* status of *I am material* suggest that one can imagine its negation. And this seems correct. I can *de se* imagine *being immaterial* by imagining having a conscious perspective on the world and picturing that at the point of origin of the conscious perspective, where I subjectively locate myself, there is no material object present. Based on this *de se* imagining I may then be inclined to judge *I could be immaterial*. But, since I am essentially material, this is false. I have fallen prey to a *de se* modal illusion.<sup>16</sup>

Can we explain such *de se* modal illusions in the same way we have explained standard *a posteriori* necessities above? Not quite. There are a number of disanalogies between the water case and the case of disembodiment, which suggest that *de se* modal illusions are not like ordinary modal illusions connected to the necessary *a posteriori*. First, the disembodiment case concerns ourselves, rather than some external object or substance. As a consequence, we typically imagine the relevant scenario from the first-person point of view. We imagine from the inside *being disembodied*. Second, the standard epistemic counterpart model does not seem straightforwardly applicable here. We have, following Yablo (2006), assumed that the epistemic counterpart situations need to include us. But whether there really is a possibility in which I am disembodied is exactly what is under discussion. To assume that the subject is present in the counterpart possibility would presuppose that there really is no modal illusion to begin with. Third, in the case of *water is H<sub>2</sub>O*, the seeming of contingency weakens on reflection. After learning that water is

---

<sup>15</sup> The Kripkean model is of course compatible with the fact that we can conceive of counterfactual situations in which we do not exist.

<sup>16</sup> At this stage, I am using the case of imaginary disembodiment merely for illustration. In §6.2, I will present reasons for thinking that it is in fact a *de se* modal illusion.

H<sub>2</sub>O and recognizing the terms's rigidity, we cannot properly conceive of water being something else anymore. In contrast, in the case of *I am a material object* the apparent contingency remains. Even when I assume that I am in fact a material being, my ability to imagine from the inside being disembodied is unaffected. We can frame this point in terms of the problem of modal epistemic friction (Vaidya and Wallner, 2021). In the case of water, the information that water is essentially H<sub>2</sub>O creates enough modal epistemic friction to render a scenario in which water is some other substance inconceivable. In the first-person case, in contrast, the assumption that we are essentially material does not generate sufficient modal epistemic friction to make the disembodiment scenario inaccessible for inside imagination. These points apply *mutatis mutandis* to other cases of *de se* imagination and *de se* modal illusions. They suggest that *de se* cases deserve their own analysis.

## 6 *De Se* Illusions of Possibility

If the case of imaginary disembodiment is indeed illusory, it is an instance of a *de se* illusion of possibility: the scenario seems possible but, as described, it is not a real possibility. We can distinguish two sub-varieties of this type of illusion. First, there are *de se* illusions of possibility about personal ontology. They concern what type of entity we could be. Second, there are *de se* illusions of possibility about personal persistence. They concern what changes we are able to survive. The two different types of illusion are produced by different features of *de se* imagination. Illusions about personal ontology are due to the *content* of *de se* imagination. They can arise because the imaginer is (typically) not part of *de se* imagination's content, i.e. Subjectless Content. Illusions about personal persistence are in part due to the structure of *de se* imagination. They can arise because in imagining a persistence case from the inside one successively ascribes a whole sequence of experiential perspectives to oneself, i.e. Successive Self-ascription.

### 6.1 *De Se* Illusions about Personal Ontology

I will first explore *de se* illusions of possibility about personal ontology. According to the proposed picture of *de se* imagination, a *de se* imagining does not automatically offer evidential support for the existence of corresponding situations involving the imaginer. The reason is that the imaginer does not explicitly feature in its content (Subjectless Content). For instance, the content of my first-person imagining of *being happy* is the set of centered possible worlds with happy center-individuals. It is not part of the content that this individual has to be me. The different centered worlds in the content contain different center-individuals, and if I am not in fact happy, I will not be among them.

I will assume with orthodoxy that metaphysical possibilities are represented by regular uncentered possible worlds. I will furthermore assume an orthodox Kaplanian semantics for *de se* modal sentences (Kaplan, 1979). An utterance of the sentence 'I could be F' by S at context c is true iff there is an uncentered possible world accessible from c according to which S is F. We can then ask: which (uncentered) metaphysical possibilities does a given *de se* imagining gives us access to? There is a relatively straightforward connection between a centered content and

a corresponding set of uncentered possibilities, since centered possible worlds are nothing but regular uncentered worlds with a marked individual and time. One can derive the uncentered possibilities determined by a centered content simply by removing the marking from each centered world. To illustrate, the content of the *de se* attitude *I am happy* is the set of centered worlds with happy center-individuals:  $\{ \langle i, t, w \rangle : i \text{ is happy at } t \text{ in } w \}$ . This determines a set of uncentered worlds whose elements each contain a happy individual  $\{ \langle w \rangle : \exists x (x \text{ is happy in } w) \}$ , corresponding to the existential statement *someone is happy*. This link between centered and uncentered content allows us to maintain the general connection between *de se* imagination and possibility.

The crucial upshot is that *de se* imaginings do not typically provide evidence for possibilities involving the imaginer. Instead, they only support possibilities in which *someone* (potentially distinct from her) is the way she imagines herself being, as this is the uncentered modal content of the imagining. Hence, *de se* imagination does *not* license a singular modal statement about myself: *I could be F*. It only licenses a weaker *existential* judgment: *someone could be F*. In many cases, the fact that there is a gap between what is strictly speaking supported by *de se* imagination and the modal claim we acquire in practice is harmless, since the claim may simply happen to be true. But sometimes the gap matters. It is worth summarizing these points:

(Non-veridical De Se Imagination) *De se* imagining *being F* does not provide justification for the *de se* modal judgment *I could be F*. There is a justificatory gap between *de se* imagining *being F* and the *de se* modal judgment *I could be F*.

