

A Philosophically Inexpensive Introduction to Twin-Earth

Bryan Frances

I say that it's philosophically inexpensive because I think it is more convincing than any other Twin-Earth thought experiment in that it sidesteps many of the standard objections to the usual thought experiments. I also briefly discuss narrow contents and give an analysis of Putnam's original argument.

THE THOUGHT EXPERIMENT

When I was a child I sometimes played a joke on my older sister Leslie. My best friend Greg and I would say 'yes' when we meant *no* and we'd say 'no' when we meant yes. Obviously, this caused much confusion in Leslie until she caught on.

It's easy to see that the English language could have been different in that 'no' had always meant yes and 'yes' had always meant *no*. There's nothing magical in the symbols 'no' that they must mean something negative! If that had happened, if the words had always had meanings opposite to what they actually are, then when someone responded to the question 'Is there milk in the refrigerator?' with 'Yes' they would really mean *no*. There's nothing controversial or odd about that.

We just imagined a possible world in which the meanings of 'yes' and 'no' were switched from their real, actual meanings. Now I want you to imagine another possible world, one that switches the meanings of two other words.

As things *actually* stand, in our (real) world, the terms 'walleye' and 'sauger' (pronounced SAW-GER) each stands for a common North American game fish, just like trout, salmon, cod, etc. You can trust me on this fact, as my father lives in Minnesota, fishes quite a bit, has caught lots of both kinds of fish, which are of course not the same, and insists on telling me about his fishing adventures. The two fish species are very similar to the eye, just like gold and "fool's gold" or elm trees and beech trees. Only an expert can tell the two kinds of fish apart (usually by the pattern of scales, if I remember right). Now imagine an alternative possible situation (or "world") in which 'walleye' and 'sauger' are switched (just like how we imagined a world in which 'yes' and 'no' were switched). So in this imaginary world 'walleye' is used to pick out sauger and 'sauger' is used to pick out walleye. So when someone in that world says 'Walleye are big' she has really said that 'Sauger are big'. This is just like the 'yes'-'no' story.

Now suppose that in this imaginary world my father Ron had read an article in an authoritative fishing magazine that contained the sentence 'Walleye get bigger than sauger'. Clearly, what that sentence means in that world is that sauger get bigger than walleye (because the meanings of 'walleye' and 'sauger' are switched in that imaginary situation). So when the author wrote 'Walleye get bigger than sauger', she was expressing the idea that sauger get bigger than walleye. This is a false idea, but it's an idea all the same. Suppose also that Ron believed what the article said. So in that imaginary situation he believes what the author said, namely that sauger get bigger than walleye—although he expresses his belief the same way as the author does in that world, with the sentence 'Walleye get bigger than sauger'.

Now pretend that in this fantasy world Ron tells me his opinion. He says to me 'Walleye get bigger than sauger', thereby echoing the fishing article author. As before, what he means with that sentence is that sauger get bigger than walleye. That's what he and the author believe—that sauger get bigger than walleye.

Let's say that in that world I am agreeable about such matters, and so I take my father at his word. That is, I come to accept his belief. Like him, I believe that sauger get bigger than walleye and I express this belief just like everyone else does in that world, with the sentence 'Walleye get bigger than sauger'.

Now let us leave the realm of imagination and return to the real world. A few years ago when I lived in Minnesota Ron actually told me that walleye get bigger than sauger. And since the only language he knows is English, he used the sentence that actually expresses that idea—the sentence 'Walleye get bigger than sauger'. So as things actually stand today, I firmly believe that walleye get bigger than sauger. In sum:

In the Imaginary World:

'Sauger' picks out walleye

'Walleye' picks out sauger

'Walleye get bigger than sauger' means that
sauger get bigger than walleye.

Ron (my father) believes that sauger get bigger than walleye.

Ron expresses that belief with 'Walleye get bigger than sauger'.

Ron says to Bryan 'Walleye get bigger than sauger'.

Bryan believes that *sauger get bigger than walleye*.

In the Actual World:

'Sauger' picks out sauger

'Walleye' picks out walleye

'Walleye get bigger than sauger' means that
walleye get bigger than sauger.



Ron believes that walleye get bigger than sauger.
Ron expresses that belief with 'Walleye get bigger than sauger'.
Ron says to Bryan 'Walleye get bigger than sauger'
Bryan believes that *walleye get bigger than sauger*.

So in the two worlds I believe different things. That's because it's as plain as day that the belief that sauger get bigger than walleye—my imaginary belief—is distinct from my belief that walleye get bigger than sauger—my actual belief. One belief is just the reverse of the other. One says that A is bigger than B and the other says that B is bigger than A.

