

## NOW, ME

My topic is the state of mind of people who utter sentences containing indexicals; sentences like "I am here now". It is a subject much discussed in recent years but invariably it has been addressed as an issue in philosophical semantics: as a problem about what indexical words mean<sup>1</sup>. I want here to pursue a different and, I hope, more productive approach. To see things fresh, we'll start from scratch. If, at first, it seems as if am re-inventing wheels, bear with me. Sometimes that is the only way to get anywhere.

## PSYCHOLOGY IN TWO DIMENSIONS

It has been clear since the discovery of Twin Earth<sup>2</sup> that there must be two different determinants of the propositional attitudes. Our Twin Earth doppelgangers are atom for atom identical with us. Any psychological theory— any theory about how sensory inputs get translated into behavioral upshots— true of us must be true of them. And yet, because of our different environments, we and our twins have different propositional attitudes. Apparently, these differences in what we think and think about, must be irrelevant to psychology. How can we reconcile this with our folk psychological conviction that people do what they do because of what they think?

Suppose that we one day have to hand The One True Psychological Theory of Human Beings. Like any theory it will have a proprietary vocabulary— a way of carving up the world into psychological natural kinds— and a finite set of laws. While we do not know what those laws are, we can be sure they will entail any number of nomological regularities of the form:

$$(L) \quad (\forall x)(\ ^1x \text{ at } t_1 \supset \ ^2x \text{ at } t_2)$$

---

<sup>1</sup> For an overview see Braun, David, "Indexicals", *The Stanford Encyclopedia of Philosophy (Fall 2001 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/fall2001/entries/indexicals/>>.

<sup>2</sup> Putnam, Hilary, "The Meaning of 'Meaning'," in Keith Gunderson, ed., *Language, Mind and Knowledge*, Minnesota Studies in the Philosophy of Science, VII, Minneapolis: University of Minnesota Press, 1975; reprinted in Putnam, *Mind, Language and Reality: Philosophical Papers Volume 2*, Cambridge: Cambridge University Press, 1975.

where the ' $\Psi$ 's name individuals' *total psychological states*; everything that can be said about them in the terms of the Theory.

The lesson of Twin Earth is that individuals might share the same total state but believe different things. What propositional attitudes one has depends upon extra-psychological facts about the world in which one realizes that state and about one's location in time and space in that world. I'm going to call these locations in possible worlds *situations*<sup>3</sup>. I write:

in  $\sigma$  in  $\Psi$ :  $x$  believes  $p$

to say that, in a certain situation  $\sigma$ , someone in state  $\Psi$  would believe  $p$  (among, presumably, many other things). In Putnam's story the situations and states are such that.

In  $\sigma$  in  $\Psi$ :  $x$  believes rain is H<sub>2</sub>O.

In  $\tau$  in  $\Psi$ :  $x$  believes rain is XYZ.

It might be tempting to gloss all this by treating psychological states as functions from situations onto attitudes, but this would be a mistake. Functions are identical if they yield the same values for arguments, but, as we shall see, psychological states can be different in psychologically potent respects even though the differences don't show up as differences in what subjects think or think about.

If we were perfect epistemic engines, our psychological states would be realized by true beliefs in every situation in which we found ourselves. Alas we know that isn't so and the situation dependence of the attitudes reveals another source of error. A psychological state realized by true beliefs in one situation might be realized by false beliefs in another. Here is Oedipus insisting, "My Mom is not my wife". The proposition he believes is false; indeed, necessarily false by some ways of reckoning. But Oedipus is not irrational or unfathomable, only badly situated. Surely things could

---

<sup>3</sup> By which I do not mean what Perry and Barwise call a "situation", but rather what I say here. Alas, every viable noun-- "context", "scenario", "locale"-- has already been appropriated as someone's technical term.

be arranged, behind the scenes, over on Twin Earth, so that Twin Oedipus is married to a woman who is not his wife, without it making any difference to what is in Twin Oedipus's head.

Twin Earth, so constructed, is what Dan Dennett, would call one of Oedipus's "notional worlds". That is, worlds or, better, (since Twin Earth is in our possible world), *situations* in which Oedipus's psychological state would be realized by true beliefs.<sup>4</sup>

Dennett noticed that, for the purposes of explaining behavior, what matters is not what propositions subjects actually believe, but what they would believe in their notional situations. This makes sense from an evolutionary point of view. Presumably, true belief has survival value. But evolution can operate only on what is inside our heads and what we believe and whether it is true is largely determined by what's outside. However clever we may be, we may still, like Oedipus, end up badly situated. On the other hand, it would be a poor psychological design that allowed us to get into psychological states that could not be realized by true beliefs in *any* situation. Thus we might expect that the psychological states of well adapted organisms would be realizable by true beliefs in at least *some* situations. Moreover we should expect those organisms to behave and think in ways that were appropriate to those notional situations. As I will put it, your notional situations are *the sorts of situations you think are in*.

