

The Demonstrative Model of first-person thought

Daniel Morgan

Published online: 26 September 2014

© The Author(s) 2014. This article is published with open access at Springerlink.com

Abstract What determines the reference of first-person thoughts—thoughts that one would express using the first-person pronoun? I defend a model on which our ways of gaining knowledge of ourselves do, in much the way that our ways of gaining knowledge of objects in the world determine the reference of perceptual demonstrative thoughts. This model—the Demonstrative Model of First-Person Thought—can be motivated by reference to independently plausible general principles about how reference is determined. But it faces a serious objection. There seems to be an explanatory rival to it with a great deal of plausibility. The rival is The Simple Rule Model, which says that first-person thoughts are governed by the simple rule that any token of one will be about the person whose token it is. I pinpoint a crucial unclarity in The Simple Rule Model, hinging on how we understand the notion of a rule, and argue that no version of the Simple Rule Model is both plausible and a genuine explanatory rival to the Demonstrative Model. I also provide an argument that the Demonstrative Model is extensionally adequate.

Keywords First-person thought · Token-reflexive rule · Knowledge-maximization · Introspection · Quasi-proprioception

Suppose I think “I am hungry”—i.e. I think a thought that I would express by uttering that sentence. The object of that thought will be myself. Why is that? What determines the reference of our first-person thoughts—the thoughts that we express using sentences that contain the first-person? On the view defended here, our ways of gaining knowledge of ourselves do. We have a multitude of very rich ways of gaining knowledge of ourselves—e.g. proprioception, kinaesthesia, introspection.

D. Morgan (✉)
University College, London, UK
e-mail: daniel.morgan@ucl.ac.uk

The view defended is that these ways of gaining knowledge are what enable us to latch onto ourselves in first-person thought.

An analogous view about the mechanisms for reference-determination is highly plausible for some other kinds of thought. Suppose I am looking directly at a bottle placed before me on a table. I think “That bottle is small”. We may stipulate that I do not possess any name that refers to the bottle, and do not use any descriptions it uniquely satisfies to single it out. Nevertheless, I can in fact single the bottle out. Intuitively, that is because vision—a very rich way of gaining knowledge of objects in one’s environment—enables me to latch onto the bottle in thought. On the view about first-person thought defended here, it is similarly the case that when I think a first-person thought, my various ways of gaining knowledge of the object my thought is about—myself—are what enable me to latch onto that object in thought.

Just so as to have a label, I shall refer to the view I defend about how the reference of first-person thoughts is determined as *The Demonstrative Model*. The label is natural since the view implies that the story about how the reference of first-person thoughts is determined is parallel to the most natural story to tell about how the reference of at least one kind of thought that can be expressed using a demonstrative expression—our example is a visual demonstrative thought—is determined. But it is no part of what *The Demonstrative Model* claims that “I” is really a demonstrative expression (perhaps meaning the same as “this body” or “this mind”), or that first-person thoughts are really a kind of demonstrative thought. No assimilation of one kind of thought to a different kind of thought is intended. Relatedly, it would not matter for my purposes if it somehow turned out that, actually, the right story to tell about how the reference of thoughts like my “That bottle is small” thought is determined does not mention vision. My focus here is just on the question of how the reference of first-person thoughts is determined. The comparison between first-person thoughts, as I think we should conceive of them, and certain demonstrative thoughts, as it highly natural to conceive of these, is purely an expository device.

In Sect. 1, I provide some initial motivation for my account. *The Demonstrative Model* addresses a question about how the reference of a particular kind of thought is determined. To answer this kind of question we need to appeal to some general principle about how thoughts’ reference is determined. A variety of such principles have been proposed. I argue that any of the most plausible candidate principles can be used to make a *prima facie* case for *The Demonstrative Model*. That means that the defence of *The Demonstrative Model* need not rest on the details of how delicate, and highly general, issues about reference-determination are resolved.

In Sect. 2, I address the objection that that there is an alternative, much less contentious model available of what determines the reference of first-person thoughts. This alternative model—*The Simple Rule Model*—says that first-person thought are governed by the simple rule that any token of one will be about the person whose token it is, i.e. the person who produces the token. If *The Demonstrative Model* is inconsistent with *The Simple Rule Model*, the argument goes, then *The Demonstrative Model* should be rejected. But if *The Demonstrative Model* is consistent with *The Simple Rule Model*, then it is redundant. So *The Demonstrative Model* is at best redundant. *The Simple Rule Model* occupies such a

hegemonic place in the literature on first-person thought that this kind of argument can seem like a devastating objection to *The Demonstrative Model*.¹

My response to it is to pinpoint a crucial unclarity in *The Simple Rule Model*. We can distinguish a robust and a non robust way of understanding what the model says, corresponding to a robust and non robust way of understanding the notion of a rule that it invokes. On the robust notion of a rule, what the model says is not plausible. On the non-robust notion of a rule, what the model says is plausible. But it fails to address the explanatory question addressed by *The Demonstrative Model*. So it turns out that there is no version of *The Simple Rule Model* that is both plausible and explanatory. One large block to accepting *The Demonstrative Model* is thus dislodged.

