

The case for externalism

January 13, 2004

1	Externalism	1
2	Putnam's 'Twin-Earth Argument'	2
3	Evaluating Putnam's argument	3
3.1	What does it mean to say that meaning determines reference?	3
3.1.1	Shifting the context	3
3.1.2	Shifting the circumstance of evaluation	3
3.1.3	The solution: double indexing and relativizing the thesis that meaning determines reference to a context and circumstance	3
3.2	Twin earth, context, and circumstance	5
3.2.1	Option 1: Does 'water' shift reference with respect to different circumstances of evaluation?	5
3.2.2	Option 2: Is 'water' an indexical?	5
3.3	The prospects of divorcing meaning from reference entirely	6
3.3.1	Unavailability of classical explanations of reference	6
3.3.2	The connections between meaning, what is said, and truth-value	6
3.3.3	The connections between meaning, the belief expressed, and the truth of beliefs	6
3.3.4	Skepticism about reference and skepticism about truth	7
3.4	Consequences for 'psychological' or 'internalist' views of meaning	7
3.5	Consequences for 'internalist' views of belief	8
4	Burge's 'Tharthritis Argument'	8
5	Evaluating Burge's argument	9
5.1	Do A and B understand 'arthritis'?	9
5.2	Does sameness of belief content assure sameness of belief truth-value?	9
6	Kripke on names	9

1 Externalism

'Externalism' is the doctrine that certain important properties associated with mind and language are, in the following sense, externalist properties:

A property of x is an *externalist property* \equiv an intrinsic duplicate of x could have failed to have that property.

What is an intrinsic duplicate? This is not an easy question to answer precisely, but it's not hard to get a grip on the idea. Intrinsic duplicates of you have all the properties that you have except those that depend on objects other than you; 'intrinsic duplicate' is often glossed as 'molecule for molecule duplicate.'

It is uncontroversial that there are externalist properties; the property of being to the left of me is an obvious kind of example, since an intrinsic duplicate of whoever is to my left could have failed to have been to my left.

What is controversial is whether there are any *mental* externalist properties. Beginning in the early 1970's, a series of now classic arguments convinced many philosophers that at least some mental properties, such as *beliefs*, are externalist. This is a very controversial result, since intuitively the contents of your mind depend only on you.

We will consider three classic arguments for externalism.

2 Putnam's 'Twin-Earth Argument'

The Twin-Earth thought-experiment. A slightly modified version of the implicit argument can be stated as follows:

- P1. If two words apply to different things (i.e., have a different reference), then they have a different meaning.
 - P2. At the time of the spaceship landing, the word 'water' on Earth and Twin-Earth apply to different things. (At this time, we suppose that scientists on earth have discovered that the substance they call 'water' is H₂O, and the scientists on Twin-Earth have discovered that the substance they call 'water' is XYZ.)
 - P3. The reference of the word 'water' did not change on either Earth or Twin-Earth between 1750 (before the chemical constitution of water was discovered) and the present; in both scenarios, the term applies to the same things in 1750 and at present.
 - P4. In 1750, the word 'water' referred to different things on Earth and Twin-Earth. (P2 & P3)
 - P5. In 1750, the word 'water' *meant* different things on Earth and Twin-Earth. (P1 & P4)
 - P6. In 1750, the speakers of English (on Earth) and Twin-English (on Twin-Earth) were intrinsic duplicates of each other.
-
- C1. Intrinsic duplicates can mean different things by words; hence meaning so-and-so by a word is an externalist property.

Putnam's failed attempt to isolate externalism in language, and save an internalist conception of 'psychological states.' The problem posed by 'bridge principles' connecting thought and language. The irrelevance of the division of linguistic labor to this first example of Putnam's.

When a property of an agent is externalist, we can always ask: what objects or properties other than those intrinsic to the agent does possession of that property depend upon? In the case of this argument, the answer seems to be: facts about the physical constitution of objects in the environment of the agent.

3 Evaluating Putnam's argument

3.1 *What does it mean to say that meaning determines reference?*

P1 is a key premise in the argument; without it, we can only derive the conclusion that *reference* rather than *meaning* is an externalist property. And this is not a very surprising thesis at all.