As folk psychologists we are adept at understanding one another in this way. If we are told that someone believes the proposition that Jocasta Jocasta, we have no way of telling what he might do. On the other hand if we know that someone *thinks they are in* a situation of the sort Twin Oedipus is actually in, we understand his psychological state, know how he might have acquired it and how he is likely to respond to what happens next.

---

<sup>4</sup> Dennett, Daniel, "Beyond Belief," in Andrew Woodfield, ed., *Thought and Object: Essays on Intentionality*, Oxford: Oxford University Press, 1982; reprinted in Dennett, *The Intentional Stance*, Cambridge: MIT Press, 1987.

What happens next in the Oedipus story is that a group of messengers arrive telling tales. Listening to them, Oedipus and Twin Oedipus will change their common psychological state until each them comes to think he is in a situation in which it is immensely probable that his mother is his wife. At that point, in a transition described by laws like (L), the penny drops. Oedipus and Twin Oedipus move into a new psychological state which is realizable by true beliefs only at situations in which they have committed incest. Oedipus's actual situation has become his, and Twin Oedipus's notional one. And both behave appropriately to *that* situation; they go ape, as anyone would if he thought he was in a situation in which he had married their mother.

In the course of the story, Oedipus and Twin Oedipus acquire new beliefs given what they see and hear. Assessed in terms of propositional content, they learn very different things. The messengers hither are telling the awful truth about Jocasta; over yon, they are telling ugly lies about Jocasta's twin. Still, the Oedipi's sensory inputs will be psychologically indistinguishable and will have identical effects on their psychological states and notional situations. For the purposes of explaining and predicting behavior, the most productive way to characterize empirical uptake is not in terms of what propositions it is evidence for or what propositions it might lead a subject to believe, but rather in terms of its effects on the subject's notional situations. By this accounting we can say that in our story both Oedipi make the same *inferences* based upon the same *information*. Of course, they end up with different beliefs. In Twin Oedipus's case: necessarily false beliefs; but that is beside the psychological point.

## **WHO, ME?**

Perhaps the situation dependence of the attitudes should not surprise us. After all, if we were given a description of the total physical state of planet Earth in the language of pure physics, we could infer that the Earth orbits some star, but could not say which star that was. 'Sol' is not a term of physical theory and Sol's haecceity is surely not a natural physical kind. Sol and Twin Sol are

atom-for-atom identical and if you instantaneously swapped planets neither Earth would behave differently.<sup>5</sup> What star a planet orbits about depends on its physical state but also on where it is situated. Likewise, we may say, what an individual thinks about depends upon his psychological state and situation.

But this easy comparison seems to come apart when we consider the situation *invariance* of beliefs about the self. Here is Sally. Sally admires Paris Hilton. Of course, over on Twin Earth, Twin Sally admires Twin Paris. Who Sally admires depends, in part, on her situation. But Sally is not alone in admiring Paris. It is a good bet that Paris Hilton also admires Paris Hilton. A good bet too that this is *not* an accident of her situation. We suppose that, if transported to Twin Earth, Paris would still admire herself, not her twin. We suppose this because we suppose that Paris is *vain*.

Vanity is more than self-admiration. Suppose that one day Sally gets a glimpse of someone she thinks is as beautiful as Paris. "I wish I looked that good!", she says to herself. And she gets her wish, because she has been looking in a mirror without realizing it. Sally admires Sally but she isn't vain. On the contrary, she is lacking in self esteem to the point of neurosis. The difference between Sally and Paris is that Sally's self admiration is an accident of her situation. In a situation where there was no mirror but a Sally look-alike behind a pane of glass, Sally would admire the look-alike, not herself. In the sort of situation Sally *thinks she is in*, she is not admirable.

Sally's accidental self admiration is no sin. Vanity, the sin, is a matter of being in a psychological state which would be realized by self admiration in *any* situation.<sup>6</sup>

"x is vain" =<sub>df</sub> " $(\exists s) x \& (\forall s)(in\ s : x\ admires\ x)$ "

---

<sup>5</sup> From which some philosophers would no doubt conclude that astronomy has no business ascribing relational properties to planets.

<sup>6</sup> Almost certainly, the True Psychological Theory will entail that some psychological states cannot be had in some situations. Properly speaking, then, generalizations like this over situations and states should be restricted to situations in which a total state was nomologically possible. I will ignore this complication in what follows.

Psychological states which are otherwise very different in the attitudes they realize might nevertheless embody vanity. Some people are vain about their looks, some about their intellect. What such states have in common (and what makes them sinful) are the broad kinds of odious behaviors which they typically produce : the preening, the writing of autobiographies. Our store of folk wisdom about the perils of vanity is, I suggest, evidence that implicit quantification over psychological states is part of folk psychology.

