Can There Be Indexical Thoughts?

Derek Ball
Arché, University of St Andrews

Following influential work by John Perry (2000), Hector-Neri Casteñeda (1966; 1968), David Kaplan (1977, §XVII), and others, the consensus in philosophy of mind and language is that *indexical thoughts*[1] play an essential role in our cognitive economies.[2] But the consensus is mistaken. I will argue that there can be no indexical thoughts of the sort that could account for the Perry data – not, at least, for thinkers whose minds are implemented computationally, as ours plausibly are.

The paper is structured as follows. In the first section, I describe the examples that have been used to motivate the idea that there are indexical thoughts. The second section describes the computational/representational theory of mind, and articulates two conceptions of what indexicality in thought might amount to on this theory. The next two sections argue that neither conception is tenable, and thus that there can be no indexical thoughts. I conclude by replying to some objections.

1 Perry's Data

That so many philosophers have been convinced of the importance of indexical thought is a tribute to the power of Perry's examples. Perhaps most famous among these is the messy shopper:

Messy Shopper "I once followed a trail of sugar on a supermarket floor [...] seeking the shopper with the torn sack to tell him he was making a mess. With each trip around the counter, the trail became thicker. But I seemed unable to catch up. Finally it dawned on me. I was the shopper I was trying to catch. I believed at the outset that the shopper

Draft of 11 October 2010

[1] To avoid terminological confusion, let me stipulate from the beginning that I am using the term 'thought' to pick out mental representations, rather than the contents of such representations. This differs from the terminology of Kaplan and Perry, who use 'thought' to pick out contents. Kaplan and Perry would agree that indexicality is not a matter of content. But since they hold that there are indexical modes of presentation, they hold that there are indexical thoughts in my sense of the term. For more discussion, see sections 1 and 2 below.

[2] Among the few notable dissenters are Stalnaker (1981) and Millikan (1990). I touch on Millikan's view in section 5 below.

> with a torn sack was making a mess. And I was right. But I did not believe that I was making a mess. [...] My change in beliefs seems to explain my change in behavior." (1979, p. 27)

Perry drew two morals from the case. First, beliefs that we naturally express with certain indexicals can have different behavioural consequences than truth-conditionally equivalent beliefs expressed without indexicals. Thus Perry's behavior changed when he formed the belief that he would express by saying "I am making a mess": he stopped circling the counter and went to fix his torn sack.

Second, Perry claimed that the relevant beliefs which are naturally expressed with indexicals are not equivalent to any belief expressed without indexicals (e.g. Perry (1979, p. 29); see also Perry's discussion of Rudolf Lingens in Perry (1977, p. 17)). In particular, one can know all of the 'non-indexical' information while remaining ignorant of the 'indexical' information. Thus Perry might have believed all along that Perry is making a mess, or that the only bearded philosopher in a Safeway store west of the Mississippi is making a mess, or even (on seeing his own reflection in a mirror) that *he* is making a mess, as long as he failed to realise (as he would say), "*I* am John Perry/the only bearded philosopher in a Safeway store west of the Mississippi/he (the man in the mirror)".

Perhaps the most famous example in support of this point comes from David Lewis:

Two Gods "[Two gods] inhabit a certain possible world, and they know exactly which world it is. Therefore they know every proposition that is true at their world. Insofar as knowledge is a propositional attitude, they are omniscient. Still I can imagine them to suffer ignorance: neither one knows which of the two he is. They are not exactly alike. One lives on top of the tallest mountain and throws down manna; the other lives on top of the coldest mountain and throws down thunderbolts. Neither one knows whether he lives on the tallest mountain or on the coldest mountain; nor whether he throws manna or thunderbolts [...] (The trouble might perhaps be that they have an equally perfect view of every part of their world, and hence cannot identify the perspectives from which they view it.)" (1979, p. 139)

Lewis's gods know all of the information about their world, except that information that they would express by saying "I am the god on the highest mountain" or "I am the god on the coldest mountain". So the information

that they would express using "I" is not equivalent to any information that is naturally expressed without indexicals.

Many of the famous cases involve the first person. But this is not essential. As Perry pointed out, "now" beliefs are equally apt to change behavior:

Absent-Minded Professor "[A] professor, who desires to attend the department meeting on time and believes correctly that it begins at noon, sits motionless in his office at that time. Suddenly, he begins to move. What explains his action? A change in belief. He believed all along that the department meeting starts at noon; he came to believe, as he would have put it, that it starts *now*." (1979, p. 28)

And we can easily duplicate *Two Gods* with respect to times:

Temporal God A god inhabits a certain possible world, and she knows exactly which world it is. Therefore she knows every proposition that is true at her world. Insofar as knowledge is a propositional attitude, she is omniscient.[3] Still I can imagine her to suffer ignorance: she does not know what time it is. For she exists over a period of many years. She does not know whether it is now 9:00am GMT January 23, 1987, or 10:32pm GMT March 18, 2010. (The trouble might perhaps be that she has an equally perfect view of every time of her world, and hence cannot identify the perspective from which she views it.)

It takes only a little ingenuity to generate similar cases involving "here" (cf. Perry (1979, p. 28)), and cases involving other indexicals (perhaps "left" and "right") may also be possible.

Surely Perry (and Lewis, and Kaplan) were right that there is a phenomenon here that deserves an explanation. And it appears to be the same phenomenon in the first-personal cases and in the temporal cases. So any account of *Messy Shopper* and *Two Gods* must also be able to handle *Absent-Minded Professor* and *Temporal God*.