But you might think that there are at least two kinds of obvious counterexamples to the thesis that meaning determines reference, and so two ways to block Putnam's externalist argument.

3.1.1 *Shifting the context*

Consider the following example where meaning seems not to determine reference: two individuals, *A* and *B*, utter the sentence 'I am Canada's greatest philosopher.' In at least one good sense, 'I' means the same thing in both utterances; and yet in the first utterance, 'I' refers to *A*, whereas in the second utterance it refers to *B*. So you might think that here meaning fails to determine reference.

3.1.2 *Shifting the circumstance of evaluation*

Some expressions refer to different things depending on the state of the world; definite descriptions are good examples of this. Consider, for example, 'the tallest man now alive.' This expression refers to someone different in 2004 than it did in 1904; but, it seems, it *means* the same thing in both cases. So it seems that again we have a case of two utterances of an expression with the same meaning, but different reference.

3.1.3 *The solution: double indexing and relativizing the thesis that meaning determines reference to a context and circumstance*

The solution to these apparent failures of the thesis that meaning determines reference is to concede that the meaning of an expression is only one factor which determines

its reference on a given occasion of use: also relevant are the *context* in which the expression is used, and the *circumstance* with respect to which it is to be evaluated. Context and circumstance are sometimes called *indices*.

To see why we need indices to figure in an account of reference, consider the sentence ‘If my parents had been 9 feet tall, then I would have been the tallest person in North America.’ Suppose we ask: what does ‘I’ refer to here? The answer is that, apart from a context of usage, ‘I’ does not refer to anything at all. We need to supply an index — in this case, the context in which the sentence is used — to determine the reference of ‘I.’ So suppose that I utter this sentence in seminar. Call that context of utterance SEM. Relative to SEM, ‘I’ refers to me, Jeff Speaks.

The example sentence above — ‘If my parents had been 9 feet tall, then I would have been the tallest person in North America’ — also shows that we need a second index, beyond the context in which the sentence is used, in order to determine reference. For suppose, for purposes of argument, that this sentence is true. What it says is that, had something been the case — my parents being a certain height — then something else would have been the case — I would have been the tallest person in North America. Another way of putting roughly the same point is that if my parents had been 9 feet tall, then ‘I am the tallest person in North America’ would (given the way we use this sentence) have been true.

What is required for ‘I am the tallest person in North America’ to be true? It seems that ‘I’ and ‘the tallest person in North America’ must refer to the same object. We already know what the reference of ‘I’ is — because our utterance was made in SEM, the reference of ‘I’ is Jeff Speaks. But what is the reference of ‘the tallest person in North America’? We know that it must also be Jeff Speaks if our example sentence is to be true. But then — and this is the key point — the reference of ‘the tallest person in North America’ is not determined by finding out who the tallest person is in SEM, the context of utterance (since I am not the tallest person in North America). Rather, it seems, the reference is determined by finding out who the tallest person in North America is *in the scenario in which my parents are 9 feet tall*.

The conceptual distinction corresponding to the formal distinction between context and circumstance is the distinction between (i) features of the way an expression is used which determine how it is to be interpreted, and (ii) the scenario we must ‘look at’ to determine the truth of what is being said.

In many cases, (i) and (ii) — context and circumstance — are the same state of the world. In a simple sentence like ‘I am in Montreal’, uttered in seminar, the context and circumstance relevant to the utterance are the same. ‘I’ refers to me because I uttered the sentence in SEM, and the sentence is true because, in SEM, I am in Montreal. What the above example shows is that certain ‘operators’ — such as ‘Possibly ...’ and ‘If it had been the case that ...’ — can cause context and circumstance to come apart.

(For more on double indexing and the semantics of indexicals, the best article to read is David Kaplan’s “Demonstratives.”)