Paris's vanity might seem paradoxical at first. We saw in Sally's case how self-admiration can be situation dependent. What makes Paris' self admiration situation *independent*? If Paris's psychological state guarantees that she always admires herself, then it must be *nomologically* necessary that her state be realized by admiration of her and no one else. But how can that be? It seems implausible to suppose that "admiration of Paris" is a psychologically natural kind. The laws of nature are no respecters of persons, even celebrities.

The puzzle goes away when we note that laws of nature can dictate circumstances in which things must stand in relations to themselves. Take the astrophysical analogy. The Earth and Twin Earth are in the same physical state. But part of what constitutes the state of each planet is that it stands in certain relations to itself. Each is in orbit, hence each is undergoing acceleration. For  $x$  to be accelerated is for  $x$  to have a different velocity at  $t_1$  than  $x$  has at  $t_2$ . Nothing can be in precisely the same physical state as the Earth unless it stands in this relation to itself.

And so it seems with some psychological states. For another illustration, consider a victim of one of David Kaplan's examples<sup>7</sup>. Jones rounds a corner one day and sees someone whose pants are on fire. "My God! That man's pants are on fire!" he exclaims and starts heading for the fire

---

<sup>7</sup> "Demonstratives", in Joseph Almog, , John Perry, and HowardWettstein, 1989. *Themes from Kaplan*. (Oxford: Oxford University Press. 1989) , pp. 481-563.

extinguisher. On the way he feels an unwelcome, southern warmth. He realizes that he has been looking in a mirror. "My God! My pants are on fire!" he yells and then jumps in the lake.

Obviously, Jones undergoes a significant change of mind in the course of the episode. What puzzled Kaplan was how to describe it. Apparently, the difference can't be captured by considering Jones' attitudes *de re*. After all, both before and after he notices the mirror, Jones is such that Jones believes his pants are on fire. On the other hand there is no *dictum* with which we can readily capture Jones state of mind *de dicto*. At the point in the story at which Jones would say:

“His pants are on fire!”

he doesn't believe *his* pants are on fire and doesn't jump in the lake. It's when he says,

“My pants are on fire!”

that he jumps in the lake, but at no point did Jones believe my pants are on fire. He's never met me.

The moral philosophers have drawn from cases like this is that the objects of belief are not propositions— or aren't *just* propositions— but something else: “characters”<sup>8</sup>, “senses”<sup>9</sup>, “concepts”<sup>10</sup>, some other kind of “narrow contents”<sup>11</sup> or “properties”<sup>12</sup>. On these accounts the difference between Jones before and after is a matter of his coming to believe a special sort of *ego-centric* character, sense, concept, etc..

My answer is different. I think that the objects of belief are just propositions. We can not explain Jones's change of mind in terms of his coming to believe new propositions. But, as we saw in Oedipus's case, what propositions people believe is never more than incidentally relevant to the explanation of their behavior. What propositions you believe is only symptomatic of your

---

<sup>8</sup> Kaplan, op. cit

<sup>9</sup> Cf. Frege "Thoughts." In Frege (ed. B. McGuinness, trans. P. Geach and R.H. Stoothoff), *Collected Papers on Mathematics, Logic, and Philosophy*, (Oxford: Blackwell 1984), pp. 351-72 or “John Perry. "Frege on Demonstratives." *Philosophical Review* 86 ( 1977). , pp. 474-97.

<sup>10</sup> Cf. Christopher Peacocke, *A Study of Concepts* ( Cambridge, MA: MIT Press , 1992).

<sup>11</sup> Brown, Curtis, "Narrow Mental Content", *The Stanford Encyclopedia of Philosophy (Spring 2007 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2007/entries/content-narrow/>>.

<sup>12</sup> David.Lewis, "Attitudes *De Dicto* and *De Se*." *Philosophical Review* 88 (1979), pp. 513-43 .

psychological state. Like Oedipus, Jones has made a discovery and, like Oedipus, his recognition centrally consists, not in his acquiring new beliefs, but in his coming to have his old beliefs in a new psychological state.

Psychological states are different if<sup>13</sup> they would be realized by different attitudes in the same situations. By this measure there is a clear difference in Jones before and after his self recognition. Initially, Jones believes that Jones is in danger, but that he believes this about himself is due to features local to his situation. Had there been no mirror but rather some Jones look-alike with smoking pants, Jones would have believed that the look-alike was in danger. Compare this with the state Jones acquires after he feels the heat. This new state is realized by (among other attitudes) a belief that Jones is in danger. But now this belief is self-centered in the way Paris's admiration was. At that point, even if there *were* a look-alike and no mirror, Jones would still believe that he, *himself*, was in danger. It's upon acquiring this self-centered psychological state that he jumps in the lake.

