

De Se Beliefs, Self-Ascription, and Primitiveness

Florian L. Wüstholtz
University of Fribourg

Abstract

De se beliefs typically pose a problem for propositional theories of content. The Property Theory of content tries to overcome the problem of *de se* beliefs by taking properties to be the objects of our beliefs. I argue that the concept of self-ascription plays a crucial role in the Property Theory while being virtually unexplained. I then offer different possibilities of illuminating that concept and argue that the most common ones are either circular, question-begging, or epistemically problematic. Finally, I argue that only a primitive understanding of self-ascription is viable. Self-ascription is the relation that subjects stand in with respect to the properties that they believe themselves to have. As such, self-ascription has to be primitive if it is supposed to do justice to the characteristic features of *de se* beliefs.

Keywords: *de se*, self-ascription, property theory, essential indexical

I INTRODUCING THE PROBLEM OF *DE SE* BELIEFS

De se beliefs are beliefs about oneself such as *My legs are crossed* which are typically expressed in language using the first-person pronoun. As such, *de se* beliefs are different from other beliefs about oneself such as Valentina Tereshkova believing of herself *The first woman in space is brave*. Furthermore, *de se* beliefs come with specific features such as their connection to self-knowledge, intentional action, perception and other first-personal phenomena. Now, the problem of *de se* beliefs is to differentiate clearly between *de se* beliefs and other beliefs that subjects have about themselves. This enterprise proved especially problematic in classical theories of propositional content such as the so-called ‘Doctrine of Propositions’ (cf. Castañeda 1999a; Perry 1979), which can be characterised by two features that are relevant here:

1. Belief is a relation between a subject and an object/content.

2. The object of a belief is a proposition with an invariable truth-value.¹

The question for the proponent of the Doctrine then is: can we accommodate *de se* beliefs with all their characteristics within the Doctrine? David Lewis argues in his ‘Attitudes *De Dicto* and *De Se*’ (1979) that this is not possible because a problem arises for the Doctrine from endorsing feature 2. If we take propositions with invariable truth-values as the objects of beliefs, we are incapable of distinguishing between a subject believing something of herself in the *de se* way and her believing something of herself in the non-*de se* way. He forcefully illustrates this problem in his well-known case of the two gods:

Consider the case of the two gods. They inhabit a certain possible world, and they know exactly which world it is. Therefore they know every proposition that is true at their world. Insofar as knowledge is a propositional attitude, they are omniscient. Still I can imagine them to suffer ignorance: neither one knows which of the two he is. They are not exactly alike. One lives on top of the tallest mountain and throws down manna; the other lives on top of the coldest mountain and throws down thunderbolts. Neither one knows whether he lives on the tallest mountain or on the coldest mountain; nor whether he throws manna or thunderbolts. (Lewis 1979: 520–521)

From Lewis, who endorses feature 1., we thus get the following argument against the Doctrine: We want to be able to distinguish between *de se* beliefs and non-*de se* beliefs. If belief is a relation between a subject and a truth-conditional propositional object, then we cannot draw this distinction. This is because there are *de se* beliefs and non-*de se* beliefs which the Doctrine analyses as relations to the *same* proposition despite being distinct beliefs. To illustrate this: If Alpha believes *I am standing*, the (singular) proposition she is related to in believing is <that Alpha is standing>. The alternative first-personal proposition <that I am standing> is not available for the Doctrine, because it does not come with an invariable truth-value in the required sense. If Alpha entertains that first-personal proposition, it is true iff Alpha is standing and if Beta is entertaining it, it is true iff Beta is standing. Hence, the truth-value of such a first-personal proposition is variable—it might be true in Alpha’s case and false in Beta’s. The problem is that, if Alpha entertains the non-*de se* belief *Alpha is standing*, she is also related to the (singular) proposition <that Alpha is standing>. We are thus left with two distinct beliefs which

¹ What does it mean for a proposition to have an ‘invariable truth-value’? In the present context, it is supposed to capture that propositions are such that they are generally true in a given possible world. In this way, propositions are not ‘relativised’ with respect to a given possible world. Accordingly, the Doctrine’s propositions can be true in one world and false in another, but they cannot be true in one part of a possible world (e.g. for Alpha) and false in another part of the same possible world (e.g. for Beta). Perry and Castañeda include further features as characteristic of the Doctrine, which make the feature of invariability stronger and more explicit. For the present purposes of analysing the concept of self-ascription, these additional ones are not relevant since the main target is feature 2.

are analysed as a relation to the same object. And since the propositional object of belief is the only individuating feature of beliefs, we have two distinct beliefs which cannot be distinguished within the Doctrine. Hence, the Doctrine is false. This does not imply that all theories of propositional content are incapable of solving the problem of *de se* beliefs and hence false. For instance, Perry defends a two-dimensionalist solution where he distinguishes between belief state and propositional belief content in a fashion similar to Kaplan's (1989) distinction between character and content.

