Updating as Communication*

SARAH MOSS
University of Michigan

On many traditional theories of belief, your belief state is represented by an assignment of credences to propositions, or sets of possible worlds. If you are rational, your credence distribution will be a probability measure. Traditional theories of belief fit with a standard Bayesian theory of rational belief change: on learning a proposition, you must update your belief state by conditionalizing your credence distribution on the proposition you learn. That is, you must update by assigning 0 credence to those worlds incompatible with what you learn, and re-normalizing your credence distribution over the remaining worlds.

Following Quine 1969, Lewis 1979 argues that we should instead represent your belief state by an assignment of credences to sets of *centered worlds*: world-time-individual triples. For instance, if you have .5 credence that it is 3:00 pm, your belief state should be represented by a measure that assigns .5 to the set of centered worlds with that time coordinate. Unlike traditional theories of belief, Lewis's theory does not fit with a standard Bayesian theory of rational belief change. For instance, Bayesian conditionalization preserves certainties. If you update by conditionalizing on the set of centered worlds you learn, it follows that if you are ever certain that it is 3:00 pm, you must always remain certain that it is 3:00 pm. But clearly this is not what rationality requires. If we agree with Lewis about how to represent belief states, we must develop another set of principles governing rational belief change.

In this paper, I develop a procedure for rationally updating credence distributions over sets of centered worlds. I argue that rational updating can be factored into two steps. Roughly speaking, in forming an updated credence distribution, you must first use information you recall

---

* EDITOR'S NOTE: This paper won the Young Epistemologist Prize for the Rutgers Epistemology Conference held in 2011.

from your previous self to form a hypothetical credence distribution, and then change this hypothetical distribution to reflect information you have genuinely learned as time has passed. In making this proposal precise, I argue that your recalling information from your previous self resembles a familiar process: agents' gaining information from each other through ordinary communication.

The updating procedure I develop relies on relationships between two kinds of sets of centered worlds: *de se* and *de dicto* propositions. I will define *de dicto* propositions to be boring sets of centered worlds: sets of world-time-individual triples such that if one triple is in the set, so is every other triple which shares its world coordinate. *De se* propositions are sets of centered worlds that are not *de dicto* propositions. *De dicto* propositions are entirely about what the world is like, while *de se* propositions are also about where you are in the world. In §1, I introduce a claim about how *de se* contents of attitudes are related to *de dicto* propositions. In §2, I use this claim to solve a puzzle about imagination.

The discussion in §1–2 provides the foundation for a unified theory of communicating and updating beliefs. In §3, I describe how agents communicate *de se* beliefs. In §4, I argue that rational updating begins with a similar process. I then describe the rest of a complete procedure for rationally updating credences in *de se* propositions. Finally, in §5, I apply my theory to particular cases. The last of these cases is easily handled by my theory, while presenting a problem for the theory of updating in Titelbaum 2008. This problem for Titelbaum highlights an important feature of rational updating that any successful updating procedure must recognize.

1. *De Se* and *De Dicto* Contents

In giving a theory of how you should update your *de se* beliefs, it is helpful to understand how the contents of those beliefs are related to various *de dicto* propositions. In what follows, I will argue that the following claim is theoretically and intuitively attractive: given a *de se* proposition, there is a *de dicto* proposition that is equivalent with that *de se* proposition, given what you believe. In more precise terms:

> (PROXY) Given a *de se* proposition, there is a *de dicto* proposition such that for any centered world compatible with what you believe, that centered world is in the former proposition just in case it is in the latter.

Some semantic theories of attitude ascriptions help us find *de dicto* propositions equivalent with contents of *de se* attitudes by endorsing

the following claim: that speakers use first-person indexicals to self-ascribe attitudes with *de dicto* contents.[1] For example, suppose Kaplan sees himself in a mirror, without realizing that he is seeing himself. Looking at the mirror, Kaplan sees that his pants are on fire, without realizing that his own pants are on fire. In recounting his experience, suppose Kaplan says:

(1) I expected that I would be rescued.

Kaplan can truly utter (1), even though he was not aware of being in danger when he looked at the mirror. In this respect, (1) differs from (2):

(2) I expected to be rescued.

Unlike (1), reports such as (2) are true only if the ascribee has a self-directed attitude. Here is one useful explanation of this contrast: reports such as (2) ascribe attitudes with *de se* contents, while reports such as (1) ascribe attitudes with *de dicto* contents.[2] The content of the expectation ascribed by (2) is a set of centered worlds where the center is rescued. But the content of the expectation ascribed by (1) is a set of centered worlds that is characterized not by any property of the center, but by some property of the person Kaplan sees. Since (1) and (2) ascribe expectations with different contents, these ascriptions can have different truth conditions.

This semantic account of the contrast between (1) and (2) fits well with (PROXY). Suppose that (1) ascribes a *de dicto* attitude while (2) ascribes a *de se* attitude. Kaplan believes that the person he sees is not himself, so the content of the attitude that (1) ascribes is not equivalent with the content of the attitude that (2) ascribes, given what he believes. But normally when a speaker utters (1) and (2), the content of the *de dicto* attitude that (1) ascribes will be equivalent with the content of the *de se* attitude that (2) ascribes, given what she believes. These contents will still be distinct propositions. In particular, only one will be a *de se* proposition. But in normal cases, the centered worlds at which the contents differ in truth value will not be among the centered worlds compatible with what the speaker believes.

In other words: even in normal cases with no identity confusion, you have a third-personal way of thinking about yourself. On the

---

[1] Morgan 1970 and Lakoff 1972 were among the first to highlight third-personal readings of embedded first-person pronouns. In setting up my examples, I use a case developed by Kaplan in the late 1970's and familiar from Kaplan 1989.

[2] For recent semantic proposals developing this explanation, see von Fintel 2005, Percus & Sauerland 2003, and Stephenson 2009.

semantic account just sketched, this way of thinking about yourself gives rise to *de dicto* attitudes that you use first-person indexicals to self-ascribe. The contents of these attitudes are equivalent with the contents of your *de se* attitudes, given what you believe. Or in fewer words: your *de dicto* attitudes are equivalent with your *de se* attitudes, given what you believe. Furthermore, you retain your normal third-personal attitudes about yourself, even if you acquire other attitudes about yourself upon failing to recognize yourself in a mirror. Even in an identity confusion case, your normal third-personal *de dicto* attitudes are equivalent with your *de se* attitudes, given what you believe. So whether or not you are in an identity confusion case, there is a *de dicto* proposition that is equivalent with any given *de se* proposition, given what you believe.

