

Fitting Attitudes *De Dicto* and *De Se*

JASON TURNER

December 8, 2009

According to orthodoxy, the contents of propositional attitudes are, well, propositions. To believe, desire, hope or fear is to stand in a certain relation to a proposition. On an alternative account, developed and defended by David Lewis (1979) and Roderick Chisholm (1979 and 1981), the contents of the attitudes are not propositions but *properties*: to believe, desire, hope or fear is to stand in a certain relation to a property.

Call this alternative the *property theory*. Its chief advantage over orthodoxy is its superior handling of the irreducibly first-personal *de se* attitudes. Not knowing I am looking in a mirror, I see a man with an inkspot on his collar and say to myself, "His collar is stained." I may believe that Jason Turner's collar is stained — perhaps, thanks to selective amnesia, I remember what Jason Turner looks like but not that I am he — but, until I realize it is a mirror I am looking in, I still lack the belief that *my* collar is stained. What I lack is a *de se* belief.

If I already believe the proposition that Jason's collar is stained, what new entity do I believe when I realize that I am the one with the stained collar? A property, says the property theorist: the property of having a stained collar. For someone to believe *de se* of himself that he is *F* is for him to stand in the belief relation to the property *F*. But the property theorist need not posit a disjunctive account: properties for *de se* attitudes, propositions for *de dicto*. For whenever orthodoxy says the content of our attitude is *P*, the property theorist says the content of our attitude is the property of being such that *P*.

Property theory isn't the only theory that handles the *de se* attitudes well.¹ But it does have some particularly attractive features lacked by its rivals. First, when different people have the intuitively 'same' *de se* belief, property theory can say that there is a single object — a property — that is the common object of their beliefs. If David Hume believes *de se* that he is David Hume, and I believe *de se* that I am David Hume, then we both believe the same thing: the property of being David Hume. However, since Hume instantiates this property and I don't, he is right and I am deluded.

Second, property theory can distinguish *de se* and *de dicto* beliefs that are intuitively different. Hume's *de se* belief that *he* is Hume and his *de dicto* belief that Hume is Hume are different: when he believes *de se* that he is Hume, he believes the property of being Hume. But when he believes *de dicto* that Hume is Hume, he believes the property of being such that Hume is self-identical — a

¹See, e.g., Boër and Lycan 1986 and Perry 1993 for alternatives.

very different property altogether. After all, everybody instantiates the second, but only Hume instantiates the first.²

But despite these advantages, not all is well in paradise. Property theory faces the *non-existence* problem (Markie, 1984; Nolan, MS). It comes in two forms. First, the *de se* problem: Sadsack desires that he not exist. The property theory says that he desires the property of non-existence. But, necessarily, nothing has the property of non-existence — every world is as bereft of non-existent objects as it is of round squares — so property theory must say his desire is not satisfied in any world. This result is clearly wrong; there are plenty of worlds where Sadsack doesn't exist, so there should be plenty of worlds where his desire is satisfied.

Next, the *de dicto* problem: I wish for world peace. According to the property theory, I wish the property of being such that there is world peace. If everybody lives in peace and harmony in a world w_1 whereas world w_2 is rife with war, my wish is intuitively fulfilled in w_1 and not in w_2 , even if I exist in neither. Insofar as my pacifistic desires are concerned, I ought to favor w_1 over w_2 . But if I am absent from both worlds, I instantiate the property of being such that there is world peace in neither. So the property theory must say my wish is satisfied in neither, that w_1 is not favored over w_2 in virtue of my wish.

The trouble has nothing to do with the particular nature of desires. Suppose the clever arguments of, say, an early Peter Unger (1979) have befuddled me into believing that I do not exist. I am wrong — I *do* exist — but we can meaningfully ask which worlds my (actually mistaken) belief is true in. We cannot say my belief is true in those worlds where its content is true, for its content is not a proposition but a property, and properties are neither true nor false. So we will likely say my belief is true in those worlds where I instantiate its content. But I instantiate that content — the property of non-existence — in *no* worlds, so my belief is true in no worlds, including worlds where I do not exist. (Likewise for *de dicto* beliefs: my belief that there is *not* world peace ought to count as true even relative to some worlds where I do not exist. But if I do not exist in those worlds, and if my belief's being true in a world is a matter of my instantiating its content in that world, we will not be able to say this.)