The folk psychological explanation of what went on with Jones would subsume his behavior under a general law:

(L1) If someone believes himself in danger he will attempt to flee.

Where jumping in the lake is the flight behavior appropriate for this threat. But if we tried to frame this regularity as:

$$(\forall x)(x \text{ believes } x \text{ is in danger} \supset x \text{ flees})$$

we get a falsehood since Jones satisfies the antecedent at the very beginning of the story and is not then tempted to take the plunge.

Here again, we need to generalize over states and situations, thus:

$$(L1) \quad (\forall x)((\exists s) x \text{ \& } (\forall s)(\text{in } s : x \text{ believes } x \text{ is in danger})) \supset x \text{ flees}$$


---

<sup>13</sup> But not only if. See below.

This describes a psychological state such that anyone in that state would believe he was in danger, no matter what his situation, and says that anyone in such a state will flee. As we would expect of a folk psychological generalization, (L1) is an instance of a rationalizing principle. If someone satisfies (L1) they will be in danger in all their notional situations and if you *think you are in* a situation in which you are in danger, the rational thing to do is to flee. At the outset of the story this isn't Jones' psychological state. At that point, his psychological state could only be realized by true beliefs in situations in which his pants were not on fire. At that point, he does not think he is in a dangerous situation.

(L1) is, admittedly, a little cumbersome. Let us introduce a notational device to simplify matters. Let, ' $xRy$ ' represent any relational psychological property of  $x$ 's, for example:

$x$  admires  $y$

or

$x$  believes  $y$  is in danger.

Letting ' $x$ ' range over total psychological states, we abbreviate as follows:

" $xRx^*$ " =<sub>df</sub> " $(\exists x)(x \& (\forall y)(\text{in } y \text{ : } xRy))$ "

Tagging variables with an asterisk parallels the English convention of augmenting first person pronouns, with forms like 'himself', 'herself', 'his own' and so on. The folk psychological law (L1) now becomes:

(L1)  $(\forall x)(x \text{ believes } x^* \text{ is in danger} \supset x \text{ flees})$

What makes Jones flee, we may now say, is that he comes to believe that he\* (he *himself*), is in danger. That Jones believes he\* is in danger entails that Jones believes Jones is in danger, but it also says something more about his psychological state. And for the purposes of psychological prediction and explanation it is always psychological states that are relevant.

These asterisked variables are, of course, meant to recall Hector-Neri Castaneda's "quasi-indicators".<sup>14</sup> Castaneda introduced them as a device for marking the special status of pronouns like 'himself and 'he himself' within psychological verbs. For Castaneda, quasi-indicators constituted a semantic problem: they seemed not quite variables and yet not quite names; the constructions they yielded seemed neither wholly referentially transparent nor simply opaque. On our analysis, quasi-

Seeing a guy with burning pants is apt to make Jones behave excitedly and seeing the agitated behavior of the guy with the burning pants might lead Jones to conclude that the guy he is looking at believes himself to be on fire. If Jones does draw this conclusion, then Jones is mistaken, because it is false that the guy (i.e., Jones) satisfies:

$$(1) \quad (\exists x)(x \text{ believes } (\forall y)(\text{in } y \text{ : } x \text{ believes } x \text{ is on fire}))$$

Take away the mirror, substitute someone else with burning pants and it is *that* guy Jones will believe is on fire and that guy Jones will believe has a self-centered belief about being on fire. All this notwithstanding the fact that, in his actual situation, it is true that:

Jones believes Jones is on fire.

and:

Jones believes Jones believes Jones is on fire.

That (1) is false about Jones is no obstacle to Jones's believing it is true about him. In that event we might say:

$$(2) \quad \text{Jones believes he believes he* is on fire.}$$

But now we have a problem, for our asterisk convention leaves it open to interpret (2) in different ways depending upon how we unpack its implicit quantifiers. (2) is ambiguous; ambiguous in precisely the ways:

$$(2') \quad \text{Jones believes he believes his own pants are on fire.}$$

is ambiguous.

How many ways is that? Well, first, there is the difference between believing, of someone, that he believes himself to be on fire versus believing, of someone, that they believe you, yourself, are on fire. This is the difference between:

$$(3) \quad (\exists y)(x \text{ believes } (\exists z)(z \text{ believes } (\forall w)(\text{in } w \text{ : } y \text{ believes } y \text{ is on fire}))))$$

and

(4)  $(\exists x)(x \text{ and } (\forall y)(\text{in } \text{in } : (\exists y)(x \text{ believes } y \text{ believes } x \text{ is on fire})))$

(3) describes Jones when he says:

"That man over there believes that his pants are on fire"

whereas a Jones who satisfies (4) would say:

"That poor fool thinks my pants are on fire"

Neither (3) or (4) should be confused with the case where Jones satisfies:

$(\exists x)(x \text{ \& } (\forall y)(\text{in } \text{in } : x \text{ believes } x \text{ believes } x \text{ is on fire}))$

Which describes a Jones who would say:

"I believe I believe my pants are on fire."