It is natural to think that the solution must have something to do with indexicality, and since we are dealing with issues of cognitive significance and the explanation of behavior, it is natural to think that it must be our *thoughts* – our mental representations, such as beliefs and desires – that are indexical.

[3] A certain sort of A-theorist might hold that we cannot fully characterize the actual world without specifying what time it is; given that it is actually, say, 8:00, then any possible world where it is 9:00 is distinct from the actual world. This view would have implausible consequences; for example, it would entail the truth of "Necessarily, it is actually 8:00". I therefore set it aside.

This was Kaplan's view. He proposed that different uses of a single indexical word share a common *character*; that is, a semantic rule that determines the referent of the word in the context. He then argued that the cognitive significance of a thought is a matter of its character (Kaplan, 1977, p. 530). On this view, the reason that the thought Perry expressed by 'I am the messy shopper' was cognitively distinct from the thought that he expressed by 'John Perry is the messy shopper' is because the two thoughts have different characters; in particular, the 'I' thought has a context-sensitive character. So the 'I' thought is indexical.[4]

I will argue that this explanation cannot be correct if thinking is a matter of computation. The next section explains the computational theory of mind.

2 Cognition and Computation

Jerry Fodor famously argued that cognitive psychological theories typically involve computational processes that act on mental representations (Fodor (1975, 1981, 1989, 1994, 2008); see also Newell (1980)). Let me begin by discussing the metaphysics implicit in this conception of mind.

Mental representations are particulars; for example, they might be particular patterns of neural firing. Like all particulars, they can be grouped according to various respects of similarity. For example, a particular mental representation might be a belief, a belief that water is wet, an r-fiber firing, etc. These groupings are *types*, and particular representations are *tokens* of these types.

How does a computational account of the mind work? Consider a simple example: what is required by a computational account of of our ability to perform inferences in accordance with modus ponens? At a very rough first pass, such an account would postulate a mechanism M such that, for any representation types X and Y, if M is presented with a token of X, and a token of the form $X \supset Y$, then M outputs a token of Y; otherwise, M outputs nothing. (For example, perhaps it 'writes' this token in the thinker's 'belief box'.)

What features must a mechanism have if it is to perform this task? A mechanism like M must be sensitive to type identities and differences among representation tokens and their constituents. In the case at hand, it must be sensitive to the fact that the token X that constitutes the first premise of the

[4]This was not Lewis's view, although the form of context-sensitive thought that he postulates is equally problematic. See section 5 below.

inference is of the same type as the token X that constitutes the antecedent of the conditional second premise. If it were not so sensitive, it would be liable to output Y on the basis of X and $Z \supset Y$, or to output nothing on the basis of X and $X \supset Y$. In effect, it could not distinguish valid instances of modus ponens from invalidities, and so could not meet the description given in the previous paragraph.

The same point applies to atomic components of mental representations. For example, a mechanism responsible for the inference from *a is F* and *a is G* to *something is both F and G* must be sensitive to the fact that the token *a* that is a component of the first premise is a token of the same type as the *a* that is a component of the second premise, lest it be prone to draw a conclusion from *a is F* and *b is G*.

There may be many ways of typing mental representations; for example, natural kind concepts are a type of representation, as are representations tokened on a Tuesday. But the types that a computational reasoning mechanism is sensitive to must correspond to the content of these representations in the sense that such a mechanism must not treat representations with different contents as being of the same type.[5] If it did, it would be liable to produce invalid reasoning. For example, a mechanism that treated all natural kind concepts as of the same type would be disastrous in reasoning about natural kinds. (The mechanism may be sensitive to types that are *more* finely-grained than content. One way of glossing Millian talk of modes of presentation is that modes of presentation are the types that computational mechanism are sensitive to.)

Thus the following is a constraint on computational/representational theories of mind:

Type Sensitivity Constraint (TSC) There is some way of typing mental representations such that (i) if two representations differ as regards their content, then they are of different types; and (ii) computational mechanisms such as those postulated to account for reasoning are sensitive to type identities and difference among representation tokens.

Call the types referred to by TSC *computational* types.

[5] This fact is independent of one's views of the nature of content. If one holds a Fregean view of content, presumably one will think that one ought not infer (without some further premise) from the claim that Hesperus is bright and the claim that Phosphorus is a star to the claim that something is bright and a star. So a computational mechanism ought not to treat *Hesperus* representations and *Phosphorus* representations in the same way.

I have associated the style of computational/representational account under discussion with Fodor, but let me be more explicit: something like TSC is a constraint on any plausible computational/representational theory. The many controversial and idiosyncratic aspects of Fodor's views are irrelevant to TSC. For example, theorists who hold that mental representations are map-like (Braddon-Mitchell and Jackson, 1996, ch. 10), or take the form of mental models (Johnson-Laird, 1983), all share a commitment to something like TSC. Similarly, other well-known views of Fodor are irrelevant: nativists and empiricists share this commitment, as do atomists and holists.

In what follows, I will argue that the TSC poses a problem for the claim that there are indexical thoughts. If that is correct, then there is a problem with reconciling indexical thoughts with any broadly computationalist theory. This is important because it is not obvious that any theory that is not broadly computationalist could admit indexical thoughts either; certainly it could not use them to account for the Perry cases. Theories that deny that there are mental representations have no room for indexical representations. And theories that deny that thoughts are structured, or that the structure of thoughts plays a role in the explanation of reasoning and behavior, seem to have little room to admit that indexicality is playing a role in accounting for the Perry phenomena.