The property theory's trouble stems from its account of what it is for an attitude to *fit* a world. The objects of attitudes are by anybody's lights representational: they depict the world as being a certain way. To have a desire with a content is to want the world to be the way depicted by the content; to have a belief with a content is to think the world is the way depicted by the content. If these contents are propositions, they are egalitarian in their representation:

²See Lewis 1979 for both of these examples.

they depict the world as viewed from nowhere. If they are properties, they are egocentric: they depict the world as viewed from the subject's perspective. But either way, being representational, they can depict accurately or inaccurately. When an attitude's content accurately represents a world, we say it *fits* that world. A belief is true, a desire satisfied, a wish fulfilled or a fear realized according to a world when the thing believed, desired, wished or feared fits that world; it is false, unsatisfied, unfulfilled or unrealized according to that world otherwise.

Orthodoxy holds that an attitude fits a world if and only if its propositional content is true at that world. Property theory must say something else, and the most natural thing for it to say is that an attitude fits a world if and only if the attitude's content is instantiated by the subject of the attitude in that world. In other words, it is natural for property theory to subscribe to:

The Natural Account of Fit: If a subject *S* has an attitude with content *F*, *S*'s attitude fits a world *w* if and only if *S* instantiates *F* in *w*.

The non-existence problems come from combining the Natural Account with:

Property Actualism: For any object *x*, property *F*, and world *w*, if *x* instantiates *F* in *w*, then *x* exists in *w*.³

If we endorse both claims, we will be forced to say that a subject's attitudes can only fit worlds where the subject exists.

This is unacceptable. What is to be done? We might reject Property Actualism. It is, at least, a contestable thesis.⁴ But it would be nice to solve the problem without making any substantive claims about the metaphysics of modality; it would be nice to have an account of fit compatible with Property Actualism. The thing to do, then, is to revise our account of fit. But what account shall we put in its place?

Consider for a moment a problem faced by certain philosophers in the orthodox camp. Suppose someone accepted the following theses:

Propositional Fragility: For any object *x*, proposition *P*, and world *w*, if *x* is a constituent of *P*, then if *P* exists in *w*, *x* also exists in *w*.

³For extended discussions of Property Actualism, see Plantinga 1982, pp. 11–15 and Fine 1985, pp. 194–202. Plantinga calls this "Serious Actualism," but I, like Fine (pp. 194–195), cannot see what is so frivolous about other sorts of actualisms.

⁴We might instead follow Timothy Williamson (e.g. 2002) in denying that anything has contingent existence, recapturing ordinary intuitions about what can and can fail to *exist* as really being about what can and can fail to have a special property of *concreteness*. Then everything exists necessarily, although ordinary tokens of "exist" really express a quantifier restricted to concrete things. Sacks's desire that he not exist is then *really* a desire that he not be concrete.

Alethic Actualism: For any proposition P and world w , if P is true in w , then P exists in w .

Suppose she thought further that propositions are “Russellian” in the sense that, say, Leibniz is a constituent of the proposition that Leibniz doesn’t exist. She would then be forced to say that the proposition that Leibniz doesn’t exist is true in no world. Such a theorist won’t want to say that an attitude’s fitting a world is a matter of its content being true in that world; if she did, she would be forced to say that, for instance, Newton’s desire that Leibniz not exist is satisfied in no world at all, including worlds that Leibniz does not inhabit.

To account for fit, she may instead borrow a distinction between two kinds of world-indexed truth from Kit Fine:

One should distinguish between two notions of truth for propositions, the *inner* and the *outer*. According to the outer notion, a proposition is true in a possible world regardless of whether it exists in that world; according to the inner notion, a proposition is true in a possible world only if it exists in that world. We may put the distinction in terms of perspective. According to the outer notion, we can stand outside a world and compare the proposition with what goes in the the world in order to ascertain whether it is true. But according to the inner notion, we must first enter with the proposition into the world before ascertaining its truth. (1985, p. 194)⁵

Say that a proposition is true *in* a world iff it is true there according to the inner notion, and true *at* a world iff it is true there according to the outer notion. To be true in a world w is to be there, in w , accurately representing how w is from the inside. To be true at a world w is to be *here*, with us, accurately representing how w is from the outside.⁶ Since the proposition that Leibniz doesn’t exist does in fact exist here in the actual world, and accurately represents, from the outside, the goings-on of Leibniz-less worlds, our orthodox friend avoids trouble if she simply says an attitude fits a world by virtue of having a content true at that world.