Now comes the clincher: ***there's nothing preventing me from being physically identical in the two worlds***. E.g., I need never see, smell, or touch either fish in order to come to have beliefs about them, so my experiences with the fish themselves can be identical (since there are none!). In addition, my fishing conversations with my father can be exactly the same physically:

Bryan: "Dad, which fish gets bigger?"
Ron: "Walleye get bigger than sauger!"
Bryan: "Okay. Whatever you say. Walleye get bigger than sauger."

So although I hear the same *words* from my father in the real and fantasy worlds, I acquire different *beliefs*—his beliefs. In fact, all my encounters with 'walleye' and 'sauger' can be identical in the two worlds.

The thought experiment is intended to prove two theses:

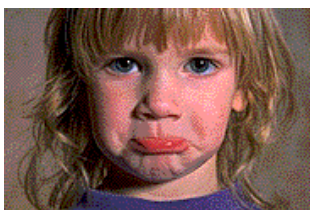
Linguistic Anti-Individualism:

Bryan is physically identical in the two worlds but the *meanings* of his words are different. So *what we mean by our words isn't fixed by what's going on in our brains or body*.

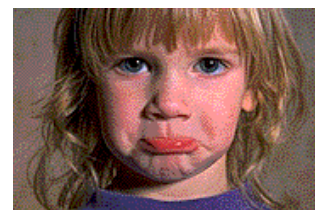
Mental Anti-Individualism:

Bryan is physically identical in the two worlds but his *beliefs* are different. So *what we think or believe isn't fixed by what's going on in our brains or body*.

When philosophers talk about anti-individualism they mean the mental one.



Thus: ***physical duplicates***—people who have *exactly* the same physical qualities, brain processes, sensory experiences, utterances, etc.—can have ***different*** thoughts (e.g., me in



the actual world and me in the imaginary world). Throughout their lives they experience or have *identical* visual fields; they utter *identical* words with *identical* pronunciations; they see the *very same* objects at the *exact* same times from the *very same* perspectives, etc. Surely everything in their respective and physically identical local environments seems precisely the same to them from the inside, from their own cognitive perspective. But how could this be compatible with them thinking different things? A very surprising result.

NARROW CONTENT

So perhaps I have different thoughts in the two worlds. More precisely: although in the actual world I believe that walleye get bigger than sauger, in the counterfactual world I believe that sauger get bigger than walleye. In the actual world my belief is true whereas in the counterfactual world my belief is false—even though the things the beliefs are about (the two kinds of fish) are identical in the two worlds.

Does this mean that I have different thoughts in the two worlds? Well, they surely differ in their *truth conditions*. Here's what that means. My imaginary belief is true just in case sauger get bigger than walleye; that's the condition in the world that will make that belief true. But my actual belief is true just in case walleye get bigger than sauger. So those two beliefs are true under different conditions. That's what we mean by saying that they differ in their truth conditions.

So if the Anti-Individualist is right then the thoughts definitely differ in one important way. So any theory that tries to understand a thought's truth conditions internally is doomed to fail. Fair enough. Still, one might think that the two beliefs don't differ in any psychologically important way. The properties of my beliefs that are important when it comes to understanding me do not differ across worlds.

Here's an analogy that motivates this position. You probably do not believe, like I do, that baseball was invented in 1863. But pretend that you do. Let's say that you're right; baseball was invented that year. Your evidence is that you read it in a couple baseball history books and have heard it orally from experts as well. So you *know* that baseball was invented in 1863. Now consider an alternative possible world in which your language is exactly the same (no word switching or other funny stuff) but baseball was invented in 1864, not 1863. Everything else about baseball is the same; it just started one year later by the very same people and in the same way. For some reason, the experts have made an error and have logged the invention of baseball at 1863. So the books you read in this possible world have the exact same words—with the exact same meanings—in both worlds. The only difference is that in the real world the books are right whereas in the alternative world they're wrong.

So in both worlds you believe (justifiably so) that baseball was invented in 1863. The actual belief amounts to knowledge whereas the counterfactual belief does not. But if you're interested in understanding me, you won't pay any attention to this difference! The fact that in just one world I *know* that baseball was invented in 1863 is not at all to the point if you're interested in understanding my psychology and behaviour. The knowledge difference has nothing to do with me; it concerns only the

world outside of me. The knowledge difference is, we might say, *explanatorily irrelevant* when it comes to understanding me.

So we might say that in light of anti-individualism the truth conditions of our thoughts are explanatorily irrelevant as well. *Just as in the case of knowledge*, the difference in my beliefs has nothing to do with me; it concerns only the world outside of me.