2 THE PROPERTY THEORY AND THE NECESSITY OF SELF-AScription

As an alternative to the propositional strategy, Lewis offers us a solution to the problem which contains two elements. First, there is the *semantic aspect* which concerns the nature of the objects of beliefs. He proposes that the objects of beliefs are properties, understood as sets of possible individuals (Lewis 1979: 515). Accordingly, the property 'standing' is the set of all standing individuals and the property 'being such that a is F' is the set of all the individuals which inhabit a possible world where a is F. So, while the Doctrine took *propositions* to be the objects of beliefs, Lewis defends a 'Property Theory of Content' in arguing that we should exchange them with *properties*.²

The obvious first question is whether such a move is capable of reproducing the logical role that propositions played within the Doctrine. Lewis thinks it does and defends two essential theses concerning the relation between properties and propositions:

- (I) There corresponds a property to every proposition.
- (II) There are some properties to which there correspond no propositions.

Thesis (I) together with feature 1. entails that every individuation of beliefs in terms of propositions can be equally successfully done in terms of properties. If we understand propositions coarsely truth-conditionally, e.g. as sets of possible worlds (cf. Lewis 1979: 515), then they can 'cut' logical space only along possible world borders—just like the Doctrine's propositions. The proposition <that Alpha is standing> throws all possible worlds into either of two baskets. The ones where it is true are thrown into the 'true' basket and the others into the 'false' basket. But since Lewis is so liberal about the nature of properties, he can gerrymander a property that corresponds to every such proposition.³ In the case of the proposition <that Alpha is standing>, we can use the property that is possessed by every individual in a world where it is true, namely 'inhabiting a

² Such a theory is also independently elaborated and defended by Chisholm (1981) and has recently been revived by Feit (2008, 2012). The differences between these theories are marginal and insignificant for the present discussion since they all struggle with the concept of self-ascription.

³ Lewis being an extreme liberalist with regard to properties also manifests itself in his *On the Plurality of Worlds*, where he writes that 'the abundant properties may be as extrinsic, as gruesomely gerrymandered, as miscellaneous disjunctive, as you please. They pay no heed to the qualitative joints, but carve things up every which way' (Lewis 1986: 56).

world where Alpha is standing’ or ‘being such that Alpha is standing’. This property similarly throws things into one of two baskets. But this time it refers to individuals and not possible worlds and the two baskets are labelled ‘instantiates F’ and ‘does not instantiate F’. In the end, things will end up in the same baskets for every corresponding property–proposition pair because every individual which instantiates F also inhabits a world where the relevant proposition is true and *vice versa*. Hence, properties are capable of playing this logical role of the Doctrine’s propositions.

Thesis (II) together with feature 1. entails that some individuations of beliefs in terms of properties cannot be equally successfully done in terms of propositions. In other words: properties can ‘cut’ up logical space in a more fine-grained way than propositions. For instance, while everyone in this world has the property of ‘being such that Alpha is standing’, only some individuals in this world have the property of ‘standing’ (e.g. Alpha, lamps, tables, etc.). So, there are some properties that can distribute things into our two logical baskets in ways that are not available to the Doctrine’s propositions. This implies that it is possible to logically differentiate in a more fine-grained way using properties than it would be possible using propositions. This is not a particularly surprising result since properties are generally such that they need not be instantiated in every individual of a given possible world. On the other hand, the Doctrine’s propositions are such that they are true for every individual of a given possible world.

Lewis can therefore support his negative argument against the Doctrine with the positive argument for the Property Theory: The Doctrine’s propositions are coarsely truth-functional, hence their truth-value does not differ within possible worlds. Properties, on the other hand, are not generally instantiated in every individual of a given world, hence their instantiation may differ within possible worlds. Thus, properties can differentiate in a more fine-grained way than propositions. As we saw before, *de se* beliefs require a more fine-grained individuation that is unavailable within the Doctrine. In contrast, the Property Theory provides the tools for such an individuation. If Alpha believes *I am standing*, she is related to the property ‘standing’ whereas if she believes *Alpha is standing*, she is related to the distinct property ‘being such that Alpha is standing’. And hence, the Property Theory can solve our problem.

This sounds as if the semantic move from propositions to properties is enough to solve our problem. However, we can argue that this semantic aspect of the Property Theory is incapable of solving the problem *on its own* because the semantic aspect is insufficient for drawing all the necessary distinctions we are interested in. It is true that we can distinguish between Alpha’s belief *I am standing* and her belief *Alpha is standing* using property content—something that was not possible within the Doctrine. While the first belief has the property ‘standing’ as its content, the second has the property ‘being such that Alpha is standing’ as its content. But this alone does not suffice to clearly distinguish *de se* beliefs from non-*de se* beliefs in general.

Let me explain: Lewis accepts that we want to analyse the distinction between believing *de dicto* and believing *de se*. The semantic aspect tells us that believing *de dicto*

amounts to being related to a property such as ‘being such that a is F’ whereas believing *de se* amounts to being related to the property ‘F’. But this semantic difference does not yet account for the specific features of *de se* beliefs. In either case the subject is related to a property of some kind. But what makes the former *de dicto* and the latter *de se*? Why is it that Alpha believes of herself *in the de se way* that she is standing if the object of her belief is the simple property ‘standing’? And why is it that Alpha does not believe of herself *in the de se way* that she is standing if the object of her belief is the more complex property ‘being such that Alpha is standing’? There is of course the intuitive appeal of the idea that believing *I am standing* amounts to me simply being related to the property ‘standing’. But intuitions are not arguments, so the Property Theory needs additional support.