Even though the connection between *de se* imagination and corresponding first-person modal statements is not reliable (at least this is so in philosophically relevant cases see §8.1), we have also seen that there *is* a straightforward connection between *de se* imagination and metaphysical possibility. *De se* imagination does give us justification to accept a corresponding *existential* modal claim:

(Possibility<sub>existential</sub>) *De se* imagining *being F* does provide justification for the modal judgment:  
 $\diamond \exists x Fx$ .<sup>17</sup>

A *de se* illusion about personal ontology can arise when we base a judgment about what kind of entity we could be on a *de se* imagining. According to Subjectless Content, the content of *de se* imagination does not involve us. Consequently, the fact that we can conceive of the content as obtaining does not show that the corresponding scenario is a possibility for us. As Non-veridical De Se Imagination states, *de se* imagination is not a reliable guide to first-person modal judgments. However, there is nevertheless a strong, potentially erroneous, modal seeming that *de se* imagination *does* inform us about what is possible for us. For one thing, we connect ourselves to the individual at the center of the imagined scene by imaginatively self-ascribing being the center-individual. Further, our way of reporting *de se* imaginings also supports the impression that the situation directly involves us. We typically report its content using the first-person pronoun: e.g. ‘I am imagining a situation in which *I* am skiing down a slope’. On our analysis, this is a misdescription, similar to misdescribing an XYZ-world as a ‘water’-world. We can see the real

<sup>17</sup> This claim can fail for impossible predicates such as: *being made from water but not containing H<sub>2</sub>O*.

situation clearer when we report the imagining using a subject-control PRO construction, such as ‘I am imagining PRO<sub>s</sub> skiing down a slope’. This report makes it somewhat easier to recognize that the content of *de se* imaginings does not represent the imaginer herself. Together, the fact that we are imaginatively self-ascribing being the center-individual and the fact that we can naturally report our *de se* imaginings using the first-person pronoun help explain why *de se* imagination *seems* to support possibilities for us.

(De Se Modal Seeming) *De se* imagining *being F* seems to provide justification for the *de se* modal statement *I could be F* and typically disposes one to judge that the statement is true.

In our actual modal reasoning, we are usually unaware of the gap between the uncentered modal content of *de se* imagination and the truth conditions of the modal judgments which *de se* imagination *seems* to support. We therefore mistakenly assume that *de se* imagination does support singular modal judgments about ourselves. (Of course, this judgment can be overruled, e.g. by conflicting metaphysical commitments). This can lead to *de se* illusions of possibility. I may try to *de se* imagine being F. If it is possible for a conscious subject to be F, then I will typically succeed. But my metaphysical nature may not permit my being F, i.e. there may be no possible world where I, NN, am F. Nonetheless, because of the associated *de se* modal seeming I may judge that *I could be F*. I have succumbed to a *de se* modal illusion.

The account can explain why the *de se* modal seeming persists, unlike the modal seeming in the case of water. The reason is that the semantic features which generate the illusion are different in both cases. In standard cases of *a posteriori* necessities the relevant semantic feature is rigidity. Rigidity can be uncovered by reflecting on our intuitions about how to describe the relevant scenarios. Each of relevant scenarios contained a watery liquid and we were in a position to recognize that only scenarios in which the watery liquid has the chemical structure H<sub>2</sub>O are properly described as ‘water’-possibilities. In contrast, the semantic feature which generates the illusion in the *de se* case is the fact that the subject is not represented in the content of the *de se* attitude. This feature cannot as easily be unveiled by reflecting on how to describe possible *de se* scenarios. In fact, we have seen that a standard way of describing these cases involves the first-person pronoun and is therefore a main cause of the illusion. Further, since this feature is more elusive and more theoretically removed from ordinary speakers’ intuitions, it is harder to recognize. Consequently, the illusion is more persistent in the *de se* case.

The fact that the imaginer is not represented in the content of *de se* imagination also explains why *de se* imagination is unaffected by our assumptions about actuality. Since the imaginer herself does not enter the picture to begin with, her assumptions about what she is like in actuality have no influence on the imaginability of the corresponding *de se* scenarios. Assuming that I am in fact material therefore leaves my ability to *de se* imagine being disembodied untouched. This explains why, in the *de se* case, the relevant essentialist knowledge fails to generate the required modal epistemic friction to render the relevant scenarios inaccessible.

The proposed account of *de se* modal illusions about personal ontology is founded explicitly on the specific features of *de se* imagination. Hence, it does not lead to a general skepticism about our access to metaphysical modality. In fact, by endorsing Possibility<sub>existential</sub>, it preserves the connection between imagination and possibility and can therefore be considered a complementary

addition to the Kripkean picture. Let us now apply the picture to a concrete case. I will discuss the case of Dualism as a plausible case of a *de se* illusion about personal ontology.

## 6.2 Dualism as a *De Se* Modal Illusion about Personal Ontology

Materialists about personal ontology claim that we are material entities. Dualists claim that we are immaterial beings (“we” refers to human persons). Materialism is motivated both by broadly scientific considerations and by common-sense. Science does not give us evidence for the existence of concrete immaterial entities, and supports instead the claim that the physical world is causally closed. In addition, common sense has it that we have a spatial position, a certain weight, etc. Both strongly suggest that we are material. When considering other material objects, such as rocks or mountains we have the pre-theoretic intuition that these objects are *essentially* material. Consider a particular rock, call it ‘Rocky’. Could Rocky, this very object, have been immaterial? This supposition seems bizarre.<sup>18</sup> If ordinary material objects are essentially material, why should we be different? Assuming that we are seems like *ad hoc* special pleading. You might reply that we are indeed special: we are *minded and conscious* material objects. This distinguishes us from ordinary material objects such as rocks and mountains (but not from other animals). But, one may ask in response, why should this difference matter? It seems to matter precisely because we can imagine our mental life being instantiated without a material basis, and we have seen that the probative value of this fact is dubious.

If we are indeed essentially material, the statement *I am material* is necessary. Arguably, it is not *a priori*. Our universe could have turned out to be the way Descartes thought it is. So if true, *I am material* is an *a posteriori* necessity. But even if we assume that we are essentially material, we can nevertheless imagine *being immaterial* from the inside. I may then be tempted to judge: *I could be immaterial*. But since I am necessarily material, this judgment is based on a *de se* modal illusion.<sup>19</sup>

We have seen above that the Kripkean epistemic counterpart model is not straightforwardly applicable to such *de se* modal illusions. But we now have an alternative account for such cases. According to this alternative, what is at the core of the illusion is the fact that the imagined situation does not really contain an immaterial version of myself. I am not being presented with an impossible situation in which I am disembodied. Rather, I am strictly speaking merely picturing the existence of an immaterial soul (from the inside), and by self-ascribing being that soul I take on its perspective in imagination. It is tempting to report the imagined scenario by saying: ‘I am imagining that *I* am disembodied’, and to take it as evidence for a possibility that genuinely involves me. But, taken literally, this would be a misdescription of the imagined scene.

