

## Twins, Swampmen and Our Interpretive Practice

I'd like to talk today about the apparent clash between two sets of intuitions, which for short I'll label "Twin" and "Swamp". The two can be characterized roughly as follows.

**Twin:** If there were a world just like ours, a 'Twin Earth', but whose 'water' was made up of a substance different from ours (say, XYZ rather than H<sub>2</sub>O), then the people in that world would mean something different than we do by "water". In particular, their 'water' claims would be true or false of XYZ, while ours would be true or false of H<sub>2</sub>O. Indeed, even if I had a perfect duplicate on Twin Earth (supposing that we were both ignorant of chemistry) we would mean different things by "water".<sup>1</sup>

**Swamp:** Suppose lightning strikes a dead tree in a swamp; I am standing nearby. My body is reduced to its elements, while entirely by coincidence (and out of different molecules) the tree is turned into my physical replica. My replica, the Swampman, moves exactly as I did; according to its nature it departs the swamp, encounters and seems to recognize my friends, and appears to return their greetings in English. It moves to my house and seems to write articles on radical interpretation. No one can tell the difference, since the Swampman can speak English and thinks just what I would have thought in his place.<sup>2</sup>

Both intuitions are fairly robust, but it has become increasingly common to think that the two, in spite of their *prima facie* plausibility, cannot coexist.

In his Presidential address to the American Philosophical Association, Donald Davidson famously denied the intuition behind **Swamp**, and after admitting that "no one can tell the difference" between the Swampman and himself, he still insisted:

But there *is* a difference. My replica can't recognize my friends; it can't *recognize* anything, since it never cognized anything in the first place. It can't know my friends' names (though of course it seems to), it can't remember my house. It can't mean what I do by the word 'house', for example, since the sound 'house' it makes was not learned in a context that would give it the right meaning -- or any meaning at all. Indeed, I don't see how my replica can be said to mean anything by the sounds it makes, nor to have any thoughts.<sup>3</sup>

According to Davidson, since the Swampman lacks the relevant type of causal history, it initially has no contentful thoughts at all. Cases like **Twin** are taken to show that a term's causal history determines what it means, and so if something's current use of a term lacks such history, it must be meaningless.

---

<sup>1</sup> The best known development of this intuition, see Putnam, H. (1975). "The Meaning of 'Meaning'". Mind Language and Reality. H. Putnam. New York, Cambridge University Press: 215-271.

<sup>2</sup> The example, though certainly not the intuition about it, is Davidson's. (See Davidson, D. (1987). "Knowing ones own mind". Proceedings and Addresses of the American Philosophical Association. pp. 443-44.)

<sup>3</sup> Ibid. pp. 443-44.

Davidson's claim that the Swampman meant nothing by his words and had no contentfull thoughts has not, however, been universally accepted. The intuition that the Swampman has thoughts is very robust. For instance, it would be very difficult to insist that since the Swampman is initially a non-thinking creature, we can't hold it responsible for what it does. Indeed, many took Davidson's paper to show instead that, since the Swampman clearly had thoughts and meant things by his words, 'historical' accounts of meaning and content must be mistaken. That is to say, we should hold on to **Swamp** even if it means giving up on **Twin**. As Jerry Fodor eloquently put it:

One might, like Davidson, treat these consequences of the historical account of content as bullets that one's intuitions must learn to bite. But I think that they should be viewed rather as serious embarrassments for Davidson's causal history kind of externalist semantics.... Scissor cuts paper, Swampman eats Twins .... I think the unbiassed intuition is that Swampman thinks all sorts of thing that Davidson does; that it's Wednesday, and that radical interpretation is possible, and that water is wet, for example. I think this is because, although he lacks Davidson's causal history, Swampman shares Davidson's dispositions, and it's the *counterfactuals* that count for content, just as informational theories claim. True, no 'water' token of Swampman's has ever been caused by water. But Swampman means *water* by "water" for all that. The reason he does is that it's water that *would* cause his "water" tokens in the worlds that are closest to one that Swampman actually lives in. Roughly, a thought means what *would* cause it to be tokened in nearby possible worlds.<sup>4</sup>

As he puts it later, "it is only demonstrative thoughts whose content is determined by their *actual* etiology. Everywhere else, Twins included, it's the counterfactuals that count."<sup>5</sup> Because of this Fodor thinks that we should understand intentionality "in terms of subjunctive rather than actual causal history."<sup>6</sup> Such a subjunctive account allows Swampman to have contentfull thoughts (indeed, the same thoughts as Davidson), but it also has the consequence of having Davidson mean the same thing by his words as, say, his Twin-Earth counterpart.<sup>7</sup> In short, the intuitions revolving around **Twin** suggest that what we mean by our terms is determined by their causal history, while another set of intuitions revolving around **Swamp** suggests that such causal history is unimportant,

---

<sup>4</sup> Fodor, J. (1994). *The Elm and the Expert*. Cambridge, MIT., pp. 117-18.