Presumably, the Property Theorist would want to pin down the reason in the *nature* of these specific kinds of properties which the subject is related to in believing. The general idea behind this rescue attempt is that some properties individuate on the level of individuals and others on the level of possible worlds. So, in some cases we have a property that picks out a set of possible worlds, and in other cases a property that picks out individuals and cuts through possible worlds. Simple properties such as ‘standing’ are of the latter kind. Since they pick out individuals, the subject is able to locate herself in logical space on the level of individuals. Accordingly, having such a property as the object of one’s belief results in a *de se* belief. More complex properties such as ‘being such that Alpha is standing’ are such that they pick out possible worlds. As such, the subject is only able to locate herself in logical space on the level of possible worlds. This is the problem that the two gods are facing. Hence, having such a property as the object of one’s belief is insufficient for making the belief *de se*. The resulting proposal is that subjects entertain genuine *de se* beliefs iff they are related to a property which individuates on the level of individuals.

So far, so good. Now, consider the difference between believing *Beta is happy* and believing *I am such that Beta is happy*. Admittedly, the latter is a strange belief to have, but it nonetheless seems to be a genuine *de se* belief. After all, it is typically expressed in language using the first-person pronoun and would be reported in indirect speech using the *quasi-indicator* ‘she★’ (Castañeda 1966).⁴ And it also seems reasonable to hold that the former is not a *de se* belief. If that is correct, then the semantics of the former ought to differ from the semantics of the latter, if we want to generally distinguish *de se* beliefs from *de dicto* beliefs on the basis of their content. However, the Property Theory

⁴ Bermúdez (1998: 3) writes: ‘This is the sort of self-reference whose natural linguistic expression is the first-person pronoun “I”, because one cannot use the first-person pronoun without knowing that one is thinking about oneself’. This indicates the general *de se* status of belief reports involving the first-person pronoun. Furthermore, the *de se* belief *I am such that Beta is happy* should be reported in indirect speech with the use of the quasi-indicator ‘she★’ while the non-*de se* belief *Beta is happy* can only be reported with the use of the regular pronoun ‘she’ such as in the sentence ‘Alpha believes that she is such that Beta is happy’. This is because she self-ascribes a property without having self-knowledge. This clarification was thankfully suggested by an anonymous reviewer.

tells us that the content of the *de se* belief *I am such that Beta is happy* is the property ‘being such that Beta is happy’—the same property which is the content of the *de dicto* belief *Beta is happy*. But then again we have two beliefs with the same property content that ought to be distinct. Therefore, the semantic aspect alone cannot account for the distinction between believing *de dicto* and believing *de se*. We need something more.⁵

At this stage then, the Property Theory is only marginally better off than the Doctrine. This is where the second element of Lewis’s solution kicks in—the *epistemic aspect*. Lewis argues that the relation between a subject and a content in believing *de se* is that of *self-ascription*. Accordingly, subjects self-ascribe properties when they entertain *de se* beliefs.⁶ First off, we need to distinguish self-ascription from the more general garden-variety way of ascribing something *de re* to an individual. When a subject believes *I am F*, she believes something of herself in the special characteristic first-person way of *de se* beliefs by self-ascribing ‘F’. On the other hand, when a subject believes *a is F*, she ascribes a property ‘F’ to an individual ‘a’. In the case of self-ascription we are dealing with a two-place relation between a subject and a property. In contrast, *de re* ascription is a three-place relation between a subject, an individual and a property. It is important to note that self-ascription cannot just be the special case of ascription where the individual ‘a’ is identical to the ascribing subject. Alpha can ascribe the property ‘standing’ *de re* to herself without *ipso facto* believing that she herself is standing—i.e. without self-ascribing the property ‘standing’.

Is there more that we can say about the concept of self-ascription? Surprisingly, Lewis has very few illuminating things to say about self-ascription. There are only a few passages where something like an explication is provided. Let us look at two of them:

We can and we do have beliefs whereby we locate ourselves in ordinary time and space; whereby we self-ascribe properties that don’t correspond to propositions; and *whereby we identify ourselves as members of subpopulations* whose boundaries don’t follow the borders of the worlds. (Lewis 1979: 519; italics mine)

Wishing that cyanoacrylate dissolved in acetone is *wishing to have the property* of inhabiting a world where it does. (Lewis 1979: 529; italics mine)

⁵ The Property Theorist might object that we can still distinguish the natures of these properties using my earlier proposal. In the *de se* case, the property ‘being such that Beta is happy’ individuates on the level of individuals whereas in the *de dicto* case it individuates on the level of possible worlds. However, such a move would severely beg the question for it remains mysterious why these two identical properties individuate in different ways in different contexts.

⁶ The label ‘epistemic’ points to the connection between the relation of self-ascription between a subject and a content and the possibility of self-knowledge. A subject can know that she herself is F (in the *de se* way) on the basis of her self-ascribing the property F. This epistemic aspect of Lewis’s theory has thus far been largely overlooked (some recent exceptions are Cappelen and Dever 2013; Holton 2015; Liao 2012).