What would it take for there to be indexical thoughts? Indexicality is a property of representation types. A necessary condition on a type being indexical is that it must be context-sensitive; that is, if different tokens of that type have different contents. But this is not a sufficient condition: different tokens of the type *words that contain a 'c'* have different semantic values, but words that contain a 'c' are not all indexicals. Indexicals are different in that this context-sensitivity is built into their semantics: there is a semantic rule (roughly, a Kaplanian character) that determines the content of a token representation based on the context in which it is tokened. In the case of *words that contain a 'c'*, there is no common rule that fixes their contents in different contexts; *words that contain a 'c'* is not an interesting type from the point of view of semantic theory.

It may seem obvious that there are indexical thoughts. But there is a tension between TSC and the existence of indexical thoughts. For suppose that there were mental indexicals; that is, that there is some psychologically interesting type of representation T that is such that different tokens of T can have different contents. Now either T is, or corresponds to, a computational

type, or it does not. But each suggestion faces a serious difficulty:

1. If different tokens of T are *not* of the same computational type, then the claim that T is a semantically relevant indexical type is unmotivated. The claim is little more plausible than the view that the natural language names 'Anna', 'Bill', and 'Carla' are all manifestations of a single indexical type. In both cases, the view that there are different, non-context-sensitive types is far more natural.

2. But if different tokens of T *are* of the same computational type, then TSC is violated, since tokens with different content will be of the same computational type. For example, if there is no computational difference between different tokens of some mental *that*, we should be disposed to infer from *that is a cat* and *that is a dog* to *something is both a cat and a dog*. But we have no such dispositions.

In what follows, I pursue each suggestion in turn. The challenge for the proponent of the first view will be to argue that some feature of the Perry cases should make us regard some component of the relevant thoughts as sharing a semantic rule that fixes their referent in a context. The challenge for the second view is to provide a mechanism that accounts for the facts about our cognitive lives, which is also consistent with regarding the various tokens as of a computational type. I argue that neither challenge can be met.

3 Different Computational Types

A theorist who opts for the first horn must argue that there is a type of mental representation, which is not itself a computational type, but is instead a determinable of which multiple computational types are determinates. Moreover, these determinates must share a semantic rule that determines their content in the context in which they are tokened.

In this section, I discuss two possible motivations for the view that there is such a determinable.[6] First, it might be held that the various 'indexical'

[6]There is one other motivation that I do not discuss in detail because it does not bear on the Perry cases. Some theorists, following Tyler Burge (1977), have claimed that *de re* beliefs are indexical. This claim is typically motivated by the fact that thinkers need not possess uniquely identifying descriptions of objects that they think of *de re*; instead, the object of a *de re* thought is determined partially in virtue of non-conceptual (e.g. causal) relations that hold between the thought and the world (Burge (1977, p. 346); see also (Bach, 1987, pp. 12-13)). I doubt that the claim that this sort of environment-dependence is a matter

thoughts should be grouped together because they share a common, privileged relation to indexical language. I argue that they do not share such a relation, and that even if they did, this would not be much reason to think that the thoughts themselves are indexical. Second, some might hold (following Perry) that the various 'indexical' thoughts can be grouped together because they share a functional role. I argue that they do not share a functional role.

3.1 Indexical Thoughts and Indexical Language

It might be claimed that we should regard certain computational types as falling under a common indexical type because they bear some privileged relation to indexical language. This view might be developed in four ways:

1. Mental states involving representations of these types can correctly be *reported* using indexical language. For example, certain beliefs about oneself should be regarded as indexical because they are correctly reported using 'I'.

2. Mental states involving representations of these types can correctly be reported *only* using indexical language. For example, certain beliefs about oneself should be regarded as indexical because they cannot correctly be reported in English without using 'I' (or perhaps 'he*').

3. Mental states involving representations of these types can correctly be *expressed* using indexical language. For example, certain beliefs about oneself should be regarded as indexical because they are correctly expressed using 'I'.

4. Mental states involving representations of these types can correctly be expressed *only* using indexical language. For example, certain beliefs about oneself should be regarded as indexical because they cannot correctly be expressed in English without using 'I'.

Each of these developments faces very serious difficulties:

Against (1). Almost any belief can be reported using indexical language under some circumstances. For example, suppose that Mike has a belief that he would express by saying, "Bob is tall." Bob could accurately and

of indexicality can be sustained. If externalism about mental content is true, then all, or very nearly all, mental representations depend for their content on some aspect of their environment. This does not entail that all thoughts are indexical. A more plausible view is that the role of the environment in these cases is pre-semantic.

felicitously report Mike's belief by saying, "Mike believes that I am tall." But this is no reason to suppose that the concept Mike expresses with "Bob" is indexical.

Against (2). Castañeda argues that certain beliefs can only be reported in the third-person using his 'he*' device (which he regards as equivalent to 'he himself') (1968, p. 442); on this view, if Bob has a belief that he would express by saying, "I am a philosopher", I could make a fully adequate report of this belief in English only by saying, "Bob believes that he* is a philosopher." A similar view of first-person reports would hold that if I have a belief that I would express by "I am tall", I can best report this belief only by saying, "I believe that I am tall".