Our problems parallel those of the imagined orthodox theorist. She had an egalitarian representation — a proposition — that accurately represented

⁵I’ve presented Fine’s distinction as between two different ways to be true at a world, but his distinction is officially between two ways to be true *simpliciter*. The subtle differences between that official view and our unofficial gloss here won’t matter in what follows, though.

⁶Some (e.g. Plantinga, 1985, esp. pp. 341–344) find the distinction tendentious as it stands, in need of philosophical elucidation. I attempt to elucidate it elsewhere (Turner, 2005); I here take it for granted.

worlds from the outside, but could not do so from within. We have an egocentric representation — a property relativized to a subject — that accurately represents worlds from the outside, but cannot do so from within. The pair ⟨Sadsack, non-existence⟩ represents, egocentrically, from the perspective of Sadsack, the non-existence of Sadsack. It thus accurately represents, from the outside, any Sadsack-deprived world. But it cannot represent the non-existence of Sadsack from *within* a Sadsack-deprived world, for those worlds do not have the Sadsack perspective for it to represent *from*.

If a proposition that accurately represents, from the outside, the goings-on of a world w is called “true at” w , then if an object/property pair likewise accurately represents, from the outside, the goings-on of w , we ought to say that the object instantiates the property “at” w . If I do not exist in w , then I instantiate non-existence *at* w , for w is a world where I am as the property of non-existence represents me: namely, not. Yet I do not instantiate non-existence *in* w , for that would require me and the property to be inhabitants of w related by instantiation.

It will no doubt be objected that, without a more detailed account of property instantiation at worlds, this is just picture thinking. So here is a more detailed account. The property of non-existence and the proposition that I don’t exist are intimately bound up with each other: in order to get the proposition that I don’t exist, we just need to add me to the property of non-existence. So we can say that I instantiate non-existence at exactly those worlds at which the proposition created by combining me and non-existence is true.⁷ And in general, whenever a property and a proposition are tied together in this way, so that the proposition P is the result of combining an object x and a property F , then x instantiates F exactly at worlds at which P is true.

More precisely: let a *propositional function* be a function from objects to propositions. Imagine we had access to an ideal “language of metaphysics” with a predicate for every property and a singular term for every object. (We’ll let it be an infinitary language, too, just for good measure.) Then for any property F , there would be a predicate Π that expressed F . There would also be, for any variable α and object x , a proposition expressed by $\lceil \Pi(\alpha) \rceil$ when α is assigned to x — and, thanks to the compositionality of language, that proposition would be a function of F and x . So if Π expresses F there will be a unique propositional function that returns, for any object x , the proposition that is expressed by the

⁷If we accept Propositional Fragility, Alethic Actualism, and the contention that I’m a constituent of the proposition that I don’t exist, we’ll have to say I instantiate non-existence at worlds where the proposition that I don’t exist is true with respect to the outer notion. I will speak of truth “at” worlds hereafter, with the understanding that, if the inner/outer distinction turns out to be unnecessary, truth at a world just is truth in a world.

sentence $\lceil \Pi(\alpha) \rceil$ when α is assigned to x . Call this the *propositional function for* F , p_F . The proposal, then, is that an object x instantiates a property F at a world w if and only if $p_F(x)$ is true at w .

We can use at-a-world instantiation to fix the property theory's account of fit:

The Revised Account of Fit: If a subject S has an attitude with content F , S 's attitude fits a world if and only if S instantiates F at that world.