So what about my beliefs is explanatorily relevant to understanding me and my behaviour? It must be some psychological property other than truth conditions; call it *narrow content*. The term 'narrow' is used to reflect the idea that the property in question is internal to me, so I have it in every world in which I'm physically duplicated. The term 'content' is used because whatever it is, it is very much like truth-conditional content. The task: produce a theory of narrow content and improve the argument given above for its existence.

WORDS, CONCEPTS, EXTENSIONS, PROPERTIES, REFERENTS

By keeping the following distinctions in mind you can avoid making lots of annoying mistakes when thinking about this Twin-Earth stuff.

We have our word 'dog', which is a linguistic string of three symbols. If I write on the wall 'Tom is a dog, a dog larger than our cat Fred', I have used the word 'dog' twice. We say that the single word 'dog' has two occurrences in that sentence on the wall. The occurrences are called tokens. The word itself, which has two occurrences there and many elsewhere, is called the word type. Just to make things complex, we should distinguish between word tokens and uses of tokens. If I write 'yes' on a card and walk around with it, flashing it to people I want to say 'yes' to, then I have one token—the ink pattern on the card—that has many uses, where the uses are the individual flashes of the card.

Then there is our concept of a dog. This is a mental thing. It is part of thoughts, the thoughts about dogs. Much of the philosophy of mind, including this course, is devoted to figuring out what concepts are. Then there is the concept's extension, which is the set of dogs: the set of all things that "fall under" the concept.

In addition to those is the species, the dog species, which might be thought of as the referent of 'dog' and is somehow made up of all the individual dogs. For our purposes we can think of the species as a property.

Now we move on to Putnam's original argument. I assume you have already read it.

PUTNAM'S MASTER ARGUMENT

Here is the initial part of Putnam's famous argument:

- A. The linguistic string 'water is wet' is used on both Earth and Twin-Earth. In order to distinguish the two cases, let's pretend that 'water' as used by Oscar₁ on Earth with H₂O, has an invisible subscript, so it may be written (with anti-invisible ink) 'water_E' ('E' for 'earth'). Similarly, pretend that 'water' as used by Oscar₂ on Twin-Earth with XYZ, has an invisible subscript so it may be written 'water_{TE}' ('TE' for 'twin-earth').
- B. The extension of 'water_E' differs from the extension of 'water_{TE}'.
- C. Now for the sake of argument adopt **Assumption (II)**: Contents (i.e., intensions) determine extensions; so if the extensions of words *A* and *B* differ, then the contents of *A* and *B* differ.
- D. By (B) & (C) the content of 'water_E is wet', call it C_E, differs from C_{TE}, the content of 'water_{TE} is wet'.
- E. Oscar₁ (on Earth) believes C_E. Similarly, Oscar₂ (on Twin-Earth) believes C_{TE}.
- F. **Uniqueness Assumption**: Oscar₁ has just one content here; and the same holds for Oscar₂. This assumption is in Putnam's 'assuming that *A* has just *one* meaning for Oscar in each world' near the bottom of p. 221.
- G. Thus, by (D)-(F) Oscar₁ and Oscar₂ differ in their contents: Oscar₁ has just C_E, Oscar₂ has just C_{TE}, and C_E ≠ C_{TE}.
- H. Oscar₁ and Oscar₂ are molecularly identical.
- I. **Narrow Assumption**: If Oscar₁ and Oscar₂ are molecularly identical, then they are narrowly psychologically identical—identical in all narrow psychological states. That is, all narrow states are supervenient.
- J. Thus, by (H) & (I) Oscar₁ and Oscar₂ are narrowly psychologically identical.
- K. Thus, by (G) & (J) C_E and C_{TE} are not narrow states. This contradicts Methodological Solipsism or what Putnam calls **Assumption (I)**, the thesis that contents are narrow states.

Now for the rest of the argument:

- L. The argument (A)-(K) is valid.
- M. The premises are (A), (B), (C), (E), (F), (H), and (I). So if they're true, the conclusion (K) is true too.
- N. (A), (E), and (H) are co-stipulations supposed to be unproblematic. He's probably right about that: we can stipulate that in this possible world all three are true at the same time.