Again, the analysis explains why the disembodiment illusion persists even on the assumption that we are essentially material. This assumption does not prevent us from picturing the above scene, since we are not part of the scene to begin with. It thereby also explains why the cor-

---

<sup>18</sup> As Kripke (1972) reminds us, this counterfactual possibility needs to be distinguished from the epistemic possibility that Rocky might turn out to be immaterial, e.g. a hologram. We are here assuming that Rocky is in fact material.

<sup>19</sup> For a related analysis of this case in a two-dimensional semantic framework see (Weber, 2021).

responding knowledge about essence does not create enough modal epistemic friction to make the disembodiment case unavailable for *de se* imagination.

We can also accommodate other *de se* illusions about person ontology, such as fairy tale scenarios where we are turned into a conscious tea pot, along the same lines. When imagining this, we may come to endorse the modal judgment *I could be a conscious tea pot*. This, too, is a *de se* modal illusion. If we are indeed biological organisms, we could not be artefacts. Nevertheless, we can imaginatively self-ascribe being a conscious tea pot and take on its perspective from the inside, since all *de se* imaginability requires as an anchor is the existence of a corresponding conscious viewpoint, and we may allow for a far-fetched possible world containing a tea pot with an inner life.

### 6.3 *De Se* Illusions about Personal Persistence

The second philosophically interesting type of *de se* illusion of possibility concerns personal persistence. We have seen in §2 that a vast range of persistence scenarios appears survivable from the inside. From the first-person perspective our survival seems unconstrained by physical, psychological, or immaterial facts. It also appears unconstrained by particular facts about our identity. A further point that requires explanation is the fact that we can *de se* imagine with equal ease *ceasing* to exist at any point during a given scenario.

Only some of these observations can be explained with recourse to the fact that *de se* imagination's content is subjectless. This feature arguably accounts for the fact that persistence from the first-person perspective is not constrained by facts about who or what type of thing we are. Just as it explains that we can imagine being Napoleon or being a Cartesian soul for a time, so it can explain that we can picture living Napoleon's entire life, or having the whole career of a Cartesian soul. But we need further explanatory resources, since it does not show why survival from the inside also seems detached from the relations that plausibly ground the persistence of objects such as Napoleon or Cartesian souls.

The required additional explanatory tool concerns the *structure* of *de se* imagination. Above, we have modelled imaginary persistence cases using *sequences* of centered contents, i.e. Sequential Self-Ascription. An ordinary *de se* imagining consists in simulating a short-lived experiential perspective. Imaginatively self-ascribing a sequence of experiential perspectives amounts to simulating a whole *stream of experiences*. Since the imager simulates and self-ascribes the entire stream of experiences, it is presented to her as the experiential stream of single enduring individual. Crucially, this is so even when the simulated experiences belong to different individuals. We can summarize this point as follows:

(Persistence) In *de se* imagining a temporally extended episode, the subject successively self-ascribes a sequence of experiential perspectives which appears as the stream of consciousness of a single persisting individual.

A *de se* imagining of a persistence scenario presents a segment of the career of one or several individuals from the inside. One pictures what it would be like to undergo the vicissitudes the individuals are involved in. As the imager herself occupies a continuous phenomenal perspec-

tive throughout the imaginary episode, the corresponding stream of consciousness appears as the inner life of a single individual. The imagining therefore gives rise to the belief that the imagined vicissitudes are survivable, even when the centre-individuals whose experiences constitute the stream are not bound together by the relation required for personal persistence.

(De Se Modal Seeming<sub>persistence</sub>) Imaginatively self-ascribing a sequence of centered contents  $\langle\langle i, t_1, w \rangle; \langle i', t_2, w' \rangle \dots; \langle i'', t_n, w'' \rangle\rangle$  seems to provide justification for the *de se* modal statement *I could survive the imagined process stretching from  $t_1$ – $t_n$*  and typically disposes one to judge that the statement is true.

To illustrate, consider a variant of the ‘combined spectrum’ scenario from (Parfit, 1984). Imagine being a subject, call them ‘o’ for original, whose cells are successively replaced with cells of Greta Garbo. Suppose that as o’s cells are being replaced, her psychology starts to resemble more and more that of Garbo. Call the initial temporal stage  $o_{100}$ ;  $o_{99}$  represents the temporal stage where 1% of cells are replaced, and so on. According to Parfit, it is clear that the final stage  $o_0$ , a perfect psychological and physical duplicate of Garbo, and the initial stage  $o_{100}$  are not stages of one and the same person, since they lack the required physical and psychological continuity.

But when one imagines from the inside being o and undergoing the replacement procedure, there is a strong intuition that one persists through the process. One imagines experiencing a continuous stream of experiences, starting with  $o_{100}$ ’s perspective. One pictures what it would be like to become physically and psychologically more and more similar to Garbo, and finally being completely like her. We can represent the content of the imagining in simplified form as follows:  $\langle\langle o_{100}, t_1, w \rangle; \langle o_{99}, t_2, w \rangle; \langle o_{98}, t_3, w \rangle; \dots \langle o_1, t_{99}, w \rangle, \langle o_0, t_{100}, w \rangle\rangle$ . Because one imagines that each stage in the process is part of a continuous stream of consciousness, it seems clear that one experiences and therefore survives the entire process. Successively self-ascribing the different centered contents gives rise to the modal seeming: *I could survive from  $t_1$ – $t_{100}$* .

Are there constraints on which conscious individuals can constitute the content of an imaginary stream of consciousness? We have seen that they do not need to be stages of the same person. Should they perhaps still e.g. follow a smooth path through space-time? No, persistence as imagined from the inside is not restricted by spatio-temporal continuity. As mentioned before, we can imagine being a time traveller who jumps from one space-time location to another. Or we can imagine instantly switching from one person’s perspective to another, e.g. as a spirit inhabiting different people’s bodies. Equally, psychological continuity is not a constraint on persistence as imagined from the inside. Again, the only substantial restriction seems to be that we can simulate the successive phenomenal perspectives.