As we can see, Lewis is extremely elusive on the details with regard to the work that self-ascription is supposed to do and how exactly we are supposed to analyse the concept of self-ascription. According to these passages, it either involves an identification as a member of some group (i.e. the one whose members have the property in question), or the belief of having a specific property, or taking oneself to have some property.⁷

The result from the argument above was that the semantic aspect *alone* is incapable of solving our problem, because merely standing in a relation to a property instead of a proposition in no way enables us to clearly distinguish between *de se* beliefs and non-*de se* beliefs. In believing *Alpha is standing* and *I am standing* Alpha might be related to a property in both cases, but that does not help us because being related to either of these properties does not account for the first-person way of thinking characteristic of *de se* beliefs. From this we concluded that it must be the concept of self-ascription that does the work.⁸ However, in order to understand the work that self-ascription is doing, we need to have a clear grasp of the concept in play. Unfortunately, Lewis himself does not illuminate the concept in any way, so we have to do it for him.

However, before discussing three different interpretations of self-ascription, it is important to note that Property Theorists generally take all beliefs to be ultimately *de se* in nature. So, when Alpha believes *I am standing*, she self-ascribes the property ‘standing’ which is tantamount to a *de se* belief. And when she believes *Alpha is standing*, she self-ascribes the property ‘being such that Alpha is standing’ which again is a *de se* belief since it consists in the self-ascription of a property. The motivation for this move is to obtain a uniform treatment of the belief relation and to avoid distinguishing between the relation of self-ascription in the case of *de se* beliefs and some other relation for other beliefs. The principal worry about this general reduction to the *de se* is that it fails to do justice to the characteristic epistemic features of (genuine) *de se* beliefs. For instance, in virtue of her *de se* belief *I am in pain*, Alpha presumably is in a position to know that she is in pain. This is not the case for *de dicto* beliefs. In virtue of her belief *Beta is happy*, Alpha is neither in a position to know that Beta is happy nor that she is such that Beta is happy—which is one and the same thing—despite her presumably self-ascribing the property ‘being such that Beta is happy’. We will see this problem resurface in the discussion of some of the interpretations of self-ascription below.

⁷ Feit (2008: 16–17) provides a restatement of these points without providing a further elucidating analysis: ‘I shall use the term “self-ascription” for the relation between subjects and the properties that they believe themselves to have. (...) The self-ascription relation is in an important sense necessarily reflexive. To self-ascribe a property is to ascribe it *to yourself* and not to any other thing. (...) So it is impossible for one person to [self-ascribe] a property to somebody else’. Also, there is a further passage at the end of Lewis’s paper which discusses the concept of self-ascription. I will discuss this passage in more detail in section 4.

⁸ Holton (2015: 400) provides a similar but less forceful argument. Cappelen and Dever (2013: Chapter 5) provide a different argument to the same conclusion based on other semantic considerations.

3 THE REGRESS PROBLEM

The first and seemingly obvious interpretation of the passages above is the following. A subject self-ascribes ‘F’ iff she identifies herself with a certain subpopulation of individuals which are F. Identifying oneself in this way amounts to believing something of oneself in the first-person way—i.e. believing to be a member of a group. If Alpha believes *I am standing* then she identifies herself with the group of standing things. And this again amounts to her believing *I am a member of the group of standing things*. It is rather obvious that such an interpretation of the concept of self-ascription is hopeless. For, if we interpret the relevant passages as involving such a new *de se* belief, then we are lead into an infinite regress:

- (1) Every *de se* belief of the form *I am F* involves self-ascription of some property ‘F’.
- (2) Self-ascription of ‘F’ amounts to identifying oneself as a member of a subpopulation of things which are F.
- (3) Identifying oneself as a member of a subpopulation of things which are F is a *de se* belief of the form *I am F**.
- (4) Hence, it involves self-ascription of ‘F*’.
- (5) ...

We can immediately see that the simple self-ascription of ‘F’ would require the subject to self-ascribe infinitely many properties of greater and greater complexity. This is an undesirable result. So, Lewis’s concept of self-ascription should not imply that an additional *de se* belief is required for the self-ascription to go through. If that were the case, we would simply shift our problem to the next level *ad infinitum*. Hence, the argument shows that our first interpretation of self-ascription fails and we need to reject premise (2) as the most straight-forward interpretation of what Lewis has in mind. Self-ascription does not require the subject to identify herself with a subpopulation by way of a *de se* belief of the form *I am a member of the group of F-things*.

4 THE IRREDUCIBILITY PROBLEM

The second interpretation follows a common understanding of Lewis’s position. The source of this interpretation is a small passage at the end of the paper where Lewis claims that self-ascription is ‘ascription of properties to oneself under the relation of identity’ (Lewis 1979: 543). What it means for a subject to self-ascribe a property is for her to ascribe a property *de re* to herself under the relation of identity. It seems that he is here invoking a kind of reduction of self-ascription to *de re* ascription. This sounds initially promising because we already have a grasp on what it means for a subject to ascribe a property *de re* to an individual and self-ascription would then just be a special form of ascription under certain conditions.

The condition that Lewis has in mind is that the ascription is made ‘under the relation of identity’. He thus proposes the two-part interpretation that self-ascription can be explained by invoking the notion of a *de re* belief about oneself together with the *relation of identity*. Now what does it mean for a subject to ascribe a property ‘under the relation of identity’? First off, we are faced with an obvious problem. We already saw in section 2 that the *mere identity* of ascribing subject and ascribed individual does not result in a self-ascription since the subject can ascribe ‘F’ *de re* to herself without *ipso facto* having a *de se* belief of the form *I am F* despite the fact that the ascribing subject and the *de re* ascribed individual are identical. So what else could Lewis have had in mind? We can use this first observation to distinguish between a weak and strong reading of the identity conditions:

Weak reading

Any instance of self-ascription implies that the ascribing individual is identical to the ascribed individual.