The latter view faces a challenge from those who prefer to refer to themselves in the third person. When Bob Dole reports, "Bob Dole believes that Bob Dole is hungry," the indexicalist must argue that he has misreported his belief, or at least reported it less than fully adequately. Perhaps this conclusion can be accepted. But we seem to be able to imagine a linguistic community that has no first-person indexical term, or indeed, no indexical terms at all. Would members of this community be unable to report their own beliefs?[7]

It might be objected that such a community would have no ability to distinguish between the beliefs of the amnesiac who knows that Rudolf Lingens is Rudolf Lingens but does not know who he* is, and the beliefs of the former amnesiac who has realized that he* is Rudolf Lingens. But this need not be the case. The community could have a convention whereby each person has a special name that is only used where we would use 'I' (or 'he*') in English. Perhaps all and only names that start with the letter 'C' are used in this way. We can then imagine the amnesiac stipulatively introducing one of these special, first-personal name for himself; suppose he chooses 'Christopher'. Then he might express his discovery by exclaiming, "Christopher is Rudolf Lingens!", and we might report it by saying, "He once did not realize, but now knows, that Christopher is Rudolf Lingens." Given the 'C' convention, these reports are unambiguously of the sort at issue in the Perry cases.

An indexical-less community might have other conventions. Castañeda appended his '*' device to an English indexical. But it is hard to see why one could not append it to an ordinary proper name, with the understanding

[7]Perry (1979, pp. 28-29) seems to endorse the claim that speakers like Bob Dole do fail adequately to report key beliefs. (His example is de Gaulle.) But Perry does not consider the sort of conventions I discuss in the next two paragraphs.

that attitudes reported using the '*' must be self-conscious or *de se*. Given this convention, we might report Lingens as realizing that Lingens is Lingens*. Again, the report seems perfectly adequate, despite the fact that it contains no obvious indexical expression.

Against 3. It does not in general seem to be the case that expressibility using indexical language is a marker of indexicality of the relevant beliefs. For example, it seems unlikely that there are uniquely second-person beliefs. (What would the character of such beliefs be? What is the analogue of an addressee for belief?) But some beliefs are correctly expressed using 'you'. Thus the fact that a belief can be expressed using indexical language does not show that the belief itself is indexical in nature.

Against 4. The cases developed in the argument against 2 generalize against 4.

I conclude this section with a more general point. It is not clear that the connection between thought and language is as close or as direct as the view assumes. Absent some strong reason to think that we can read semantic features of natural language sentence into the thoughts those sentences express or report, we should be skeptical of the claim that indexicality in thought is grounded in indexicality in language.

3.2 Perry and Functional Role

John Perry famously emphasized the role of beliefs attributed using indexical terms in explaining behavior:

> Now consider all the good-hearted people who have ever been in a supermarket, noticed sugar on the floor, and been ready to say "I am making a mess." They all have something in common, something that leads us to expect their next action to be that of looking into their grocery carts in search of the torn sack. Or consider all the responsible professors who have ever uttered "The department meeting is starting now." They too have something important in common; they are in a state that will lead those just down the hall to go to the meeting, those across campus to curse and feel guilty, those on leave to smile. (Perry, 1979, p. 39)

This suggests that we should regard thoughts expressed with 'I' as a type because they share a common functional role. In particular, Perry suggests that indexical thoughts play a special role in causing action. An alternative

proposal, developed by Perry (1980) and Recanati (1993) has it that mental indexicals are linked to particular sources of information. Each likens mental indexicals to a mental file or "repository of information received in a certain way" (Recanati, 1993, p. 88).

But there are well-known arguments (developed with particular force by Schiffer (1987)) against the claim that in general, there is some functional role that is shared by all tokens of a given mental state type. We can bring out the force of these arguments by considering a closely related case. Hilary Putnam (1975) famously argued that natural kind terms are indexical. The present proposal would provide one motivation for this view: although the belief that my twin-Earth doppelgänger expresses with the sentence 'Water is wet' is of a different computational type than the belief I express with a homophonic sentence, these beliefs share a functional role, and thus should be construed as of a common indexical type.

So far, the proposal has some plausibility. What has not even begun to be established, however, there is some functional role in common to *every* belief that water is wet; in particular, we have so far no reason to believe that in general, beings who are not atom-for-atom duplicates of me have a belief that shares an interesting functional role with mine. We thus have as yet seen no reason to think that all beliefs that water is wet are tokens of a common indexical type. In order to establish this, it must be shown that my belief that water is wet shares a functional role with Helen Keller's, expert chemists', individuals' under the Burgean misconception that gin is a kind of water, intelligent aliens' with very different sensory apparatus and means of behavior, etc. In each case, the beliefs in question would differ in their relations to sensory inputs, behavioral outputs, or inferential relations to other mental states. Any functional role weak enough to encompass the beliefs of all these beings would likely encompass many others: for example, my belief that gin is wet. This functional role thus cannot be linked to character on pain of assigning too many representations the same character.

This style of argument generalizes to Perry's case. Against the claim that sharing a functional role is *necessary* for being of the same type, we can note that beings with very different perceptual and behavioral capacities, as well as very different background beliefs and dispositions to behavior, can exhibit Perry-like *de se* mental states. Thus even though belief that I express by 'I am hungry' shares a precise functional role with the belief my atom-for-atom twin expresses with the same words, it does not share a precise functional role with the belief that a being with very different sensory apparatus would

express with the same words. Nor does it share a precise functional role with beings who have very different potential for behavior (for example, intelligent coral with little capability of movement).