This discussion highlights an important intuition that is ultimately independent of any semantics of attitude ascriptions: you can always think about yourself in the same way you think about any other person, and in the same way other people think about you. Here is an example where no ascriptions are uttered: suppose Kaplan has gotten lost in his sea kayak and calls you for directions. He describes the coastline and you look at maps, but you still cannot figure out where he is paddling. There are two possible worlds compatible with your beliefs as you talk to Kaplan: one where Kaplan is in Bellingham, and one where Kaplan is in Seattle. Just the same possible worlds are compatible with what Kaplan believes. Neither you nor Kaplan can figure out which of the worlds is actual. And so Kaplan is not only ignorant of *de se* facts. He is also ignorant of the same *de dicto* facts as you. Suppose Kaplan is in Bellingham and someone exactly like him is in Seattle, and suppose Kaplan introduces a name for himself as he is talking to you:

(3)   Let 'Dr. Demonstrative' name myself.

In one possible world compatible with what you and Kaplan believe, Dr. Demonstrative is in Bellingham and some other guy is in Seattle. In another world compatible with your beliefs, Dr. Demonstrative is in Seattle and some other guy is in Bellingham. Each possible world corresponds to two centered worlds: one centered on Bellingham and one centered on Seattle. Since Kaplan believes that he himself is Dr. Demonstrative, he can rule out exactly two of these four centered worlds. That is why his *de se* belief that he himself is in Bellingham is equivalent with his *de dicto* belief that Dr. Demonstrative is in Bellingham, given what he believes. The same goes for all of his *de se* attitudes.

Considering the case where Kaplan utters (3), we can see that an even stronger moral holds: there is a *de dicto* proposition equivalent

with any given *de se* proposition, given merely what you believe with certainty. For instance, Kaplan could always have some shred of doubt about whether he is the man whose pants are on fire, or even about whether he is David Kaplan. Contrast this with your immediate conviction, on uttering (1) and (2) in a normal case, that if one expectation is satisfied then the other will be. Similarly, on uttering (4), you cannot doubt that your expectation is about yourself:

(4) I expect that I will be rescued.

In just this sense, you are always certain about which person is yourself. Given what you believe with certainty, your *de dicto* beliefs about that person will be equivalent with your *de se* beliefs about yourself. Since Kaplan can be certain that he is Dr. Demonstrative, his Dr. Demonstrative beliefs are equivalent with his *de se* beliefs, given what he believes with certainty.

To sum up so far: the *de se* expectation that Kaplan ascribes using (5) is not equivalent with the expectations he ascribes using (6) and (7), given what he believes with certainty:

(5) I expect to be rescued.

(6) I expect that the man whose pants are on fire will be rescued.

(7) I expect that David Kaplan will be rescued.

But other *de dicto* attitudes about yourself are equivalent with your *de se* attitudes, given what you believe with certainty. For instance, the kind of expectation ascribed in (5) is normally equivalent with expectations ascribed using first-person indexicals (as in (4)) and expectations ascribed using names introduced with first-person indexicals (as in (8)):

(8) I expect that Dr. Demonstrative will be rescued.

In particular, (5) and (8) will ascribe equivalent expectations as long as Kaplan remains certain that he himself is Dr. Demonstrative, as he is when he first introduces the name. The same results hold for your *de se* attitudes about your temporal location. In addition to thinking about yourself from an impersonal perspective, you can think about your temporal location from an atemporal perspective. Just as with your impersonal thoughts about yourself, the resulting *de dicto* attitudes are equivalent with your *de se* attitudes, given what you believe with certainty.

The content of the belief that Dr. Demonstrative is in Bellingham is a *de dicto* proposition. It is natural to wonder exactly which worlds it contains. Or in a familiar Stalnakerian idiom, one may simply wonder what the world is like according to Kaplan when he believes that Dr. Demonstrative is in Bellingham. If Kaplan forgets that he himself is David Kaplan, he may have the belief that Dr. Demonstrative is in Bellingham without having the belief that David Kaplan is in Bellingham. Hence it seems our approach should distinguish the former belief from the latter, and one may wonder what distinguishes these beliefs.

None of these questions specifically concern *de se* attitudes. For instance, one may analogously wonder what worlds are contained in the content of the belief that Hesperus is bright, or what the world is like according to someone who believes that Hesperus is bright, or what distinguishes Hesperus beliefs from Phosphorus beliefs. The ensuing debate is familiar. If we say the content of the belief that Hesperus is bright is just the set of worlds where Venus is bright, we seem to neglect differences between Hesperus beliefs and Phosphorus beliefs. But if we say the content is some descriptively identified set of worlds, we run afoul of familiar anti-descriptivist injunctions. The resolution of this debate will determine how we characterize the content of the belief that Dr. Demonstrative is in Bellingham. In other words, distinguishing Dr. Demonstrative beliefs and David Kaplan beliefs is just another instance of Frege's puzzle. Any theory of the latter will address the questions raised above.

Several solutions to Frege's puzzle are compatible with (PROXY). For present purposes, I will remain neutral between them. I am not aiming to solve Frege's puzzle, but simply to show how it subsumes a number of puzzles concerning *de se* attitudes. (PROXY) is only incompatible with theories according to which we cannot adequately characterize the contents of Hesperus and Phosphorus beliefs unless we say that these contents are *de se* propositions.[3] In that case, nearly all of our beliefs will have *de se* contents, including Dr. Demonstrative beliefs. In order to make my framework compatible with this kind of theory, one must distinguish between deeply and superficially *de se* contents, and read my claims about *de se* contents as claims about deeply *de se* contents.

(PROXY) is attractive in part because it provides for an elegant and unified account of several complicated and disparate phenomena. So far I have argued that (PROXY) fits well with a simple semantics of attitude ascriptions, and that it provides for a natural account of what

---

[3]   For instance, Chalmers says that the epistemic intension of an indexical is a *de se* proposition. One might argue that 'Hesperus' is an implicitly indexical expression, and conclude that the epistemic intension of 'Hesperus is bright' is a *de se* belief content. See Chalmers 2002 and Chalmers 2003 for further discussion.

happens in cases where you think about yourself in the same way other people think about you. In upcoming sections, I put (PROXY) to work in solving a puzzle about imagination, and in saying how agents communicate and update *de se* beliefs. These applications provide further reason to accept (PROXY) itself.