Let N be the property of non-existence. The proposition $p_N(\text{Sadsack})$ — the proposition that Sadsack doesn't exist — is true at some worlds, so Sadsack instantiates N at some worlds, so his desire is satisfied by some worlds. Likewise, if W is the property of being such that there is world peace, $p_W(\text{Jason})$ is just the proposition that there is world peace (p_W delivers the same proposition for *any* argument). It is true at some worlds where I do not exist, so I instantiate W at some worlds where I do not exist, so my wish is fulfilled at some worlds where I do not exist.

An objection: "Compositionality does guarantee you a propositional function for non-existence, but it's not the one you say it is. Let's distinguish *predicative* from *impredicative* propositions. Predicative propositions say of something that it has some property; impredicative ones do not. The proposition that Fido is brown, for instance, says that Fido instantiates brownness and is predicative. That Fido is non-brown is also predicative and says that Fido instantiates brownness's complement. But the impredicative proposition that it's not the case that Fido is brown doesn't ascribe non-brownness to Fido, but merely withholds brownness from him.

"Something similar happens when you apply p_N to Sadsack. Doing so gives you the predicative proposition that Sadsack is non-existent, which says that Sadsack is a certain way (namely, non-existent) and therefore cannot be true even *at* worlds where Sadsack does not exist. Admittedly, there is another, impredicative, proposition — the proposition that it's not the case that Sadsack exists — which is true at Sadsack-less worlds, but that's not the result of combining Sadsack and the property of non-existence. Your problem is that you have confused the propositional function associated with the open sentence 'It's not the case that x exists' with the one associated with the complex predicate 'is non-existent.'⁸

"Likewise, when you apply p_W to yourself, you get the proposition that you are such that there is world peace, which says that you are a certain way (namely, inhabiting a peaceful world) and therefore cannot be true even *at* worlds where you do not exist. Admittedly, there is another, impredicative proposition — the

⁸Cf. Plantinga 1985, pp. 345–347.

proposition that there is world peace — which is true at Jason-less worlds, but that’s not the result of combining you and the property of being such that there is world peace. You have confused the propositional function associated with the open sentence ‘ x is such that there is world peace’ with the constant function from individuals to the proposition expressed by ‘There is world peace.’”

I do not know what to think of the alleged distinction between predicative and impredicative propositions. So I let the objector have it: let’s suppose that the property of non-existence, whatever that may be, is not associated with a function that takes us from (say) Sadsack to the proposition that it’s not the case that Sadsack exists. We still have the propositional functions themselves; let *them* be the objects of the attitudes. If Sadsack cannot desire the property of non-existence on pain of having no worlds satisfy his desire, let him desire the propositional function associated with the open sentence “It’s not the case that x exists,” and say that his desire is satisfied in exactly the worlds at which the proposition delivered when that function is applied to him is true.

More precisely: let a *property** be a propositional function p where, for some sentence ϕ and variable α (of our idealized, infinitary language), $p(x)$ is the proposition expressed by ϕ when α is assigned to x , for any x . Say that a property* p is *instantiated** by x iff $p(x)$ is true. An object x *instantiates** a property* *in* a world w iff x and p are inhabitants of w related by instantiation*, and x *instantiates** p *at* a world iff $p(x)$ — the proposition actually delivered by p for x — is true at w . We can now say that the objects of attitudes are properties* and that these attitudes fit worlds when and only when their subjects *instantiate** their contents at those worlds.⁹

In this case, we say that Sadsack desires the property* associated with the open sentence “ x does not exist”, and he *instantiates** this property* at worlds at which its value for Sadsack — the proposition expressed by “ x does not exist” when x is assigned to Sadsack — is true. Likewise, we say that I desire the property* associated with the (closed) sentence “There is world peace”, and I *instantiate** it at worlds at which its value for me — the proposition expressed by “There is world peace” when x is assigned to me, otherwise known as the proposition that there is world peace — is true.

Our modifications have sacrificed none of the advantages of property theory. If Hume and I each believe *de se* of ourselves that we are Hume, then we each believe the same thing: the property* associated with the open sentence “ x is

⁹Some may worry that functions are modally fragile and therefore unfit candidates for attitudes’ contents. But this comes from assuming functions are sets of ordered pairs. We may instead construe functions as a special case of relations, understood as of a kin with properties. Then there is no reason to think that our new functional contents are any more fragile than our old property ones.