Suppose that this worry can be answered, and that there is some functional role that is shared by all those who think, "I am hungry," regardless of their perceptual and behavioral capacities, and their background intellectual and behavioral dispositions. It is most unlikely that such a functional role would be *sufficient* for being of a common type. For example, one could imagine a being who has a parasite that moves in a way that causes its host to feel sensations much like the ones I feel when I am hungry. The being is aware that it has the parasite, and is disposed to judge on the basis of this feeling that the parasite is hungry. (Suppose that its species does not normally need to eat, and does not normally feel hunger.) When the being feels this sensation, it eats, which feeds the parasite and stops its movement. It seems that the being's judgement that the parasite is hungry could have quite a similar functional role to my judgement that I am hungry. Indeed, its functional role seems far more similar than the "I am hungry" beliefs of the intelligent coral mentioned above. Nonetheless, we should not construe them as of a common indexical type.

The challenge for the functionalist proposal will be to describe a functional role that *all and only* the allegedly indexical representations of a certain sort share. This must be done if the view is to show that all of the allegedly indexical representations are of a common type. Perhaps some clever functionalist can develop a theory that solves these problems. But the prognosis is not at all good.

Of course, all of this is not to deny that many of those who are prepared to utter "The meeting is starting now" will share the sort of behavioral dispositions that Perry mentions. The point is simply that not all similarities in behavior are to be explained by postulating mental states of the same type. Suppose I want to eat apples, and you want to eat oranges. We will engage in many similar behaviors: going to grocery stores, proceeding to the produce section, giving money to cashiers, chewing, and so forth. But surely there need be no interesting, psychologically relevant type of belief or desire in common between us. Why, then, postulate a common type on the basis of the behavioral similarities in Perry's cases?

4 The Same Computational Types

We have seen no reason to think that the aspects of the Perry cases discussed so far are linked to indexicality, and no prospect of vindicating indexicality in thought via the first option. What of the second option? The second option would advocate abandoning or modifying the TSC, and claiming that different tokens of the same computational type can differ with respect to their contents. The basic problem is this: if different tokens of an indexical representation are of the *same* computational type, then we should be liable to errors in reasoning that we in fact do not commit. For example, if there is no computational difference between different tokens of some mental *that*, I should be disposed to infer from *that is a cat* and *that is a dog* to *something is both a cat and a dog*; or if there is no computational difference between tokens of some mental *now*, I should be disposed to infer from *now is 2:00* and *now is 2:01* to *some time is both 2:00 and 2:01*.[8]

In order to avoid this sort of consequence, the proponent of mental indexicals would have to modify the TSC. One possibility involves changing the computational mechanism to treat indexicals differently than non-indexicals. For example, the inference mechanism M might handle an indexical NOW by behaving in the following way: if it is presented with a token of X, and a a token of the form $X \supset Y$, then: (i) if X and Y do not contain NOW, M outputs a token of Y; (ii) if X contains NOW, M outputs a token of Y if and only if the token of X and the token of $X \supset Y$ were formed at the same time; (iii) otherwise X outputs nothing. Then the TSC could be modified in something like the following way:

Type Sensitivity Constraint* (TSC*) There is some way of typing mental representations such that (i) if two representations differ as regards their content, then they are of different non-indexical types, or of the same indexical type; and (ii) computational mechanisms such as those postulated to account for reasoning are sensitive to type identities and difference among representation tokens of non-indexical types, and sensitive to type identities and differences as well as other associated information among representation tokens of indexical types.

A mechanism of the sort under consideration could work only if it records

[8] There may be some analogue of indexicality at pre-conceptual levels of representation. For example, Pylyshyn (1989) develops a theory of spatial perception that involves what he describes as 'indexical references'; since the 'indexicals' in question are not stored in memory and are not involved in inference, there is little danger of error. In any case, this sort of representation is irrelevant to the Perry phenomena.

the time at which each NOW-involving representation is formed. Each token NOW would have to be linked to a particular time. These links might be implemented in various ways; for example, we could imagine representations prefixed with a sort of temporal quantifier or operator (AT 2:00, THE MEETING STARTS NOW), or representations that link the NOW to a time via a sort of parenthetical (THE MEETING STARTS NOW (2:00)). No doubt there are other possibilities. But once it is admitted that NOWs must be linked to representations of particular times, we seem to have no reason to think that NOW is indexical. In our first example, NOW looks like a bound variable; in the second, NOW (2:00) looks like a name for 2:00. There is no apparent context-sensitivity here, and modifying the TSC to the more complex TSC* looks unwarranted.

More generally, we can put the point as follows. Indexical concepts of the sort under consideration in this section will not result in invalid inferences only if different tokens of the same computational type have different causal and inferential roles. For example, suppose that at 2:00pm I form two beliefs: "Now it is 2:00pm" and "Now it is raining", and at 3:00pm I form two more beliefs: "Now it is 3:00pm" and "Now it is not raining". Roughly, I should be able to infer from these that it was raining at 2:00, and that it was not raining at 3:00. That is to say that the two token *now*'s in the beliefs I formed at 2:00 'go together'; I should be able to combine them in inference with each other, but not with the token *now*s I formed at 3:00.[9]

But now it looks like for the purposes of computation, we have two distinct types: the *now*s tokened at 2:00 are type-distinct from the *now*s tokened at 3:00. Thus it seems that we have regressed to the first option. In general, any computational system that has the resources to avoid making the bad inferences in question will end up treating the alleged indexicals as of distinct types.