- O. (B) is established by Kripke's arguments regarding natural kinds. So it's true. Better yet: instead of using H₂O and XYZ, we can use aluminium/molybdenum (p. 225) or elm/beech (pp. 226-7). Or we could use gold and iron pyrites, as on p. 124 of *Naming and Necessity*. Or walleye/sauger.
- P. Thus, since we've decided that (A), (B), (E), and (H) are all okay, the only premises left are (C), (F), and (I).
- Q. The Uniqueness Assumption, (F), is supposed to be unproblematic. It isn't, at least not anymore. Dual content theorists, who believe that in addition to wide, truth-conditional content there is narrow content, would say that each Oscar has a belief type with two contents.
- R. The Narrow Assumption, (I), is supposed to be unproblematic. But it isn't, as you'll see below.
- S. So, since Putnam has concluded that every premise except (C) is true, he concludes that if (C) is true, then (K) is true. That is, if assumption II is true, then assumption I is false. That is, either assumption I or assumption II is false; you can't have both. He later expressed a preference for (C) over (K), i.e., II over I.

QUESTIONS AND ANSWERS

In the rest of this essay I try to explain some of the basics of the material that I expect you might find puzzling. Here are questions addressed below.

1. ***Please explain extensions and intensions. In particular, what does assumption II, 'intensions determine extensions', mean and why does anyone think it's true? And what about the reverse claim, that extensions determine intensions?***
2. ***Please explain narrow, wide, and supervenient. Is being narrow the same as being supervenient? If not, are some narrow states not supervenient? Or are some supervenient states not narrow?***
3. ***What does this stuff about narrow, wide, and supervenient have to do with Putnam's Narrow Assumption: If two people are molecularly (physically) identical, then they have the same narrow belief states?***
4. ***What's the difference between externalism and anti-individualism?***
5. ***Please explain how on earth meaning could not be in the head. Surely meaning and thought content is mental, something in my mind, something having to do with me, not my physical environment.***
6. ***Burge's externalism is really complicated. What does it mean, really? And why does he think it's true?***

Now on to the answers!

1. Please explain extensions and intensions. In particular, what does ‘intensions determine extensions’ mean and why does anyone think it’s true? And what about the reverse claim, that extensions determine intensions?

Roughly put, the extension of a word or thought or concept is the thing or set of things it applies to. So the extension of ‘gold’ is all the gold stuff. The intension is the meaning. The big problem here is figuring out what meaning or intension really is.

Putnam says that intensions determine extensions. That’s his assumption II. Here’s what that means:

Suppose you have words A and B. Suppose further that the extensions of A and B are not the same. Then the intensions aren’t the same either (that’s the assumption). So if you know that the intensions of two words are identical, then their extensions must be identical as well; that is, intensions fix (or “determine”) extensions.

So if the extensions of ‘gold’ and ‘iron pyrites’ differ, then the intensions of those two terms differ. Since their extensions really do differ, their intensions really differ. Since the extensions of ‘Current Prime Minister of England’ and ‘Current Queen of England’ differ, their intensions differ.

The same holds for beliefs or thoughts instead of words:

Suppose you have thoughts A and B. Suppose further that the extensions of A and B are not the same. Then the intensions aren’t the same either.

Thus, if your thought that gold melts at 1200°C has a different extension than your belief that iron melts at 1200°C, then those thoughts differ in intension.

So: extensions differ $\rightarrow$ intensions differ (i.e., if extensions differ, then intensions differ). But what about the other way round? Is Putnam also saying that: intensions differ $\rightarrow$ extensions differ?

Just because $X \rightarrow Y$ does not *automatically* or necessarily mean $Y \rightarrow X$ —as you all know quite well from elementary logic (e.g., ‘If you won the World Series, then you won the playoffs’ is true but ‘If you won the playoffs, then you won the World Series’ is false). And there is good reason to think that ‘extensions differ $\rightarrow$ intensions differ’ is false. Surely the meanings—in some sense of ‘meaning’—of ‘creature with a heart’ and ‘creature with a kidney’ differ. That is, they differ in intension. But they don’t differ in

extension: they pick out the same set of animals.¹ Same for 'gold' and 'element with atomic number 79': they differ in intension but not in extension (in a natural way of understanding extension).²

So the principle "intensions differ → extensions differ" looks false for words. Putnam knew that. He also knew that that principle probably fails for thoughts. For instance, the thought that gold is a metal differs in intension from the thought that the element with atomic number 79 is a metal—even though those two thoughts have the same extension. After all, lots of people had the first belief centuries ago but they didn't have the second belief. So the beliefs differ in intension.

"Yes, but how on earth are we supposed to know whether either any of this is right? What are extensions and intensions really? Unless I have some kind of firm grasp of what these ideas are it's hopeless for me to try to evaluate what Putnam, Crane, and everyone else are saying! I can't evaluate what I don't understand."

Right. So let's discuss extensions and intensions some more.

