Imagination, Inside and Out

Dilip Ninan[*]

Abstract

This paper presents an account of the distinction between imagining 'from the inside' (imagining *doing* or *experiencing* something) and imagining 'from the outside' (imagining *that* one does or experiences something). I first show that an independently motivated semantic account of imagination reports supports the view that the inside-outside distinction just is the *de se* - *de dicto* distinction, where the latter distinction is understood along the lines of Lewis (1979). I then consider an objection to this view, along with an alternative to it. I argue that the original view can evade the objection and is preferable to the alternative. I close with a brief discussion of imagination and possibility.

1 Introduction

There are two ways that an imaginative episode can be about oneself. The distinction between imagining 'from the inside' and imagining oneself 'from the outside' can be introduced by considering the following imagination reports (both inspired by Lakoff (1972)):

(1) I imagined that I was Brigitte Bardot and that I kissed me.

(2) I imagined that I was Jesus and that I forgave me all my sins.[1]

In (1), one imagines being Bardot from the inside, and one imagines oneself as the 'kissee' from the outside. The examples show that a single imaginative episode can be both from the inside and from the outside. Thus, the inside-outside distinction does not partition the space of imaginings; rather, it is best thought of as a distinction between two ways one can represent oneself in an imagining. By "imagining from the inside", I shall simply mean however it is one imagines being Bardot in (1); by "imagining from the outside", I shall simply mean however it is one imagines the person kissed in (1).[2]

---

[*]dninan@mit.edu

[1]The original sentence in Lakoff (1972) is *I dreamt that I was Brigitte Bardot and that I kissed myself.* The 'dream' version of (2) appears in Heim (1994).

[2]The inside-outside distinction appears, in different guises, in many places in the literature. See, e.g, Williams (1966), Nagel (1974, n.11), Peacocke (1985), Walton (1990, 28-35), Shoemaker (1994), Velleman (1996), Hill (1997), Martin (2002) and Higginbotham (2003).

A more familiar way of drawing the inside-outside distinction uses the characteristic imagery associated with each type of imagining. Consider the difference between imagining skiing down a steep hill and imagining *that* you are skiing down the same hill. In one case, you imagine what it would be like to actually be there, on the hill. The imaginative scenario is viewed through one's imaginary ski goggles, and includes a view of the tips of one's skis and the white snow before oneself; it may be accompanied by imagining being cold and full of adrenaline. In the other case, you take a god's-eye view on the scenario and 'see' yourself zooming down the hill, as if you were watching a film of yourself skiing.

Thinking about imagery makes the distinction vivid, but it is potentially misleading. I can imagine from the inside being Napoleon, sitting in a pitch-black and completely quiet room by visualizing a featureless black expanse; I can also imagine that the universe contains nothing by visualizing a featureless black expanse. The former imagining is from the inside, the latter is not. I can also imagine the visual experience of someone skiing down a hill, without imagining actually being the one doing the skiing, controlling the movements of the skier's body. I can imagine the *view from* the driver's seat without imagining *being in* the driver's seat.

There is a larger question looming in the background: what is the relationship between imagery and imagination? We won't have much to say about this, but the previous paragraph indicates some of the subtleties involved. Instead, let us stick with our characterization of the inside-outside distinction by thinking about the two ways in which one represents oneself in imaginings like the ones reported in (1) and (2).

The primary aim of this paper is to argue that imagining from the inside is imagining a *centered possibility*, in the sense of Lewis (1979). There are several reasons for being interested in the inside-outside distinction and in the claim that imagining from the inside is imagining a centered possibility. I will briefly mention two.

First, the inside-outside imagination distinction has played an important role in various philosophical arguments. Foremost in my mind is the role the distinction has played in discussions of personal identity. For example, in his paper "The Self and the Future" (Williams 1970), Williams discusses a now-famous puzzle about personal identity, a puzzle which (one might argue) is generated by first imagining a hypothetical case from the outside and reaching one conclusion about it, and then imagining it from the inside and reaching another, apparently incompatible, conclusion about it. One might hope that a better understanding of the inside-outside distinction would throw light on Williams's puzzling case.[3]

---

[3]The inside-outside imagination distinction has also played a role in arguments in the philosophy of mind, and so a better understanding of the distinction may help us evaluate these arguments. I have in mind, in particular, the method of explaining away the apparent contingency of the mind-body relationship that is urged in Nagel (1974, n.11) and Hill (1997); and the argument in Martin (2002) for the claim that perceptual experiences lack representational content.

Second, in light of the relationship between imaginability and possibility, the fact that one can imagine from the inside being someone else may support Lewis's contention that one person could have been another (Lewis 1983a, 1986). If the content of one's imagining from the inside that one is Fred is the set of *centered worlds* in which the center is Fred, then this account of imagining from the inside dovetails nicely with Lewis's analysis of modality in terms of quantification over centered worlds.[4] (I will return to these matters in §4.)

The paper is organized as follows. I first introduce the *de se - de dicto* distinction and Lewis's (1979) account of it (§2.1), and then discuss the possibility that the inside-outside distinction just is the *de se - de dicto* distinction. My primary argument for this view (which I call *the Simple View*) is that it is supported by an independently motivated semantic theory (§2.2). I then consider an objection to the Simple View, and offer a reply to the objection on behalf of the Simple View (§3.1). I then consider a possible alternative to the Simple View, but I argue that the Simple View is preferable to the alternative (§3.2). I conclude in §4 with a discussion of the relationship between imaginability from the inside and centered possibility.