Here is one way to see the problem. Suppose that at any given time we have the ability to form a belief that we would express by "Now is the time". There are two *prima facie* plausible descriptions of this ability. The first is the one under assessment in this section: there is some computational type that has different contents in different contexts. The second is that there is

[9]Someone might think that indexical now-beliefs are essentially fleeting: one can believe that now is 2:00 at some given time, but one cannot (or at least ought not) retain this belief. This might avoid the sort of problem discussed here in some cases. But some now-beliefs are retained: I might believe all winter that it is now the winter of our discontent. Moreover, it is not clear that the view could establish the existence of mental indexicals. I discuss this proposal further in section 5 below.

a mechanism that is capable of generating tokens of distinct computational types. Given the arguments of this section, it looks like any computational mechanism that is in a position to avoid serious errors in reasoning will be better described by the second account.

4.1 Thinking in English?

Even if there is reason to doubt that there are indexicals in some language of thought, we know that there are indexicals in spoken languages. And it is obvious that in some sense we sometimes think in the same language we speak. Suppose that on some occasion I find the English sentence "Now it is 2:00pm" 'running through my head'. Does this not show that there are indexical thoughts?

There are preliminary reasons to doubt that this could explain the Perry phenomena. The Perry phenomena can occur even with non-linguistic creatures: for example, a dog might wag her tail at her own reflection, failing to recognize that it is she herself. Nonetheless, there seems to be a problem: if some of our beliefs are 'in' natural language, doesn't this show that there are indexical beliefs after all? I will argue that even if there is an attenuated sense in which these beliefs count as indexical, they are not *essentially* indexical; they are reducible to non-indexical beliefs. To see why this is so, we will have to look carefully at what it would mean for someone to believe in English, given a computationalist perspective.

The ability to learn language is part of our innate cognitive endowment, but we are not born speaking any particular language. English and its ilk must be learned. Similarly, we are not born *thinking in* a language like English. Young children think in a language of thought that is not merely an internalized spoken language. Call this language of thought *mentalese*.

The basic computational mechanisms that constitute our minds must manipulate mentalese representations. So, for example, a creature might have a simple practical reasoning module that, when presented with a P representation in the 'desire box' and a $Q \supset P$ representation in the 'belief box', produces a Q representation in the 'intention box', where P and Q range over sentences of mentalese. We have good reason to think that there can be no indexicals in mentalese. But suppose that this creature is an English speaker. What would it take for entertained sentences of English to enter this story about beliefs and desires?

There are two possibilities: the English sentences might be translated into sentences of mentalese, or there might be a distinct mechanism for han-

dling English. Neither possibility supports the idea of essential indexicality. If sentences of English are simply translated into mentalese, then given that there are no indexicals in mentalese, indexicality must not be an essential feature (for the purposes of belief and desire) of the relevant English sentences. For example, it might be that sentences of English are 'marked' with the time that they are recorded: AT 2:00 "The meeting starts now". In the translation, any occurrences of "now" are replaced by the recorded time. This proposal is clearly no different from the attempt to rescue the Same Computational Type view considered above. No trace of indexicality remains.

The hypothesis of a dedicated mechanism for dealing with sentences of English does not substantially change the situation. Such a mechanism must be implemented in mentalese. Since there are no indexicals in mentalese, it is hard to see how indexicality could play an essential role in the mechanism. (Indeed, it is rather hard to see how such a mechanism could work except by in effect translating English into mentalese.)

An analogy might help clarify this picture of 'thinking in English'. Computers can only run programs that are written in their machine language. But most humans cannot read or write programs in machine language. In order to make computer programming easier, other programming languages have been developed. Sometimes, programs written in these other languages are translated (or compiled) into machine language. But in other cases, programs are *interpreted*. In this case, no machine language version of the entire program is ever generated. Instead, another program, the interpreter, is directly run by the computer. The interpreter reads in each instruction in the program to be interpreted, and translates it into machine language as needed.

A traditional motto of the computational theory of mind is that the mind is the software of the brain. To stretch this metaphor: some of the software of the brain must be written in the machine language of the brain; that is, it must be encoded in a way that can directly affect the brain. But there could also be interpreted language of the brain. For example, some of our beliefs might be stored natural language sentences.[10] But such beliefs would have to be interpreted – that is, in effect, translated – into the machine language. And since there are no indexicals in the machine language, the final story

[10] A related alternative is that our beliefs are not a part of the software of the brain at all, if that is to mean that they are a sort of program that is executed by the brain; instead, belief are a sort of data source that is read by software. My discussion could easily be adapted to this view.

can be told in an indexical-free way. So there is no prospect for essential indexicality here.

5 Objections and Replies

Objection 1: Perhaps your arguments show that there can be no indexical 'now' or 'that' thoughts. But they do not show that 'I' is not indexical. There is a potential problem with invalid inferences involving a particular indexical only if there is a possibility of context shifts within the reasoning of a single thinker that change the reference of that indexical. For example, THAT is problematic because demonstrating different objects can shift its reference mid-inference. But the context could not shift in such a way as to make any given person's 'I'-concept shift its reference. So there is no problem about the indexicality of 'I'.