In Sect. 3, I address the question of whether *The Demonstrative Model* is extensionally adequate, focusing on two kinds of influential problem case in which *The Demonstrative Model* appears to conflict with the claim that any token first-person thought will be about the person whose token it is.² One response to such cases is to deny that the claim that any token first-person thought will be about the person whose token it is constitutes an extensional adequacy constraint on an account of first-person thought. The most influential defender of *The Demonstrative Model*, Gareth Evans, pursued this response. He insisted that it was possible, in empirically remote but metaphysically possible circumstances, for first-person thought tokens to fail to refer to anything at all.³

In my view, this is the wrong response. *The Demonstrative Model* will only be believable if it can give a treatment of the problem cases that enables us to respect the constraint that any token first-person thought will be about the person whose token it is. I show that giving such a treatment is possible. In giving so, it is necessary to accept, pace Evans, that introspection—which we can think of, minimalistically, as our direct way of gaining knowledge of our mental properties—can play exactly the same kind of role in fixing the reference of first-person thoughts as, for example, proprioception and kinaesthesia do. In appealing to introspection, we crucially—and, I argue, legitimately—expand *The Demonstrative Model*'s set of resources for dealing with problem cases. With the expanded set of resources in place, *The Demonstrative Model* is seen to have more plausibility than Evans's pioneering defense of it did.

I think a fair, if slightly startling, summary of the attitude amongst contemporary theorists of the first-person to *The Demonstrative Model* is that it is a perennially attractive position for which a decisive refutation exists—so, a siren-song position. In my view, the felt attractiveness of the position is an expression of dissatisfaction with *The Simple Rule Model* alternative. Section 2 articulates a justification for this

¹ The argument is central in Campbell (1994). The idea that first-person thoughts are governed by a simple rule is also prominent in O'Brien (2007) and Peacocke (2008). It is the transposition to the domain of thought of the Kaplanian orthodoxy according to which the first-person in language is governed by the rule that any utterance of it will refer to the utterer. Nothing I say in this paper is incompatible with this being the right thing to say about the first-person in language.

² The problem cases are discussed in Campbell (1994), and developed further in O'Brien (2007).

³ For Evans's defense of *The Demonstrative Model* see Evans (1982, pp. 205–266).

dissatisfaction by explaining why there is no version of *The Simple Rule Model* that is both plausible and explanatory. The supposedly decisive refutation of *The Demonstrative Model* is that it cannot respect the extensional adequacy constraint that any token first-person thought will be about the person whose token it is. Section 3 shows that this supposedly decisive refutation depends on underestimating *The Demonstrative Model's* resources for dealing with problem cases. Sections 2 and 3 thus jointly constitute a pincer-movement which, if successful, tips the balance decisively against *The Simple Rule Model* and in favor of *The Demonstrative Model*.

1 Reference-determination principles and *The Demonstrative Model*

Suppose that a subject is looking at a bottle on a table and this leads them to think “That bottle is small”. Why is the subject’s thought about just that bottle and no other? “Because that bottle is the bottle the subject is looking at” may be a perfectly good answer to this question. But there is a natural follow up question. Why is the thought about the bottle the subject is looking at? What general principle about how thoughts’ reference is determined is there that illuminates why the fact that a particular bottle is the thing being looked at can contribute to making it, in addition, the thing being thought about?

Giving satisfying answers to such reference-determination questions has proven difficult. There is a surfeit of candidate answers each of which has some plausibility. Perhaps there is some general acquaintance constraint on reference, and the reason the subject’s thought is about the bottle is that looking at the bottle is a way of being acquainted with it.⁴ Perhaps there is a general constraint on reference according to which reference tends to maximize knowledge, and looking at the bottle contributes to making it the thing thought about because looking at something is a way of gaining knowledge of that thing. Perhaps there is a general constraint on reference according to which reference tends to maximize justification, internalistically construed, and looking at the bottle contributes to making it the thing thought about because looking at something is a way of gaining justification, internalistically construed, of that thing.⁵ It is not easy to adjudicate between these principles. Many of the examples one might give to motivate one’s favorite generalization about how reference is determined in fact fall just short, failing to provide more support for one’s favorite principle than they do for nearby rivals. On the other hand, it is unappealing to reject reference-determination questions about thoughts as unanswerable. That would be to accept that it is just a brute fact that certain thoughts have the reference they have, and that there is nothing further to be said about what makes this the case.

Fortunately, from the point of *The Demonstrative Model* many of the controversies about how reference-determination questions are answered do not

⁴ This is the kind of principle Evans accepts, and uses to motivate *The Demonstrative Model*. For criticism of acquaintance requirements, see Manley and Hawthorne (2012).

⁵ For a defense of the knowledge-maximization principle see Williamson (2007). Williamson also discusses and criticizes the internalist justification-maximization principle. For discussion and criticism of Williamson on knowledge-maximization see McGlynn (2012) and Martin (2009).

matter. For any one of the principles described in the previous paragraph may be used to make a *prima facie* case for *The Demonstrative Model*.

