

Reflection on the Twin Earth thought experiment by Hilary Putnam

From Helen De Cruz (ed.) *Philosophy Illustrated: 42 Thought Experiments to Broaden Your Mind*, OUP 2021.

Antti Kauppinen, University of Helsinki

Unless you're Finnish, you probably don't know what the word "vesi" means. But what exactly would it take for you to learn its meaning? One natural thought is that you come to know it when you acquire the ability to tell whether something is "vesi" on the basis of some of its features. I can teach you that: "vesi" is the clear and potable liquid that, among other things, fills the lakes and oceans on our planet. Now, in one sense, you know what it means.

Sometimes, though, we say that the meaning of a word is what we talk about when we use it.

In the case of "vesi", that's water, or H_2O , as science tells us. Putting these together, we get an appealing picture, according to which there are two aspects of meaning. First, there's something like a description in our minds (like "clear, potable liquid") and second, there's whatever it is that fits the description (in this case, H_2O). Frege called them "sense" and "reference" [see Chapter 23, *The morning star is the evening star*].

This is a neat story, but Hilary Putnam thinks it's all wrong. In his most famous thought experiment, he asks us to imagine visiting a distant planet in our galaxy that happens to be almost exactly like the Earth. Indeed, the only difference is that on Twin Earth, the watery stuff in the lakes is *not* the same as on Earth, but a substance Putnam labels XYZ, which happens to share all the surface properties of H_2O . If we take a swim on Twin Earth and say "Boy, this water sure feels nice after the quantum jump!", we're mistaken, because the stuff isn't water, although it seems just like it.

This is a big deal, because it follows that what we have in our minds doesn't determine what we're talking about – "meanings just ain't in the head," as Putnam's famous slogan goes. But where are they, then? Why is that when I use the word "water", I refer to H_2O and not to XYZ, even if I can't tell the difference between the two substances? Here Putnam appeals first to what he memorably calls the "linguistic division of labor": at least when it comes to *natural kind terms* like "water" or "elm", I tend to defer to experts, whom I treat as authorities on the actual nature of the stuff I'm talking about. Putnam himself confesses he doesn't know the difference between a beech and an elm. Nevertheless, the two words have a different meaning in his mouth, because when he says "beech", he means to talk about the tree that botanists recognize as beech, and likewise for "elm". So meanings are *social*.

The second crucial thing for Putnam is that reference depends on the environment in which we acquire the term. When we here on Earth talk about "water", we mean to talk about the very substance that we can point to and that we causally interact with when we drink it or bathe in it. On Twin Earth, the clear liquid in the lakes is not the *same stuff*, and that's why it's not water, though it has the same appearance. Meanings are in this sense also *environmental*, at least for natural kind terms. That is, they depend on the nature of the things that actually causally influence our thought and talk.

This has many interesting consequences. Consider metaphysical questions about necessity and possibility. It used to be common to think that all necessarily true propositions were *conceptual* or *analytic* truths: you couldn't possibly come across a false belief that amounts to knowledge, because it is part of the very concept of knowledge that only truths can be known. This is something we can learn *a priori*, without empirical investigation, just by reflecting on

our concepts – and indeed *only* that way. In contrast, our concept of water is the same one our ancestors used before modern chemistry was born, so no matter how hard we reflect, we can't learn that water is H₂O. Instead, it's empirical science that reveals that the watery stuff around here is H₂O. But suppose that Putnam is right and our term "water" refers *in any possible situation* to the watery stuff around here. In that case, "water is H₂O" doesn't just happen to be true – rather, water is *necessarily* H₂O. This is philosophically groundbreaking, because it seems we've learned something about the *nature* or *essence* of water by conducting an *a posteriori* empirical investigation.

OK, in the case of water, this may not sound that exciting, but what about "knowledge" or "rightness"? Some have thought that they, too, are natural kind terms, so that we can discover their nature via scientific investigation into the things that as a matter of fact causally regulate their use.

But hold on. Suppose that the term "morally right" plays the same role in the social life of Twin Earth as it does in ours – for example, they approve of what they call "right" actions, feel guilt for the "not right" ones, and teach children to do what's "right". However, they don't call "right" quite the same things most of us do. They say it is "right" to sacrifice the innocent when it makes many people happy and "not right" to give your own children preference over strangers. It turns out they call "right" only actions that result in as much total welfare as possible, even if it involves using someone as a mere means or rules out personal relationships. Should we conclude that just like they talk about XYZ when they use "water", they talk about maximizing welfare when they talk about "right"? If they did, there would be no straightforward disagreement between my Twin Earth counterpart and me when he says that "It's right to sacrifice the elderly if it helps the economy enough" and I say "No,

that's not right!". But as Mark Timmons and Terence Horgan point out, that's implausible. Rather than talking past each other, we genuinely disagree. And that suggests that while Putnam may be right about how natural kind terms like "water" work, we shouldn't rush to conclusions about other kinds of expression, like moral terms.

Discussion Questions

1. If Putnam is right, it is not *metaphysically possible* for water to be anything other than H₂O, if it indeed is H₂O. Nevertheless, it is easy enough to *imagine* that water had some other chemical composition. As philosophers like Frank Jackson and David Chalmers (2001) put it, it is *epistemically possible* that water is something else – it is not ruled out by what we can know just by reflecting on our concept. But consider again morality. Can we imagine, even in a story, that smothering a baby because its crying interrupts one's Zoom call is morally right? (See Gendler 2000.) If not, what does it tell us about the difference between natural kinds and morality?
2. Putnam says at one point that natural kind terms turn out to be covertly *indexical*, like the words "I" and "here". Why do you think that is?
3. When Twin Antti and I encounter watery stuff, we will engage in the very same sort of reasoning and action, and will both jump in the lake for a refreshing swim, for example. Some philosophers, like Jerry Fodor (1987), argue that this shows that what we believe and want must after all be the same thing – how else could we end up *doing* just the same things? – in spite of our having acquired our terms in different environments. The claim is that our thoughts must have at least some content that does *not* depend on the environment. But does this follow?

Recommended reading:

Chalmers, D. and F. Jackson, 2001, “Conceptual Analysis and Reductive Explanation”, *Philosophical Review*, 110: 315–61.

Fodor, Jerry A. 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. MIT Press.

Gendler, Tamar Szabó, 2000, “The Puzzle of Imaginative Resistance”, *The Journal of Philosophy*, 97(2): 55–81.

Kripke, S. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Horgan, Terence and Timmons, Mark 1991. “New Wave Moral Realism Meets Moral Twin Earth.” *Journal of Philosophical Research* 16 (1991): 447-65.

Glossary items

Natural kind term: A word that purports to pick out a set of things that “cuts the nature at the joints”, because the members are grouped together independently of human interests or reactions. Thus, “water”, “electron”, and “alligator” are natural kind terms, while “boat” or “cool” are not.