Suppose you have a word W such as 'dog'. Ask yourself what entities in the world that word W is true of; ask yourself this question 'Is that thing a W?' of every thing in the universe.³ The collection of all the things for which the correct answer to the question is 'yes' is the extension of the word W. For instance, 'dog' is true of, well, all the dogs, because when you point to something and ask 'Is that thing a dog?' you get to answer affirmatively for dogs only. 'President of the USA' is true of George Bush Junior, Bill Clinton, George Bush Senior, Ronald Reagan, Jimmy Carter, and all the other people who have held that office.

But this still doesn't settle everything, even for the terms we just looked at. Everyone says that 'gold' and 'element with atomic number 79' have the same extension. But do they? Consider the two tests: 'Is that thing gold?' and 'Is that thing the element with atomic number 79?'. The **first** test seems to pick out just gold nuggets (pieces of gold). The **second** test picks out not gold nuggets but something like the *property* of being gold. A single piece of gold is *not* the element with atomic number 79. For the piece is specific or particular; the element is a general thing, not limited to that piece of gold alone. Perhaps the second test picks out something like a natural *structural pattern*: the structural pattern that's common to everything that has 79 protons in its nuclei. Of course the structural pattern shows up in pieces of gold only; that's the connection between 'gold' and 'the element with atomic number 79'. Perhaps to

¹ Actually, I read somewhere that this isn't exactly right, even in nature. But just pretend that the two sets of creatures are identical or come up with a better example.

² Personally, I find the traditional view of extension—which I'm dutifully articulating here—to be pretty suspect, even if all creatures with hearts are creatures with kidneys and vice versa. For extension is supposed to capture—or be identical with—reference, and it seems to me that whereas 'creature with a heart' refers *in part* to the property of having a heart (or perhaps to hearts themselves), 'creature with a kidney' does not. So reference is very different from extension, in my view.

³ Some words don't have extensions at all; e.g., 'of', 'have', 'lots', 'the', 'the current King of France'.

be more accurate we should say that 'gold' and something like 'composed of the element with atomic number 79' are coextensional (i.e., have the same extension). Or, perhaps better, we can say that the extension of a word is the set of things that the term *ultimately* applies to. For it is reasonable to say that both 'gold' and 'the element with atomic number 79' both *ultimately* apply to just gold nuggets.

But what are intensions? After all, philosophers use 'intension' in a purely philosophical sense, something not found in ordinary life. *It's an invented, technical, artificial term of jargon.* What is it supposed to mean?

This is much harder. ***It's fair to say that this notion is the target of all the debate.*** Since that debate is so deep and difficult, it's not surprising that there is little one can say confidently about intension. Roughly put, the intension of a term is its meaning but not its reference—but the notion of meaning other than reference is as muddled as any notion. Even so, the intuitive idea of intension has already been given above. The belief that gold is a metal differs in some clearly meaningful way from the belief that the element with atomic number 79 is a metal—even though those two thoughts have the same ultimate extension. That difference in meaning is a difference in intension. And the *belief* that Mary's tall, the *doubt* that Mary's tall, the *suspicion* that Mary's tall, the *denial* that Mary's tall, and the *realisation* that Mary's tall all have the same meaning—the same intension. After this amount of explanation everything else is pretty controversial.

You can view the whole individualism/anti-individualism debate as a discussion meant to elaborate on these initial remarks about intension and how it relates to extension and belief. So it's impossible to *start* the discussion out by saying "This is exactly what intensions are. Now let's look at some views about them". Instead, we start with our vague idea of what meaning is—as something intimately related to but over and above extension—and try to make it less vague and muddled by thinking through the Twin-Earth thought experiments. Often the goal in philosophy is to take some quite muddled but absolutely central notion—goodness, meaning, truth, reality, possibility, beauty, knowledge, responsibility, reason—and examine it to see what kind or kinds of less muddled diamonds are to be found in that rough.

2. Please explain narrow, wide, and supervenient. Is being narrow the same as being supervenient? If not, are some narrow states not supervenient? Or are some supervenient states not narrow?

You and your physical twin—a person completely physically identical to you from the skin in your entire life—have all the same brain processes at the same times. Since our brain seems to control our mental life— isn't that what science has shown?—you'd think that your physical twin would have all the same thoughts as you. That intuition is the idea that *brain states fix or determine mental states*. Another way to say that: mental states *supervene on* brain states. Here's a definition: mental states *supervene on* brain states means, by definition, that physical twins have the same mental states.

That's supervenience. (There are actually many kinds of supervenience, but we can ignore the variety.) Narrowness is different. Imagine a universe that to Chris seems just like ours. But suppose that

contrary to everything Chris believes her whole life has been a sham: there is this all-powerful demon who is creating the illusion of a physical world external to her. In actuality her body is in deep space moving exactly as she thinks it is—so it's a normal human body—but there are no other physical objects anywhere. The demon is feeding her body/mind sensory signals to make it look to her as if there is an Earth and people and everything else.