What we imagine from the inside is the unfolding of a conscious perspective over time. An account of personal persistence that ties our survival to such streams of consciousness should therefore best be able to capture the intuitions generated by *de se* imagination. This is exactly what Dainton and Bayne (2005) have recently suggested in their phenomenalist version of Lockeanism. At the heart of Phenomenalism lies the observation that it seems impossible to imagine from the inside both having a continuous stream of consciousness and ceasing to exist somewhere in the middle (Dainton, 2008). Phenomenalists maintain that the explanation for this is

that our survival is metaphysically grounded in streams of consciousness. The account of *de se* imagination proposed here has an alternative, non-veridical explanation. It predicts that *de se* imagination will deliver the same verdict as Phenomenalism, even if our survival is not in fact grounded in phenomenology. This does of course not show that Phenomenalism is false. But it does suggest that we should not be too impressed by Phenomenalism's ability to capture our intuitions about persistence, and arguably undermines a main source of support for the position.<sup>20</sup>

Let us next consider the observation that we can imagine from the inside *failing* to survive at any point throughout a given scenario. Take again Parfit's case from above. I can imagine surviving the entire process. But I can also imagine that my existence ends at  $t_{50}$ , i.e. when exactly half of my cells have been replaced. I can equally imagine surviving until  $t_{51}$ , when 49% of my cells are preserved, or until  $t_{73}$  when exactly 27% of the original cells remain, etc. In the latter two cases, the resulting picture of personal persistence would be less principled, but this does not render the cases unimaginable.

The explanation for this datum is, I propose, that the decision of whether to include a certain center-individual and her experiential perspective in an imaginary stream is up to the imaginer. For instance, I can decide to terminate the imaginary stream at  $t_{50}$  and not include  $o_{49}$ . My decision is not mandated by the underlying properties of the individuals. From the inside, persistence is demarcated by the boundaries of the imaginary stream of consciousness—when our inner movies end, we end. And we are the directors of our inner movies and can stop the movie at will. In practice, our decision may often be guided by our assumptions about the metaphysics of personal persistence, especially when we engage in the imaginative exercise in the context of a philosophical discussion. But this guidance is optional. In principle we can terminate the imaginary stream at any point whatsoever. This explains the otherwise puzzling observation that we can both imagine surviving and failing to survive qualitatively identical scenarios.<sup>21</sup>

In the case of *de se* illusions about personal ontology, I suggested that *de se* imagining *being F* licenses an existential modal claim, *someone could be F*, which corresponds to the imagining's uncentered modal content. Do the imaginary *de se* persistence cases likewise license existential modal claims about persistence such as *someone could survive such-and-such a process*? Arguably, such existential statements are all we need in debates about personal persistence; for instance, it is enough to show that *some human person or other* can survive body switching to establish that physical continuity is not necessary for our persistence. The answer is: no, imaginary *de se* persistence cases do *not* license corresponding existential modal claims. The reason is that the individuals which figure in relevant sequences of centered contents do not have to be stages of the same person. In other words, it is not guaranteed that a given imaginary *de se* persistence case determines a set of uncentered metaphysical possibilities that contain a persisting person who survives the

---

<sup>20</sup> Dainton (2008) offers additional arguments for Phenomenalism, based on the metaphysical structure of streams of consciousness.

<sup>21</sup> A similar observation has been made by (Kipper, 2016). Kipper's explanation for this observation is that our concept of the self has a conditional structure. It primarily aims to refer to an immaterial substance, but may refer to a material entity if no immaterial substance turns out to be available. I believe that the above explanation is superior, since we can imagine that our immaterial substance persists while we ourselves cease to exist (perhaps our soul then becomes the locus of someone else's consciousness). And we can also imagine, as noted above, surviving a replacement of souls.

imagined process.

To conclude, let me give two short clarificatory comments. First, the account maintains that persistence from the inside is imagined as a continuous stream of experiences. Is this not an idealization? Can we not imagine scenarios that involve interruptions in consciousness? For instance, I can imagine having a car accident, being hurried into the operating theatre, anesthetized and operated on, and then picture waking up in the hospital bed. True, we can imagine cases like this. But they plausibly involve a shift in perspective. The period of unconsciousness is not itself imagined from the inside. It can be merely assumed as a fact about the scenario. Or when visually depicted, it is viewed from an outside perspective. When one imagines from the inside waking up, one then switches back from the third-person perspective to the first-person point of view.

Second, the discussion may suggest that *de se* imagination is purely experiential. Not so. The centre-individuals in the content of a *de se* imagining represent both the phenomenal and the non-phenomenal aspects of how one imagines oneself being. For instance, in the above case, the centre-individuals not only have experiences of what it is like to have a car accident, they are all having a car accident.

#### 6.4 Lockeanism as a *De Se* Modal Illusion about Personal Persistence

We can use this framework to defuse a central approach to personal identity. If Materialism about personal ontology is true, we are material objects. A natural suggestion is that we are biological organisms, i.e. human animals.<sup>22</sup> It seems then furthermore plausible that our persistence conditions are those of animals and that we survive as long as our biological functioning is maintained.<sup>23</sup> The case against the biological approach to personal persistence and for a psychological one is primarily based on intuitions about the possibility of body-switching (Locke, 1694; Shoemaker, 1984).

When one imagines a body-switch from the inside, one simulates a stream of consciousness that is first being attached to one body, and later to a different body. Consider Locke's famous example of prince and cobbler (Locke, 1694). Imagine from the inside first being the prince and then being the cobbler. Assume that the scenario involves prince and cobbler existing in possible world  $w$  and stretches from time  $t_1 - t_4$ . The imagining's content can be represented as the following sequence of centered worlds:  $\langle \langle \text{prince}, t_1, w \rangle; \langle \text{prince}, t_2, w \rangle; \langle \text{cobbler}, t_3, w \rangle; \langle \text{cobbler}, t_4, w \rangle \rangle$ .<sup>24</sup> One imaginatively self-ascribes first being prince at  $t_1$ , then being prince at  $t_2$ , then being cobbler at  $t_3$ , and lastly being cobbler at  $t_4$ , simulating a stream of consciousness consisting of the prince's perspectives up to  $t_2$ , and continuing with the cobbler's perspective up to  $t_4$ .

---

<sup>22</sup> I am endorsing Animalism mainly for illustrative purposes. The general materialist stance I am advocating here is compatible with alternative views, e.g. the view that we are bodies (Thomson, 2008), or brains (Nagel, 1986).