2. Two Ways of Imagining

Suppose that it is 3:00 and you are teaching class, and while you are teaching, I ask you to imagine that it is 5:00. There are two very different ways you might respond. For instance, you might play along by saying either of the following:

(9)   Then I am in my kitchen, starting to make dinner.

(10)   Then my watch is wrong, and all of us must be strangely confused to be here so much later than usual.

Once you decide to respond in one of these ways, it is clear how you should go on with what you are imagining. Either you imagine that two hours have passed and your day has proceeded normally, or you imagine that someone has played a practical joke on you and your students. These responses involve very different kinds of imaginary scenarios. The acceptability of either response raises a puzzle: what distinguishes these two ways of imagining that it is 5:00?

In both cases, when I ask you to imagine that it is 5:00, you comply by imagining a certain *de se* proposition. In particular, all centered worlds compatible with what you imagine are in the set of centered worlds whose time coordinate is 5:00. But what you imagine in each case is distinguished by whether you also imagine a certain *de dicto* proposition. Outside the pretense, you actually believe the *de se* proposition that it is 3:00. Furthermore, you have a *de dicto* belief equivalent with this *de se* belief, given the propositions that you actually believe with certainty. In other words, you believe that a certain time is 3:00. The *de dicto* content of this belief is central to our solution of the puzzle.

The different ways of imagining that it is 5:00 are fundamentally separated by whether what you imagine is consistent with what you actually believe. In the case where you imagine as in (9), you not only imagine the *de se* proposition that it is 5:00, but also the *de dicto* proposition that a certain time is 3:00. In the case of (10), this *de dicto* proposition is not part of what you imagine. In other words, there is an extra constraint on the worlds compatible with what you imagine in

(9): the *de dicto* proposition that a certain time is 3:00 holds in all these worlds.

Our natural responses to (9) and (10) support my characterization of the difference between these ways of imagining. For instance, it is natural to say that when you accept (9), you are imagining that *some time has passed*. If you are imagining that the actual current time has already passed, you may freely imagine that it is 5:00, while imagining that you correctly identified the actual current time as 3:00. In this case, your *de dicto* belief that a certain time is 3:00 is true at worlds compatible with what you imagine. By contrast, it is natural to say that when you accept (10), you are imagining that *the actual current time is not what you thought it was*. In this case, your *de dicto* belief that a certain time is 3:00 is not true at worlds compatible with what you imagine.

The same puzzle arises for several attitudes besides imagining. For example, there are two natural ways to suppose the *de se* proposition that it is 5:00, corresponding to two indicative conditionals:

(11) If it is 5:00, then I am in my kitchen, starting to make dinner.

(12) If it is 5:00, then my watch is wrong, and all of us must be strangely confused to be here so much later than usual.

Here the puzzle is to say why both of these very different conditionals can be acceptable. Let us agree with Ramsey 1931 that 'if *p*, would *q*' is acceptable to those who accept *q* after "adding *p* hypothetically to their stock of knowledge" (248). Both (11) and (12) can be acceptable because there are different ways to add the *de se* proposition that it is 5:00 to your stock of knowledge. In particular, as you suppose that it is 5:00, you may or may not continue to accept the *de dicto* proposition that a certain time is 3:00. If you retain your *de dicto* belief, you will accept the consequent of (11). If you give up your *de dicto* belief, you will accept the consequent of (12).

So far I have distinguished ways of imagining and supposing centered contents. The distinctions I have drawn are related to the distinction between belief updating and belief revision often cited in literature on *de se* belief change.[4] In order to accept the consequent of (11), you must update on the antecedent as if some time had passed. In order to accept the consequent of (12), you must instead revise your current beliefs. In both updating and revising, you give up some *de se* beliefs. Updating and revising are distinguished by whether you also

---

[4]   See Katsuno & Mendelzon 1991 for an influential introduction.

give up certain *de dicto* beliefs that your old *de se* beliefs were equivalent with. If you retain your *de dicto* beliefs, you are updating. If you give them up, you are revising.

Retaining *de dicto* beliefs equivalent with your *de se* beliefs is what unifies several attitudes: imagining as in (9), supposing as in (11), and updating rather than revising. I hope to have forestalled the objection that your *de dicto* belief that a certain time is 3:00 is trivial, by arguing that whether you retain such *de dicto* beliefs grounds substantive differences in ways of imagining and supposing propositions. I also hope to have forestalled the objection that your belief that a certain time is 3:00 is really a *de se* belief about what time it is, since you imagine the same *de se* contents about what time it is in (9) and (10), while you imagine the content that a certain time is 3:00 only in the former case. To sum up: our puzzle about imagining gives us reason to think that there are non-trivial *de dicto* beliefs equivalent with your *de se* beliefs, given what you believe with certainty. Specifically, positing such *de dicto* beliefs yields a simple and intuitive solution to our puzzle. In what follows, I argue that *de dicto* beliefs also play an important role in a simple unified theory of *de se* communication and updating.

3. Learning from Other Agents

In §4, I develop a theory of how agents should maintain and modify their *de se* beliefs as time passes. On this theory, part of updating resembles another instance of the transmission of centered information: interpersonal communication. Communicating agents may exchange beliefs, even though they distribute their credence over disjoint propositions, namely sets of centered worlds with distinct person coordinates. Similarly, an agent may retain beliefs, even though at different times she distributes her credence over disjoint propositions, namely sets of centered worlds with distinct time coordinates.

Lewis says that many belief contents are *de se* propositions. But these *de se* propositions cannot always be what is conveyed in communication. For example, suppose Kaplan believes that his own pants are on fire, and when he tells his sister what he believes, she comes to believe just this same centered proposition. Then his sister would come to believe that her own pants were on fire. This is the same centered proposition that Kaplan believes. But obviously, it is not the information that Kaplan should have conveyed to his sister, in telling her what he believed. Instead she should have come to believe some other *de se* propositions, such as the set of centered worlds where the center has a brother whose pants are on fire. The same goes for the transmission of centered information across times. Suppose I express one of my beliefs

by saying 'it is Monday' and one day later I remember this belief. Then I should not come to self-ascribe the property of being located on Monday, but the property of being located on Tuesday.

These examples illustrate a *prima facie* tension between two intuitive ideas. On the one hand, we may favor a Stalnakerian "package delivery" model of communication. The package delivery model says that when I communicate my beliefs to you, there is something that I believe, that I convey, and that you come to believe. On the other hand, Lewis suggests that I believe *de se* propositions. But when I communicate my beliefs, you do not come to believe the same *de se* propositions that I believe. Instead you come to believe other *de se* propositions, ones that I don't believe.