Jones was Kaplan's example. Kaplan's solution to his puzzle was to see belief as having two components. On Kaplan's account, one believes a proposition *under a character*. A character is a function from contexts to propositions. The difference between Jones before and after is that Jones believes the proposition that Jones is in danger under different characters. The character that gets Jones jumping in the lake is of a special egocentric sort whose value is the proposition that the believer is in danger, no matter what the context. Not so the character under which Jones initially believed the proposition.

Why put things my way rather than Kaplan's? Three reasons. The first is ontological: we don't have to posit characters or any other sorts of representations as objects of belief. True, we quantify over psychological states, but we can say (have said) what sort of things psychological states are. Ontologically speaking, they are just a kind of property, properties that we are already committed to if we believe psychological theory is possible.

The second gain has to do with explanation. We could recast (L2) as:

(Lk) Whenever one believes one is in danger under an egocentric character, one flees.

But why should egocentric characters have this property? Do not say that it is because entertaining these characters always yields a belief that one is in danger. Why should that ever get Jones jumping in the lake? Before he noticed the mirror Jones already believed he was in danger. Why didn't he jump then? Kaplan stipulates that the difference in characters will correlate with differences in behavior. But stipulation is not explanation

Our explanation is simply that Jones didn't flee at first because had no *reason* to. We evaluate reasons by examining notional situations. Unless you believe you\* are in danger, the sort of situation you think you are in is not dangerous for you.

The third gain is that, rather than just stipulating that some states are egocentric, we can explain how self-centered states come about.

How did Jones come to believe that he\* was in danger? What converted him from pity to panic? We know it had to do with what he experienced. When he felt the heat, as opposed to just seeing the smoke, he came to believe his\* pants were on fire. That is only psychological common sense. But how does it work?

It isn't enough just to say that feeling the pain will make Jones believe he is in pain. As soon as Jones saw that the man in the mirror (i.e. Jones) had burning pants, Jones concluded immediately that he (i.e., Jones) was in pain. But that only made him rush all the more frantically for the fire extinguisher. Clearly, it would take a belief that he\* is in pain to get Jones to realize that his\* pants are on fire. But Jones believing that he\* has a pain is another self-centered belief and what we were trying to explain was how such beliefs arise.

And yet, we are a little closer to our goal. Jones initial belief that he is in pain may arise from inferences he makes about the likely effect of burning pants. But his belief that he\* is in pain comes about after he *feels* the pain. Our approach helps us to understand why the psychological

states caused by feeling the fire as opposed to just seeing it are different. We can understand that difference if we suppose that it is a contingent fact about human beings that:

$$(L3) \quad (\forall x)(x \text{ feels a pain} \supset x \text{ believes } x \text{ is in pain})$$

I conjecture that (L3) is a psychological law.<sup>15</sup> To have a pain is to have a psychological property. Indeed “being in pain” is as paradigmatic a candidate for a psychologically natural kind of state as ordinary language is likely to offer us. Having the pain is going to be a component of Jones's total psychological state wherever he is situated. Which means that, if (L3) is true, then any sensation of pain will always be conjoined with a belief, by the sufferer, that he is suffering no matter what his situation. That is for any psychological state  $\psi$ :

$$(\forall x) (\psi : x \text{ is in pain} \supset (\forall \psi) (\psi : x \text{ believes } x \text{ is in pain}))$$

So if (L3) is true, so is:

$$(\forall x)(x \text{ has a pain} \supset x \text{ believes } x^* \text{ has a pain.})$$

Thus Jones gets his asterisk: anyone who feels a pain will believe he\* is in pain. Given (L3), having a pain leads you to believe you\* are in pain, just as, given Newtonian laws, force leads to acceleration.

## KNOW THYSELF

Consider John Perry's puzzle about Hemison<sup>16</sup>. Hemison began as biographer of David Hume, but he has gone crazy. He now thinks he *is* Hume. He's crazy but no fool. He knows (suppose) as much about Hume as Hume knew about himself. He acts and thinks exactly as Hume might have done were Hume to have been magically transported to the late twentieth century.

---

<sup>15</sup> Or nearly enough a law for philosophical purposes. Of course, many philosophers have thought (L3) necessarily true. Whether or not this is so does not affect my argument here.

<sup>16</sup> Perry, *op. cit.*

Hemison's psychiatrist is not crazy but is a fool; Hemison has convinced the psychiatrist that he, Hemison, is Hume.