<sup>23</sup> Animalists, qua Animalists, are not committed to the claim that we have biological persistence conditions (Bailey, 2015; Thornton, 2016).

<sup>24</sup> This is again a simplification. Unless my imagination is maximally specific, each step in the sequence will correspond to a *set* of centered worlds, rather than a single centered world.

We can picture switching bodies without difficulty. We now also have an explanation as to why this *seems* to support the judgment *I could switch from this body to another one* while maintaining that body switching is impossible for us. The imagined situation seems possible, since the neutrally described content of the imagination does indeed correspond to a possible situation. It simply involves two material bodies with their respective conscious perspectives. If the imagination world does not contain a switching perspective, what explains our *seeming* to switch from one body to another in imagination? This can be explained with recourse to the attitudinal relation—the switch occurs at the level of self-ascription. We switch from imaginatively self-ascribing prince’s perspective to imaginatively self-ascribing cobbler’s perspective. Since we simulate the entire stream of consciousness from  $t_1$ – $t_4$  ourselves, it seems to us that we, as a single subject of consciousness, are experiencing the unfolding of the entire imaginary stream and therefore that we could survive the imagined ‘body switch’ (De Se Modal Seeming<sub>persistence</sub>).

This provides a non-veridical explanation for *de se* modal seemings connected to body-switching cases. Is there reason to think that they are indeed illusory? If Materialism about personal ontology is indeed correct, the claim that *I could switch from my body to a different body* may well be false. In the Materialist’s eyes I simply *am* a certain material object. Switching bodies would consist in me first being one material object and then another one. “But when the matter is put this way, it is evident that your belief is simply impossible, a violation of the very well established modal principle that a thing and another thing cannot become a thing and itself.” (van Inwagen, 1997, p. 310)<sup>25</sup> From the materialist’s perspective, we successively imagine being two distinct people by amalgamating their separate perspectives into one unified stream. We then misdescribe the imagining as one in which a single persisting subject genuinely switches from one body to a different body. We are subject to a *de se* illusion about personal persistence.

## 7 De Se Illusions of Impossibility

So far, we have considered *de se* illusions of possibility. *De se* imagination also gives rise to *de se* illusions of impossibility. These concern genuine possibilities which are inaccessible from the first-person point of view. They are grounded in the experiential character of *de se* imagination. As noted before, *de se* imagination requires a conscious perspective (Consciousness) which we are able to simulate (Accessibility). We cannot imagine from the inside being unconscious or having radically unfamiliar experiences. Taking *de se* imagination at face value would predict that the following statements are true of us: *I could not be unconscious*, *I could not have bat-type experiences*. Assessed purely from the inside, these verdicts may seem plausible. It is, however, relatively easy to see that they are false. We are often are unconscious (and actuality entails possibility). Similarly, we could have had experiences rather unlike the ones we are in fact having.

While it is not difficult to see how *de se* imagination can lead to *de se* illusions of impossibility, another question remains. Why are *de se* illusions of impossibility less persistent than *de se* illusions of possibility? Few are inclined to endorse the above impossibility claims. Plau-

<sup>25</sup> Materialist perdurantists can still make sense of body-switching. According to Perdurantism, I first overlap some of this body’s temporal parts, and then some of another body’s temporal parts (van Inwagen, 1997).

sibly, the answer lies in the different logical structure of possibility and impossibility claims. Possibility claims are positive existential claims, whereas impossibility claims are negative existential claims. A single possibility verifies a given possibility claim. In contrast, a single counter-possibility *falsifies* a given impossibility claim. And it is not hard to find counter-possibilities for the above claims. We know from experience (e.g. through testimony) that we are often unconscious. Equally, it is relatively easy to imagine from a third-person point of view a scenario where e.g. evolution has endowed humans with biological sonar. The apparent verdicts delivered by *de se* imagination are therefore easily overruled in these cases.<sup>26</sup>

To sum up. I have argued that we should not in general trust the offerings of *de se* imagination. It is in one sense overly liberal and generates illusions of possibility. In another sense, it is too restricted and produces illusions of impossibility. I have presented an account of how the specific features of *de se* imagination can lead to the two types of illusions. And I have also made the case that the view is compatible with maintaining a link between imagination and possibility. Before concluding, I will respond to potential objections.

## 8 Objections and Replies

### 8.1 Mundane *De Se* Modal Knowledge

*Objection:* We have a large amount of mundane *de se* modal knowledge. For instance, I know that *I could have skied down this slope*. The non-veridical approach renders this fact mysterious.

*Reply:* The concern about the unreliability of *de se* imagination does not automatically extend to mundane *de se* modal knowledge. First, the relevant modal belief may not be based on *de se* imagination at all, but may be supported by other sources such as general modal principles (e.g. I know that it is possible for entities of type T to be skiing down a slope and that I am of type T) or memory (e.g. I remember skiing down a similar slope as a child). Second, even where imagination is involved, the judgment may depend on third-person imagination. Third, and most importantly, the account predicts that there is reliable connection between *de se* imagination and possibility in ordinary cases. When I imagine from the inside skiing down a slope, I will typically imagine a human person skiing down a slope and picture how things feel from her point of view. If the relevant modal predicate, *possibly being F*, applies to me in virtue of being a typical human being, it does not matter whether the person I imagine really is me or someone else. Therefore, *de se* imagination will not lead us astray in such run-of-the-mill cases. In many philosophically interesting cases, on the other hand, this reasoning does not apply. When considering the possibility of being Napoleon or of having different parents, the identity of the imagined subject does matter. And when considering the possibility of disembodiment, I am not picturing an ordinary human person but rather a disembodied soul, and it is questionable whether *any* human person could be such an entity. In such cases *de se* imagination can be unreliable. Somewhat ironically, therefore, *de se* imagination is untrustworthy in precisely the kind

---

<sup>26</sup> Not all cases of seeming *de se* impossibilities can be so easily resolved. I have argued in other work (Weber, 2024) that a more persistent *de se* illusion of impossibility concerns *indeterminate personal survival*.

of cases in which we have relied on it most heavily in philosophy.

## 8.2 *De Re* Imagination about Oneself

*Objection:* The proposed account of *de se* modal illusions misdescribes the relevant imaginings. When assessing whether a certain situation is a possibility for me, I do not simply imagine the situation from the inside and picture experiencing what the center-individual experiences, rather, I explicitly assume that it is me, NN, who is the subject of that situation.