It is not hard to resolve this tension with notions we already have at hand. There is something that Kaplan believes, that he tells his sister, and that his sister comes to believe. That content is a *de dicto* content equivalent with the *de se* proposition that his own pants are on fire, given what he believes with certainty. In coming to believe this proposition, Kaplan's sister does not come to believe that her own pants are on fire. Of course, she may acquire several *de se* beliefs of her own. For instance, she may infer that she herself has a brother whose pants are on fire. But the "delivered package" of the Stalnakerian model is a *de dicto* proposition. Just as we can use indexicals to self-ascribe *de dicto* beliefs, we can use indexicals to convey *de dicto* information.

This theory fits with the Lewisian framework, while respecting our intuitions about the identity conditions of contents conveyed in conversation. Stalnaker 2008 worries that the Lewisian framework conflicts with our intuitions about individuating contents:

> Lewis's account distinguishes contents that ought to be identified. If Rudolf Lingens tells you that he is sad, or that he is Rudolf Lingens, and you understand and accept what he says, then it seems that the information you acquire is the same information he imparted. (50–1)

But Lewis can accommodate this intuition, while still taking belief contents to be sets of centered worlds. If Lingens tells you that he is sad, he conveys a *de dicto* proposition equivalent with the content of his *de se* belief that he himself is sad, given what he believes with certainty. This proposition is something that Lingens believes, that he conveys, and that you come to believe. Our judgment that we should identify what you and Lingens believe reflects the fact that you both believe this *de dicto* proposition. Stalnaker also worries:

> [Lewis] identifies contents that ought to be distinguished. What I believe when I believe that I was born in New Jersey is something

about myself, something different from what my fellow New Jersey natives believe about themselves. What I tell the waiter when I tell him that I will have the mushroom souffle is different from what you tell the waiter if you decide to have the same thing. (50)

But Lewis may respond that when Stalnaker believes that he was born in New Jersey, he believes a *de dicto* proposition equivalent with the content of his *de se* belief that he himself was born there, given what he believes with certainty. His fellow New Jersey native believes a different *de dicto* proposition. Our judgment that we should distinguish what Stalnaker and his fellow New Jersey native believe reflects the fact that they believe different *de dicto* propositions. Similarly, our judgment that we should distinguish what you and Stalnaker tell the waiter reflects the fact that you convey different *de dicto* propositions to the waiter, even if you use the same indexicals when you order.

This discussion suggests a simple theory of the role your *de se* beliefs play in communication. Each *de se* proposition you believe is equivalent with some *de dicto* proposition, given what you believe with certainty. This sort of *de dicto* proposition is something you convey to your audience, and something they come to believe. Furthermore, your audience already has some *de se* beliefs about their relation to you. So they also come to believe some *de se* propositions: the consequences of their standing *de se* beliefs and their acquired *de dicto* information.[5]

Suppose we are standing in a line. I see that I am just behind you, but I have no idea how many people are ahead of you. Suppose you believe a *de se* proposition: that you yourself are fourth in line. This proposition is equivalent with some *de dicto* proposition, given what you believe with certainty. If you say 'I am fourth in line' to me, then that sort of *de dicto* proposition is what you convey to me and what I come to believe. Furthermore, I already have some *de se* beliefs about my relation to you: that I myself am just behind you in line. So I also come to believe a *de se* proposition: that I myself am fifth in line. So when we communicate, I gain *de se* beliefs: not your beliefs, but the consequences of my standing *de se* beliefs and my acquired *de dicto* information.

The theory I am defending is incompatible with a certain understanding of why we need to use *de se* propositions to represent mental states. Consider an example from Perry 1977: the amnesiac Lingens is lost in the Stanford library. He reads many books, but nevertheless "still won't know who he is, and where he is, no matter how much knowledge he piles up, until that moment when he is ready to say,

---

[5]   This is a theory of how agents normally communicate. See Egan 2005 for arguments that speakers use epistemic modals to directly convey *de se* propositions.

'*This* place is aisle five, floor six, of Main Library, Stanford. *I* am Rudolf Lingens'" (710). Some understand the Lingens example to demonstrate the following: Lingens fails to know some information relevant to his location. He knows every relevant *de dicto* proposition. Hence we must use *de se* propositions to characterize his ignorance.

The theory of communication I defend suggests that we should reject this moral of the Lingens case. Lingens is ignorant of *de dicto* propositions relevant to his location, namely those equivalent for him with his *de se* beliefs about his location. This is just the sort of proposition that you would convey to him if you were to resolve his remaining ignorance by saying to him, "you are Rudolf Lingens." Suppose he echoes you, incredulously, "I am Rudolf Lingens?" The package delivery account of this exchange says that you have used an indexical to communicate a *de dicto* proposition to Lingens, one that you knew and he did not. This account of your exchange is compatible with our pretheoretical intuitions about the Lingens case. Our intuitions suggest that Lingens fails to know some information about where he is located. But our intuitions do not legislate that Lingens' *de se* ignorance must be unaccompanied by *de dicto* ignorance. The point illustrated by the Lingens case is not that Lingens knows every *de dicto* proposition relevant to his location, but that his own biography could not tell Lingens some *de dicto* propositions that you could use an indexical to tell him. And it may well be impossible to communicate the relevant *de dicto* propositions without using either an indexical or an expression like 'Dr. Demonstrative' that is introduced using an indexical. But it is a fallacy to infer that the proposition in question is not a *de dicto* proposition after all. As explained in §1, indexicals and names like 'Dr. Demonstrative' may be used to communicate *de dicto* information. Some readers may remain suspicious of Dr. Demonstrative beliefs on the grounds that one cannot give descriptive or non-circular specifications of their contents. But this fact does not call for any special explanation. Similar claims hold for just about any *de dicto* proposition. It is true that 'Dr. Demonstrative' as introduced by Kaplan does not have the same cognitive import as any other name already in his vocabulary. But 'Hesperus' does not have the same cognitive import as any other name he has for Venus, either. Again, the *de dicto* contents posited by (PROXY) pose just the same challenges for theories of content as ordinary *de dicto* contents expressed with proper names. The case of Rudolf Lingens is no harder than any other instance of Frege's puzzle.