2 The Simple View

2.1 Lewis on *de se* attitudes

According to what I shall dub *the Simple View*, the inside-outside distinction just is the *de se - de dicto* distinction. Thus, before we can formulate and assess the Simple View, we need to introduce the *de se* phenomenon and the theory of *de se* attitudes and attitude reports that we're going to adopt here.

The phenomenon of *de se* belief is best introduced by way of an example:

> John is looking at himself in the mirror, but he doesn't realize that it is he who he sees. The arm of the man in the mirror is bleeding, i.e. John's arm is bleeding. John thinks to himself, *That guy is bleeding*, and also thinks to himself, *I'm glad I'm not bleeding*.

Examples of this sort are easily multiplied.[5] In the above example, John is in a belief state he would express by saying, *That guy is bleeding*; he is not in a belief state he would express by saying, *I am bleeding*. Following Lewis (1979), we will call the former type of belief a *de dicto* belief, and the latter type a *de se* belief. How should we account for this distinction?

According to the modal theory of attitudes, the content of an attitude like belief is a set of possible worlds.[6] What does that mean? What is it to believe according to this account, e.g., that Sam is a priest? In thinking about this question, it is useful to think about what it is for a possibility to be compatible with a belief. If I believe that Sam is a priest, then my belief is incompatible with

---

[4]The centered worlds machinery will shortly be introduced in more detail.

[5]See Lewis (1979) and Perry (1977, 1979) for more examples.

[6]This theory was introduced in Hintikka (1962).

possible situations in which Sam does not exist. But my belief is compatible with possible situations in which Sam lives in Vermont, since priests can live in Vermont.

Think of all the worlds compatible with everything I believe—my 'belief worlds'. This is the set of worlds which I haven't ruled out as failing to obtain; as far as I'm concerned, I might be in one of these worlds. According to the modal theory of attitudes, to believe that Sam is a priest is for it to be true in all my belief worlds that Sam is a priest. Belief is necessity across doxastic alternatives. If none of my belief worlds are worlds in which Sam is a minor, then I believe that Sam is not a minor. If Sam lives in Vermont in some of my belief worlds but not in others, then I neither believe that Sam lives in Vermont nor do I believe that he doesn't; both those possibilities are compatible with what I believe. A proposition (i.e. a set of worlds) is the content of one my beliefs just in case the set of my belief worlds is included in that proposition.

Thus, according to the modal theory of belief, when John thinks, *That guy is bleeding*, the proposition that that guy is bleeding is true in all of John's belief worlds.[7] Since that guy is John, the proposition that John is bleeding is true in all of John's belief worlds. But since John also thinks *I'm glad I'm not bleeding*, it would also seem that the proposition that John is bleeding is false in all of John's belief worlds. But which is it: is that proposition true in all of John's belief worlds or isn't it? The modal theory seems unable to distinguish beliefs that need to be distinguished.

Lewis (1979) shows how to generalize the modal theory of attitudes to account for the *de se - de dicto* distinction. On (a terminological variant of) Lewis's picture, we switch from thinking of belief as necessity over those possible worlds compatible with what one believes to thinking of it as necessity over the *centered worlds* compatible with what one believes. A centered world is thought of as a possible world with a designated individual (*the center*).[8] A set of centered worlds is a *centered proposition*. A centered world is compatible with what I believe just in case I think I might be the center in that world. My centered belief worlds, then, are the set of world-individual $\langle w, x \rangle$ pairs such that, as far as I can tell, I might be x in w.

How does Lewis's theory deal with the mirror example? John doesn't take himself to be bleeding; since he thinks to himself, *I'm glad I'm not bleeding*, it is false in all his centered belief worlds that the center inhabits a world in which the center is bleeding. On the other hand, since he does think, *That guy is bleeding*, it is true in all his centered belief worlds that the center inhabits a world in which that guy (i.e. John) is bleeding. So John believes the centered proposition $\{\langle w, x \rangle$: John is bleeding in $w\}$, but doesn't believe the centered proposition $\{\langle w, x \rangle : x$ is bleeding in $w\}$.[9] That is, John's centered belief worlds

---

[7] A proposition Γ is true at a world w just in case $w \in \Gamma$.

[8] A centered world is really a pair of a world and a time-individual pair, where the latter object is referred to as *the center*. But we can ignore this extra complexity since we will not be talking about the temporal properties of attitudes. The centered worlds terminology is due to Quine (1969).

[9] In this notation, w a is variable ranging over worlds, and x a variable ranging over indi-

are contained in the first, but not in the second, set. This gives us the result we want: John believes that the person who is in fact him is bleeding, but John doesn't believe that *he* is bleeding.

Following Egan (Forthcoming), we can say that the second centered proposition – the one John does believe – is *boring*, since for all worlds w and individuals x, y in w, $\langle w, x \rangle$ is in that centered proposition just in case $\langle w, y \rangle$ is. A boring centered proposition doesn't distinguish between world-mates; it makes a demand on the world (unless it is the necessary, or the empty, centered proposition), but not on the center. A boring centered proposition is a *de dicto* content, and corresponds to exactly one ordinary proposition (set of worlds). A *non-boring* centered proposition is one that isn't boring. It is a *de se* content, and it doesn't correspond to any ordinary proposition.