3.2 *Twin earth, context, and circumstance*

How is this relevant to Putnam's argument for externalism? The opponent of externalism about meaning wants to reject P1: the inference from the fact that 'water' on Earth refers to H₂O while 'water' on Twin-Earth refers to XYZ to the conclusion that the word 'water' has different meanings in the two scenarios. The above points give us two ways in which the opponent of externalism might do this:

1. He might claim that 'water' is like 'the tallest man alive' — maybe it picks out different things with respect to different circumstances of evaluation. If this were true, then maybe 'water' could have the same meaning on Earth and Twin-Earth while it picks out different things with respect to the two worlds.
2. He might claim that 'water', like 'I', is an indexical, or context-sensitive expression. If this were true, then 'water' would have the same meaning on Earth and Twin-Earth in the sense in which 'I' has the same meaning when used by two different people, despite the fact that, like 'I', it would have different references in the two cases.

3.2.1 *Option 1: Does 'water' shift reference with respect to different circumstances of evaluation?*

Option 1 is the less plausible of these options. If Option 1 were true, then if someone on Earth were to say, 'Twin-Earth is full of water', she would be speaking truly. But in this scenario it seems that the sentence would be false.

3.2.2 *Option 2: Is 'water' an indexical?*

A lot of recent work has gone into defending and arguing against the thesis that 'water' is an indexical, which, like 'I', picks out different references when used in different contexts. A few arguments against treatments of 'water' as an indexical:

1. If 'water' were an indexical of one normal sort, then its meaning would be fixed by the context of utterance. (This would be so if, for example, its meaning were 'the clear drinkable liquid around here.')
- Then it seems that if someone from Earth visited Twin-Earth on a spaceship, and said, 'There's lots of water around here,' she would speak truly — this follows from the fact that she would be speaking in the same context of utterance as Twin-Earthlings, who *could* speak truly by uttering this sentence. But it seems that what the Earthling says would be false out of her mouth, context of utterance notwithstanding. (Not all indexical treatments of 'water' fall to this objection; consider, for example, one on which 'water' means the same as 'the clear drinkable liquid prominent in the region where I learned the term.')
2. Substitution failures of supposed indexical analyses in propositional attitudes.
3. The claim that competent speakers should be able to recognize context-sensitive expressions as such.

3.3 *The prospects of divorcing meaning from reference entirely*

The above two options provide a way for the opponent of externalism to reject the externalist argument while still keeping the link between meaning and reference. But even if we reject this path, there is a more radical route for the opponent of externalism to take: he might claim that there is no direct link at all between meaning and reference.

Following are some considerations which seem to count against such a view.

3.3.1 *Unavailability of classical explanations of reference*

We need some explanation of how words, like ‘I’ when used by me, come to refer to what they refer to. A natural explanation is that we associate a meaning with these words, and that this meaning determines a reference. If we give up the thesis that meaning determines reference, we also give up this kind of classical explanation of how reference is determined.

3.3.2 *The connections between meaning, what is said, and truth-value*

Part of the difficulty in trying to resolve issues in the philosophies of mind and language is the difficulty in finding common ground concerning the meaning of terms like ‘meaning.’ Here is a plausible principle which tries to isolate the sense of ‘meaning’ relevant to the argument:

Two sentences mean the same thing \equiv the two sentences, uttered in the same context, make the same claim about the world

To this we can add the following principle:

If two sentences (in a context) make the same claim about the world, then one cannot be true while the other is false.

It is plausible that the first principle picks out one important pre-theoretic sense of ‘meaning’, and that the second principle is true. But the opponent of the idea that meaning determines reference must reject one of these principles.

3.3.3 *The connections between meaning, the belief expressed, and the truth of beliefs*

We can also reconstruct Putnam’s argument as an argument for the thesis that beliefs are externalist without making use of the principle that meaning determines reference, but instead making use of a principle connecting the contents of beliefs with their truth values. Consider the following principles:

Sincere assertions of sentences with the same meaning express beliefs with the same content.

If two individuals have beliefs with the same content, then, if all the relevant facts about the world are the same, one cannot be true if the other is false.

To clarify the idea of the content of a belief, we can appeal to a (rough) principle similar to the above:

Two beliefs have the same content \equiv they represent the world as being the same way

Again, one who wants to block the conclusion of Putnam's argument must deny one of these principles.