Suppose the reason that looking at a bottle enables reference to it in thought is that looking at something is a way of being acquainted with it. That raises the question of why looking at something counts as a way of being acquainted with it. Presumably the answer is that it does because vision is a rich and direct epistemic relationship. If so, then proprioception should also count as a way of being acquainted as well. For it too is an extremely rich and direct epistemic relationship. But if proprioception is a way of being acquainted with something, then it seems eminently suited to being a relationship in virtue of which one satisfies the acquaintance requirement when one thinks about oneself using a first-person thought like “I have crossed knees”. After all, first-person thoughts are always about oneself and, at least at first glance, the only person one ever has proprioceptive awareness of is oneself. Exactly the same case could be made for introspection’s being a way of being acquainted with oneself. It may be in the end that *The Demonstrative Model* must be rejected. But, if the principle of acquaintance is correct, there is a *prima facie* case for it.

Suppose instead that the reason that looking at a bottle enables reference to it in thought is that a knowledge-maximization principle is correct and looking at something is a way of gaining knowledge of that thing. On this supposition, being a source of knowledge is precisely what qualifies vision to fix the reference of visual demonstrative thoughts’ reference. But proprioception and introspection have precisely the same qualification to fix the reference of first-person thoughts—they too are sources of knowledge. And if proprioception and introspection are what fix the reference of first-person thoughts, that means that *The Demonstrative Model* is correct. It may be in the end that *The Demonstrative Model* must be rejected. But if the knowledge-maximization principle is correct, there is a *prima facie* case for it.

Suppose instead that the reason that looking at a bottle enables reference to it in thought is that a justification-maximization principle is correct and looking at something is a way of gaining justification of that thing. On this supposition, being a source of justification is precisely what qualifies vision to fix the reference of visual demonstrative thoughts’ reference. But proprioception and introspection have precisely the same qualification to fix the reference of first-person thoughts—they too are sources of justification. And if proprioception and introspection fix the reference of first-person thoughts, that means *The Demonstrative Model* is correct. It may be in the end that *The Demonstrative Model* must be rejected. But if the justification-maximization principle is correct, there is a *prima facie* case for it.

Having provided some *prima facie* motivation for *The Demonstrative Model*, I turn to the arguments that have led many theorists of first-person thought to reject it.

2 The Demonstrative Model and the Simple Rule Model

We can use an intuitive contrast between first-person thoughts and visual demonstrative thoughts as a way of introducing and motivating the dominant alternative to *The Demonstrative Model*.

The intuitive contrast is this. Even in advance of any particular story about how vision enables us to latch onto objects in thought—e.g. in advance of a story that appeals to knowledge-maximization or to justification-maximization or to a principle of acquaintance—we know that this is something vision does. We can describe cases in which there is really nothing but the fact that a subject can see an object that could explain how they are latching onto that object in thought—cases in which the subject has no name to pick out the object, and does not rely on any description in picking it out. Merely taking as fixed that a particular thought is a thought of a kind that can be expressed by a demonstrative expression, and that it is the thought of a particular subject, nothing at all follows about which object the thought will be about. So there is always an explanatory question about why the thought has the object that it does, and not some different object. In some cases, vision—or perhaps some more demanding vision-involving relationship, like the relationship of visual attention—will be all that one could reasonably appeal to in answering that question.

Contrast this with the case of first-person thought. Merely taking for granted that a particular thought can be expressed by the first-person, and that it is a thought of a particular subject, something does follow about which object the thought is about. It follows that the thought is about the subject themselves. This is guaranteed by what seems to be the indisputable fact that any token first-person thought will be about the person whose token it is, i.e. the person who produces the token. Since this fact alone determines who a particular first-person thought will be about, there is no *need* for, for example, proprioception to do the kind of singling out work that vision does in the case of visual demonstrative thoughts.

The metaphorical picture we have, in the case of a visual demonstrative, is that of a first arrow shooting out at an object—this arrow represents vision. The first arrow is then followed by a second arrow—this arrow represents thought. The second arrow tracks the first. But, in the first-person case, we already know what trajectory the arrow representing thought must describe. It must describe a loop. So, the suggestion is, there is no need to regard the arrow representing thought as tracking the progress of any first arrow—e.g. one representing proprioception or introspection. If there happens also to be a first arrow that describes a loop that is interesting. But we could say what the right trajectory for the second arrow is without reference to it, or to anything else.

John Campbell has forcefully argued that the fact that any first-person thought will be about the subject undermines the need for anything like *The Demonstrative Model*. He writes:

Only with reluctance should we take this step, which can seem so easy and indeed inescapable, of supposing that there is a use of the first person on which its reference is fixed not by the token-reflexive rule [the rule that any token of the first person will be about the producer of that token] but in some other way. Many of the most distinctive phenomena involving the first person are straightforwardly explained by its being governed by the rule. Once we leave the rule behind, problems swarm upon us. Most immediately and simple is the question of whether the alternative method of reference-fixing, whatever it is,

is guaranteed to yield the same results as the token-reflexive rule in what references it finds for particular uses of the first person. If there is this guaranteed coincidence, then it is not apparent what advantage there can be in shifting to the new method. If the coincidence is not guaranteed, then we have opened up the possibility of someone using the first person to refer to someone other than himself. But this would not be recognizable as a use of the ordinary first person. (1994, p. 124)

The token-reflexive rule (or “simple rule”, as Campbell also calls it, and as I am following him in calling it) Campbell explicitly speaks of here applies to the first person—a particular linguistic device. But it is clear that Campbell thinks that something analogous holds for first-person thoughts. The analogous rule for first-person thoughts is the rule that any token first-person thought will be about the producer of that token. Campbell in fact allows the categories of thought and language to blur somewhat. But it turns out to be crucial that we not allow these two categories to blur, if we are properly to evaluate the choice between *The Demonstrative Model* and *The Simple Rule Model*—the model that says that the reference of first-person thoughts is fixed by the rule that any token first-person thought will be about the producer of that token.