2.2 The semantics of imagination reports

In light of this theory of *de se* attitudes, we can now formulate the Simple View as follows:

> Simple View
>
> Inside
> S imagines from the inside in w that she is F just in case in all the $\langle w', x' \rangle$ compatible with what S imagines at w, x' is F in w'.
>
> Outside
> S imagines from the outside in w that she is F just in case in all the $\langle w', x' \rangle$ compatible with what S imagines at w, S is F in w'.

In the simplest cases, the Simple View entails that an imagining from the inside has a non-boring centered proposition as its content and that an imagining from the outside has a boring centered proposition as its content. But since an imagining can be both from the inside and from the outside (recall (1) and (2)), this characterization will only be accurate for the simplest cases.

Why think the Simple View true? The Simple View has an initial plausibility: imagining from the outside is a way of regarding oneself third-personally, whereas imagining from the inside involves taking on the first-person perspective. So it is natural to think of the inside-outside imagining distinction as a difference between *de se* and *de dicto* contents. But a more robust motivation for the Simple View comes from the semantics of imagination reports.

Often when philosophers write about the distinction between imagining from the inside and imagining from the outside, they emphasize that it is the distinction between imagining *doing* or *experiencing* something and imagining *that* one is doing or experiencing something (where the former is intended to capture imagining from the inside). For example, consider this passage from Walton (1990, 29):

---

viduals in w.

5

> Imagining from the inside is... a form of self-imagining characteristically described as imagining *doing* or *experiencing* something (or *being* a certain way), as opposed to merely imagining *that* one does or experiences something or possesses a certain property.

This suggests that the standard way to report an episode of imagining from the inside is to use sentences like (3) and (4):

(3) I imagined skiing down the hill.

(4) Bernie imagined catching a touchdown pass.

The content of Bernie's imaginative episode reported in (4) is whatever is expressed by the gerundive complement *catching a touchdown pass.*

I said sentences like (4) are the *standard* way to report episodes of imagining from the inside. Why this qualification? Because it seems to me that such imaginings can also be reported by sentences in which *imagines* takes a *that*-clause. That is, the imaginings reported in (3) and (4), could also be reported by using (5) and (6), respectively:

(5) I imagined that I was skiing down the hill.

(6) Bernie imagined that he was catching a touchdown pass.

It seems to me that sentences like (5) and (6) admit both inside and outside readings.[10] In contrast, if (3) and (4) do not unambiguously report imaginings from the inside, then the inside reading is, at the very least, highly preferred. I will proceed to theorize as if (3) and (4) unambiguously report imaginings from the inside.[11]

We begin by giving an analysis of (3) and (4). According to a well-known theory, the subject of (3) and (4)'s lower clauses is *PRO*, a phonologically null pronoun that is *controlled* by the subject of the higher clause.[12] So the real structure of (3) is: *I imagined PRO skiing down the hill.*

An observation due to Morgan (1970) is that subject-controlled *PRO* gives rise to unambiguously *de se* readings. To see this, consider the following variation on our original mirror case:[13]

> Situation: John is looking at himself in the mirror, but he doesn't realize that it is he who he sees. The arm of the man in the mirror is bleeding, i.e. John's arm is bleeding. Knowing that those around him are a helpful sort, John thinks to himself, *That guy will get help.*

---

[10] In saying this, I do not want to commit myself to the claim that (5) and (6) are ambiguous; the two readings may reflect the fact that they are sensitive to context.

[11] This assumption is in keeping with the limited literature on the topic. See Higginbotham (2003); Lakoff (1972, 640 - 41) is also relevant.

[12] Not every gerundive phrase has *PRO* as subject of course; but standard diagnostics (e.g. idiom and dummy-*it* tests) reveal that *imagines* is a subject-control, rather than a raising, verb.

[13] This is a modified version of a case in von Fintel (2005).

> But not believing that his own arm is bleeding, John doesn't believe
> that he himself will get help, i.e. he doesn't think to himself, *I will
> get help.*

> (7) John expects that he will get help. (true under one interpreta-
> tion)

> (8) John expects (PRO) to get help. (false)

In the above situation, John's thought is not an *I*-thought: it is a thought about
someone who he takes to be someone else, but who is in fact himself. John has
a *de dicto* expectation, but lacks the corresponding *de se* expectation, and this
is why (7) is a true description of the situation (at least on one of its readings),
and (8) a false description.

Why does subject-controlled *PRO* give rise to unambiguously *de se* readings?
The analysis of attitude verbs and *PRO* that we adopt is due to Anand & Nevins
(2004) and von Fintel (2005).[14] (This theory is couched in the more general
framework for intensional semantics found in von Fintel & Heim (2004)). In
this framework, semantic values are given relative to a context and an index,
and we take both the context and the index to be world-individual pairs, i.e.
as (what we are calling) centered worlds.[15] Putting a person coordinate in the
index is what allows the semantics to represent *de se* content.

The semantic value of (subject-controlled) *PRO* relative to a context c ($=
\langle w_c, x_c \rangle$, x_c the speaker of c) and an index i ($= \langle w_i, x_i \rangle$) in this framework is
the person coordinate of the index:

$$[\![\text{PRO}]\!]^{\langle w_c, x_c \rangle, \langle w_i, x_i \rangle} = x_i{}^{16}$$

This gives us the following semantic value (intension) for *PRO to get help*:

$$\lambda \langle w, x \rangle. [\![\text{PRO to get help}]\!]^{\langle w_c, x_c \rangle, \langle w, x \rangle} = \lambda \langle w, x \rangle.\ x \text{ gets help in } w.$$

That phrase expresses a non-boring centered proposition, which is a *de se* con-
tent.[17] So (8) is true just in case John stands in the relation of expectation to
that non-boring centered proposition.