Reply: The objection is strictly speaking correct, but the point that mental indexicality cannot explain the Perry phenomena still stands. There are two facts to be noted. First, as I pointed out in section 1, Perry-style cases involving 'now' are just as convincing as cases involving 'I'. Since the 'I' phenomena and the 'now' phenomena appear identical, it would be quite strange if the 'I' cases are to be explained by indexicality and the 'now' cases are not. Second, although an indexical 'I' would not be apt to produce bad inferences in the manner of an indexical 'now', it is hard to see what motivation we would have for thinking 'I' an exception to the TSC. What argument could show that my thoughts about myself share a semantics with your thoughts about yourself?

Objection 2: You have granted the possibility of indexical 'I' thoughts. But Perry argued that 'here' could be reduced to 'I' and 'now' ('here' = 'the place I am now') (Perry, 1979, p. 38). And Lewis showed that 'now' can also be reduced to 'I', if we accept the claim that 'I' picks out an instantaneous temporal part. So 'I' is the only indexical we need.

Reply: It may be that 'here' can be reduced to 'I' and 'now'. But the proposed reduction of 'now' to 'I' runs afoul of my arguments. For if we hold that 'I' picks out an instantaneous temporal part, then context-shifts within a single course of reasoning can shift the reference of 'I'. For example, someone might be inclined to reason as follows: 'I am the instantaneous temporal part at 2:00', 'I am the the instantaneous temporal part at 2:01', 'Therefore, some instantaneous temporal part exists at both 2:00 and 2:01'.[11]

[11] A presentist would have other options; for example, 'now' might be reduced to 'the existing time'. But this reduction is not indexical.

Objection 3: Lewis develops his account not as a claim about the reference of the word 'I', but as a claim about the nature of propositions. On his view, propositions have truth values at *centered worlds*, rather than worlds simpliciter. Similarly, the temporal version of the Perry phenomena can be developed with examples that exploit tense rather than temporal pronouns. (For example, consider the difference between "The meeting will happen" and "The meeting is happening".) But one plausible account of tense involves operators on propositions that have truth values at worlds and times, rather than anything like indexicals (Kaplan, 1977). Your arguments do not tell against these non-indexical contextualist accounts.[12]

Reply: The arguments do tell against non-indexical contextualist accounts. Any mechanism that implements a deductive inference must perform only those inferences that preserve truth. This becomes problematic once propositional truth is relativized in the way envisioned by the objection. For example, some instances of the form: *p, therefore p* will be invalid if truth is relativized to times. (Suppose, for instance, that I am sitting at some time t, and I think at t that I am sitting. On the view envisioned by the objector, this thought will be true relative to t. At $t + 1$ I stand up. I then infer from my previous thought that therefore, I am sitting. This thought will be false relative to $t + 1$.)

How can a mechanism for deductive inference perform only those inferences that preserve truth? That is, how can it distinguish valid from invalid instances of (for example) I AM SITTING, I AM HUNGRY, therefore SOMEONE IS BOTH SITTING AND HUNGRY? I see no way it can do so without recording information about the time a particular representation was formed. For example, if each instance of I AM SITTING or I AM HUNGRY is 'stamped' with the time at which it was formed, the the computational mechanism can 'put together' only those representations that were formed at the same time.

But we have seen this style of view before: it is much like the attempt to rescue the Same Computational Types strategy by linking each token NOW to a representation of a particular time, which we rejected in the previous section. The present view fails for the same reason. Once we have explicit representations of the time at which each token representation is formed, what reason do we have for thinking that these representations vary in truth value across times? After all, on this view the thought: I AM HUNGRY,

[12] For general discussion of this style of view, see Lewis (1980) and MacFarlane (2009). For arguments against treating tense in this way, see King (2003).

marked as created at 2:00, will be treated just like the thought: AT 2:00, I AM HUNGRY. Since there is no reason to think of the latter as context-sensitive,[13] there is no reason to think of the former as context-sensitive.

Objection 4: There is a strong intuition that the 'I' thoughts involved in Perry cases are *private*, in the sense that they can be entertained only by a single thinker (Frege, 1997, pp. 332-3). Similarly, perhaps, the 'now' thoughts are *fleeting*, in the sense that they can be entertained only at a single moment. Indexicality offers an explanation of this fact: if there are indexicals in mentalese, then there are some sentences of mentalese that express a particular content only in a particular context. For example, if mentalese 'I' is an indexical, then I cannot get in a context where I AM HUNGRY has the content that Barack Obama is hungry; only Obama can be in such a context. So if there are indexical thoughts, then there is a way of thinking that Obama is hungry that only Obama can perform.

Reply: I am aware of no compelling argument for the view that there are private or fleeting thoughts. But even if we grant it, the indexicalist account does not follow. We are familiar with the idea of thoughts that can be grasped only by some thinkers; for example, it is often alleged that natural kind thoughts can only be grasped by those who have causally interacted with their referents, and that certain thoughts about phenomenal experiences can only be grasped by those who have undergone those experiences. It is not typically concluded that these thoughts are indexical. Why should thoughts that can only be grasped by a single person be different?

Objection 5: My belief that I am hungry is cognitively distinct from my belief that NN is hungry; even though these beliefs are true in all of the same worlds, they seem to represent the world in different ways. A mark of this is the apparent fact that one might give a complete list of the objective facts about the actual world without using any indexical language, and someone might know all of the truths on the list without being in a position to rule out many possibilities about who she is and what time it is. (This is brought out vividly by the (Two Gods) case.) The best explanation of this fineness of grain will involve indexicality.