Hume". Since this function takes Hume to a true proposition and me to a false one, he instantiates* this property and I do not. That is, his belief is true and mine is false.

Likewise, if Hume believes both *de se* that he is Hume and *de dicto* that Hume is Hume, he believes two different properties*: the function associated with the open sentence "x is Hume" and the function associated with the closed sentence "Hume is Hume". The latter property* is a constant function, taking everyone to the true proposition that Hume is self-identical; everybody instantiates* it. But the former property* takes everybody but Hume to the false proposition that they are Hume; as a result, only Hume instantiates* it.

But our objector is nonetheless dissatisfied. "You started out to defend the property theory, but now you have done no such thing," he says. "You have rather given up on property theory, replacing it with this property* theory instead. But properties* aren't properties. Your so-called 'defense' of a theory is really a retreat."

A retreat perhaps, but a retreat to higher, firmer ground. The objector appropriated the term "property" for himself, insisting that a property will never combine with an object *x* to deliver a proposition that can be true at *x*-less worlds. We, in an effort to be as ecumenical as possible about the metaphysics of properties, let him have the term. Now he complains that the theory we defend does not deserve its name. Fair enough; let us defend property* theory instead. Property* enjoys all the advantages of property theory, and gets around the non-existence problem besides. We can have propositional attitudes with genuinely egocentric contents, call them what you will. Our project was never to defend that any conception anyone has ever had of properties can be used as contents of attitudes, but merely that attitudes with egocentric contents — which we called "properties" before our interlocutor took that name unto himself — can be made to fit even those worlds where their subject doesn't exist. That we have done.¹⁰

REFERENCES

- Boër, Steven E. and William G. Lycan (1986). *Knowing Who*. Cambridge, Mass.: MIT Press.
- Chisholm, Roderick (1979). "The Indirect Reflexive." In Cora Diamond and Jenny Teichman (eds.), *Intention and Intentionality: Essays in Honour of G. E. M. Anscombe*, 39–53. Brighton: The Harvester Press.

¹⁰Thanks to Neil Feit, Shieva Kleinschmidt, Daniel Nolan, and Ted Sider for helpful comments, discussion, and encouragement.

- (1981). *The First Person: An Essay on Reference and Intentionality*. Minneapolis, Minn.: University of Minnesota Press.
- Fine, Kit (1985). "Plantinga on the Reduction of Possibilist Discourse." In J. E. Tomberlin and van Inwagen (1985), 145–186. Reprinted in Fine (2005, 176–213).
- (2005). *Modality and Tense*. Oxford: Oxford University Press.
- J. E. Tomberlin and Peter van Inwagen (eds.) (1985). *Alvin Plantinga*. Dordrecht: D. Riedel.
- Lewis, David (1979). "Attitudes *De Dicto* and *De Se*." *The Philosophical Review* 88: 513–543. Reprinted in Lewis (1983): 133–156.
- (1983). *Philosophical Papers*, volume 1. Oxford: Oxford University Press.
- Markie, Peter J. (1984). "*De Dicto* and *De Se*." *Philosophical Studies* 45(2): 231–237.
- Nolan, Daniel (MS). "Selfless Desires." *Philosophy and Phenomenological Research* Forthcoming.
- Perry, John (1993). "The Problem of the Essential Indexical." In *The Problem of the Essential Indexical and Other Essays*. Oxford: Oxford University Press.
- Plantinga, Alvin (1982). "On Existentialism." *Philosophical Studies* 44: 1–20.
- (1985). "Replies to my Colleagues." In J. E. Tomberlin and van Inwagen (1985), 313–396.
- Turner, Jason (2005). "Strong and Weak Possibility." *Philosophical Studies* 125: 191–217.
- Unger, Peter (1979). "I Do Not Exist." In *Perception and Identity*, 235–251. Ithaca, N.Y.: Cornell University Press.
- Williamson, Timothy (2002). "Necessary Existents." In *Logic, Thought, and Language*, 233–251. Cambridge: Cambridge University Press.