The claim that the word “I” is governed by the simple rule is fairly plausible (though it is not uncontroversial). It is fairly plausible that the word “I” is governed by the simple rule because it is fairly plausible that the word “I” is governed by a *convention* whose content is the content of the simple rule. Rules in language just are conventions.

The exactly parallel suggestion, at the level of thought, would be that first-person thoughts are governed by the convention that any token first-person thought token will be about the subject who produces that token. But that just seems wrong. Thoughts, unlike words, are not typed purely syntactically. Thoughts do not get their objects in virtue of being governed by conventions. To be a first-person thought is already to have a certain semantic profile.

It might be replied that the objection assumes an overly robust notion of a rule. On a less robust understanding, to claim that there is a rule to the effect that any token first-person thought will be about the subject who produces that token is not to invoke a convention or anything similar. It is just to claim that the pattern of reference of first-person thoughts can be *correctly described* by saying that any token first-person thought will be about the subject whose thought it is.⁶

⁶ Might there be a third option for how to conceive of rules in thought—a conception of them on which they are robust but on which they are also not to be identified with conventions? A more comprehensive discussion would take into account the defense of the idea that first person thought is governed by the Simple Rule—and that thoughts (or, at any rate, concepts) in general are governed by rules—that is developed in Peacocke (2008). However, it is unclear what conception of rules Peacocke is offering. It may be that he is treating the notion of a rule as an explanatory primitive. But the question still arises of why we must accept that thoughts are governed by rules, construed as primitive. It is not just part of commonsense that they are. Peacocke argues that the claim that first-person thoughts are governed by the Simple Rule can earn its keep by explaining other first-person phenomena—e.g. the phenomenon of what he calls “fully self-conscious uses” of the first-person. But, on examination, the detailed explanations of first-person phenomena Peacocke gives seem to require only that the Simple Rule be true—i.e. that it be

At this point though, the defender of *The Demonstrative Model* can turn the tables on *The Simple Rule Model*. Campbell's thought was that, if an alternative reference-fixing story to *The Simple Rule Model* is compatible with *The Simple Rule Model*, then there is "no advantage to shifting to the new method". But, on the non robust way of understanding *The Simple Rule Model* we have been forced to consider, this becomes far less clear. The point of *The Demonstrative Model* is to *explain* why first-person thoughts have the references that they do. If we agree that it is correct to say that any token first-person thought will be about the subject who produces that token, then the task for *The Demonstrative Model* will be to explain why it is that any token first-person thought will be about the subject who produces that token. So, *The Demonstrative Model* will aim to explain precisely what *The Simple Rule Model* takes for granted. If *The Demonstrative Model* is successful in this aim it should be preferred to *The Simple Rule Model*. It should be preferred even if both models make exactly the same predictions about which things "particular uses of the first person" refer to.

The task of the next section is to work out whether *The Demonstrative Model* really can make the same predictions about which things "particular uses of the first person" refer to as *The Simple Rule Model* does. If it cannot, then *The Demonstrative Model* will not be acceptable—since it will not even be extensionally correct. But it is worth pointing out that, even before we have settled that question, an important conditional conclusion has been established.

The question of how our thoughts engage with things that exist independently of them has always been regarded as philosophically pressing. It is also an important constraint on how we conceive of the mental. To take an example of the philosopher who is closest at hand here, Campbell himself in other work argues that perceptual consciousness must be what enables perceptual demonstrative thought to engage with the world, and he argues that the first task for any theory of perceptual consciousness is to make it intelligible how it can play this role.⁷ But first-person thought is equally a kind of thought in which one engages with something—oneself—that is independent of the thought in which that engagement consists. So, on the face of it, there seems to be just as pressing a question about how first-person thought manages to engage with its objects.

It turns out though, once we have distinguished first-person reference in thought and first-person reference in language, that *The Simple Rule Model* does not engage with this question. Properly understood, *The Simple Rule Model's* content is exhausted by the observation that any token first-person thought will be about the producer of that token. So the model does not address the explanatory question of why first-person thought has the pattern of reference that it has. It is not obvious though why shirking that explanatory question should be thought any more acceptable in the case of first-person thought than it would be in the case of

Footnote 6 continued

the case that any first-person thought will be about the thinker. They do not seem in addition to require that this be true *because* there is some robust rule whose content it is that any first-person thought will be about the thinker.

⁷ See Campbell (2002).

perceptual demonstrative thought. So, unless *The Demonstrative Model* turns out to be correct—or unless some hitherto unsuspected third option does, which seems unlikely—there seems to be an embarrassingly central gap in our understanding of the engagement between mind and world.

This adds urgency to the question of whether *The Demonstrative Model* can be defended against the crucial charge that it makes the wrong predictions about which things certain first-person thoughts refer to.