What to say about Hemison? On the one hand, we want to say that Hemison and Hume are, psychologically, very much alike; after all, they believe the same things about themselves. On the other hand, we can't just say they believe "the same things" since virtually everything Hemison believes about himself is false. In some sense the psychiatrist believes the same falsehoods as Hemison, but, then again, not quite the same, since the psychiatrist doesn't believe he, himself, is Hume.

Perry's solution, like Kaplan's, involved treating belief as having two components. To have a belief is to "entertain" a sense in a context; senses are functions from contexts onto propositions believed. Hemison and Hume entertain the same senses and this is what constitutes their psychological identity, but because they entertain them in different contexts (Hume's head *vs.* Hemison's) they believe different propositions. Hemison and the psychiatrist accept the same propositions but by way of entertaining different senses, hence their psychological differences.

I think Perry's solution, like Kaplan's, raises more questions than it answers.<sup>17</sup> What are senses and how are they "entertained"? How does a sense fix an object in a context? Is it a *contingent* fact that a sense picks out the objects in the contexts that it does and if so what would happen if those facts were different? Why should a sense which has the *semantic* property of yielding the entertainer in any context entertained be paired with the *psychological* trait of promoting varieties of self-centered behavior? Why for example, why do we expect someone who entertains the sense of 'I am beautiful'— to behave vainly? Don't say it's because such a person believes they are beautiful. Remember self-loathing Sally believed that proposition, albeit under the sense of

---

<sup>17</sup> To be fair, it should be noted that these questions are problems for *any* version of the Representational Theory of Mind.

‘That woman is beautiful’. What is different about these senses that yields these different behaviors?

It seems to me we do better by practicing two dimensional psychology rather than two dimensional semantics. *Ex hypothesi*, Hemison and Hume are in the same psychological state. The psychologist is in a different state from both, but happens to believe the same propositions as Hemison. Hume’s situation (being called ‘Hume’, authoring *The Treatise &c.* ) is the sort of situation Hemison thinks he is in and his behavior is wholly appropriate to that notional situation.

Hemison believes he is Hume but the fact that Hemison’s beliefs are about Hume is not wholly the upshot of his psychological state. Transposed to Twin Earth, Hemison would be believing he was Twin Hume. Hemison does not believe he is Hume\*. On the other hand, we expect that Hume, *wherever* he were situated, would there believe that he was Hume and no one else. We expect this not because we suppose that Hume was specially self aware, but because it is plausible to suppose that everyone satisfies:

$$(L4) \quad (\exists x)(x \text{ believes } x = x) \ \& \ (\forall x)(x \text{ believes } x = x)$$

Certainly this is so of anyone prepared to say “I am me”. Which is why we suppose that Hume believes that he is Hume\*, indeed that he believes he\* is.<sup>18</sup> Hemison, like Hume, believes he is himself in every situation, but that means only that Hemison believes he\* is Hemison\*. The difference arises from the uncomplicated fact that Hemison is identical with Hemison, not Hume.

David Lewis offers us a divine puzzle.

Consider the case of the two gods. They inhabit a certain possible world, and they know exactly which world it is. Therefore they know every proposition that is true at their world.

Insofar as knowledge is a propositional attitude, they are omniscient. Still I can imagine

---

<sup>18</sup> And we can construct cases in which Hume simultaneously believes that he\* is Hume\* and he is not. Thus: Hume unwittingly pointing at the mirror insisting “He is not Hume, I am!”.

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.<sup>19</sup>

Lewis tells us that even seeing a map marked “You are here” wouldn’t help the Gods, since they see everything, from every perspective, equally well.

The case is puzzling because it shows one can believe nothing but the truth and the whole truth but still be ignorant. What is it that the gods don’t know?

Lewis's answer is that the objects of belief are not propositions but *properties*. To believe anything is to ascribe oneself a property. Each of the gods ascribes to himself the property of inhabiting a world where certain propositions are true, e.g., the proposition that Jehovah lives on the highest mountain and Zeus lives on the coldest. Still, neither Zeus or Jehovah is prepared to ascribe to himself the property of living on the highest mountain or the property of living on the coldest.

This is elegant. But the down side is that to accept this account we must take the relation of "ascribing oneself a property" — belief *de se*— as primitive. Which is to say that we must not press for an explanation, e.g. of the difference between Jehovah's ascribing to himself the property of being called ‘Jehovah’ and Jehovah's ascribing to himself the property of being at a world where Jehovah is called ‘Jehovah’. It's hard not to feel that this takes as primitive the very difference we wanted explained.