We assume the attitude verb *expect* shifts the index, so that we get the
following truth condition for (8):

---

[14] Chierchia (1989) offers a different way of connecting the interpretation of these gerunds
and infinitives to Lewis's theory of *de se* attitudes. My own argument depends less on the
specifics of the compositional semantics and more on the following claims: (i) that the gerun-
dive complement of *imagines* in a sentence like (3) is headed by subject-controlled *PRO*; (ii)
that in such sentences, *PRO* is correlated with unambiguously *de se* readings; and (iii) that
observations (i) and (ii) can be implemented in a way consistent with Lewis's theory.

[15] Normally, both the context and the index would also include a time parameter. But since
we are ignoring the temporal properties of these attitudes, we can ignore the time parameter.

[16] The semantic interpretation brackets "$[\![\ ]\!]$" denote a three-place function that takes
expression-context-index triples into semantic values.

[17] Centered propositions can be thought of either as sets of centered worlds or as character-
istic functions thereof; the two formulations are essentially equivalent and I will use both.

$[\![$John expects PRO to get help$]\!]^{\langle w_c,x_c\rangle,\langle w_i,x_i\rangle} = 1$ iff

$[\![$John expects$]\!]^{c,i}(\lambda\langle w,x\rangle.[\![$PRO to get help$]\!]^{\langle w_c,x_c\rangle,\langle w,x\rangle}) = 1$ iff

All the $\langle w',x'\rangle$ compatible with what John expects in w_i are such that $[\![$PRO to get help$]\!]^{\langle w_c,x_c\rangle,\langle w',x'\rangle} = 1$ iff

All the $\langle w',x'\rangle$ compatible with what John expects in w_i are such that x' gets help in w'.

But (8) comes out false in the relevant scenario, because it is not true in all the $\langle w',x'\rangle$ compatible with what John expects that x' gets help in w'.

The complement of *expects* on the intended reading of (7), however, is a boring centered proposition:

$\lambda\langle w,x\rangle.[\![$that he will get help$]\!]^{\langle w_c,x_c\rangle,\langle w,x\rangle} = \lambda\langle w,x\rangle.\ g_c(he)$ gets help in w.

where the semantic value of *he* is given by a contextually determined variable assignment g_c.[18] In the present context, g_c assigns John to *he*. So (7) will be true just in case John stands in the relation of expectation to that boring centered proposition, i.e. to that *de dicto* content. This gives us the following truth-condition for (7):

$[\![$John expects that he will get help$]\!]^{\langle w_c,x_c\rangle,\langle w_i,x_i\rangle} = 1$ iff all the $\langle w,x\rangle$ compatible with what John expects in w_i are such that John gets help in w.

Since John *does* expect that the person who is in fact identical to John will get help, (7) is true. Thus, this semantic theory gets the desired result: (7) is true in this scenario and (8) is false.

We now have an independently motivated account of subject-controlled *PRO* which we can apply to imagination reports. Recall our earlier observation that imaginings from the inside are standardly and unambiguously reported by sentences like (3):

(3) I imagined skiing down the hill.

We assume that the semantic value (intension) of the gerundive complement *PRO skiing down the hill* is given as follows:

$\lambda\langle w,x\rangle.[\![$PRO skiing down the hill$]\!]^{\langle w_c,x_c\rangle,\langle w,x\rangle} = \lambda\langle w,x\rangle.\ x$ is skiing down the hill in w.

And this gives us the following truth condition for (3):

$[\![$I imagined PRO skiing down the hill$]\!]^{\langle w_c,x_c\rangle,\langle w_i,x_i\rangle} = 1$ iff in all the $\langle w,x\rangle$ compatible with what x_c imagined in w_i, x skis down the hill in w.

---

[18]Thus, we should really be treating contexts as world-person-assignment triples.

On this account, the content of a gerundive complement with a *PRO* subject is again a non-boring centered proposition. Since a gerundive complement with a *PRO* subject is the standard way to unambiguously report an episode of imagining from the inside, imagining a non-boring centered proposition just is imagining from the inside. This completes our linguistic argument for the Simple View.

3   An objection and an alternative

3.1   *De se* imagining from the outside?

In spite of this argument in favor of the Simple View, there is a potential problem for it, a problem which takes the form of a counterexample. The putative counterexample is a case of imagining from the outside that is also, in some sense, *de se*.

Consider Perry's character, Rudolf Lingens (Perry 1977, 21-22). Lingens is an amnesiac and doesn't know who he is. He is lost in the Stanford Library, and, we shall suppose, has stumbled upon a book entitled *The Life of Rudolf Lingens*. As Lingens reads the book, he learns more and more facts about Rudolf Lingens. But he's still in the dark about who he is. Let us suppose that Lingens puts the book down momentarily and begins to daydream.

We can distinguish two possible daydreams Lingens might have:

Situation 1   Lingens imagines being Brigitte Bardot and kissing the person he was just reading about, i.e. that he is Bardot and that he is kissing Lingens. Someone asks him what he's doing and he replies, "I'm imagining that I'm Brigitte Bardot and that I'm kissing Rudolf Lingens."