The case of the two gods introduced in Lewis 1979 presents a special sort of challenge to (PROXY), since Lewis attempts to stipulate that the gods in his thought experiment are omniscient with respect to *de dicto* but not *de se* propositions. In response to a similar objection, Stalnaker

2008 defends the claim that "ignorance about where one is in the world is always ignorance about what possible world is actual" (cf. §3.3). In short, Stalnaker responds that one should not take for granted the possibility of the case Lewis describes. Suppose that one of the gods says "let 'Dr. Deity' name myself" and you immediately tell him that Dr. Deity is on the tallest mountain. Then either he learns some *de dicto* proposition from you, or he did not have any *de se* ignorance to start with. I will not rehearse the details of Stalnaker's argument here, but I endorse the same response on behalf of my own theory. But I would add that the advantages of (PROXY) far exceed its enabling an elegant model of communication, as (PROXY) is also presupposed by simple and powerful theories of attitude ascriptions, content individuation, imagination, indicative supposition, and as I will argue, updating. Given his pragmatic defense of modal realism, Lewis should appreciate that the theoretical fruits of (PROXY) may be substantial enough to justify our acceptance of the claim itself.[6]

The existence of ignorance does not force us to use *de se* propositions to represent mental states. But the theory I am defending is compatible with a second motivation for introducing *de se* propositions, namely that such propositions play a distinctive role in our cognitive economy. Perry 1979 argues that we must mention a certain kind of belief to explain why someone spilling sugar stops pushing his own shopping cart through the supermarket. Typically one may start to explain an action by saying what beliefs and desires motivate that action. Suppose that Kaplan and I both believe that Dr. Demonstrative is spilling sugar and both want the spilling to stop. It still remains to be explained why I am motivated to wave and point my finger, while Kaplan is motivated to inspect the cart he is pushing. I have argued that Kaplan has *de dicto* beliefs that are equivalent for him with his *de se* beliefs. But I have not argued that such *de dicto* beliefs obviate reference to *de se* beliefs in explaining action.

4. Learning from Your Previous Self

Giving a theory of how agents with *de se* beliefs communicate illuminates how agents maintain and modify their *de se* beliefs over time. The model of updating I develop relies on an intuitive notion of *genuine learning*. Everyone recognizes that as you sense that time is passing, you should change your credences to reflect your awareness of your changing temporal location. And your opinions about exactly how much time has passed should influence how you update. But ordinarily as time passes, you are not merely sitting in a hypothetical black box, keeping

---

6   For readers looking for a more conciliatory approach: one may simply restrict (PROXY) to agents that can have *de dicto* thoughts about themselves and their temporal location. The resulting theory of updating is still substantial in scope.

track of the minutes as they pass by. You have experiences that make you more informed than your previous self, imposing novel constraints on your credences. In other words, you genuinely learn information.

In hypothetical black box updating, you form beliefs on the basis of information you get from your previous self. Getting information from your previous self is just like getting information from other agents. Each *de se* proposition you used to believe is equivalent with some *de dicto* proposition, given what you used to believe with certainty. This sort of *de dicto* proposition is something you can currently believe. Furthermore, you currently have some *de se* beliefs about your relation to your previous self. So you can also currently believe some *de se* propositions: the consequences of your current *de se* beliefs and your old *de dicto* information.

Suppose you used to believe a *de se* proposition: that it was the fourth of the month. This proposition is equivalent with some *de dicto* proposition, given what you used to believe with certainty. This sort of *de dicto* proposition is something you can currently believe. Furthermore, you currently have some *de se* beliefs about your relation to your previous self, e.g. that your current self is located one day later. So you can also currently believe a *de se* proposition: that it is the fifth of the month. Just as an agent may have certain *de se* beliefs once she acquires *de dicto* beliefs from other agents, you may have certain *de se* beliefs once you recall the *de dicto* beliefs of your previous self.

Genuine rational updating happens in two steps. First you update as if you were in a black box. Then you conditionalize your resulting credences on what you genuinely learn. I have just described the first step of updating. In order to describe a more complete updating procedure, I will point out three respects in which the story so far has been idealized, and I will say how these idealizations can be removed.

4.1. Credences

So far I have talked about modifying beliefs, rather than credence distributions. But my aim is to develop a general theory of how agents maintain and modify *credences*. Fortunately, an appropriately sophisticated theory of interpersonal communication can again serve as our guide. In making an assertion, you can do much more than simply convey certain *de dicto* beliefs to me. If you say 'John smokes' to me, then I should believe that John smokes. But if you merely say 'John *might* smoke' to me, then you merely propose that I should believe that John might smoke. On some recent theories of modals, this means I should give at least some credence to the proposition that John smokes. Similarly, if you say 'if John smokes, then Mary drinks' to me, then

I should give high conditional credence to the proposition that Mary drinks, conditional on the proposition that John smokes. If you say 'it is .9 likely that John smokes' to me, then I should give .9 credence to the proposition that John smokes. By making assertions, you propose that my credences satisfy some constraint, presumably one that your credences already satisfy.[7]

The analogy with updating extends: in black box updating, your current credences should satisfy constraints that your past credences used to satisfy. Earlier I said that *de dicto* beliefs are what you convey in conversation and recall from your previous self. But in fact what you convey and recall are constraints on your credences in *de dicto* propositions. Suppose you used to give .9 credence to a *de se* proposition: that it was the fourth of the month. Given what you used to believe with certainty, this proposition is equivalent with some *de dicto* proposition, to which you also gave .9 credence if your credences were probabilistically coherent. If you are updating in a black box, you should currently give .9 credence to that same *de dicto* proposition.

Black box updating is like communication: it as if your previous self could talk to you and thereby propose constraints on your *de dicto* credences. Only unlike cases of real communication, there is no limit to the amount of information your previous self can convey. It is as if your previous self proposes that your current *de dicto* credences satisfy every constraint that they did before. So in a hypothetical black box updating case, a case where no genuine learning occurs, all of your *de dicto* credences should stay just the same.

4.2. Conditional Credences about Your Relation to Your Previous Self

So far when talking about how your previous *de dicto* beliefs should influence your current *de se* beliefs, I have talked about your beliefs about your relation to your previous self. But in fact you have more complicated opinions about your relation to your previous self. In particular, your credences about how much time has passed between you and your previous self are conditional in nature. For example, suppose you recently looked at a clock that read 2:00, but you think the clock may be an hour slow. Suppose you also know that time passes more quickly as the afternoon wears on. Then you might currently believe that if it was indeed 2:00 earlier, four minutes have passed since you looked at the clock. But if it was 3:00, five or six minutes may have passed. In this way, your opinions about how much time has passed are conditional credences. They are conditional on *de dicto*

---

7   See Swanson 2006 and Yalcin 2007 for developed theories that relate asserted contents to constraints on credences.

propositions, such as the *de dicto* proposition you would have used 'it is 2:00' to convey when you were looking at the clock.