I think we can do better than this. As Lewis tells the story it would make no difference to the gods if they switched places and *modus operandi*. If so, then there must be at least two situations which are indistinguishable from the divine perspective:

---

<sup>19</sup> Lewis, *op. cit.*

	Zeus	Jehovah
In <sub>1</sub>	named 'Zeus' throws thunderbolts on coldest mountain	named 'Jehovah' throws manna" on warmest mountain
In <sub>2</sub>	named 'Jehovah' throws manna on warmest mountain	named 'Zeus' throws thunderbolts on coldest mountain

Transposed from <sub>1</sub> to <sub>2</sub>, the Gods would, given godly omniscience, believe nothing but the truth about that situation. For example, in <sub>2</sub>, Jehovah would believe correctly, there, that he was not named 'Jehovah'. But, if that is so, then it is false that:

$$(\forall x)(\forall y)(\text{in } y \text{ in } \Psi^J x: x \text{ believes } x \text{ is named 'Jehovah' } )$$

where ' $\Psi^J$ ' is Jehovah's total psychological state. That means, in our terms, that Jehovah does not believe 'Jehovah' is his\* name or, *mutatis mutandis*, that he\* is on the warmest mountain or that he\* throws manna. This explains why, while Jehovah may thunder: "I am me!" and "Jehovah is Jehovah", we should not expect him to say "I am Jehovah". He doesn't have sufficient *reason* to. His notional situations include <sub>1</sub> and <sub>2</sub>. So, in the sort of situation *he thinks he is in* 'Jehovah' may or may not be his name. This, in turn, is why we want to say that Jehovah does not *know* what his name is. Jehovah knows all the true propositions but he lacks *information*. Information that would exclude <sub>2</sub> from his notional situations.

How could the gods figure out who they are? Lewis doesn't tell us. On his account, beliefs *de se* are either irreducibly basic or are inferred from other beliefs *de se*. But we have already seen how it might come about.

Let one of our gods, Jehovah say, get a pain—the smallest twinge will do—and suppose that divine psychology is enough like human so that the gods satisfy (L3). That would suffice to get a pained god to believe he\* was in pain, just as it did with Jones. Of course, since Zeus is omniscient, Zeus will believe that Jehovah is in pain the moment that becomes true. But because Jehovah's pain is part of

Jehovah's psychological state, Jehovah's belief that he is in pain will be self-centered: he would have it no matter what his situation. In contrast Zeus's belief that Jehovah is in pain won't be centered at all. Transplant Zeus to a situation in which some *other* pained deity is on the highest mountain and Zeus would be believing *that* god was in pain.

Having a pain would let Jehovah *know* who and where he\* was. Knowledge is true, *informed* belief.

### NOW WHAT?

Once upon a time— December 31, 1999, to be exact— Smith was informed by a reliable source that an attempt would be made on his life the next day, shortly after midnight. Not one to take chances, Smith set his trusty alarm clock to go off at midnight. He spent an anxious evening trying to divert himself with other matters, glancing uneasily at the clock every few minutes. Finally, the alarm went off. Smith ran off to hide under the bed.<sup>20</sup>

As before, our problem is to characterize a change of mind: the change prompted by the ringing alarm. From a folk psychological point of view the case is unproblematic: Smith believes he will be in danger at midnight and believes the alarm will go off then. So, when he hears the alarm go off, he believes he is in danger then and he flees. But how is a theory of the attitudes to account for this? What is the difference between believing that the alarm *has* gone off as opposed to believing that it *will* go off; between believing one *is* in danger as opposed to believing that one *will be*?

It seems that Smith's flight is an instance of the sort of behavior described by (L2). But we cannot simply amend (L2) to:

$$(L2') \quad (\forall x)(\forall t)(x \text{ believes } x^* \text{ is in danger at } t \supset x \text{ flees at } t)$$

---

<sup>20</sup> Cf. Perry *The Problem of the Essential Indexical*, *Nous* 13, no. 1 (1979), pp. 2-21.

Suppose it is midnight. Smith believes, of midnight, that he is in danger then. Midnight is such that Smith has believed it a dangerous time ever since he first heard the assassination threat. But if Smith's clock is slow he won't flee at midnight. He won't flee until the alarm goes off. Apparently, believing you are in danger *now* involves more than believing, of the present moment, that you are in danger then.

It is clear how the right answer *must* go. The right answer must be that there are different ways one can think about a time. A time can be thought of as past, present or future. All day, Smith has believed midnight will be a dangerous time, but when the alarm goes off at midnight he comes to think of midnight in a new, present tensed, way. Something like this must be the right answer. It's only common sense. But how should we translate this answer into the terms of a philosophical psychology?

Philosophers in the Fregean tradition will want to say it is matter of Smith's coming to *represent* a time with an special sort of indexical "sense" or "character" (etc.). Lewis will say it is matter of Smith's ascribing himself the property of being in danger as opposed to being in a world in which he is danger at a certain time.

I think we can do better than this. The crucial thing to notice about Smith's case is there are *two* temporal factors involved in the story. There are the times when he has a belief and the times his belief is about. The difference between Smith, before and after, turns upon the way in which his psychological states, before and after, connect the time he has these states with the times his beliefs are about.