Now if that's the way things really are for Chris, then what mental states can she have? Could she be jealous of anyone? Seems not: there's no one to be jealous of. Could she agree with anyone? Seems not: there's no one to agree with. Those are *wide* (or “extrinsic”) mental states: states you cannot have if there is nothing else in the universe. Narrow (or “intrinsic”) states are the opposite: a person could have them even if she were the only thing that existed in the whole universe. Presumably, the mass of your body is a narrow state, as is your centre of gravity, overall electromagnetic field, and some other physical states.

Could Chris believe that Florence is beautiful? If so, then that belief is narrow because Chris is in an otherwise empty universe. This issue is a bit harder to figure out:

On the one hand, there's no Florence (in this possible world) for her to have any beliefs about. That makes it seem as though such a person couldn't believe that Florence is beautiful. Furthermore, our belief is true just in case Florence—that city—is beautiful. Whether or not Chris's belief is true does not depend on Florence in any way—Florence has never even existed in her universe. And how could she ever acquire any concept of Florence? In her world there is no Florence to learn about; so how on earth could she acquire any thoughts about it? So she doesn't believe that Florence is beautiful.

On the other hand, just because someone believes that object X has characteristic C doesn't mean that X exists. People still believe the Fountain of Youth exists; that doesn't mean that there is such a thing. And besides, Chris could be physically identical to you. And she (he, whatever) could have all of your sensory states as well. So in some sense everything seems the same to you and Chris. So doesn't Chris in some sense believe that Florence is beautiful—just like you do? After all, everything seems the same to her. Her Florence belief is totally confused but isn't the same belief as yours all the same? Can't we say that *her* Florence exists and she thinks it's beautiful? Why not say that you and Chris have the same Florence concept, but whereas yours refers to a real city her does not? So she does believe that Florence is beautiful (just not *our* Florence).

Those two sets of considerations are intuitive but come to conclusions that contradict one another; so something has to give. Since Chris is physically identical to you, *if* she doesn't have the Florence belief *then* there are physical twins only one of which has that belief. That would mean that that belief doesn't supervene. And that would mean that anti-individualism (that is, anti-supervenience) is true. So we've proven (a) if she *does* have the belief then the belief is *narrow*, and (b) if she *doesn't* have the belief then anti-individualism is true.

That's narrowness and supervenience, but what about wideness?

If she has the Florence belief, it's narrow; so can we also say that if she *doesn't* have it, it's wide? That is, if Chris doesn't have the Florence belief does that prove that the belief is wide in addition to being non-supervenient? Maybe not. If Chris doesn't have the belief then we've proven that a *physical duplicate of you*, a non-scientist, in an empty universe cannot have that belief, but we haven't proven that *anyone* in an empty universe cannot have that belief. In order for the Florence belief to be wide it has to be impossible for *anyone* to have it in an empty universe; but all we have proven with anti-supervenience is that it's impossible for a *physical duplicate of you* to have it in an empty universe. For all we know, if a person was the greatest genius ever and lived a zillion years, then even if she were in an empty universe (empty besides her) she could eventually come up the Florence belief. After all, if she's a real genius, smarter than any human who has ever lived, and she lives a billion years, then maybe through some chain of bizarre yet incredibly creative scientific and philosophical reasoning she'd be able to conceive of Florence. It's hard to know how we could know that such a person is impossible—for who knows the bounds of genius?

Thus, even if a belief doesn't supervene (because physical twins need not share it), it might still be narrow: some nearly divine genius could have it in an empty universe. That is: there *might* be beliefs that are narrow but not supervenient. Now we haven't *proven* that there really are such beliefs! All we've shown is this: just because you know that a belief is anti-supervenient doesn't mean you can *automatically* infer that it is wide. That is, it's not a matter of definition or logic alone that an anti-supervenient state has to be wide. Another way to put it: the inferences 'If anti-supervenient then wide' and the logically equivalent 'If narrow then supervenient' could conceivably be wrong.