In practice, your opinions about your relation to your previous self are given by conditional credence distributions. Earlier your *de dicto* credences were defined on an algebra generated by some partition of atomic *de dicto* propositions. For any such *de dicto* proposition, you have a credence distribution over *de se* propositions, given that *de dicto* proposition. For example, conditional on your having looked at the clock at 3:00, you may give .5 credence to five minutes having passed and .5 credence to six minutes having passed. Conditional credence distributions like these are more precise models of your opinions about your relative location in time.

In black box updating, your credences are entirely determined by two elements: your previous credences in *de dicto* propositions, and your current conditional credences about your relation to your previous self. First your previous credences determine how much credence you give to any given *de dicto* proposition. Then your conditional credences determine how you distribute that credence among all *de se* propositions entailing that *de dicto* proposition.[8] This uniquely determines a credence distribution over both *de dicto* and *de se* propositions. If your previous opinions and your innate sense of time passing were your only sources of information, your rationally updated credences would be determined in just this way.

4.3. Genuine Learning

Once we understand how you should update in a black box case, describing a complete procedure for rational updating is straightforward. In ordinary cases, your later credences are not only informed by your previous opinions. They must reflect what you genuinely learn as time passes, information that makes you smarter than your previous self. The combination of your previous *de dicto* credences and conditional *de se* credences is a hypothetical credence distribution, representing how you should have updated if you had not genuinely learned anything. In order to arrive at the updated credences you really should have, you must conditionalize this hypothetical credence distribution on what you genuinely learn.[9]

---

[8] Over time, your *de dicto* credences are defined over an increasingly fine-grained algebra, as you come to assign credence to *de dicto* propositions equivalent with the contents of your later *de se* attitudes. So when you distribute credence among *de se* propositions, you are effectively also distributing credence among the *de dicto* propositions that are equivalent with those *de se* propositions, given what your later self believes.

[9] This may involve updating by simple conditionalization, Jeffrey conditionalization, or more complicated procedures *à la* §5 of Diaconis & Zabell 1982.

It is important to notice that the first step of updating results in a merely hypothetical credence distribution. For example, it may be that you are always genuinely learning information, so that you never have credences informed only by your own sense of time passing. Rationally updated credences may always be the product of both your black box credences and what you genuinely learn.

Distinguishing steps of updating that use different kinds of information allows us to more easily recognize how those steps of updating are related to other processes. The first step of updating is analogous to communication. If you have opinions about how you are related to a speaker, she may convey *de dicto* information that constrains your *de se* credences. If you have opinions about how you are related to your previous self, your previous *de dicto* credences may constrain your current *de se* credences in just the same way. The second step of updating is simply conditionalizing on what you learn. In a sense, we have found that conditionalization is the correct procedure for updating *de se* credences. It is just that we must be careful that we are conditionalizing the correct object on what you learn: not your previous credences, but a hypothetical modification of them.

5. Discussion

I have given a framework that organizes and highlights various features of the updating process. The most dramatic consequence of my framework is that the process of rational updating can be entirely factored into two steps: generating hypothetical credences informed only by your previous opinions and your sense of time passing, and conditionalizing these credences on what you genuinely learn. In other words, two kinds of information inform your later credences. There is information you gain from your innate sense of time passing, and there is genuinely learned information that makes you more informed than your previous self. I have argued that these different kinds of information should play different roles in rational updating.

Comparing updating with communication gives us a new way of understanding abnormal updating cases. For example, consider the case of Shangri La from Arntzenius 2003: a fair coin toss determines the way you travel to Shangri La. If the coin lands heads, you go on a path by the mountains; if tails, you go on a path by the sea. If you travel on the mountain path, nothing special happens when you arrive at Shangri La. But if you travel on the path by the sea, your memory is erased upon your arrival and replaced by a memory of traveling on the mountain path. Intuitively, even if you travel on the mountain path, you should have .5 credence when you get to Shangri La that the coin

landed heads. This is a case of abnormal updating: once you arrive in Shangri La, you can no longer be sure that you traveled on the mountain path, because you can no longer trust your apparent memory.

One virtue of my theory is that it easily makes sense of cases of abnormal updating. On my theory, such cases are understood by analogy to cases of abnormal communication. Cases where you forget something are like cases where you fail to hear or understand your interlocutor. Cases where you think your memory is anti-reliable are like cases where you think someone is lying to you. Cases where you are not sure whether your memory is anti-reliable are like cases where you are not sure whether someone is lying. In each instance, our theory of the latter sort of case may inform our theory of the former. Helpfully, theories of normal communication can be naturally extended to cover cases of lying. For instance, if you are certain that your interlocutor is lying to you, then intuitively you should come to believe the negation of whatever she says. If you merely have some credence that your interlocutor is lying, you should believe a weighted compromise of whatever she says and its negation. Finally, consider the following case: you ask John whether he is a spy. You are certain that John will tell you the truth if he is not a spy and that he will lie if he is a spy. Here you should believe a weighted compromise of John's assertion and its negation, where your credence in his assertion is informed by your independent evidence about whether he is a spy. The case of Shangri La is just the same. Once you arrive at Shangri La, you are certain that your memory of your journey is reliable if you traveled by the mountain path and that it is anti-reliable if you traveled on the path by the sea. So when you consult your recent memory to see if you traveled on the path by the sea, it is as if you are consulting John to see if he is a spy. As in the communication case, your credence in the answer should be informed by your independent evidence, namely your trustworthy memories from before your journey.

The theory of updating I defend is more modest than some alternative theories. I accept the information you get from your previous self as a primitive input in updating. Since I have not given a general prescription for generating this information, my theory is best understood as a framework within which more detailed updating procedures can be developed. I also accept as primitive the distinction between information you gain from your sense of time passing, and genuinely learned information. Fortunately, we have sufficient intuitive grasp of this distinction for my theory to issue verdicts about particular cases, and I can argue for my theory on the grounds that it yields better verdicts about such cases than competing theories. Other theories of updating generally do not distinguish the kinds of constraints on

credences that are inputs to updating. In particular, other theories generally do not recognize information you gain from your sense of time passing as a primitive input to an updating procedure. They simply accept as an input that your later self meets certain conditions, such as being certain of the *de se* proposition that it is 5:00, without recording whether you arrived at this certainty by getting additional evidence or by independently keeping track of how much time had passed. For instance, Titelbaum 2008 calls the inputs to his updating procedure "extrasystematic constraints," and says only that they "represent rational requirements derived from the specific details of the story being modeled" (560).

