

Externalism v. Internalism

A property is **internal** if whether or not an object has the property depends only on features "internal" to that object.

A property is **external** if whether or not an object has the property depends in part on features "external" to that object.

Examples of internal properties: being round, being made of atoms, containing a chemical substance.
Examples of external properties: being within 50 feet of a barn, being a mosquito bite, being a footprint.

Internalism about X: the view that X is an internal property.

Externalism about X: the view that X is an external property.

E.g. one could be an externalist or internalist about the property of having justification for one's beliefs.

We're going to ask if we should be internalists or externalists about intentional properties. Recall, these are properties like

- being something that is thinking about Paris,
- being something that hopes it will win the lottery,
- being something that is a conventional sign that help is on the way,

and so on. But if we are internalists or externalists about intentional properties we have to be able to answer the question: "internal or external to *what*?"

Remember there are two principal bearers of intentional properties: bits of language, and mental states (like belief). For mental states the following questions seem reasonable (where "sub-dermal" means "beneath the skin"):

Is the property of being in a certain intentional mental state internal or external to the sub-dermal state of the person bearing the mental state?

Are the intentional properties of a certain utterance internal or external to the utterance considered along with the sub-dermal state of the person making the utterance?

The answer to the first question will determine whether or not we are an **internalist** or **externalist about mental content**, and the answer to the second will determine whether or not we are an **internalist** or **externalist about linguistic content**.

Externalism about Linguistic Content: Putnam and Twin-Earth

Imagine a planet called "Twin Earth" which is identical to our world in every way, except that it contains a chemical substance we'll label XYZ wherever there is H₂O on Earth. XYZ is indistinguishable from H₂O in normal (non-experimental) circumstances. Now just as people call the watery stuff around us "water" here on Earth, so too people on Twin Earth call the watery stuff around them "water". In fact, in the present, it is clear that "water" for us means the stuff with chemical substance H₂O and for the people on Twin Earth "water" means the stuff with composition XYZ.

Now imagine what life was like in the early 18th century, before the discovery of the chemical composition of water, both here on Earth and on Twin Earth. Take a character, Oscar here on Earth at that time. He has a "twin" on Twin Earth, who we'll call "Twin Oscar" (we call him that, but of course his friends on Twin Earth just call him "Oscar").

Questions: What is the referent (extension) of the word "water" as used by Oscar? What about Twin Oscar?

Alleged answers: The extension of "water" as used by Oscar is (roughly) collections of H₂O molecules. The extension of "water" as used by Twin Oscar is (roughly) collections of XYZ molecules.

If you doubt this we can give an argument for these conclusions from premises like:

The extension of "water" now on Earth is given by collections H₂O molecules.
The extension of "water" now on Earth is the same as was in the 17th century.

The truth of the first premise should be clear. To see the truth of the second premise imagine (a little anachronistically) that we knew of chemical structure in the 17th c. before we knew the structure of water. One scientist says to another "I bet that water is nothing but H₂O molecules". The bet is accepted. Some experiments are performed and it comes to light that the watery stuff around the scientists is made of H₂O. Wouldn't the first scientist win the bet?

If we accept the answers above it seems that *we must accept externalism about linguistic content*.

This is because:

Twin Oscar and his uses of "water" are "internally" indistinguishable from Oscar and his uses.
Twin Oscar's use of "water" has different intentional properties than Oscar's uses of "water".

Interlude: Kripke

Kripke's causal-historical theory of reference also seems to entail externalism about the linguistic content of names. Consider someone who could have acquired the name "Joe" by one of two causal chains leading back to different people. The difference in the causal origins of the name has no immediate effect on the person who acquires the name. Yet depending on the "external" chain, the referent of the name "Joe" is different.

Externalism about Mental Content: Burge

Tyler Burge gave a similar kind of thought experiment to argue for externalism about mental content.

Jane suspects she has arthritis because of a chronic pain in her thigh. She doesn't (and couldn't) because arthritis is a condition of the joints. Burge claims that when Jane says "I have arthritis in my thigh" she not only says something false, but *expresses a false belief*—a belief about arthritis. (Indeed, she might even concede that.)

Now imagine a counterfactual scenario where Jane and her ailment are the same, but the medical community came to use the term "arthritis" not merely to apply to ailments of the joints, but also of the thighs. Let's call this disease "tharthritis". Burge claims first, that in such a scenario, Jane would have no beliefs about arthritis (any more than Jane in our world had beliefs about tharthritis). Jane has beliefs about *tharthritis*. And when Jane says "I think I have arthritis" she would not only say something true, but *express a true belief*—a belief about tharthritis.

If Burge is right to describe the case this way, *externalism about mental content* must be true.

Burge's example shows a different potential kind of dependence of language and belief on external factors: not on the causal origins of words or their referents, but on the state of the linguistic community in which one is embedded.

Internalist Rejoinders

(A) Are the intuitions correct?

Mightn't it be the case, for example, that Jane actually *does* believe she has tharthritis in both scenarios Burge describes. She reveals false beliefs in the actual world by saying "I have arthritis", namely the belief that "arthritis" refers to a disease like tharthritis. But that won't show externalism about mental content is true.

(B) Aren't there good *arguments in favor of internalism* about (say) mental content?

Don't we have a special privileged "access" to our own thoughts, and what they are about? But we don't have a special access to the external facts allegedly determining the content of our thoughts. So doesn't that mean internalism must be true?