Situation 2   Lingens imagines being Brigitte Bardot and kissing himself. Someone asks him what he's doing and he replies by uttering the Lakoff sentence, "I'm imagining that I'm Brigitte Bardot and that I'm kissing me."

There appears to be some sort of *de se* - *de dicto* difference between how Lingens imagines the 'kissee' in Situations 1 and 2. One is tempted to say that in Situation 1, Lingens imagines the 'kissee' in a *de dicto* way, but that in Situation 2, Lingens imagines the 'kissee' in a *de se* way. But since Lingens imagines himself as the 'kissee' from the outside (he must be, since he's imagining being Bardot from the inside), it looks like in Situation 2, Lingens imagines himself (*qua* person kissed) in a manner that is both *de se* (in some sense) *and* from the outside. But since the Simple View identifies *de se* imagining with inside imaginings, it appears that this case is a counterexample to the Simple View. If that's right, then since Situation 2 is clearly possible, the Simple View must be wrong.

A defender of the Simple View might respond as follows: The putative counterexample does not show that anything the Simple View says is false. It simply points to a phenomenon that the Simple View says nothing about—the phenomenon, whatever it is, that is exhibited by Lingens's daydream in Situation 2. But it is still open to a defender of the Simple View to attempt to explain the difference between Situations 1 and 2 without saying that the two imaginings have different contents.

The way to develop this reply is to find a difference between Situations 1 and 2 without positing a difference in the contents of the imaginings that occur in each of those Situations. A natural strategy is to try to locate the difference in what Lingens *believes* in each of those Situations. Here is what I propose: In Situation 1, Lingens's utterance may be thought to express a belief which we might describe as follows:

Belief 1 All the $\langle w, x \rangle$ compatible with what Lingens believes in Situation 1 are such that x imagines from the inside in w being Bardot and kissing Lingens.

Now, in Situation 2, Belief 1 may be true. But whether or not that is so, Lingens arguably believes something in Situation 2 that he doesn't believe in Situation 1. This can be specified as follows:

Belief 2 All the $\langle w, x \rangle$ compatible with what Lingens believes in Situation 2 are such that x imagines from the inside in w being Bardot and kissing x.

That is, in Situation 2 (but not in Situation 1), Lingens takes himself to be an x such that x is imagining being Bardot and kissing x. This, the defender of the Simple View claims, is the main difference between Situations 1 and 2. In Situation 2 Lingens has a belief about his imagining that he doesn't have in Situation 1.

And, of course, all of this is consistent with the the claim that the contents of the *imaginings* in Situations 1 and 2 are the same. In both Situations, the following is true:

> All the $\langle w, x \rangle$ compatible with what Lingens imagines in Situations 1 and 2 are such that x is Bardot and x kisses Lingens in w.

(Note that Lingens's self-locating belief in Situation 2 is true, according to this view. Since Lingens is imagining from the inside being Bardot and kissing Lingens, $\langle \alpha, \text{Lingens} \rangle$ is in $\{\langle w, x \rangle : x$ imagines being Bardot and kissing x in $w\}$ (where α is the world of Situation 2). Lingens's self-locating belief is true because Lingens is imagining being Bardot and kissing Lingens.)

3.2 An alternative

The reply isolates a difference between Situations 1 and 2, but is it the right difference? And is it the only difference? One might be inclined to answer in the

negative, on the grounds that, in Situation 2, the kissee is *presented to* Lingens in a manner different from the way in which the kissee is presented to Lingens in Situation 1. This thought may motivate us to look for an alternative to the Simple View. In this subsection, I will consider one such alternative, but I'll argue that it is inferior the Simple View. My hope is that comparing the two views will lead us to think the Simple View is at least on the right track, even if it isn't the whole story.

In seeking a fresh view of the inside-outside distinction, it may be helpful to reflect for a moment on the clause Inside:

> Inside
> S imagines from the inside in w that he is F just in case in all the $\langle w', x' \rangle$ compatible with what S imagines at w, x' is F in w'.

One question about this claim that immediately arises is: what is it for a centered world to be compatible with what an agent *imagines*? When we introduced Lewis's account of the *de se*, we gave an account of what it is for a centered world to be compatible with what an agent *believes*; but this does not help us in answering the question about imagination.

Our answer to the imagination question came via our linguistic argument for Inside. We first observed that inside imaginings were reported by sentences like (3) and (4) in which *imagines* takes a gerundive complement. We then noted that an independently motivated semantic theory connected such sentences to *de se* contents. So we answered the question of what it is for a centered world to be compatible with what an agent imagines by looking at how the language works.

But suppose we hadn't taken the semantics of imagination reports as our starting point. Then it would have been up to us to say what it is for a centered world to be compatible with what an agent imagines. We could have introduced the notion of a centered imagination world by saying that a centered world $\langle w, x \rangle$ is compatible with what an agent S imagines just in case S imagines from the inside being x in w. But note that we might have instead introduced the notion of a centered imagination world as follows: a centered world $\langle w, x \rangle$ is compatible with what an agent S imagines just in case S imagines in the *de se* outside way that she is x in w. On the intended reading of the *definiens* here, if amnesiac Lingens imagines in the *de se* outside way that he is seeing a sunset, then all the $\langle w, x \rangle$ compatible with what Lingens imagines are such that x is seeing a sunset in w.