3. What does this stuff about narrow, wide, and supervenient have to do with Putnam's Narrow Assumption?

Here's the assumption again: If two people are physically identical, then they have the same narrow belief states. Another way to say it: physical duplicates have all the same narrow states. Yet another way: all narrow states are supervenient; i.e., if a state is narrow, then it's supervenient. We've just seen, in the last sentence of the previous paragraph, that this claim—if narrow then supervenient—might be true, but it could be false as well. One would have to argue for it. Putnam doesn't defend it; so there is an undefended claim in Putnam's argument. But Burge defends it indirectly. In effect, Burge argues that virtually all beliefs are not narrow: believing that water is wet implies that either there's water, or cohorts with water beliefs, or the believer is some kind of creative genius ("not indifferent towards or ignorant of the nature of water"). This is his externalism. Notice that the last bit about indifference and ignorance leaves open the possibility discussed above: a genius who comes to have water beliefs even though all alone in an empty universe. But *other* remarks of his in "Other Bodies" express his view that such a genius would fail to obtain water beliefs.

We've questioned the inference 'If narrow then supervenient', but what about 'If supervenient then narrow'? That second inference is solid. Here's the proof. Consider some supervenient state S that you actually have. Now imagine your physical duplicate in an empty universe. Since she's your physical

duplicate and S supervenes on physical makeup, she has S. So she has S in an empty universe. So by definition S is narrow.

To sum up:

supervenient $\longrightarrow$ narrow

wide $\longrightarrow$ non-supervenient

non-supervenient $\nrightarrow$ wide

narrow $\nrightarrow$ supervenient

With the crossed arrows I mean that it's not a logical truth that the thing on the right must be true if the thing on the left is true. That is, the left doesn't entail the right. Even so, perhaps beliefs are narrow and supervenient anyway.

4. What is the difference between externalism and anti-individualism?

Anti-individualism focuses on supervenience: it says that beliefs are anti-supervenient. Externalism is a different claim: it says beliefs are wide. Thus, if externalism is true, then beliefs are wide; and we have just seen above that if a belief is wide, then it's non-supervenient—which would mean that anti-individualism is true. Thus, *logic alone* proves that if externalism is true, then anti-individualism is true. That is, anti-individualism is entailed by externalism. But the reverse doesn't hold: anti-individualism doesn't entail externalism.

5. Please explain how on earth meaning could not be in the head. Surely meaning and thought content is mental, something in my mind, something having to do with me, not my physical environment.

Suppose I present a word to you to see if you understand it. I say 'heedless'. Do you understand it? If you did, then presumably you grasped its meaning right there after you read it. That's something you did with your mind. It has nothing to do with the physical environment. The meaning of a word is just something in one's mind. Where *e/se* could it be?

And of course our minds are our brains. There might be ghosts, gods, and goblins, but we humans are utterly physical. At the very least, our beliefs and thoughts are physical. And if that's right, then the only place to locate meaning is in the head. Where *e/se* could it be, really?

Furthermore, surely Chris has trains of thought just like we do. After all, for all you know you are in Chris's situation! Everything she experiences you experience—all the same visual, auditory, tactile, olfactory, and gustatory sensations are duplicated. So she has concepts and thoughts.

Utterly reasonable. If those arguments don't move you at all, then you don't understand meaning. Even so, they're just initial, fairly pre-reflective, thoughts about meaning. Consider Chris again. When she "hears" the word 'Florence' something goes on her mind. Everything seems the same to her as it seems to you (assuming she's living the same physical life you are). But look: does she really believe, as you do, that *Florence* is beautiful? How on earth could she have got a concept of Florence for goodness sake? Whatever concept she expresses when she utters 'Florence' is something she picked up from the demon, not anyone who's ever been in contact with Florence. Remember Kripke's point about reference: roughly speaking, what you refer to with a word is the object that the word was wedded to in the beginning. Chris's word 'Florence' was certainly not wedded to Florence at any time at all (there being no Florence or anything else in her universe).

Even if externalism and anti-individualism are true, we can still say that each act of understanding in your mind—conceived as a token, not a property—is something wholly contained in your mind (and if we're physicalists we'll say it's in your brain too). The anti-individualist point is just that even though you and Chris have the same act of understanding a word, you are grasping different meanings. Consider another example. You've been jealous before. Suppose you were jealous of Alex. Obviously, Chris felt all the same feelings as you did. But did her thoughts and feeling amount to jealousy? It seems not: there's no one for her to be jealous *of*. We can say that she has some mental state that is a crucial part of the mental property of being jealous. That's fine. But it just ain't jealousy. The anti-individualist is saying the same thing about Chris's thought expressed by her uses of 'Florence is beautiful': perhaps she has something going on her mind that is also going on in your mind when you think that Florence is beautiful. She is thinking *something*, but she isn't thinking your thought. Perhaps she is thinking a part of your thought. We'll see later in the course that there are importantly different ways to construe this "part".