3 Is the demonstrative model extensionally correct?

Lucy O'Brien pithily sums up what she takes to be the major problem for *The Demonstrative Model* as follows:

“I” has a special propensity to refer and to stick with its object of reference, even though the waves of information might seem to want to buffet it this way and that. (1995, p. 241)

O'Brien's observation gestures at two kinds of problem case for *The Demonstrative Model* which she and Campbell both discuss—cases in which *The Demonstrative Model* might appear to make the wrong prediction about which thing a first-person thought is about.⁸ The first kind of case is the Anscombe case in which many of one's ordinary routes to knowledge of oneself are suspended, because one is inside a sensory deprivation tank.⁹ The second kind of case is one in which one is wired up so that experiences of a sort that are generally caused by the condition of one's own body are in fact caused by the condition of someone else's body—e.g. a case of “quasi-proprioceptive” experience in which I am wired up to someone else's knees.¹⁰ I will deal with the two kinds of problem case in turn, and, in doing so, flesh out the version of *The Demonstrative Model* I think we should accept in a bit more detail.

3.1 Anscombe's sensory deprivation tank

Anscombe famously noted the following possibility:

Now imagine that I get into a state of ‘sensory deprivation’. Sight is cut off, and I am locally anaesthetised everywhere, perhaps floated in a tank of tepid water; I am unable to speak to or touch any part of my body with any other (1975, p. 58)

Why might this case seem problematic for a defender of *The Demonstrative Model*? Well, in the sensory deprivation tank, one is receiving very little, if any, sensory

⁸ For Campbell's discussion, see (1994, pp. 125–126). For O'Brien's see (2007, pp. 38–47).

⁹ See Anscombe (1975, p. 56).

¹⁰ The notion of quasi-proprioception is modeled on Shoemaker's notion of quasi-memory. See Shoemaker (1970).

information. So, it might seem, one has no basis for knowledge of oneself at all. In that case, *The Demonstrative Model* will predict that singling oneself out with a first-person thought is impossible. But this is the wrong result. One will surely at least be able to *attempt* to think a first-person thought in the sensory deprivation tank. And—unlike in the case of visual demonstrative thoughts—it seems there is a guarantee that any attempt to think a first-person thought will succeed in netting an object. Evans denied the existence of this guarantee. He argued that it was based on an unjustified inference from the observation that any attempt to use the word “I” to refer was bound to succeed in referring. But following Evans in this denial feels like a huge cost.¹¹

There is an obvious, but often strangely overlooked, gap in the case that the sensory deprivation tank case represents even a *prima facie* problem for *The Demonstrative Model*. Granted, in a sensory deprivation tank one is receiving little, if any, sensory information. But it just does not follow that one has no basis for knowledge of oneself. After all, much of one’s knowledge of oneself is non-sensory. I can introspect in a sensory deprivation tank—that is, I can rely on my direct way of gaining knowledge of my own mental properties. A defender of *The Demonstrative Model* who appeals to a knowledge-maximization principle can insist that even inside the tank first-person thought is possible because introspection is a rich source for knowledge of oneself that one retains even inside the tank. Similarly, a defender of *The Demonstrative Model* who appeals to justification-maximization can insist that even inside the tank introspection provides one with a source of justified belief about oneself. A defender of *The Demonstrative Model* who appeals to an acquaintance requirement on reference can say that one satisfies the acquaintance requirement by introspecting one’s mental state.

The appeal to introspection is clearly a helpful expansion of the resources available to the defender of *The Demonstrative Model*. It may be wondered though whether the appeal is licit. Although he never did so explicitly, it is fairly natural to assume that Evans himself would have said that it is not.¹² What is less clear is why he might have said that, and whether there is any good reason to say it.

Evans’s own version of *The Demonstrative Model* is motivated by reference to a highly specific acquaintance requirement according to which thinking about an object requires one to *know which* thing one is thinking about. In the case of material objects, knowing which thing one is thinking about is cashed out in terms of *knowing where* that thing is. But, it can surely be pointed out, introspection does not give one any knowledge of where one is.

This doesn’t seem like a good reason to ignore introspection. First, it will have no persuasive force for a defender of *The Demonstrative Model* whose motivation differs from Evans’s (e.g. because they accept an acquaintance requirement that does not give a privileged place to knowledge of where one is, or because they accept a knowledge-maximization or justification-maximization principle rather

¹¹ See (1982, p. 252) for a statement of Evans’s position. See Peacocke (2008, pp. 103–110) for an argument that no first-person thought can fail to refer.

¹² Cassam (1997), who provides a very sophisticated discussion of the sensory deprivation tank case that is sympathetic to Evans, also never considers appealing to introspection.

than an acquaintance requirement). Second, the proposed restriction to faculties that provide knowledge of where one is is so theoretically constricting that even Evans himself did not adhere to it in actually giving his theory. For example, kinaesthesia—a faculty that enables one to tell that one is moving—is a faculty that Evans’s theory appeals to, and it seems a good thing that it does. Likewise, the faculty by which one tells that the surface of one’s skin is clammy is one that Evans’s theory appeals to, and it seems a good thing that it does. But neither of these faculties in any obvious sense provides one with knowledge of where one is.