<sup>5</sup> *Ibid.*, p. 119. Still others take the fact that we have good reason's to endorse historical accounts to entail that **Swamp** just shows that our intuitions are particularly untrustworthy in this area. (See Millikan, R. G. (1996). "On Swampkinds." *Mind and Language* 11(1): 13-117., Neander, K. (1996). "Swampman Meets Swampcow." *Mind and Language* 11(1): 111-129., Papineau, D. (1996). "Doubtful Intuitions." *Mind and Language* 11(1): 130-132.)

<sup>6</sup> Fodor, J. and E. Lepore (1992). *Holism: a shopper's guide*. Cambridge, Blackwell. p.237. See also: "What makes "cat" mean *cat*, according to such theories, is *not* that it *is* caused by cats, but only that it's nomologically necessary that it *would* be caused by cats under some or other circumstances that the theory is obliged to specify." (Fodor, J. and E. Lepore (1992). *Holism: a shopper's guide*. Cambridge, Blackwell. p.157.)

and that, rather, something like our *dispositions* determines meaning and content. Davidson holds on to **Twin** and gives up on the **Swamp**, while Fodor holds on to **Swamp** and gives up on **Twin**. Both choices come at a non-trivial cost for those of us who still treat intuitions as philosophically relevant, and in what follows I'll attempt to argue that one can actually hang on to both.

There is, after all, no tension between the **Twin** and **Swamp** themselves (intuitions about manifestly different cases can't 'clash'). Rather, each is incompatible with popular accounts of meaning that are taken to follow from the other. However, intuitions about cases don't *entail* particular theories of meaning, such theories are at best hypotheses used to explain or systematize the intuitions. Consequently, instead of giving up on one of the two intuitions, we should question that assumption that they have to 'clash', and look for a better theory that accommodates both.

Both Davidson and Fodor write as if what 'settles' content in their favored case determines content simpliciter. A more nuanced view may allow a number of factors to be relevant, and while some may be *more* important than others, it doesn't follow that the less important factors count for nothing. **Twin** points to the importance of our actual causal history in determining content, while **Swamp** points to the importance of subjunctive causal history, but there is no reason why a theory of content couldn't allow *both* to be relevant to what we mean.

Indeed, I think that this follows from Davidson's own methodology of drawing conclusions about meaning based upon our practice of interpretation. Davidson's own negative conclusions about **Swamp** stem from focusing too much on just one aspect of our interpretive practice rather than surveying it as a whole.

In determining what we mean, how we are (and were) disposed to apply our terms is undoubtedly treated as relevant by interpreters, though such factors may carry less weight than facts about how we have actually applied our terms in the past. The commitments associated with actual applications carry more weight than those associated with dispositions at least partially because we have a desire to be consistent, and since past use can't be changed, it makes more sense

---

<sup>7</sup> Or at least what his counterpart would mean if they were instantaneously switched.

to alter our dispositions if we want to make the total package coherent. Further, with proper names and many other terms in the language, their usefulness depends to a large extent on being applied now to what they were applied to in the past, so such past applications will often have a certain priority over current dispositions. Consequently, if one interpretation of our words is well supported by our past applications, and a rival interpretation is supported only by our dispositions, we will be more likely to favor the interpretation that is supported by actual use.

### **Accounting for Twin:**

Such an inclusive (if inegalitarian) model of what contributes to interpretation accounts fairly well for our intuitions about the Twin cases. The chemically ignorant Earthling has a history of applying “water” to H<sub>2</sub>O, and a disposition to apply the term to both H<sub>2</sub>O and XYZ. The disposition supports the interpretation of “water” as a more general kind picking out both H<sub>2</sub>O and XYZ, and while this more general interpretation is *compatible* with his actual applications of the term, it is incompatible with a fairly well entrenched belief to the effect that “water” picks out a natural kind. When faced with a conflict between this general belief about the nature of ‘water’ and our dispositions to use the term, we are inclined to treat aspects of the disposition as mistaken. Still, in spite of my being disposed to cast a wider net, my *actual* applications of “water” have just been to H<sub>2</sub>O. Consequently, an interpretation that treated “water” as picking out H<sub>2</sub>O rather than H<sub>2</sub>O or XYZ would not have to treat any of my beliefs as false or applications as misapplications. In much the same way, the ‘charitable’ interpretation of my twin would treat him as meaning XYZ by “water”. On the other hand, while the general belief that water picks out a kind may be more entrenched than some of our dispositions, it need not be more entrenched with the commitments that come from actual applications. If, say, our environment had contained a mix of H<sub>2</sub>O and XYZ, and our past use was connected to both substances, those applications would typically be more entrenched than the general belief about “water” being a natural kind, and the more ‘phenomenal’ interpretation of “water” would seem more plausible. The comparative importance of our actual

applications accounts for why my twin and I mean different things in spite of the fact that our dispositions are the same.