Strong reading

Self-ascription can be analysed in terms of ascription to oneself under certain conditions.

I already dismissed the weak reading which is echoed in Chisholm’s definition of *direct attribution*—his term for self-ascription: ‘For every x , every γ and every z , if x directly attributes z to γ , then x is identical with γ ’ (Chisholm 1981: 28). As such, it does not tell us anything new about self-ascription since we already knew that in the case of *de se* beliefs, the ascribing subject is necessarily identical with the ascribed individual. That is what it is to have a *de se* belief. So, the weak reading cannot provide us with anything new in order to explain self-ascription.

However, on the strong reading we are confronted with a kind of reduction of self-ascription to *de re* ascription which clearly goes beyond a statement of the necessary identity involved in the weak reading. Let me first explain why the strong reading actually fits in nicely with Lewis’s claim that the *de re* subsumes the *de se*, and that self-ascription is definable in terms of ascription of properties ‘under the relation of identity’. In the final paragraphs of ‘Attitudes *De Dicto* and *De Se*’, he explores the relation between the *de re* and the *de se* and argues:

So belief *de se* falls under belief *de re*. But there are other relations of acquaintance, besides identity, that a subject may bear to himself. So belief *de re* about oneself turns out to cover more than self-ascription of properties. To take an example due to David Kaplan, watching is a relation of acquaintance. I watch myself in reflecting glass, unaware that I am watching myself. I ascribe to myself, under the description ‘the one I am watching,’ the property of wearing pants that are on fire. I therefore believe *de re* of the one I am watching—that is, myself—that his pants are on fire. But I do not self-

ascribe the property of wearing pants that are on fire. Very soon I will, but not yet. So self-ascription isn't quite the same thing as ascription, *de re*, to oneself. (Lewis 1979: 543)

It is clear that Lewis is after a reduction of the form: 'A subject S self-ascribes a property F iff S ascribes F *de re* to S under conditions C'. However, it is still an open question which conditions C are needed to complete the reduction. On the basis of the quote we can distinguish at least two main readings of the conditions that need to be examined:

Strong+ reading

Self-ascription requires ascription to oneself paired with the *assumption of* identity between ascribing and ascribed individual.

Strong- reading

Self-ascription requires ascription to oneself paired with the *de facto* identity of ascribing and ascribed individual.

The problem with the strong+ reading is that it is epistemically unsound. It should not be necessary for the believing subject to think something along the lines of *I ascribe the property F to the individual a under the assumption that a is identical to me*. In such a belief the subject would have to identify an individual and attribute identity to that individual. So, if self-ascription involved the *assumption* of identity, we would have a belief which involved the identification of an individual. However, such an account of self-ascription would clash with the mere *possibility of de se* beliefs that exhibit 'immunity to error through misidentification' (Evans 1982; Shoemaker 1968).

There is a considerable amount of debate about the nature and pervasiveness of immunity (cf. Prosser and Récanati 2012). However, despite the differing opinions, any suitable explanation of *de se* beliefs must be able to account at least for the *possibility* of their immunity—even if one disagrees about the pervasiveness of the phenomenon. A plausible reason why some beliefs are immune to error through misidentification is that they do not involve an identification at all. Alpha cannot misidentify the individual she is thinking about—i.e. herself—in believing *I am F* because she does not identify an individual at all. In contrast, beliefs which are subject to identification failure are not identification-free. Alpha might misidentify Beta in her belief *Beta is F* since who she takes to be Beta is in fact someone else.

However, the strong+ reading epistemically involves such an identification because the subject has to assume that whoever she is ascribing F to is identical to herself. But if that were correct, she could fail to identify the right individual and thus her *de se* belief would not be immune to misidentification. Her epistemic basis involves some stance on the identity of the individual. The following argument makes this reasoning explicit:

- (1) *De se* beliefs might exhibit immunity.

- (2) If a belief exhibits immunity, it does not involve an identification element.
- (3) *De se* beliefs involve self-ascription of properties.
- (4) Self-ascription of properties is *de re* ascription of properties to oneself under the assumption of identity between ascribing and ascribed individual.
- (5) Ascription of properties to oneself under the assumption of identity involves an identification element.
- (6) Therefore, *de se* beliefs involve an identification element.
- (7) Therefore, *de se* beliefs cannot exhibit immunity.

The strong+ reading should be distinguished from the more subtle and plausible reading according to which the epistemic basis of a self-ascription is such that no question of identity arises. For instance, in introspection, the subject has to take the property to be instantiated by herself—for otherwise it would not be a case of introspection. But such a judgement does not involve an assumption of identity since the subject does not take an explicit stance on the identity of the individual.⁹ We can thus conclude that the strong+ reading is too strong and thus out of the question.¹⁰

Does the strong– reading fare any better? Sadly, our problem cannot be solved by assuming that *de se* ascription is a specific form of *de re* ascription paired with the *de facto* identity of ascribing and ascribed individual. As we already saw, the Property Theory entails that *any* subject having *any* kind of belief already entertains a *de se* belief and must, therefore, ascribe a property to oneself ‘under the relation of identity’. The fact that a subject ascribes a specific property to herself under the relation of identity thus cannot explain why some beliefs are *de se* and others are not.