A more general, perhaps Wittgenstein-inspired worry about the appeal to introspection is that introspection is, in some pejorative sense, not a genuine faculty. It is difficult to adjudicate this large issue, and it may be that some resolutions of it would put pressure on *The Demonstrative Model’s* appeal to introspection. Even so, there are reasons not to be overly pessimistic. The full-strength Wittgensteinian case against introspection is very strong indeed. It implies that what we would ordinarily think of as introspection based first-person ascriptions of mental states are not even candidates for knowledge, being mere expressions of the states that they seem to ascribe. Few now endorse this picture. But it is far from obvious what milder, but still substantive, claim to associated with the slogan that introspection is not a genuine faculty. Of some of the most promising candidates—e.g. the claim that it is wrong to think of introspection as a faculty that is just like vision, except that it is trained on the mind—it is far from clear why introspection couldn’t be a perfectly appropriate thing for *The Demonstrative Model* to appeal to even if it fails to be a “genuine faculty” in the defined sense.

3.2 The quasi-proprioception case

Let us say that that a subject *quasi-proprioceives* someone else’s body if and only if the following two conditions are met. (1) The subject has an experience that is subjectively indistinguishable from a genuine proprioception of their body having a certain property. (2) They are caused to have this experience by someone else’s body, not their own, actually having that property.

For example, I might be wired up to someone else’s knees so that when their knees are actually crossed, this causes me to have an experience that is subjectively indistinguishable from a genuine proprioception of my knees being crossed.

Everyone should accept that such quasi-proprioceptive experiences are metaphysically possible. Why might their possibility seem problematic for *The Demonstrative Model*? Well, in a case of quasi-proprioception, one’s experience derives from someone else’s body, not one’s own. But *The Demonstrative Model* is supposed to allow experience to play a reference-determining role in first-person thought. So, it might seem, the waves of quasi-proprioceptive experience should buffet one’s first-person thought towards the person one is actually gaining experience from. This could result in one’s first-person thought actually being about that person. Alternatively, if the other bases on which one makes first-person judgments are operating normally and buffeting one’s first-person thought in the direction of oneself, it could result in the failure of the thought to refer to anything at all. There would then be an analogy with the way a visual demonstrative thought

can fail to refer because it is anchored on visual information that, unbeknownst to the subject, derives from two different sources—e.g. two different trees whose branches have become entangled.

As O'Brien points out, neither of these results seems plausible. For both violate the constraint that any first-person thought token will be about oneself. So the defender of *The Demonstrative Model* had better be able to say something about why quasi-proprioception experiences do not buffet reference away from the subject.

The details of what they say will depend on how they motivate their model. To keep things manageable, I will consider a reply from the point of view of a defender of *The Demonstrative Model* who appeals to knowledge-maximization. It should be fairly clear that the dialectic would play out similarly for defenders of the model who appeal to a justification-maximization principle, or to a principle of acquaintance.

Critic and defender of *The Demonstrative Model* should both agree that when one judges "I have crossed knees" because one is quasi-proprioceiving the body of someone else whose knees really are crossed, that judgment (1) will be about oneself and (2) will not be knowledge. The critic may insist though out that the quasi-proprioceptive link *does* seem to put one in a position to gain knowledge of the other person. So, if we were simply assigning reference in such a way as to maximize knowledge, we would assign not oneself, but the other person, as reference. *The Demonstrative Model* (as motivated by the knowledge-maximization principle—I will also leave this implicit in what follows) therefore seems committed to predicting that the thought will be buffeted towards the other person. A defender of *The Demonstrative Model* will naturally attempt to avoid this prediction. But the critic will be suspicious that this attempt can only succeed if the constraint that any first-person thought will be about oneself is allowed to trump the unprejudiced deliverance of *The Demonstrative Model*.

I think this is a pressing objection, and it makes the quasi-proprioceptive case the harder of the two problem cases. There are different things a defender of *The Demonstrative Model* can say, a more general and a more specific thing.

The more general—and less ambitious—thing to say is that, the existence of problem cases does not establish that there is anything *especially* problematic about the idea that knowledge maximization allows our ways of gaining knowledge of ourselves to fix the reference of *first-person* thoughts. For there are analogous problem cases for the idea that knowledge maximization enables vision to fix the reference of visual demonstrative thoughts.^{13,14} Most simply, not all of our visual demonstrative thoughts are knowledge. If knowledge-maximization is to explain how vision fixes the reference of these thoughts, some fancy footwork will be required. If the fancy footwork works out in the visual demonstrative case, analogous fancy footwork may work in the first-person case. More generally, in

¹³ For example, Martin (2009) offers examples in which a knowledge-maximization principle seems to issue the wrong verdict on which object a visual demonstrative thought is about.

¹⁴ Justification-maximization and acquaintance-based principles about reference determination naturally have analogous problems.

relation to almost every kind of reference-determination question, it has proved difficult to say anything that is at once substantive and plausible. Precisely because of the difficulty of achieving plausibility, the kind of answers philosophers have been canvassing in the last decade are notably less substantial than the kind they pursued in previous decades.¹⁵ It isn't obvious there is any scope for going less substantial without giving up on reference-determination questions altogether (arguably, this is what the defender of *The Simple Rule Model* does, in relation to first-person thoughts). In that case, the existence of a problem case wouldn't stop *The Demonstrative Model* from being a promising research program.

That answer will be too mealy mouthed for some. A more ambitious approach is to deny the key premise of the critic's objection: the premise that quasi-proprioceptive experiences put one in a position to gain knowledge of the other person, the person from whom they derive. Without this premise, it does not even look as though *The Demonstrative Model* favors an assignment of reference on which one's thought is about the other person.