However, in spite of their lesser comparative weight, dispositions still must make a contribution to what we mean, since actual use typically underdetermines what our terms could pick out. For instance, consider the following variant of the Twin-Earth case. On Earth, it turns out that everything that is magenta has a certain texture causing it to reflect light in the relevant wavelength. That is to say, on Earth, “being magenta” and “having surface texture X” are coextensive. Now some people from Earth go to Twin Earth, where, due to some particularities of the elements found there, all of the magenta objects have surface texture Y. In such a case, there is comparatively little pressure to say that the people from Earth were mistaken in calling things “magenta” on Twin Earth, or that the term “Magenta” refers to different things in the two languages. This is because we consider the phenomenal properties of color to be more important than their ‘textural’ ones. In the absence of a ‘defeating’ belief, the term is assigned the wider of the two possible extensions because that is how we are disposed to apply it. Past usage alone wouldn’t settle this, dispositions are needed to supplement it.

Because dispositions ‘complete’ interpretations in this way, two people whose terms are causally connected to precisely the same objects may mean something different by their terms in virtue of being disposed to extend their use in different ways.

### **Accounting for Swamp**

Davidson seems to assimilate the constraints on interpreting the Swampman upon his creation with those constraints which would bind us if, say, Davidson was instantaneously switched with his Twin Earth counterpart. In this case, it would be a mistake to treat Twin-Davidson as meaning Memphis and water by “Memphis” and “water” (even though he would be disposed to use those terms when he saw Memphis and water) since it is clear that he actually means Twemphis and Twater when he uses those words (at least initially). The conclusion Davidson draws from this is that the correct interpretation of a speaker should be governed exclusively by his past linguistic

behavior and though his expected future behavior may affect what he can be expected to mean in the future, it does not affect what he means in the present. However, one's language faculty having no causal history need not be the same thing as its having a causal history with nothing.<sup>8</sup> Our strong inclination to attribute contentful thoughts and meaningful utterances to Swampman suggests a different interpretation of our reaction to Davidson's Twin-Earth counterpart. Namely, expected future language use can be used in interpretation but is defeasible, as in the case of Twin-Davidson, by facts about past language acquisition. In the case of the Swampman there is no past causal history, so the expected future behavior is unconstrained by any such past usage.<sup>9</sup> Swampman is disposed to apply "water" to the water in his environment, and there are no other commitments on his part that would defeat the commitment that this disposition takes on. Water is what he *will* apply the term to, and there is no reason not to treat his term as meaning just that.

After all, Davidson himself admits that, as time passes, and it interacts with items in the world, Swampman will come to have thoughts with the content it originally merely *seemed* to have.<sup>10</sup> Davidson claims to motivate his views about **Swamp** not just by reflecting on **Twin**, but by looking at how our interpretive practice relies on causal connections between words and things. However, if we look at what our practice *actually* does, we can see that it takes the associations that are observed and reads them back into the past. A radical interpreter will investigate what my *current* utterance means by seeing how I use the words in it *in the future*. It is a constraint on interpretation that we prefer interpretations that attribute constancy of meaning, since interpreting what someone means at a time would be impossible if we couldn't treat what they say after that time as being relevant.

---

<sup>8</sup> It is not as if the Swampman actually *learned* his language in some kind of void.

<sup>9</sup> Of course, one would still have to admit that the Swampman couldn't *recognize* things.

<sup>10</sup> "I should emphasize that I am not suggesting that an object accidentally or artificially created could not think; The Swampman simply needs time in which to acquire a causal history that would make sense of the claim that he is speaking of, remembering, identifying, or thinking of items in this world." (Davidson, D. (1987). "Knowing ones own mind". Proceedings and Addresses of the American Philosophical Association., p. 456). Nevertheless, some 'historical' externalists are willing to deny even this. (See, for instance, Millikan, R. G. (1996). "'On Swampkinds'." Mind and Language 11(1): 13-117., Neander, K. (1996). "'Swampman Meets Swampcow'." Mind and Language 11(1): 111-129.)

