

Twin-Earth Externalism Revisited

SIMON DIERIG
Heidelberg University, Heidelberg, Germany

The aim of this article is twofold: First, it is argued that Tyler Burge's case for externalism in the philosophy of mind, which is based on Hilary Putnam's twin-earth thought experiment, fails. Second, it is shown that a convincing argument for externalism can be nonetheless construed by relying on Putnam's thought experiment.

Keywords: Externalism; twin earth; philosophy of mind; beliefs; Putnam; Burge.

1. Burge's case for externalism

Tyler Burge has argued, using a series of thought experiments, that individualism in the philosophy of mind is misguided (cf. Burge 1979, 1982a, 1982b and 1986). Individualism is the doctrine that the individuation of our mental states depends only on what goes on inside our bodily limits: if one wishes to say under which conditions a mental state *a* is identical to a mental state *b*, one solely has to fall back on information about what happens inside the bodily limits of the persons who have the mental states in question.

The counter-thesis to individualism is anti-individualism or externalism. Proponents of the latter position claim that the individuation of our mental states depends not only on our internal states, but also on our environment. The doctrine of externalism can also be characterized as the negation of the local supervenience thesis. This thesis says that two persons must have the same mental states if their internal bodily states are identical, that is, if they have the same physical, biological, functional and dispositional properties. Externalists typically claim that this thesis is wrong.

In his essay "Individualism and the Mental," Burge used thought experiments in which the conventional meaning of certain linguistic expressions is varied counterfactually to defend the view that an individual's mental states depend upon her *social* environment (cf. Burge

1979). To show that the individuation of mental states also depends on the *physical* environment of an individual, Burge later relied on the so-called “twin-earth thought experiment,” which was developed by Hilary Putnam in his seminal essay “The Meaning of ‘Meaning’” (cf. Putnam 1975: 223–227). For his purposes, Burge modified Putnam’s thought experiment in a number of respects (cf. Burge 1982a: 100–102).

Burge imagines that Oscar assents to the sentence “There is water within one kilometer (of the Empire State Building in New York),” thereby expressing the belief that there is water within one kilometer (of the Empire State Building in New York).¹ A second person, Oscar’s twin, lives on a planet, “twin earth,” which differs from earth only in the two following respects: First, on the planet in question there is instead of H_2O a phenomenologically indistinguishable liquid with a completely different chemical composition, viz., XYZ. Second, scientists on earth know that water consists of H_2O , and scientists on twin earth know that the substance which they call “water” has the chemical composition XYZ. But Oscar and his twin do not know anything about the chemical structure of the liquid they each call “water.” For this reason one can imagine, in Burge’s view, that, inside his bodily limits, in his physical, biological, functional and dispositional properties, Oscar’s twin is indistinguishable from Oscar.²

Burge argues that Oscar’s twin does not believe that there is water within one kilometer. If he assents to the sentence “There is water within one kilometer,” he rather expresses the belief that there is twater within one kilometer, the word “twater” being a neologism which we introduce to render the twin-English word “water” into English. The expression “twater” refers to XYZ, and not to H_2O .

To support the claim that Oscar’s twin does not believe that there is water within one kilometer, Burge uses a *concept*-based argument and a *truth*-based one (cf. Burge 1982a: 109 f.).

Let me start with the concept-based argument. Burge assumes that XYZ is not water. From this he infers that Oscar’s twin never had causal contact with water. Moreover, in twin-English there is no expression we would translate with the English word “water.” Under these circumstances, Oscar’s twin was, in Burge’s view, unable to acquire the notion *water*. Accordingly, Oscar’s twin does not believe that there is water within one kilometer.

To justify this thesis, Burge in addition invokes a truth-based argument. Suppose that Oscar’s twin has beliefs about water, such as the

¹ The example is mine. In what follows, I dispense with the additions in brackets.

² Since there is H_2O in human bodies but not in twin-human ones, Oscar and his twin do not have the same physical properties. This issue can easily be resolved by invoking Burge’s aluminum example instead of his water example (cf. Burge 1982b: 284 f.). In the aluminum example, twin earth differs from earth in that there is, instead of aluminum, a metal with the complicated chemical structure ZYX that “looks like aluminum and is put to many of the uses aluminum is put to here.” (Burge 1982b: 285)

belief that there is water within one kilometer. But given that XYZ is not water, there is no water within one kilometer. The belief of Oscar's twin that there is water within one kilometer would therefore be wrong. Similar considerations show that his belief that this is water, his belief that chemical scientists in his country know about the chemical composition of water, etc. would be wrong as well. In brief, many beliefs of Oscar's twin would be wrong, whereas Oscar's corresponding beliefs are true.