My point is that, if we ignore the semantics of imagination reports, this is a matter to settle by decision. The same goes of course for the notion of a centered belief world. We could have said that a centered world $\langle w, x \rangle$ is compatible with what an agent believes just in case the agent thinks *her dog* might be x in w. This would, of course, be a useless notion; my only point is that the notion of a centered belief world is a technical one that gets its intuitive content from a stipulation about how it is to be understood. We decided how the notion of a centered belief world was to be understood, just as we are now trying to decide how to define the notion of a centered imagination world.

Of course, if we defined an agent S's centered imagination worlds by reference to S's *de se* outside imaginings, then we'd still be left with the problem of how to represent the content of an inside imagining. Thus, given that there are evidently *two* things I might mean when I say that I imagined that I did such-and-such (I might mean imagined from the inside that I did such-and-such, or I might mean that I imagined from the outside that I did such-and-such), perhaps we need a richer object in order to characterize the contents of imaginings. Rather than centered worlds with a *single* center, perhaps we need *doubly centered worlds*, i.e. triples consisting of a world and two individuals.[19]

Thus, someone who is dissatisfied with the Simple View (even after our reply to the earlier objection), could offer the following account:

New View

Inside
S imagines from the inside in w that she is F just in case in all the $\langle w', x', x'' \rangle$ compatible with what S imagines at w, x' is F in w'.

De Dicto
S *de dicto* imagines from the outside in w that she is F just in case in all the $\langle w', x', x'' \rangle$ compatible with what S imagines at w, S is F in w'.

De Se Outside
S *de se* imagines from the outside in w that she is F just in case in all the $\langle w', x', x'' \rangle$ compatible with what S imagines at w, x'' is F in w'.

An advocate of this view should offer this account of a doubly centered imagination world:

A doubly centered world $\langle w, x, x' \rangle$ is compatible with what an agent S imagines just in case it is compatible with what S imagines from the inside that he is x in w and it is compatible with what S imagines from the outside (in the *de se* way) that he is x' in w.

The relevant notion of imagining from the outside in the *de se* way that one is a person x can be fixed by way of example: when Lingens imagines himself as the person kissed in Situation 2, he imagines himself from the outside in the *de se* way.

This, of course, still won't be enough to separate the New View from the Simple View. For the advocate of the Simple View can accept the three clauses of

[19]The idea of using doubly centered worlds to analyze the inside-outside phenomenon echoes a suggestion in Lakoff (1972, 639). Lakoff suggests that one might treat the Bardot example by using counterpart theory and letting the imaginer have two counterparts, one of which is Bardot, the other of which is the imaginer. (The connection between Lakoff's suggestion and this proposal is even tighter once one realizes how closely connected centered worlds theory is to counterpart theory. See **?** for discussion.)

the New View; he would simply add the following constraint on doubly centered imagination worlds:

> A doubly centered world $\langle w, x, x' \rangle$ is compatible with what an agent S imagines only if x' is S.

So the New View should be understood as acceptance of the three clauses above plus a denial of this constraint.

We now have two views on the table, the Simple View and the New View. Which is right? I will offer two reasons for preferring the Simple View to the New View, though neither constitutes a knock-down argument. Both objections to the New View highlight the fact that, if that view were true, then imagining from the outside would be much less constrained than it actually is.

First, the Lingens example suggests that *de dicto* imagining about oneself and *de se* imagining from the outside only come apart when the imaginer is in some way in the dark about his identity. When one knows who one is, one *de dicto* imagines that one is F just in case one *de se* imagines from the outside that one is F. For example, since I know that I'm Dilip, any time I imagine from the outside that I'm seeing a sunset, it seems obvious that what I'm imagining is that Dilip is seeing a sunset. And any time I imagine that Dilip is seeing a sunset, it seems obvious that what I'm imagining is that I'm seeing a sunset. Situations 1 and 2 are distinguishable only because Lingens is an amnesiac.[20]

But as it stands, the New View leads us to think otherwise—it leads us to expect that *de dicto* imaginings about oneself could come apart from *de se* outside imaginings even when the imaginer knows who she is. For according to the New View, if I imagine *de dicto* that I am seeing sunset, then the following is true:

(9) All the $\langle w', x', x'' \rangle$ compatible with what Dilip imagines in w are such that Dilip sees a sunset in w'.

And if I imagine that I'm seeing a sunset in the *de se* outside way, then (10) is true:

(10) All the $\langle w', x', x'' \rangle$ compatible with what Dilip imagines in w are such that x'' sees a sunset in w'.

As it stands, the New View does not allow us to derive the result that if I know in w that I am Dilip, then (9) is true iff (10) is true. An adequate account of *de se* imagining from the outside should have the following consequence: when an agent S knows who she is, S *de dicto* imagines that she is F iff S *de se* imagines from the outside that she is F. But the New View doesn't yield this result; for all the New View says, (9) and (10) could differ in truth value even if I know in w that I am Dilip.

---

[20]I am using *knows* here quite loosely. All that is needed is that all of the imaginer S's centered belief worlds be centered on S.

The Simple View, of course, has no problem with this, since it doesn't posit a distinction between *de dicto* and *de se* outside imagining. Instead, it tells us that the difference between situations like Situation 1 and 2 is really a difference between what Lingens believes about his imagining in each situation. And this difference in belief is a result of the fact that Lingens is an amnesiac. So if Lingens had known that he was Lingens, Situations 1 and 2 would not have been readily distinguishable, which is the right result.