6. *Burge's externalism is really complicated. What does it mean, really? And why does he think it's true?*

Burge's Externalism: If in some possible world someone believes that K is F (for some natural kind K), then in her world either there are instances of the kind, she has cohorts who have K beliefs, or she is not relatively ignorant and indifferent about the nature of K. Here it is applied to gold:

If in some possible world someone believes that gold is yellow, then in her world either there are instances of gold, she has cohorts who have gold beliefs, or she is not relatively ignorant and indifferent about the nature of gold.

Why does he believe this? Answer: because of what he thinks about concept acquisition. We're not born having all our concepts. We acquire most of them as we learn about the world. You weren't born having the concept of uranium or Microsoft or physicalism or carburettors. Instead, you got those concepts by interacting with people who already had acquired them. This isn't to say that no concepts are innate. Maybe some are. Burge's basic idea is that many aren't innate:

In order for someone to believe that gold is yellow, for instance, she must have a concept of gold. And we aren't born with that concept. So if she has gold thoughts then she must have acquired the concept of gold.

That's pretty intuitive, right? Yes, but the complications come in when we think about *how* one might go about acquiring the concept of gold. Now either you acquire it on your own or you got it from someone. In order to do it on your own, you'd either have to be a scientist making a hypothesis, or you'd have to just come across a bunch of gold and think about it yourself. If you got it from someone, then they probably taught you the word 'gold' and told you a little bit about gold.

Here's a guess: one possesses the concept of some kind such as gold if and only if

- (i) one acquired it in something like the theoretical way scientists come up with concepts of natural kinds that for all they know have no instances (e.g., the concept of a black hole),
- (ii) one acquired it naturally and unconsciously, in something like the way each of us may (I don't know) acquire the concepts of *up*, *before*, etc.,
- (iii) one acquired it second-hand via some source with or without the presence of instances (e.g., you may have got your concept of gold from your parents showing you a wedding ring),
- (iv) one acquired it oneself by observing instances of the kind and making a linguistic baptism (e.g., 'Let's call this shade of red 'Christmas stocking red''), or
- (v) one acquired it via some ordinary, non-scientific, but partially theoretical description such as 'the plant that would result from the breeding of these two plants'.

As a matter of fact, way (iii) probably is true for every person alive today. Way (iv) is probably true for whoever came up with the word 'gold' for the first time. We can imagine a possible world in which gold is very rare and some scientist working on filling out the periodic table comes up with a concept of gold as a scientific hypothesis; that's way (i). Or, perhaps there is a planet in our universe in which there is no gold but intelligent beings much like us who end up positing gold in some scientific hypothesis. Presumably one cannot acquire a concept of gold via way (ii). It's also hard for me to see how way (v) might work for gold, but it's easier to see how it could work for plant species.

If either (i) or (v) holds, then one is not indifferent towards or ignorant of gold; if (iv) holds then there are instances of gold; if (iii) holds, then one has cohorts with gold beliefs; (ii) looks impossible for gold. Thus, *if one has gold beliefs*, then one of (i)-(v) holds; and if one of (i)-(v) holds, *then either one is not indifferent towards or ignorant of gold, there are instances of gold, or one has cohorts with gold beliefs*. But those italicised phrases amount to Burge's externalism. So he's got some reasons for holding his complicated form of externalism.

Some questions for discussion:

1. It's plausible to think that Bryan's beliefs are different in the two worlds. Why?
2. It's plausible to think that Bryan can be physically identical in the two worlds. Why?
3. "According to the anti-individualists, in the actual world Bryan believes that walleye get bigger than sauger; in the counterfactual world he believes that sauger get bigger than walleye. But that's *wrong*; he doesn't have *either* belief in either world. That's because Bryan doesn't really know what a walleye fish is; neither does he know what a sauger is. So he doesn't have either belief." Evaluate this "**Neither Belief**" response.
4. "According to the anti-individualists, in the actual world Bryan believes that walleye get bigger than sauger; in the counterfactual world he believes that sauger get bigger than walleye. That's *right*, but it doesn't mean that he has different thoughts in the two worlds. All that's been proven is that we give *different descriptions* of the *same thought*. He has the exact same thought in his mind in the two worlds; we just describe it with different sentences." Evaluate this "**Different Descriptions**" response.
5. "According to the anti-individualists, walleye and sauger are different kinds of fish. But in Bryan's amateur view they aren't; they're the same fish, and that's the relevant fact when it comes to understanding his beliefs. He would give the same description for each fish—something like 'A North American game fish that people like my father fish for in Minnesota'. So he really has the *same belief* in each world, viz. a belief that could be better expressed with 'Some North American game fish that people like Ron fish for in Minnesota is bigger than some other North American game fish that people like Ron fish for in Minnesota'." Evaluate this "**Reinterpret**" response.