Reply: The facts about fineness of grain do not entail that the representations in question are indexical. Compare Frege cases. My belief that Hesperus is bright is cognitively distinct from my belief that Phosphorus is bright; even though these beliefs are true in all of the same worlds, they seem to represent the world in different ways. Moreover, one could com-

[13] Setting aside any potential context-sensitivity of HUNGRY.

pletely characterize the actual world in such a way as to distinguish it from every other world, without using the word 'Phosphorus'; the content of the thought expressed by 'Hesperus is Phosphorus' is arguably not captured by a division in the space of possible worlds (since it is a necessary truth). Surely this does not show that the belief that Hesperus is Phosphorus is indexical.

It might be objected that given the Fregean view that facts are true thoughts, an objective description of the world that omitted the fact that Hesperus = Phosphorus would be incomplete in virtue of omitting this (objective) fact. The objection might continue that this is not the case as regards the thought that now is 2:00, or that I am NN. Though (given the Fregean individuation of facts) these are distinct from the facts that 2:00 is 2:00 and that NN is NN, they are not objective; a complete, objective description of the world could be given without mentioning them. We therefore have reason to think that it is the sort of fact expressed using indexicals.

This argument is question-begging. I have granted that my thought that now is 2:00 is cognitively distinct from my thought that 2:00 is 2:00. From the computationalist point of view, that is to say that these two token thoughts differ in their computational type. But that is to say nothing about the nature of the two computational types. Given the preceding discussion, we have at least some reason to doubt that either type is indexical in nature. And unless we are antecedently assuming that the thought that now is 2:00 is indexical, what reason do we have for believing it to be subjective? For all that has been shown, Perry cases are no different from Frege cases.[14]

6 Conclusion

The computational/representational theory of mind has no room for indexical thoughts, and hence no room for an indexicalist account of the Perry cases. A better account must be sought elsewhere.

References

Bach, K. (1987). *Thought and Reference*. Clarendon Press, Oxford.

Braddon-Mitchell, D. and Jackson, F. (1996). *Philosophy of Mind and Cognition*. Blackwell, Malden, MA.

Burge, T. (1977). Belief de re. *Journal of Philosophy*, 74:338–362.

[14]This view is developed in Millikan (1990).

Casteñeda, H.-N. (1966). 'He': A study in the logic of self-consciousness. *Ratio*, 8:130–157.

Casteñeda, H.-N. (1968). On the logic of attributions of self-knowledge to others. *Journal of Philosophy*, 65:439–456.

Fodor, J. A. (1975). *The Language of Thought*. Harvard University Press, Cambridge, MA.

Fodor, J. A. (1981). *Representations: Philosophical Essays on the Foundations of Cognitive Science*. The MIT Press, Cambridge, MA.

Fodor, J. A. (1989). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. The MIT Press, Cambridge, MA.

Fodor, J. A. (1994). *The Elm and the Expert: Mentalese and Its Semantics*. The MIT Press, Cambridge, MA.

Fodor, J. A. (2008). *LOT 2: The Language of Thought Revisited*. Oxford University Press, New York.

Frege, G. (1997). Thought. In Beaney, M., editor, *The Frege Reader*, pages 325–345. Blackwell, Oxford.

Johnson-Laird, P. N. (1983). *Mental Models*. Harvard University Press, Cambridge, MA.

Kaplan, D. (1977). Demonstratives: An essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In Almog, J., Perry, J., and Wettstein, H., editors, *Themes From Kaplan*, pages 481–564. Oxford University Press, New York.

King, J. C. (2003). Tense, modality, and semantic values. *Philosophical Perspectives*, 17.

Lewis, D. (1979). Attitudes de dicto and de se. In *Philosophical Papers Volume 1*, pages 133–159. Oxford University Press, New York.

Lewis, D. (1980). Index, context, and content. In *Papers in Philosophical Logic*, pages 21–44. Cambridge University Press, New York.

MacFarlane, J. (2009). Nonindexical contextualism. *Synthese*, 166:231–250.

Millikan, R. G. (1990). The myth of the essential indexical. In *White Queen Psychology and Other Essays for Alice*, pages 265–268. The MIT Press, Cambridge, MA.

Newell, A. (1980). Physical symbol systems. *Cognitive Science*, 4:135–183.

Perry, J. (1977). Frege on demonstratives. In Perry (2000), pages 1–26.

Perry, J. (1979). The problem of the essential indexical. In Perry (2000), pages 27–44.

Perry, J. (1980). A problem about continued belief. In Perry (2000), pages 57–76.

Perry, J. (2000). *The Problem of the Essential Indexical and Other Essays, Expanded Edition.* CSLI, Stanford.

Putnam, H. (1975). The meaning of 'meaning'. In *Mind, Language, and Reality: Philosophical Papers Volume 2*, pages 215–271. Cambridge University Press, New York. Originally published in *Minnesota Studies in the Philosophy of Science*, 7, 1975.

Pylyshyn, Z. (1989). The role of location indexes in spatial perception: A sketch of the finst spatial-index model. *Cognition*, 32:65–97.

Recanati, F. (1993). *Direct Reference: From Language to Thought.* Blackwell, Oxford.

Schiffer, S. (1987). *The Remnants of Meaning.* The MIT Press, Cambridge, MA.

Stalnaker, R. C. (1981). Indexical belief. In *Context and Content: Essays on Intentionality in Speech and Thought*, pages 130–149. Oxford University Press, New York.