A second reason for favoring the Simple View over the New View concerns an observation about imagining from the outside found in Williams (1966). Williams observes that while he can imagine from the inside being Napoleon, he finds he is unable to imagine from the outside that he is Napoleon: (Williams refers to imagining from the inside as imagination that involves 'participation imagery'):

> If the activity of imagining being Napoleon involves in any important way imagery, it is bound, I think, to involve participation imagery. Images of myself being Napoleon can scarcely merely be images of the physical figure of Napoleon, for they will not in themselves have enough of me in them... They will rather be images of, for instance, the desolation at Austerlitz as viewed by me vaguely aware of my short stature and my cockaded hat, my hand in my tunic. (Williams 1966, 43)

In this passage, Williams focuses a bit too much on imagery, but I think his underlying observation is correct. To imagine that I'm Napoleon I must imagine *being* Napoleon. It doesn't seem possible for me simply to imagine, from the outside, Napoleon, and then to stipulate that it is also true of the person in the imaginary scenario that he is me.[21] Williams's observation suggests that imagining from the outside is somehow constrained by my identity. In contrast, I can imagine from the inside being anyone I please: Brigitte Bardot, Jesus, Napoleon.

The Simple View very naturally predicts Williams's observation. On the Simple View in order to imagine from the outside that I am Napoleon all my centered imagination worlds $\langle w, x \rangle$ would have to be such that Dilip is Napoleon in w. But, given the necessity of distinctness, no centered world meets that condition, since no possible world is such that Dilip is Napoleon in it. Thus, the content of such an imagining would be the empty centered proposition.

But the New View, at least as it stands, fails to predict Williams's observation. According to that view, there should be a way for me to imagine from the outside that I am Napoleon: by imagining this in the *de se* outside way. According to that account, if all my doubly centered imagination worlds $\langle w, x, x' \rangle$ are such that x' is Napoleon in w, then I've imagined from the outside that I'm Napoleon. So nothing in the theory would lead us to expect that imagining

---

[21]It is obviously consistent with Williams's observation that Napoleon could have imagined from the outside that he was Napoleon.

from the outside is constrained in the way Williams suggests it is. The Simple View explains the phenomenon Williams is pointing to; the New View doesn't.

Thus, while there may be more to be said about the difference between Situations 1 and 2, it seems like there is something substantially right about the Simple View, something which is not captured by moving to a view with more structure like the New View.

4   Persons as Possibilities?

I close with some metaphysical speculations inspired by the foregoing discussion.

While there is much discussion in the philosophical literature about the relationship between imaginability and possibility, there is virtually none about what metaphysical consequences (if any) the inside-outside distinction has. One thought is that imagining from the inside should be taken as a guide to *centered possibility*, or possibilities for individuals, and that imagining from the outside is a guide to *uncentered possibilities*, or possibilities for worlds. If we take seriously the connection between imaginability from the inside and centered possibility, then the fact that I imagine from the inside being someone else might be thought to support the claim that I *could have been* someone else. As many philosophers have noted, this idea is not without intuitive support:

> ...my being TN (or whoever in fact I am) seems accidental... So far as what I am essentially is concerned, it seems as if I just *happen* to be the publicly identifiable person TN—as if what I really am, this conscious subject, might just as well view the world from the perspective of a different person. (Nagel 1986, 60)

> Here am I, there goes poor Fred; there but for the grace of God go I; how lucky I am to be me, not him. Where there is luck there must be contingency. (Lewis 1983a, 26)

> 'I might have been somebody else' is a very primitive and very real thought; and it tends to carry with it an idea that one knows what it would be like for this 'I' to look out on a different world, from a different body, and still be the same 'I'. (Williams 1966, 40)

For the most part, contemporary philosophers have ignored this 'very real' thought, perhaps on the grounds that it is impossible to take it literally if one adopts a possible worlds analysis of possibility. Lewis (1983a, 1986), however, endorses this thought, and argues that we can analyze the relevant possibility claims in terms of quantification over centered worlds. A Lewis-inspired proposal would look like this:

> $[\![$I could have been Fred$]\!]^{c,i} = 1$ iff there is a $\langle w, x \rangle$ accessible from $\langle w_i, x_c \rangle$ such that x is Fred in w.

(The accessibility relation is just a place-holder at the moment.) Here the possibility modal is understood as an existential quantifier over centered worlds. A centered world essentially represents a person at a world, so on this picture, the basic units of possibility are possible individuals, not possible worlds. (We still need possible worlds, of course, since possible individuals are understood as world-individual pairs.)

Of course, this analysis does not explain the claim that I could have been Fred in independently understood terms. What is the relevant accessibility relation here? Metaphysical accessibility, which relates a possible person to each of the possible persons she could have been? If that is the only answer we can give, as I suspect it is, then the analysis is to some degree circular. But in this, the centered worlds analysis is not much different from the possible worlds analysis. Possible worlds semantics analyzes *It is possible that Aristotle could have died as a child* as *There is a possible world in which Aristotle died as a child*, but the latter obviously does not explain the former in independently understood terms.

A more serious objection to the claim that I could have been Fred goes like this:

> The intuition that you could have been Fred is fairly weak. Where intuition is weak, theory rules. Our best theory of modal notions is the theory that analyzes these notions in terms of quantification over possible worlds. Since there is no possible world in which you are Fred, you couldn't have been Fred.