This echoes the general worry that the reduction of all beliefs to the *de se* threatens to undermine the special epistemic qualities of *de se* beliefs. As we saw, there is a crucial difference between Alpha believing *I am F* and her believing *Alpha is F*. If every kind of belief is ultimately *de se*, this difference remains mysterious. The original suggestion that self-ascription can distinguish between these fundamentally different kinds of belief would collapse on this picture because everything would be a self-ascription of a

⁹ Such an account of self-ascription is the basis of the primitive account hinted at in section 5 and should be seen as a natural advancement of the Lewisian theory. An anonymous reviewer has helped me to formulate this in clearer terms.

¹⁰ In fact, Holton (2015: 406) seems to take Lewis’s account to *imply* such an interpretation: ‘If this account is to work, every case of *de re* belief must be a case in which the subject self-ascribes a particular unique relation to the thing in question, that is, to the *res*. But why should we think that that is so? Could not a subject have a belief about the *res* without having any belief about the relation they bear to it?’ The question is, of course, rhetorical. There is no reason to think that a *de se* belief requires the subject to hold a belief (of any kind) concerning the kind of relation it stands in with the object of her belief. Worse still, such a consequence would render the explanation of self-ascription circular again since it would involve a new *de se* belief on the part of the subject concerning the kind of acquaintance relation she herself stands in.

property. Hence, the nature of genuine *de se* beliefs cannot be accounted for by the strong– reading as the following argument shows:

- (1) All beliefs are *de se* beliefs.
- (2) *De se* beliefs are self-ascriptions of properties.
- (3) Self-ascriptions of properties are ascriptions of properties to oneself paired with the *de facto* identity of ascribing and ascribed individual.
- (4) Hence, all beliefs involve ascription of properties to oneself paired with the *de facto* identity of ascribing and ascribed individual.
- (5) Hence, *de facto* standing in a relation of identity to oneself cannot account for the difference between *de re* ascription and self-ascription.

There is an additional oddity in the strong– reading. Let us suppose that self-ascription can be accounted for by separating the *de re* ascription to oneself on the one hand from the acquaintance relation of identity to oneself on the other. If we think about the epistemic nature of genuine *de se* beliefs, it is puzzling how this two–fold proposal is supposed to help. The first-person way of thinking that is characteristic of genuine *de se* beliefs is not exhausted nor characterised by the particular relation a subject stands in to herself—without *assuming* to stand in such a relation; for otherwise we would be having circularity troubles again. The identity relation belongs to the basis of *de se* beliefs and should not be assumed by the subject.¹¹ This defect troubles both versions of the strong reading for they are both reductive in nature. A reduction of the type ‘A subject S self-ascribes a property F iff S ascribes F *de re* to S under conditions C’ is therefore hopeless.¹²

Let us sum up the discussion of the second interpretation of Lewis’s concept of self-ascription. The idea was that self-ascription can be analysed in terms of *de re* ascription to oneself ‘under the relation of identity’. There are two ways to initially interpret this proposal. On the weak reading, it merely spells out a logical consequence of self-ascription without providing an analysis in some other terms. We saw that this reading brings nothing new to the table. On the strong reading, we should analyse self-ascription in terms of *de re* ascription to oneself under some conditions. I distinguished

¹¹ This observation is similar to one by John Campbell in his *Past, Space, and Self* (1994: 84). Reasoning from ‘I am F’ and ‘I am G’ to ‘I am both F and G’ is possible without judging that the ‘I’ in the first judgment and the ‘I’ in the second judgement are identical. In Campbell’s words, our reasoning ‘trades on identity’. If such an identity judgement were necessary, we would be lead into an infinite regress.

¹² Both the regress problem and the irreducibility problem are somewhat related to a recent argument by Simon Prosser (2015) which he also calls a regress argument. Prosser argues that proper *de se* beliefs ultimately require ‘first-person redundant representations’—a special first-personal way of representing (say) a proposition. Any purely *de re* ascription of a property or a relation to a subject would fail to account for characteristic features of *de se* beliefs if they were not based on a first-person redundant representation. Interestingly, he also notes the similarities of these arguments to arguments concerning a looming infinite regress in the first person by Ryle (1949: 186–189) and Merleau-Ponty (1962: 104–105).

two different versions of this strong reading. On the strong+ reading, self-ascription involves the epistemic assumption on the part of the subject of an identity relation between the ascribing subject and the ascribed individual. But, as my argument showed, such a reading is impossible because it is epistemically unsound. On the strong- reading, we distinguish *de se* beliefs from others by invoking the acquaintance relation of identity that we *de facto* stand in. So, on the strong- reading, self-ascription could be reduced to the concept of *de re* ascription. However, my arguments showed that such an analysis fails as well, since it is incapable of solving our problem. If all beliefs are *de se* beliefs—as the Property Theory takes them to be—then we always stand in the relation of identity to ourselves in entertaining beliefs. Therefore, that relation in itself cannot account for the difference between *de re* believing of oneself to be happy and *de se* believing of oneself to be happy. Hence, none of the proposed readings so far can account for our problem.