*Reply:* We can acknowledge that there is a type of self-directed imagination which goes beyond picturing a situation from the inside. For this type of imagination, I add an explicit identification of myself, NN, with the center-individual to the imagined content. This is an instance of the broader phenomenon of *identifying elements of the imagined scene with real-world objects* (or with objects from one's belief worlds). Call cases of imagination that include such an explicit identification '*de re* imagination'. Call the special case where I identify myself, NN, with an object in the imagined situation '*de re* imagination about oneself'.

A more comprehensive account of imagination then comprises the following three elements. First, the subject imagines a certain qualitative content. Second, the qualitative content may be perspectival and the imaginer can imaginatively self-ascribe this perspectival content. Third, the imaginer may identify elements of the imagined content with real world objects, including herself. (The succession of these steps is merely analytic and does not indicate a temporal ordering.) To accurately describe a subject's imagining we therefore have to specify: first, the qualitative content of her imagining; second, which perspective she imaginatively self-ascribes if any; and third, which elements of the imagined scene she identifies with which objects if any. An imagining is from the inside or *de se* iff the imaginer imaginatively self-ascribes a (interesting<sup>27</sup>) centered content. An imagining is *de re* iff the imaginer identifies one or several elements of the imagined content with objects from the actual world or her belief/supposition worlds. This allows for mixed cases of *de se* and *de re* imagination. For instance, I can self-ascribe being an individual skiing down a hill and in addition identify this individual with Angela Merkel. This corresponds to imagining from the inside *being Angela Merkel skiing down a hill*. We can then distinguish the following:<sup>28</sup>

1. S imagines *that someone is skiing down a hill*. [Imagining from the outside]
2. S imagines *skiing down a hill*. [Imagining from the inside]
3. S imagines that *Angela Merkel is skiing down a hill*. [De re imagining from the outside]
4. S imagines *being Angela Merkel skiing down a hill*. [De re imagining from the inside]

---

<sup>27</sup> For this notion see (Egan, 2007).

<sup>28</sup> The descriptions are to be read such that verb phrase/PRO constructions are interpreted as imaginings from the inside, whereas that-clauses are to be read as imaginings from the outside. Note that we are not trying to answer the difficult question of how to give a compositional semantics for imagination-reports; for this topic see e.g. (Ninan, 2012).

5. S imagines that *S is skiing down a hill*. [De re imagining about oneself from the outside]
6. S imagines *being S skiing down a hill*. [De re imagining about oneself from the inside]
7. S imagines *being Angela Merkel skiing down a hill while S is watching*. [De re imagining from the inside & de re imagining about oneself from the outside]

We can reformulate the initial objection within this framework as the claim that the relevant imaginings are not simply imaginings of type 2, but rather imaginings of type 5 or 6, i.e. *de re* imaginings about oneself. This raises the question of whether *de re* imagination about oneself is a reliable guide to first-person possibility. The initial discussion of Kripke's account of modal illusions in §5 and of the problem of modal epistemic friction (Vaidya and Wallner, 2021) have shown that *de re* imagination is not a reliable guide to possibility. We are often unaware of objects' essential properties, including our own ones. We may then seemingly *de re* imagine a scenario in which a certain object or person lacks some of their hidden essential properties. For instance, Kripke argued that we necessarily descend from our actual biological parents. Nonetheless, it is tempting to say that I can *de re* imagine a situation where I am Angela Merkel's and Joe Biden's child, by conceiving of a world in which Joe Biden and Angela Merkel have a child and by explicitly identifying that child with myself, NN. By coming to accept the statement *CW could be Joe Biden's and Angela Merkel's child*, I succumb to a modal illusion based on *de re* imagination.

To explain this type of modal illusion, we can resort to Kripke's original epistemic counterpart story. When I have an illusory modal seeming that *x could be F*, I am not really imagining an impossible situation where *x* is *F*. Rather, I am imagining an epistemic counterpart situation in which some other object *y* is *F* which plays the *x*-role in this situation. Since *x* is the object which in fact plays that role, we are led to misdescribe the imagining as one in which *x* is *F*. When the *de re* modal illusion concerns ourselves, we will then have to lift Yablo's requirement that we are always part of the epistemic counterpart situation. This returns us to the question of whether the model is psychological adequate. It arguably passes that test when we conceive of ourselves as the realizer of a certain role. The role may be external, such as the bearer of a certain name, who has such and such a history and visual appearance, etc., or it may be internal, such as the thinker of this\* thought-token, or a mix between the two. The model is ill-suited for cases in which we do not think of ourselves as the realizer of any such role, either external or internal. However, here we can apply the proposed account of *de se* modal illusions, as these are precisely the kinds of cases where we imagine a certain situation purely from the inside.

Acknowledging that there is an additional sense in which imagination can be directed at ourselves does not affect the overall discussion. First, in philosophically relevant cases, we are arguably often relying on *de se* imagination. Second, as we just saw, *de re* imagination about oneself does not offer reliable access to first-person possibility either.

### 8.3 *De se* attitudes and the justificatory gap

*Objection:* The proposed account of *de se* modal illusions postulates a justificatory gap between *de se* imagination and *de se* modal belief. At the same time, it assigns both *de se* belief and *de se* imagination the same type of content and the same type of subject-relative satisfaction conditions.

There seems to be no relevant difference between *de se* modal belief and *de se* imagination that could explain how the alleged justificatory gap between the two could emerge.

*Reply:* It is important to note that relevant pairs of *de se* imaginings and *de se* modal beliefs do not have the same content. Schematically, the content of the *de se* imagining is *being F*. The content of the corresponding modal judgment is *possibly being F*. Even if we adopted the veridical approach, it is not the case that a single piece of content gets transferred from *de se* imagination to *de se* belief. (If one were to directly plug in the content of the *de se* imagining into one's belief box, one would acquire the belief *I am F*.) Rather, on the veridical picture, imagining *being F* gives one a license to adopt a belief with the content *I could be F*. The non-veridical approach simply questions whether this license should in general be granted.

*Objection:* One may raise a related worry. In the case of an alleged *de se* modal illusion, the *de se* modal belief is in a certain sense inappropriate while the corresponding *de se* imagining is entirely unproblematic. What explains this asymmetry between *de se* imagination and *de se* modal belief given their close parallels in content and satisfaction conditions?