In order to illustrate how my theory works and how it differs from less modest theories, I will conclude by discussing one more case in detail.[10] Continuing the Disney princess tradition inaugurated by Elga 2000, suppose you are a mermaid and you have been given the chance to live as a human for three days. On land you lose track of time, so you are unsure whether it is Thursday, Friday, or Saturday. Say you have 1/4 credence that it is Thursday, 1/4 credence that it is Friday, and 1/2 credence that it is Saturday. Suppose that you go to sleep, and immediately upon waking up the next day, you realize that it is not yet Sunday. Intuitively, you should then have 1/2 credence that it is Friday, and 1/2 credence that it is Saturday.

The theory I defend yields this verdict. Suppose that instead of waking up to learn that it is not yet Sunday, you wake up in a black box. One day ago, you had 1/4 credence in the *de dicto* proposition that you would have used 'today is Thursday' to convey, namely that it was Thursday.[11] If you wake up in a black box, you should still have 1/4 credence in this proposition (§4.1). Furthermore, conditional on the proposition that it was Thursday, you are currently certain that it is Friday. So you must currently have at least 1/4 credence that it is Friday (§4.2). Similarly, you must have at least 1/4 credence that it is Saturday, and 1/2 credence that it is Sunday. So in a black box case, your credences about what day it is when you wake up should simply be shifted forward by one day. In the mermaid case, you should update by conditionalizing these shifted credences on what you genuinely learn when you wake up: that it is not Sunday (§4.3). Hence my theory confirms our intuition that on waking up in the mermaid case, you should have 1/2 credence that it is Friday and 1/2 credence that it is Saturday.

---

[10] See Arntzenius 2003 for a structurally similar example.

[11] I adopt the following conventions throughout this section: 'that it is Thursday' refers to a *de se* proposition, namely the set of centered worlds centered on Thursday, and 'that it was Thursday' refers to the *de dicto* proposition you would have used 'today is Thursday' to convey before waking up.

Other theories have more trouble yielding this verdict. Titelbaum 2008 gives a theory that is similar to mine in some respects. But the mermaid case presents a problem for Titelbaum. This problem is useful: it distinguishes Titelbaum's theory from mine, and illustrates the importance of distinguishing the kinds of constraints on credences that are inputs to updating. Titelbaum starts by endorsing a limited version of conditionalization that applies to any sort of algebra P over which your credences are defined:

(LC)   If you do not become uncertain of any proposition in P, then update your credence distribution over P by conditionalizing on the strongest proposition in P of which you become certain.

As Titelbaum recognizes, (LC) has counterintuitive verdicts. In particular, when you wake up in the mermaid case, you immediately learn that it is not Sunday. So you do not lose certainty in any proposition in the algebra (P1) generated by the following *de se* propositions:

(P1)   that it is Thursday, that it is Friday, that it is Saturday, that it is Sunday

(LC) says that when you wake up, you should update your (P1) credences by conditionalizing on the strongest (P1) proposition you learn: that it is not Thursday. But if you updated this way, you would end up with the wrong credences: 1/3 credence that it is Friday, and 2/3 credence that it is Saturday. The central insight illustrated by this counterexample to (LC) is that conditionalizing in order to update your credences in some propositions can get you into trouble, even when you do not lose certainty in any of those propositions. Preserving certainties is just one bad effect of updating *de se* credences by conditionalization. Failing to appropriately shift ratios of credences in *de se* propositions is another.

To ward off potential counterexamples like the mermaid case, Titelbaum qualifies (LC) by introducing a modeling rule: "ultimately we should trust the verdicts of the model whose language is a superset of the languages of all the models we have tried for a story" (580). The qualified (LC) entails constraints on how you update your credence distribution over a given algebra only if principles entail the same constraints on how you update your credence distribution over any extension of that algebra. This forestalls a counterintuitive verdict in the mermaid case. Say we extend (P1) with the *de se* proposition that one day has not passed since the first day of the mermaid case. Since you lose certainty in this proposition on waking, (LC) and other principles do not entail that your credence distribution over the extended algebra should give 1/3 credence to the proposition that it is

Friday. Hence the qualified (LC) does not entail the wrong verdict about the mermaid case.[12]

The real problem for Titelbaum arises because the qualified (LC) does not entail the right verdict about the mermaid case, either. Titelbaum does not address this problem in detail, but the following remarks contain his proposed solution:

> If the modeling language of M^+ [...] contains context-insensitive truth-value equivalents for each context-sensitive sentence at each time in the time set — we can construct a reduction of M^+ different from M whose language represents only the context-insensitive claims represented in M^+. This model will yield diachronic verdicts by (LC), and since it is a proper reduction of M^+ those verdicts can be exported back to M^+ by (PE). (579)

Here is my reconstruction of how Titelbaum aims to derive the correct verdict about how you should update in the mermaid case. In the mermaid case, your credences are defined over algebras generated by the following propositions:

(P2)   that it was Thursday, that it was Friday, that it was Saturday

(P3)   that it is Thursday, that it is Friday, that it is Saturday, that it is Sunday, that one day has not passed, that it was Thursday, that it was Friday, that it was Saturday

(P3) is a *proper expansion* of (P2) in the following sense: for every (P3) proposition, you can always find some (P2) proposition such that you are certain that it has the same truth value as that (P3) proposition. Titelbaum endorses the following principle:

(PE)   If rationality constrains your credence distribution over P, and P' is a proper expansion of P for you, then update your credence distribution over P' according to the same constraints.

And so it appears that Titelbaum can derive the correct verdict about the mermaid case.[13] First, note that on waking, you do not lose certainty in any proposition in (P2). So (LC) says that you should update your

---

12   For brevity I have omitted a detailed defense of this claim, and I have made other simplifications in recounting Titelbaum 2008 where these simplifications are incidental to my arguments. In personal communication, Titelbaum confirms that this is his intended shield against counterexamples like the mermaid case.