The defender of *The Demonstrative Model* can point out that all we indisputably have in the quasi-proprioceptive case is a causal connection between one's experience and the other person. But, not every causal connection makes for knowledge. It can of course be stipulated that the causal connection in question is a highly reliable one. But, on all but the most extreme reliabilist views, not every reliable causal connection makes for knowledge either. So it isn't just obvious that the quasi-proprioceptive experience puts one in a position to gain knowledge of the other person.

The defender of *The Demonstrative Model* can go further and give a positive argument that quasi-proprioception does *not* put one in a position to gain knowledge of the other person. It does not because quasi-proprioceptive experiences do not *present* one with the other person.¹⁶ How could they while being subjectively indistinguishable from genuine proprioceptions that present one with oneself?

The critic of *The Demonstrative Model* may counter-object as follows. Perhaps quasi-proprioceptive experiences do not present one with the other person. Perhaps it follows that those experiences do not put one in a position to gain knowledge of the other person. If both of these are correct, then the original objection about the case will have been answered. But there is the following modified objection to be made. The reason that it is so plausible that quasi-proprioceptive experiences do not present one with the other person is precisely that those experiences do not have that person as their object. The reason they do not have that person as their object is that they are experiences whose contents are first-person propositions, and any experience whose content is a first-person proposition will be about oneself. But, if it is acceptable to simply take for granted that an experience with a first-person proposition as its content will be about the subject, why is it not acceptable to take

¹⁵ I am thinking here, for example, of the programs Fodor and Millikan initiated in the 1970s, which are highly substantial in that they operate under very thorough-going naturalistic restrictions. Williamson (2007) explicitly disavows such naturalistic restrictions.

¹⁶ The same thing makes it plausible that the experiences do not provide one with justification for belief about the other person, and that they do not acquaint one with the other person.

for granted that a thought with a first-person proposition as its content will be about the subject? How can it be compulsory to address the explanatory question about thoughts that *The Demonstrative Model* addresses but permissible to ignore the exactly parallel explanatory question about experiences?

The defender of *The Demonstrative Model* can make two replies. The first reply is that it isn't obvious that experiences whose natural articulation in thought uses the first-person (such as quasi-proprioceptive experiences or genuine proprioceptive experiences) must have first-person propositions as their contents. It may be that such experiences either lack contents altogether, or have contents other than first-person propositions. There are possible explanations of why the person from whom the quasi-proprioceptive experience derives is not the object of the experience other than the one suggested by the critic of *The Demonstrative Model* in the previous paragraph. For example, there is the following function-based explanation.

In general, very many things are causally involved in the production of our experiences, first-person or otherwise. Not all of these things are the objects of our experiences. Very plausibly, facts about what *functions* our sensory faculties have are relevant to settling which among the things that are causally implicated in the production of our experiences are the objects of those experiences. For example, a particular distal object I am looking at, rather than any of the countless causal intermediaries between the object and my experience, may be the object of my visual experience in part because the function of vision is to provide awareness of distal objects.

Similarly, it is highly plausible that the function for which proprioception has evolved is to provide us with awareness of our own bodies (here, it becomes relevant that there were no instances of quasi-proprioception in the circumstances of our actual evolution).¹⁷ So, quasi-proprioceptive experiences causally derive from something other than the thing of which the faculty they are exercises of has the function of providing awareness. That is a possible explanation of why they do not present one with, or put one in position to gain knowledge of, the person from whom they causally derive that does not presuppose that they have first-person propositions as their contents.

The second reply to the modified objection the defender of *The Demonstrative Model* can make is compatible with the view that quasi-proprioceptive experiences have first-person propositions as their contents. It says that, even with that being the case, *The Demonstrative Model* can still have explanatory value. A way of seeing this is to imagine that we have been persuaded that the contents of visual experiences are visual demonstrative propositions, and are therefore identical to the contents of visual demonstrative judgments.¹⁸ Even if we have been persuaded of this, and even if we do not have a glimmer of a story about how visual experiences themselves get their contents, it still seems true to say of a particular

¹⁷ See Martin (1997) for a defense of this view.

¹⁸ This is not a view of the contents of visual experience that Campbell accept. See Campbell (2002). But the point here is not intended as *ad hominem*. It is intended to bring out what makes the *Demonstrative Model* explanatory.

visual demonstrative thought—e.g. a thought I would express by uttering the sentence “That bottle is small”—that it is about the bottle it is about because that bottle is the one I can see. For example, this seems truer than saying that it is about that bottle because that bottle is the bottle I can hear, or because it is the bottle you can see, or because it is the bottle that was an especial favorite of its maker. And, more generally, it still seems right to recognize a non-empty category of visual demonstrative thoughts—a category of thoughts that have the reference they do *because* of what the subject is looking at. Similarly, the *Demonstrative Model* says that one’s first-person thoughts are about oneself *because* their reference is fixed by first-person experiences. This can be true even if we have no worked out story about how first-person experiences themselves get their contents.