5 THE PRIMITIVENESS OF SELF-AScription

There is a final interpretation of Lewis’s concept of self-ascription available. The lessons we learned from the other two interpretations are that self-ascription needs to be capable of drawing the distinction between genuine *de se* beliefs on the one hand and non-*de se* beliefs on the other. Furthermore, a reduction of self-ascription to *de re* ascription is not possible. Instead, the new suggestion is that self-ascription is primitive and constitutes the first-person way of believing something of oneself in the *de se* way. Recently, Shen-yi Liao has argued that a primitive interpretation of self-ascription is the only one that solves our problem and thus should most plausibly be attributed to Lewis:

On this account, the identification of centres [of possible worlds] and the individuation of possible individuals are *primitive*, unable to be elucidated through other non-trivial features. (...) [T]his mystery is to be expected given the main lesson from the problem of essential indexicals: the *de se* cannot be reduced to the *de dicto*. There is something special about learning who *oneself* is that cannot be captured in learning about what features one possesses, even if that list of features is exhaustive. (Liao 2012: 313–314)

The suggestion of Liao is that any reductive account of self-ascription requires a non-primitive individuation of a given individual which runs us into the problems we encountered above. The only alternative is that self-ascription consists in a primitive individuation of a possible individual. Such an individuation is distinct from an identification of an individual in that it does not require a certain set of conditions to be fulfilled and does not import the possibility for misidentification.

There is additional support for the primitive interpretation in a passage by Chisholm (1981: 29) where he discusses the notion of direct attribution—his equivalent for self-ascription. He suggests that the most basic relation that subjects can stand in with re-

gard to the contents of their attitudes is that of ‘*considering* oneself as having a certain property’. This is the rock bottom relation that is required to have any beliefs at all, and we may call that most basic relation ‘direct attribution’ or ‘self-ascription’.

So, the primitive interpretation seems to give us the following: whenever we have a *de se* belief, we consider ourselves as having some property or other. However, we simply cannot analyse the notion ‘considering ourselves as having’ in more fundamental terms without giving a circular explanation (cf. Frank 1986; Lailach-Hennrich 2011). This can be illustrated using our second interpretation above. The suggestion was that self-ascription is to be analysed as ascription *de re* to oneself under conditions C. Accordingly, a subject self-ascribes ‘F’ iff she ascribes ‘F’ *de re* to herself under conditions C. But we can then ask: How does the subject know that conditions C are fulfilled other than by way of a further self-ascription? If self-ascription is to play the role of providing an explanation of the nature of the *de se*, it better not require *de se* beliefs in the first place. Hence, any attempt to reductively explain self-ascription is bound to fail.

Furthermore, the primitive interpretation chimes in well with the arguments already presented for the necessity of self-ascription for a solution of our problem. And Holton (2015: 403) provides us with a further argument that is supposed to show that only *primitive* self-ascription—and not some reductive variant—can provide a solution and that we cannot reduce it to or explain it in terms of some other concept:

We normally think of ascription as a two-place relation: one ascribes a property to a thing. Self-ascription would then be the special case where the thing is the self. But that won’t do the work here. If the self is just thought of extensionally, then we would have no way to distinguish the belief that one’s pants were on fire from the belief that the pants of someone, who is you though you don’t realize it, are on fire. (...) So we have to think of self-ascription as a one-place relation: one simply self-ascribes a property.¹³

Holton’s point echoes the above remarks. *De se* beliefs require some primitive non-identificational relation to the self-ascribed properties. Otherwise, we can always ask the question: How does the subject know that *she herself* is F? Even if the relevant property is as detailed as you want, there is always the possibility of believing of someone to have that property without *de se* believing oneself to have it.

It is important to note that the irreducibility of self-ascription does not make the concept mysterious. It is perfectly possible to give a non-reductive *explanation* of self-ascription in terms of other concepts. There are two important desiderata for such an account of primitive self-ascription. On the one hand, it has to do justice to the fact that *de se* beliefs involve some primitive individuation of an individual to which the property

¹³ In section 2 I characterised self-ascription as a two-place relation and *de re* ascription as a three-place relation, whereas Holton thinks that self-ascription is a one-place and ascription a two-place relation. This is only a terminological difference and has no substantial consequences.

is self-ascribed. On the other hand, this kind of individuation has to be essentially first-personal in order to result in a *de se* belief.

How could such a non-reductive account look like? A mere sketch will have to suffice for the moment. We could understand subjects as engaging with the world through their lived body (cf. Husserl 1973; Merleau-Ponty 1962). This would involve some very primitive grasp on the part of the subject concerning the question: Which parts of the body are under one's immediate control and are means of direct interaction with the world? A primitive self-ascription could then be understood as an ascription to the lived body. This could satisfy the two desiderata in the following way. Since the lived body is simply the body 'we find ourselves in', it is not necessary to identify oneself with a certain individual. And our way of engagement with the world through the lived body gives that phenomenological body a privileged first-personal epistemic position. Of course, this rough sketch would have to be developed in more detail in order to do full justice to the phenomenon of *de se* thinking.