I'm reasonably sympathetic to this line of thought. But the centered worlds analysis of modality is best seen as an *extension* of the possible worlds analysis of modality, rather than as a *replacement* of it. As Lewis (1979) points out, wherever possible worlds are needed, centered worlds can do the job, since we have a one-one correspondence between ordinary propositions and boring centered propositions. So no theoretical loss is incurred by switching from possible worlds theory to centered worlds theory.

In spite of these remarks in support of Lewis's persons-as-possibilities proposal, I admit that I still have my doubts. Imaginability is a guide to possibility, but a fallible guide. But Lewis's idea may make other philosophical problems more tractable; if it does, then that may be the best reason to accept it. My hope is that Lewis's picture may yield insights into philosophical problems concerning personal identity, contingent identity, and mental content. Any more definitive evaluation of the persons-as-possibilities picture should thus be postponed until we can evaluate its consequences for these other issues.

References

Anand, Pranav, & Nevins, Andrew. 2004. Shifty Operators in Changing Contexts. *In:* Young, Robert B. (ed), *Proceedings from Semantics and Linguistic Theory XIV*. Cornell University: CLC Publications.

Chierchia, Gennaro. 1989. Anaphora and Attitudes *De Se*. *In:* Bartsch, R., van Benthem, J., & van Emde Boas (eds), *Semantics and Contextual Expression*. Dordrecht, Netherlands: Foris.

Egan, Andy. Forthcoming. Secondary Qualities and Self-Location. *Philosophy and Phenomenological Research*.

Heim, Irene. 1994. *Puzzling Reflexive Pronouns in De Se Reports*. Handout for Bielefeld conference, March 1994.

Higginbotham, James. 2003. Remembering, Imagining, and the First Person. *Pages 496 – 533 of:* Barber, Alex (ed), *Epistemology of Language*. Oxford and New York: Oxford University Press.

Hill, Christopher. 1997. Imaginability, Conceivability, Possibility and the Mind-Body Problem. *Philosophical Studies*, **87**, 61 – 85.

Hintikka, Jaakko. 1962. *Knowledge and Belief*. Ithaca, N.Y.: Cornell University Press.

Lakoff, George. 1972. Linguistics and Natural Logic. *In:* Davidson, Donald, & Harman, Gilbert (eds), *Semantics of Natural Language*. Dordrecht: D. Reidel Publishing Company.

Lewis, David K. 1979. Attitudes *De Dicto* and *De Se*. *The Philosophical Review*, **88**, 513 – 543. Reprinted in Lewis (1983b).

Lewis, David K. 1983a. Individuation by Acquaintance and by Stipulation. *The Philosophical Review*, **92**(1), 3 – 32. Reprinted in Lewis (1999).

Lewis, David K. 1983b. *Philosophical Papers, Volume I*. New York and Oxford: Oxford University Press.

Lewis, David K. 1986. *On the Plurality of Worlds*. Oxford: Basil Blackwell.

Lewis, David K. 1999. *Papers in Epistemology and Metaphysics*. Cambridge, UK and New York: Cambridge University Press.

Martin, M.G.F. 2002. The Transparency of Experience. *Mind and Language*, **17**(4), 376 – 425.

Morgan, Jerry. 1970. On the Criterion of NP Deletion. *CLS*, **6**.

Nagel, Thomas. 1974. What is it Like to be a Bat? *The Philosophical Review*, **83**, 435 – 450.

Nagel, Thomas. 1986. *The View from Nowhere*. New York and Oxford: Oxford University Press.

Peacocke, Christopher. 1985. Imagination, Experience, and Possibility: A Berkeleian View Defended. *Pages 19 – 35 of:* Foster, John, & Robinson, Howard (eds), *Essays on Berkeley*. Oxford: Clarendon Press.

Perry, John. 1977. Frege on Demonstratives. *The Philosophical Review*, **86**(4), 474 – 497. Reprinted in Perry (1993).

Perry, John. 1979. The Problem of the Essential Indexical. *Noûs*, **13**(1), 3 – 21. Reprinted in Perry (1993).

Perry, John. 1993. *The Problem of the Essential Indexical and Other Essays*. New York and Oxford: Oxford University Press.

Quine, W.V.O. 1969. Propositional Objects. *In: Ontological Relativity and Other Essays*. New York: Columbia University Press.

Shoemaker, Sydney. 1994. The First-Person Perspective. *Proceedings and Addresses of the American Philosophical Association*, **68**, 7 – 22. Reprinted in Shoemaker (1996).

Shoemaker, Sydney. 1996. *The First-Person Perspective and Other Essays*. New York: Cambridge University Press.

Velleman, David. 1996. Self to Self. *The Philosophical Review*, **105**(1), 39 – 76.

von Fintel, Kai. 2005. *Lecture notes for Pragmatics in Linguistic Theory*. Notes for class taught at the LSA Summer Institute, MIT.

von Fintel, Kai, & Heim, Irene. 2004. *Intensional Semantics Lecture Notes*. Notes for class taught at MIT.

Walton, Kendall. 1990. *Mimesis as Make-Believe*. Cambridge, MA and London: Harvard

University Press.

Williams, Bernard. 1966. Imagination and the Self. *Proceedings of the British Academy.* Reprinted in Williams (1973).

Williams, Bernard. 1970. The Self and the Future. *Philosophical Review*, **79**(2), 161 − 180. Reprinted in Williams (1973).

Williams, Bernard. 1973. *Problems of the Self : Philosophical Papers 1956-1972.* Cambridge: Cambridge University Press.