*Reply:* The asymmetry is explained by a difference in direction of fit between belief and imagination. Belief aims at truth and has a mind-to-world direction of fit. Imagination does not. A belief is rendered inappropriate when its satisfaction conditions are not fulfilled. An imagining, on the other hand, can be perfectly appropriate even when it is not satisfied. For instance, even though I know that penguins do not fly, there is nothing objectionable in my imagining that they do. In lacking a mind-to-world direction of fit, imagination is more similar to desire. (It arguably differs from desire-type states in also lacking a world-to-mind direction of fit (see also Humberstone, 1992, fn. 25).) As a result, the rational constraints on imagination are weaker than those on belief. A *de se* believer has to ascertain that she really is the way she takes herself to be. In contrast, a *de se* imaginer does not have to do that. Arguably, there are few rational constraints on imagination beyond internal ones such as consistency (and perhaps ones related to the phenomenon of imaginative resistance). This explains why a *de se* modal belief can be inappropriate even when the corresponding *de se* imagining is not.

## 9 Conclusion

I have sketched a picture of our nature and persistence conditions according to which we are material beings whose survival is grounded in some form of physical continuity. While this picture has support from science and common sense, there are extremely compelling intuitions that pull us away from it towards Dualism and Lockeanism. The proposed account of *de se* modal illusions gives us reason to resist these forces and stick to the initial picture. But do the opposing forces really depend on first-person imagination, or could we instead replace the dialectical role of first-person imagination with its third-person counterpart?

In practice, we often rely on *de se* imagination when thinking about the self. So, if the case holds up, we should at least have a fresh look at the relevant debates. And the case plausibly reaches beyond this, since it is doubtful whether the relevant anti-materialist intuitions *could* be mobilized from a purely third-person standpoint. Judith Jarvis Thomson expresses this idea

clearly:

I doubt that people would be moved by [descriptions of body-switching] to conclude that body-switching is possible if they did not think they could imagine body-switching from inside, as it were. It is *that*, I think—the idea that one can imagine oneself switching bodies—that bears the weight [...]. (Thomson, 2008, pp. 166-167)

Is Thomson right? Consider body-switching. We just saw that judgments about personal persistence based on psychological or physical continuity can easily be overruled by intuitions concerning phenomenal continuity. When considering an associated stream of consciousness, it seems undeniable that we go where the stream goes. When taken at face value, it is difficult to avoid the impression that psychological or physical continuity on their own cannot settle the question of personal persistence. The point also suggests that what really drives the intuitions behind Lockeanism is an implicit appeal to consciousness. For instance, we implicitly conceive of the switch in the prince's psychology as accompanied by a parallel switch in his conscious perspective (Locke, 1694). When we describe the case more neutrally as a situation in which the cobbler's psychology in the morning matches, and is causally dependent on, the prince's psychology of the previous evening, and add to that the explicit stipulation that it is an open question whether or not there was a switch in conscious perspective between prince and cobbler, the case no longer elicits a clear verdict on whether a body switch has taken place. So described, the case is inconclusive. Once the implicit backing from the first-person perspective is removed, Lockeanism loses its trusted royal support.

Consider next disembodiment. It is not too difficult to imagine being disembodied from the inside. But can we imagine disembodiment from the outside? Can we imagine Joe Biden being disembodied? Perhaps we can, perhaps we cannot. The case is not clear. The proposed account has a diagnosis for the difference. In the third-person case, the object of imagination is explicitly represented in the content of imagination. How that object actually presents itself to us puts constraints on how we can represent it in our imagination. Of course, not every feature needs to be retained in imagination. For instance, even if I think about Biden as a white-haired elderly gentleman, I can imagine him as a young teenager with an orange mullet. But this has limits. I may be unable to coherently imagine an object which presents itself to me as material as immaterial. Once more, the case seems indeterminate from a purely third-person point of view.

It appears that Thomson is right: there is no decisive third-person, imagination-based path to body-switching or disembodiment. Further, in so far as we do possess residual intuitions concerning third-person cases, these may well be parasitic on corresponding first-person intuitions. For instance, when assessing whether Joe Biden could be disembodied, I may imagine being disembodied myself, decide that I could be, and then transfer this judgment to Biden, assuming that he is just like me. This leads to an additional problem. A third-person, imagination-based route to body-switching or disembodiment would need to secure that the relevant intuitions are not corrupted by the distorting influence of the first-person point of view. It is doubtful whether we can achieve a safe quarantining of our third-person intuitions in practice. And there may be a more thoroughgoing problem. Consciousness plays a crucial role for assessing the possibility of body-switching and disembodiment. Our grip on the conscious mental states of others in turn

may ultimately depend on simulating these states in our own case (Goldman, 2006), and so may be grounded in the first-person perspective.

It may therefore be safer to depart from a non-modal starting point. The case for Materialism presented above appealed to scientific considerations and to common-sense. Once the question of personal ontology is settled non-modally, we may then be able to tackle the question of personal persistence, since the available options have been limited. If we are human animals, we plausibly have biological persistence conditions. Admittedly, I have ultimately appealed to a modal claim, i.e. the idea that if we are material, then we are *essentially* material. However, the underlying reasoning was founded on intuitions about ordinary material objects, such as rocks or mountains. Arguably, there is little risk that such intuitions are distorted by *de se* imagination. A useful heuristic may therefore be that when relying on the modal realm to start from general principles not tailor-made to our own case. We may think that we are special, but from the perspective of logical space we are perfectly humdrum beings.

### Acknowledgements

For helpful comments and discussion I would like to thank Ryan Cox, Lloyd Humberstone, Frank Jackson, Thomas Sattig, Stephan Torre, Tobias Wilsch, and participants at talks at the AAP 2021 and at the Oberseminar for Theoretical Philosophy at the Eberhard Karls University of Tübingen. A special thanks goes to Wolfgang Schwarz. Research on this paper was supported by a fellowship from the Alexander von Humboldt Foundation (AUS 1217439 HFST-E) and by a fellowship from the Australian Research Council (DE220101158).