6 CONCLUSION

The concept of self-ascription plays a crucial role for understanding the nature of *de se* beliefs. Originally, the term was introduced into the debate by Lewis in his elaboration of the Property Theory of content. I argued that the mere semantic move from propositions to properties fails to solve the problem of *de se* beliefs. While properties are more fine-grained than propositions, it is still possible to have distinct beliefs with identical contents. Thus, the solution has to lie in the special epistemic relation of self-ascription. A closer inspection of the concept of self-ascription then uncovered several problems. If taken at face value, the explanation of self-ascription ends with an infinite regress. The more common two-fold interpretation according to which we need to distinguish the *de re* ascription of the property to oneself from the acquaintance relation of identity under which that ascription happens also failed. Depending on the exact reading of this two-fold approach it is either epistemically unsound or just as incapable of drawing the distinction between *de se* and non-*de se* beliefs as a purely semantic approach. I then argued that we have to understand self-ascription as a primitive epistemic relation. This result—and the arguments provided in favour of it—is not only relevant to proponents or friends of the Property Theory. The necessity of a primitive epistemic relation between subjects and the objects of their *de se* beliefs remains, even if we (say) advocate a propositional theory of content.¹⁴

¹⁴ This paper has benefitted greatly from comments by Christopher Peacocke, Gianfranco Soldati, Johannes Brandl, Pascale Anna Lötscher, Patrik Engisch, Davor Bodrozic, and an anonymous reviewer, for which I express my gratitude. It was made possible by the Swiss National Science Foundation grant 100012_156548.

REFERENCES

- Bermúdez, José Luis (1998). *The Paradox of Self-Consciousness*. MIT Press.
- Campbell, John (1994). *Past, Space, and Self*. MIT Press.
- Cappelen, Herman and Josh Dever (2013). *The Inessential Indexical*. Context and Content. Oxford University Press. DOI: 10.1093/acprof:oso/9780199686742.001.0001.
- Castañeda, Hector-Neri (1966). ‘He’: A Study in the Logic of Self-Consciousness’. *Ratio*. 8. Reprinted in Castañeda (1999b): 35–60, pp. 130–157.
- (1999a). ‘Self-Consciousness, Demonstrative Reference, and the Self-Ascription View of Believing’. In Castañeda (1999b), pp. 143–179.
- (1999b). *The Phenomeno-Logic of the I*. Ed. by James G. Hart and Tomis Kapitan. Indiana University Press.
- Chisholm, Roderick M. (1981). *The First Person: An Essay on Reference and Intentionality*. University of Minnesota Press.
- Evans, Gareth (1982). *The Varieties of Reference*. Oxford University Press.
- Feit, Neil (2008). *Belief About the Self: A Defense of the Property Theory of Content*. Oxford University Press. DOI: 10.1093/acprof:oso/9780195341362.001.0001.
- (2012). ‘Self-Ascription and Self-Awareness’. In *Consciousness and Subjectivity*. Ed. by Sofia Miguens and Gerhard Preyer. Ontos, pp. 213–230.
- Frank, Manfred (1986). *Die Unhintergebarkeit von Individualität*. Suhrkamp.
- Holton, Richard (2015). ‘Primitive Self-Ascription: Lewis on the *De Se*’. In *A Companion to David Lewis*. Ed. by Barry Loewer and Jonathan Schaffer. Wiley-Blackwell, pp. 399–410.
- Husserl, Edmund (1973). *Zur Phänomenologie der Intersubjektivität. Erster Teil*. Ed. by Iso Kern. Martinus Nijhoff.
- Kaplan, David (1989). ‘Demonstratives’. In *Themes From Kaplan*. Ed. by Joseph Almog, John Perry, and Howard Wettstein. Oxford University Press, pp. 481–563.
- Lailach-Hennrich, Andrea (2011). *Ich und die anderen. Zu den intersubjektiven Bedingungen von Selbstbewusstsein*. De Gruyter.
- Lewis, David (1979). ‘Attitudes *De Dicto* and *De Se*’. *Philosophical Review*. 88, 4. Reprinted in Lewis (1983): 133–156, pp. 513–543. DOI: 10.2307/2184843.
- (1983). *Philosophical Papers*. Oxford University Press. Vol. 1.
- (1986). *On the Plurality of Worlds*. Blackwell Publishers.
- Liao, Shen-yi (2012). ‘What are centred worlds?’ *The Philosophical Quarterly*. 62, 247, pp. 294–316. DOI: 10.1111/j.1467-9213.2011.00042.x.
- Merleau-Ponty, Maurice (1962). *Phenomenology of Perception*. Trans. by Colin Smith. Routledge.
- Perry, John (1979). ‘The Problem of the Essential Indexical’. *Noûs*. 13, pp. 3–21. DOI: 10.2307/2214792.

- Prosser, Simon (2015). 'Why Are Indexicals Essential?' *Proceedings of the Aristotelian Society*. 115, pp. 211–233. DOI: 10.1111/j.1467-9264.2015.00392.x.
- Prosser, Simon and François Récanati (eds.) (2012). *Immunity to Error through Misidentification: New Essays*. Cambridge University Press. DOI: 10.1017/CB09781139043274.
- Ryle, Gilbert (1949). *The Concept of Mind*. London: Hutchinson.
- Shoemaker, Sydney (1968). 'Self-Reference and Self-Awareness'. *The Journal of Philosophy*. 65, 19, pp. 555–567. DOI: 10.2307/2024121.