3.3.4 Skepticism about reference and skepticism about truth

One way of responding to these worries is to be skeptical about the idea of reference: meaning does not determine reference because, strictly speaking, there is no such thing as reference. (Some Chomskyans hold a view like this.) Two points to make about this view: (i) According to standard views of truth, skepticism about reference entails skepticism about truth. (ii) This is not really an objection to the claim P1 that meaning determines reference; rather, it is a rejection of premises P2-P4, which make claims about the reference of 'water.'

3.4 Consequences for 'psychological' or 'internalist' views of meaning

Putnam says, "Cut the pie any way you like, 'meanings' just ain't in the *head!*" But suppose that, following one prominent contemporary research program, you took linguistic facts to be essentially psychological facts, where 'psychological' is used in Putnam's internalist sense. Is there a conflict between this research program and the conclusion of Putnam's argument?

In one sense, yes. If that research program purports to reveal all of the linguistic facts, then if Putnam's argument is a good one, it shows that the research program cannot do what it sets out to do.

In another sense, no. Putnam's argument shows that meaning, in at least one important sense of that word, is externalist. But that does not show that there are not other interesting linguistic properties what are internalist psychological properties.

3.5 Consequences for ‘internalist’ views of belief

Putnam thinks that his argument shows that knowledge of meaning is not a ‘psychological state’ — i.e., not an internalist psychological state. So he takes his argument to have consequences for meaning, but not for belief.

But there is a problem with this kind of halfway externalism: there seem to be principles connecting meaning and belief which, if true, show that an externalist conclusion for one entails the same for the other. E.g.:

If someone sincerely endorses a sentence, he thereby gains a belief with the same content as the sentence.

If someone sincerely asserts a sentence, he thereby expresses a belief with the same content as the sentence.

4 Burge’s ‘Tharthritis Argument’

Burge’s argument for externalism also has the form of a thought-experiment, but has a very different moral and establishes a very different kind of externalism.

His argument may be restated as follows:

- P1. Agents A and B could be intrinsic duplicates, one of which lives in a linguistic community in which ‘arthritis’ means arthritis (a disease of the joints), and the other of which lives in a linguistic community in which ‘arthritis’ means tharthritis (a disease of the joints and thigh).
 - P2. If A says ‘I have arthritis in my thigh’, he expresses a false belief (since in his community, as in ours, ‘arthritis’ refers only to a disease of the joints).
 - P3. If B says ‘I have arthritis in my thigh’, he expresses a *true* belief (since in his community ‘arthritis’ refers only to a disease of the joints *and thigh*).
 - P4. The beliefs of A and B differ in truth-value. (P2 & P3)
 - P5. If two beliefs differ in truth-value, and the non-linguistic facts relevant to evaluation of the two beliefs are the same, then the two beliefs have different contents.
 - P6. By their respective utterances, A and B express *different* beliefs. (P4 & P5)
-
- C. Believing so-and-so is an externalist property, since A and B have different beliefs despite being intrinsic duplicates. (P1 & P6)

We can again ask the question we asked with respect to Putnam’s argument for externalism: what objects or properties other than those intrinsic to A and B does

possession of a certain belief depend upon? In the case of Burge's argument, the answer seems to be, not facts about the physical constitution of objects in the environment of the agent, but rather *social facts about the meanings of words* in the linguistic communities which A and B inhabit.

5 Evaluating Burge's argument

5.1 Do A and B understand 'arthritis'?

'Understands' is a vague term; in certain contexts we would certainly say that neither A nor B *really understand* 'arthritis' in their respective communities. But if they don't understand the word, why should we think that they express beliefs when they use it?

The category of 'incomplete understanding.' There is a distinction between mastering every aspect of the use of a word, or being an expert with respect to the word, and merely satisfying the criteria required to be a competent user of the expression, who can use the word in conversation to acquire and express beliefs. Kripke's examples involving names.

5.2 Does sameness of belief content assure sameness of belief truth-value?

6 Kripke on names

It may well seem that the above results are limited to rather recondite cases; Kripke's examples of the uses of names go some distance toward dispelling that impression. We will discuss some of the Kripke's examples, and consider the possibility of 'internalist' explanations of our ability to refer to and think about the individuals who are the referents of the names we use.