13   Following the passage quoted above, let M be (P1), let M^+ be (P3), and let the "reduction of M^+ different from M whose language represents only the context-insensitive claims represented in M^+" be (P2).

credence distribution over (P2) by conditionalizing on the information you learn, namely that either it was Thursday or it was Friday when you went to sleep. On the first day, you started with 1/4 credence that it was Thursday and 1/4 credence that it was Friday. So after conditionalizing, you should have 1/2 credence that it was Thursday. Second, since (P3) is a proper expansion of (P2), (PE) says that you should update your credence distribution over (P3) in the same way, by coming to have 1/2 credence that it was Thursday, and hence 1/2 credence that it is Friday. Finally, Titelbaum's modeling rule entails that since (P3) is an expansion of (P1), you should also update your credence distribution over (P1) by coming to have 1/2 credence that it is Friday.

The problem with this chain of reasoning is that it works in both directions. If the above reasoning is acceptable, then we should also be able to reason as follows: by (LC), you should update your credence distribution over (P1) by coming to have 1/3 credence that it is Friday. Since (P3) is a proper expansion of (P1), (PE) says that you should update your credence distribution over (P3) in the same way. Hence you should update your credence distribution over (P3) by coming to have 1/3 credence that it was Thursday. By the modeling rule, you should update your credence distribution over (P2) in the same way. Hence you should update your credence distribution over (P2) by coming to have 1/3 credence that it was Thursday and 2/3 credence that it was Friday, rather than becoming equally confident of these propositions.

As far as Titelbaum's theory is concerned, the algebras (P1) and (P2) stand in symmetric relations to the larger algebra (P3). So Titelbaum faces a dilemma. If (LC) applies to your credence distribution over (P2), then it should apply to your credence distribution over (P1), and (LC) will yield the wrong verdict about how to update that credence distribution. But if (LC) does not apply to your credence distribution over (P2), then Titelbaum cannot use its verdict about (P2) to derive the right verdict about how to update your credence distribution over (P1). Titelbaum has not demonstrated that he can prevent (LC) from generating counterintuitive consequences, without also preventing (LC) from generating the right verdicts about updating in particular cases. This problem with his theory is very wide in scope. In any natural case of updating, you will lose your certainty that some small amount of time has not passed. So the unqualified version of (LC) will not apply to your entire credence distribution. Instead we must always figure out how constraints on coarser credences to which (LC) applies will induce constraints on your real credences. If we do not know which coarser credences are relevant to constraining your real credences, there will be no way to say what your real credences should be.

This problem for Titelbaum's theory highlights an important feature of rational updating. (P1) contains *de se* propositions about what day it is, while (P2) contains *de dicto* propositions about what day it was on a particular occasion. Titelbaum has trouble with the mermaid case because he allows that shifts in either (P1) or (P2) credences may direct your updated credence distribution. The mermaid case illustrates that your opinions about these propositions play different roles in updating. Intuitively, your current credences about what day it is should be informed by your previous credences about what day it was, and by your current information that it is not Sunday. This means that your opinion that one day has passed is not on a par with your information that it is not Sunday. Instead your opinion about how much time has passed plays a special role, namely determining how your previous (P2) credences should inform your current (P1) credences. The moral of the mermaid case is that a theory of updating should distinguish your credences about what day it was, what day it is, and how much time has passed. This is what my theory does. Once we recognize that different credences inform your updated credence distribution in different ways, our theory will naturally yield intuitively correct verdicts about cases of rational updating.[14]

References

Arntzenius, Frank. (2003) "Some Problems for Conditionalization and Reflection." *Journal of Philosophy*, vol. 100: 356–70.

Chalmers, David. (2002) "On Sense and Intension." *Philosophical Perspectives*, vol. 16: 135–82.

——. (2003) "The Nature of Narrow Content." *Philosophical Issues*, vol. 13: 46–66.

Diaconis, Persi & Sandy Zabell. (1982) "Updating Subjective Probability." *Journal of the American Statistical Association*, vol. 77: 822–30.

Egan, Andy. (2005) "Epistemic Modals, Relativism, and Assertion." *Philosophical Studies*, vol. 133 (1): 1–22.

Elga, Adam. (2000) "Self-Locating Belief and the Sleeping Beauty Problem." *Analysis*, vol. 60: 143–47.

von Fintel, Kai. (2005) "*LSA 311: Lecture 11*." Handout from LSA 2005, Pragmatics in Linguistic Theory.

Kaplan, David. (1989) "Afterthoughts." In *Themes from Kaplan*, Joseph Almog, John Perry & Howard Wettstein, editors, 565–614. Oxford University Press, Oxford.

Katsuno, Hirofumi & Alberto Mendelzon. (1991) "On the Difference Between Updating a Knowledge Base and Revising it." *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, 387–94.

Lakoff, George. (1972) "Linguistics and Natural Logic." In *Semantics of Natural Language*, Donald Davidson & Gilbert Harman, editors, 545–665. Reidel, Dordrecht.

Lewis, David. (1979) "Attitudes *De Dicto* and *De Se*." *Philosophical Review,* vol. 88: 513–43.

Morgan, Jerry. (1970) "On the Criterion of Identity for Noun Phrase Deletion." *Proceedings from the Annual Meeting of the Chicago Linguistics Society*, vol. 6.

Percus, Orin & Uli Sauerland. (2003) "Pronoun Binding in Dream Reports." In *Proceedings of NELS 33*, M. Kadowaki & S. Kawahara, editors, 265–83. GLSA, Amherst.

Perry, John. (1977) "Frege on Demonstratives." In *Readings in the Philosophy of Language*, Peter Ludlow, editor, 693–714. MIT Press, Cambridge.

——. (1979) "The Problem of the Essential Indexical." *Noûs*, vol. 13: 3–21.

Quine, W. V. O. (1969) "Propositional Objects." In *Ontological Relativity and Other Essays*, 139–160. Columbia University Press, New York.

Ramsey, F. P. (1931) "General Propositions and Causality." In *The Foundations of Mathematics*. Routledge & Kegan Paul, London.

Stalnaker, Robert. (2008) *Our Knowledge of the Internal World*. Clarendon, Oxford.

Stephenson, Tamina. (2009) *Towards a Theory of Subjective Meaning*. Ph.D. thesis, Massachusetts Institute of Technology.

Swanson, Eric. (2006) "Interactions with Context." Ph.D. thesis, Department of Linguistics and Philosophy, MIT.

Titelbaum, Michael. (2008) "The Relevance of Self-Locating Beliefs." *Philosophical Review*, vol. 117 (4): 555–606.

Yalcin, Seth. (2007) "Epistemic Modals." *Mind*, vol. 116: 983–1026.