Consider Smith's total psychological state just before the alarm goes off. Smith believes he will be in danger, but does not believe he is in danger yet. The time he believes is dangerous is 12:00 a.m., Jan. 1, 2000. And, in more ways than one, this time is the fixed object of Smith's belief. Let him fall asleep before midnight and snooze like Rip van Winkle for 100 years. If he wakes in the same psychological state in which he went to sleep, he will still be fearing the start of 2000, still believing that time is in his future.

On the other hand, consider the psychological state which Smith is disposed to acquire as soon as he hears the alarm. In the alarmed state, Smith will believe he is in danger then. When? Whenever it is that he hears the alarm go off and acquires that state. The post alarm state is such that having it gives its bearer frightening beliefs about the time at which it is had, at the time at which it is had. A *present centered* psychological state is one which is always realized by beliefs about the time at which the subject is in that state.

Some notation will help. Situations are spatio-temporal locations in possible worlds, so every situation is associated with a unique time. Let us write:

in  $s$  at  $t$ :  $p$

when we want to say that  $p$  is true at a time  $t$  at which  $s$  is centered. This will save us from the tedium of translating what we want to say into "eternal" sentences. The psychological states that interest us here are ones that are realized by relational attitudes towards times. Such beliefs can be about times different from the times when they are had. That is, it can be true that:

at  $t_i$ :  $x$  believes  $t_j$  is dangerous.

Whether or not  $t_i = t_j$ . Someone's belief about a time is present centered if his psychological state would be realized by a belief about the time at which it is had whenever it is had. That is:

$(\exists s) x \& (\forall t) (\text{in } s \text{ at } t \text{ in } s : x \text{ believes } \dots t \dots)$

which we can abbreviate as:

$x \text{ believes } \dots t^* \dots$

Smith's pre-alarm belief that he is in danger is not present centered. Not even if he has it at the time which he believes to be dangerous. If the clock is slow, then *at* midnight, Smith will believe *of* midnight that he is in danger then. But he will not believe he is danger then\*. When the alarm rings, *whenever* that is, Smith will believe of that time, that he is in danger then\*.

When we say that someone,  $x$ , *now* believes, of a time,  $t$ , that it is past, or present, or future we are saying:

PAST:  $x$  believes it is now\* later than  $t$ .

PRESENT:  $x$  believes now\* =  $t$ .

FUTURE:  $x$  believes it is now\* prior to  $t$ .

The right way to incorporate a temporal component into (L2) is by amending it to read:

(L2)  $(\forall x)(\forall t)(x \text{ believes } x^* \text{ is in danger at } t^* \supset x \text{ flees at } t)$

(L2) is a sensible rule for rational organisms to live by. If the clock is slow, Smith will believe, at midnight, that he is in danger then, but won't flee. Why not? Because he has no *reason* to. Smith is in the pre-alarm state and in his notional situations the danger is still in his future: his psychological state would be realized by true beliefs only at times prior to the dangerous time. After the alarm, Smith's state changes. He then thinks he is in a sort of situation in which he is in immediate danger. And, sensibly enough, he flees: Any time such that you believe you\* are in danger then\*, you should flee, *then*.

How are present centered states *possible*? How can psychology insure that an attitude is about the time at which it is had? In Smith's case the explanation obviously has something to do with the alarm. The assassination threat and the trust he places in his clock have placed him in a psychological state such that:

(5)  $(\forall t)(\text{Smith believes the alarm is ringing at } t^* \supset \text{Smith believes he}^* \text{ is in danger at } t^*)$

Smith will acquire a present centered belief that he is in danger when he acquires a present centered belief that the alarm is ringing. And it is not hard to see how Smith could come by a present centered belief of the latter sort, *viz.* by *hearing* the alarm ring *then*. That is all that's needed, provided we assume that something like:

(L5)  $(\forall x)(\forall t)(x \text{ hears ringing at } t. \supset x \text{ believes } x \text{ hears ringing at } t)$

is a contingent law of human psychology. Note that (L5) is asterisk free but, if it is true, centering is inevitable. Having an auditory experience is a component of an individual's total psychological state. Thus any time, in any situation, than someone's total psychological state includes the hearing of bell-like sounds, that state will also be realized by a belief that a bell-like sound is being heard at that time. Which is to say that if (L5) is true, then so is:

$$(\forall t)(\text{Smith hears ringing at } t \supset \text{Smith believes he* hears ringing at } t^*)$$

Given that (5) is true in Smith's pre-alarm state, when Smith hears the bell, he will think he\* is in danger then\*.

If Smith's clock is correct, then we'll say he flees because he "knows what time it is". His auditory experience has given him new *information*, even though it has not lead him to believe any proposition he did not believe before. The moral, once again, is that knowledge is a matter of information, not justification.

Terrance Tomkow  
Dalhousie University