If we take future behavior to be relevant to the interpretation of present use, we can say that the Swampman *now* means *water* by “water” because, in the future, Swampman’s “water” use will be causally connected with water. The types of content Davidson must say that the Swampman cannot have until he has interacted with his environment for a few months we should actually be entitled to ascribe to him immediately.

If one insists that an agent’s past causal history is said to be *intrinsically* constitutive of content, then the absence of any such history can lead one to claim that there can be no contentfull utterances in such situations. It is only by losing sight of the fact that an agent’s past causal history gets its importance from the role it plays in our interpretive practice, and not, as it were, from any intrinsic ‘meaningfulness’, that we can come to think that nothing can be called upon to play a similar role in its absence.

### **Swamp 2:**

Still, while I’ve given an account of **Swamp** there remain swampmen who are no so easy to interpret. What, for instance, should we say about ‘Swampman2’, a duplicate of Davidson that materialized not on Earth in 1987, but 20,000 years ago upon a completely uninhabited planet (and so never could have any causal contact with Berkeley, automobiles, toasters, houses and everything else it might seem to think about).

Of course, there is less temptation to think that Swampman2 means just what Davidson does by terms like “Berkeley” or “water”, but (1) it doesn’t follow from this that he means *nothing* by such terms, and (2) there may be terms of other sorts for which he may still mean just what Davidson does.

For instance, consider the following modification of **Twin**. Not only does water have a different chemical structure on Twin Earth, but everything does. Appliances are made of different materials work according to different principles, colored objects appear as they do because of textural properties that they not found on earth, & etc. Now we might want to say that, if we were to be transported to such a planet, we would misapply proper names such as “Davidson” or

“Memphis” and kind terms such as “water” or “gold”, but it seems less intuitive to say that we would be equally mistaken in our application of terms such as, say, “house”, “toaster” or “magenta.” We would be mistaken if, for instance, the things that looked like houses were only used to keep farm animals, or what looked like a toaster was only used for executing mice, but as long as the function remained the same, the difference in microstructure would not be relevant.<sup>11</sup>

Given that this is the case, we might equally want to say that a Swampman on an uninhabited planet could have thoughts about toasters and houses, even if he couldn’t think about Memphis or Hillary Putnam.<sup>12</sup> Furthermore, even with proper names, it is not clear that Swampmans wouldn’t be in a position vis a vis “Berkeley” “Putnam” that our ancestors were vis a vis “Dionysus” or “Prometheus”. While the type of thought associated with empty names is less ‘robust’ than the kind associated with genuine referring terms, it remains a type of thought and there seems little reason to think that the Swampman cannot have it.

In conclusion, while **Twin** and **Swamp** highlight different aspects of our interpretive practice that would be incompatible if either of these aspects were mistaken for a single factor that was *essential* to meaning, there is no reason why an interpretive account can’t be ecumenical enough about what contributes to meaning to accommodate both.

2917Words

---

<sup>11</sup> For a discussion of the differences between such kinds and how our actual environment can determine what sort of kind is picked out by a given term, see my \_\_\_\_\_.

<sup>12</sup> We might, ultimately want to say the same thing about the Swampman on Earth. The Swampman is not only disposed to apply the term “Putnam” to Putnam, but he would also understand his own use of “Putnam” as being causally derived from Putnam, and, in the case of proper names, this is a belief we are not usually willing to give up. If we find that our own use does not derive from the person we are disposed to apply the name to, then we consider that application (or at least the first few) to be mistaken. On the other hand, it is far less clear that this is the case with terms like “house” or “toaster”.

### Works Cited

- Davidson, D. (1987). "Knowing ones own mind". Proceedings and Addresses of the American Philosophical Association.
- Fodor, J. (1994). The Elm and the Expert. Cambridge, MIT.
- Fodor, J. and E. Lepore (1992). Holism: a shopper's guide. Cambridge, Blackwell.
- Kripke, S. (1972). Naming and Necessity. Cambridge, Harvard University Press.
- Millikan, R. G. (1996). "'On Swampkinds'." Mind and Language **11**(1): 13-117.
- Neander, K. (1996). "'Swampman Meets Swampcow'." Mind and Language **11**(1): 111-129.
- Papineau, D. (1996). "'Doubtful Intuitions'." Mind and Language **11**(1): 130-132.
- Putnam, H. (1975). "The Meaning of 'Meaning.'" Mind Language and Reality. H. Putnam. New York, Cambridge University Press: 215-271.