The quasi-proprioception case is a difficult case for a defender of *The Demonstrative Model*. Different defenders of the model will say different things about it, depending on how they have motivated their model—e.g. by reference to a knowledge-maximization principle, or to a justification-maximization principle or to a principle of acquaintance. Even just focusing on one of these motivations—e.g. knowledge-maximization—there are different points at which the objection can be tackled. A defender of the *Demonstrative Model* just needs one of these defenses to work.

4 Conclusion

What determines the reference of first-person thoughts? The answer *The Simple Rule Model* gives to this question initially appears both crystal clear and extremely plausible. Things get murkier once we clarify that the question is a question about thought, not language, and we pause to ask what notion of rule governing thoughts the model invokes. So far, we have not seen a version of *The Simple Rule Model* that is both plausible and explanatory. Dialectically, that gives *The Demonstrative Model* its opening.

The outstanding objection to *The Demonstrative Model* is that certain problem cases show that it is not extensionally adequate. Finding the right way to deal with these problem cases is tricky enough to make it tempting to indulge in a revisionary view about what the pattern of reference of first-person thoughts is. This, in my view, is what Evans did (although he denied that what he was doing was revisionary). But we should resist the temptation—the pattern of reference of first-person thoughts is a surer thing than any philosophical theory that seeks to explain it.

Showing that *The Demonstrative Model* is extensionally adequate, even in the problem cases, is a non trivial task. Part of the difficulty is generic: *The Demonstrative Model* attempts to answer a reference-determination question, and saying something substantial and plausible about reference-determination questions tends to be difficult. Still, it is more convincing if we can say something specific to fend off the charge of extensional inadequacy. Doing so encourages us to make a few somewhat surprising argumentative moves. For example, it encourages us to give an explanatory role to introspection. It encourages us to take a particular view

about the nature of quasi-proprioception. The main pay off for making these moves is that it puts us in a position to give an explanation of the pattern of reference of first-person thought that does not involve being revisionary about what that pattern of reference is.

Another, closely related, pay-off is that we are in a position to reject what I think of as ‘first-person exceptionalism’, the idea that there is something highly exceptional about first-person thought. The idea runs through Campbell’s rich seam of work on the topic.¹⁹ According to Campbell, what is exceptional about first-person thoughts is that their conceptual role and their pattern of reference are only rather loosely connected. Visual demonstrative thoughts are the main foil he uses to bring this out. The conceptual role of visual demonstrative thoughts involves evidential sensitivity to certain experiences—visual experiences—and those very experiences also play a role in fixing visual demonstrative thoughts’ reference. By contrast, the conceptual role of first-person thoughts involves evidential sensitivity to certain experiences—experiences whose natural articulation in thought uses the first-person—but those same experiences, on Campbell’s view, play no role in fixing first-person thoughts’ reference. This observation about first-person thoughts even leads Campbell to express qualified sympathy for Anscombe’s—arch first-person exceptionalist—idea that first-person thoughts do not refer. What is right about Anscombe’s idea, in Campbell’s view, is that the ascription of reference to first-person thoughts does no explanatory work.

If this is right, then one urgently needs to answer the question of why it is that, uniquely, in the case of first-person, the ascription of reference does no explanatory work. Much of Campbell’s recent positive work on first-person thought is motivated precisely as an attempt to answer this question. If one thinks that reference just is an explanatory notion, one is likely to be pessimistic about these attempts. But if *The Demonstrative Model* is correct, the puzzle that makes them necessary is dissolved. On *The Demonstrative Model*, first-person experiences *do* fix the reference of first-person thoughts.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Anscombe, G. E. M. (1975). The first-person. In S. Guttenplan (Ed.), *Mind and language* (pp. 45–65). Oxford: Clarendon Press.
- Campbell, J. (1994). *Past, space and self*. Oxford: Clarendon Press.
- Campbell, J. (2002). *Reference and consciousness*. Oxford: Oxford University Press.
- Campbell, J. (2004). What is it to know what “I” refers to? *The Monist*, 87(2), 206–218.
- Campbell, J. (2012). On the thesis that “I” is not a referring term. In S. Prosser & F. Recanati (Eds.), *Immunity to error through misidentification: New essays*. Cambridge: Cambridge University Press.
- Cassam, Q. (1997). *Self and world oxford*. Oxford: Oxford University Press.
- Evans, G. (1982). *The varieties of reference*. Oxford: Clarendon Press.
- Manley, D., & Hawthorne, J. (2012). *The reference book*. Oxford: Oxford University Press.

¹⁹ See, for example, Campbell (1994, 2004, 2012).

- Martin, M. G. F. (1997). Self-observation. *European Journal of Philosophy*, 5, 119–140.
- Martin, M. G. F. (2009). Reupholstering a discipline: Comments on Williamson. *Philosophical Studies*, 145, 445–453.
- McGlynn, A. (2012). Interpretation and knowledge maximization. *Philosophical Studies*, 160(3), 391–405.
- O'Brien, L. (1995). Evans on self-identification. *Nous*, 29(2), 232–247.
- O'Brien, L. (2007). *Self-knowing agents*. Oxford: Oxford University Press.
- Peacocke, C. (2008). *Truly understood*. Oxford: Oxford University Press.
- Shoemaker, S. (1970). Persons and their pasts. *American Philosophical Quarterly*, 7, 269–285.
- Williamson, T. (2007). *The philosophy of philosophy*. Oxford: Blackwell Publishing.