### References

- Bailey, A. (2015). Animalism. *Philosophy Compass* 10(12), 867–883.
- Blackburn, S. (1997). Has Kant refuted Parfit. In J. Dancy (Ed.), *Reading Parfit*, pp. 180–201. Oxford: Blackwell.
- Byrne, A. (2007). Possibility and imagination. *Philosophical Issues* 21, 125–144.
- Caie, M. and D. Ninan (MS). First-person propositions. *Philosophers' Imprint*. Forthcoming.
- Chalmers, D. (1996). *The Conscious Mind*. Oxford: Oxford University Press.
- Chalmers, D. (2002). Does conceivability entail possibility? In T. S. Gendler and J. Hawthorne (Eds.), *Conceivability and Possibility*, pp. 145–200. Oxford University Press.
- Chalmers, D. (2010). The two-dimensional argument against materialism. In D. Chalmers (Ed.), *The Character of Consciousness*, Chapter 6, pp. 141–205. New York: Oxford University Press.
- Chalmers, D. (2011). The nature of epistemic space. In A. Egan and B. Weatherson (Eds.), *Epistemic Modality*, pp. 60–107. Oxford: Oxford University Press.
- Dainton, B. (2008). *The Phenomenal Self*. Oxford: Oxford University Press.

- Dainton, B. and T. Bayne (2005). Consciousness as a guide to personal persistence. *Australasian Journal of Philosophy* 83(4), 549–571.
- Descartes, R. (1998 [1641]). *Meditations and Other Metaphysical Writings*. London: Penguin.
- Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. New York: Oxford University Press.
- Hill, C. S. (1997). Imaginability, conceivability, possibility and the mind-body problem. *Philosophical Studies* 87(1), 61–85.
- Humberstone, L. (1992). Direction of fit. *Mind* 101(401), 59–83.
- Jackson, F. (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.
- Kaplan, D. (1979). On the logic of demonstratives. *Journal of Philosophical Logic* 8, 81–98.
- Kipper, J. (2016). Substance and the concept of personal identity. *Ergo* 3(1).
- Kripke, S. (1972). Naming and necessity. In D. Davidson and G. Harman (Eds.), *Semantics of Natural Language*, pp. 253–355, 763–769. Dordrecht: Reidel. Revised edition published in 1980 as *Naming and Necessity* (Cambridge, MA: Harvard University Press).
- Lewis, D. (1979). Attitudes *de dicto* and *de se*. *The Philosophical Review* 88, 513–43.
- Locke, J. (1975 [1689/1694]). The Clarendon Edition of the Works of John Locke. In P. H. Nidditch (Ed.), *An Essay concerning Human Understanding*. Oxford University Press,.
- Nagel, T. (1974). What it is like to be a bat? *Philosophical Review* 83(4), 435–50.
- Nagel, T. (1986). *The View from Nowhere*. New York: Oxford University Press.
- Nichols, S. (2008). Imagination and the *I*. *Mind and Language* 23(5), 518–535.
- Ninan, D. (2008). *Imagination, Content, and the Self*. Ph. D. thesis, MIT.
- Ninan, D. (2009). Persistence and the first-person perspective. *The Philosophical Review* 118(4), 425–464.
- Ninan, D. (2016). Imagination and the self. In A. Kind (Ed.), *The Routledge Handbook of Philosophy of Imagination*, Chapter 20, pp. 274–285. Abingdon: Routledge.
- Parfit, D. (1984). *Reasons and Persons*. Oxford: Clarendon.
- Perry, J. (1977). Frege on demonstratives. *Philosophical Review* 86, 474–97.
- Recanati, F. (2007). *Perspectival Thought*. Oxford: Oxford University Press.
- Roca-Royes, S. (2011). Conceivability and de re modal knowledge. *Noûs* 45(1), 22–49.

- Shoemaker, S. (1984). Personal identity: A materialist's account. In R. Swinburne and S. Shoemaker (Eds.), *Personal Identity*, pp. 67–132. Oxford: Blackwell.
- Shoemaker, S. (1994). The first-person perspective. *Proceedings and Addresses of the American Philosophical Association* 68(2), 7–22.
- Soames, S. (2005). *Reference and Description: The Case against Two-Dimensionalism*. Princeton: Princeton University Press.
- Soames, S. (2006). The philosophical significance of the Kripkean necessary *A posteriori*. *Philosophical Issues* 16, 288–309.
- Thomson, J. J. (2008). People and their bodies. In T. Sider, J. Hawthorne, and D. W. Zimmerman (Eds.), *Contemporary Debates in Metaphysics*, pp. 155–176. Malden, MA: Blackwell Publishing.
- Thornton, A. K. (2016). Varieties of animalism. *Philosophy Compass* 11(9), 515–526.
- Torre, S. (2021). First-person imaginings. Unpublished manuscript.
- Vaidya, A. (2008). Modal rationalism and modal monism. *Erkenntnis* 68(2), 191–212.
- Vaidya, A. J. and M. Wallner (2021). The epistemology of modality and the problem of modal epistemic friction. *Synthese* 198(Suppl 8), 1909–1935.
- van Inwagen, P. (1997). Materialism and the psychological continuity account of personal identity. *Philosophical Perspectives: Mind, Causation, and World: Metaphysics* 11, 305–19.
- van Inwagen, P. (1998). Modal epistemology. *Philosophical Studies* 92, 67–84.
- Vendler, Z. (1979). Vicarious experience. *Revue de Métaphysique et de Morale* 84, 161–73.
- Weber, C. (2021). Material people in logical space. *Australasian Journal of Philosophy* 99(3), 517–530.
- Weber, C. (2023). A puzzle about first-person imagination. *Philosophical Studies* 180, 2435–2455.
- Weber, C. (2024). Determinacy of the self and imagination from the inside. Unpublished manuscript.
- Williams, B. (1973). Imagination and the self. In B. Williams (Ed.), *Problems of the Self*, pp. 26–45. Cambridge University Press.
- Wright, C. (2002). The conceivability of naturalism. In T. S. Gendler and J. Hawthorne (Eds.), *Conceivability and Possibility*, pp. 401–439. Oxford University Press.
- Yablo, S. (1993). Is conceivability a guide to possibility? *Philosophy and Phenomenological Research* 53(1), 1–42.
- Yablo, S. (2006). No fool's cold: Notes on illusions of possibility. In M. García-Carpintero and J. Maciá (Eds.), *Two Dimensional Semantics*, pp. 327–345. Oxford University Press.