But, according to Burge, there is no reason to think that Oscar's beliefs are true while his twin's beliefs are wrong. He writes: "Their beliefs were acquired and relate to their environments in exactly parallel and equally successful ways." (Burge 1982a: 110) The supposition that Oscar's twin has beliefs about water must accordingly be abandoned.

In Burge's view, Oscar, but not his twin, believes that there is water within one kilometer. Oscar and his twin have different mental states even though they cannot be distinguished with regard to what is inside their bodily limits. The difference in their mental states is due to the fact that they are situated in different physical environments: Oscar is situated in an environment with H_2O , twin Oscar is situated in an environment with XYZ instead. Thus, a person's propositional attitudes depend not only on her internal states but also upon her physical environment.

So much for Burge's attempt to support *physical* externalism or anti-individualism. Note that I will not examine, in this essay, his defense of *social* externalism which he put forward in his article "Individualism and the Mental." For most philosophers who participated in the debate about Burge's anti-individualism think that the thought experiments contrived in this article are based on dubious intuitions.³

2. *Burge's case for externalism criticized*

In the following, I wish to go into several objections against Burge's case for physical externalism. The first objection says that Oscar and his twin only have different *de-re* beliefs (Oscar believes of water that it exists within one kilometer; Oscar's twin does not believe of water that it exists within one kilometer), but not beliefs with different contents.

In Burge's spirit one can rejoin that, in the sentences "Oscar believes that there is water within one kilometer" and "Oscar's twin does not believe that there is water within one kilometer," the expression "water" occurs obliquely. In other words, one cannot always substitute the term "water" in these sentences with a co-extensional one without altering the truth value of the whole sentence. In Burge's view, oblique

³ The so-called "arthritis thought experiment", for example, is based on the dubious intuition that one can say of a person that she believes (wrongly) that she has arthritis in the thigh, although arthritis is by definition an inflammation of the joints. This intuition is dubious because its correctness is questioned by many philosophers (cf., for example, Crane 1991: 15–22, and Bach 1994: 264–270).

occurrences of expressions in content clauses specify the content of the mental state ascribed. Accordingly, Oscar and his twin have mental states with different contents.⁴

The second objection says that XYZ is water after all, or in other words, the English expression “water” refers to H₂O as well as to XYZ. This is because, first, the expressions “water” and “colourless, transparent, drinkable etc. liquid which boils at 100 °C” are synonymous, and, second, XYZ obviously belongs to the extension of the latter expression. Accordingly, Oscar’s twin had causal contact with water and, therefore, could have acquired the notion *water*. In short, no compelling argument has been advanced against the claim that Oscar as well as his twin have beliefs about water.

According to Burge, the premise of this objection is mistaken: XYZ is not water. He writes: “I shall not argue for this view because it is pretty obvious, pretty widely shared, and stronger than arguments that might be or have been brought to buttress it.” (Burge 1982a: 100)⁵

Let us grant, for the sake of the argument, that XYZ is not water, that Oscar’s twin therefore never had causal contact with water and that there is no expression in twin-English which we would translate with the English expression “water.” Does it follow that Oscar’s twin could not have acquired the notion *water*, as Burge presumes?

One evidently cannot invoke so-called “folk psychology” to support this inference. The *theory* of concept acquisition presupposed by Burge demands a justification – which he remains short on.

Burge’s truth-based argument for externalism is in even worse shape than his concept-based argument for externalism. In his view, Oscar’s twin does not have false beliefs about water but true beliefs about twater because Oscar’s and his twin’s “beliefs were acquired and relate to their environments in exactly parallel and equally successful ways.” (Burge 1982a: 110) Oscar’s twin is as successful with his beliefs as Oscar. Can one infer that their beliefs must have the same truth values? I do not think so. For the truth values of some of their beliefs could be irrelevant for their success.

Burge may rejoin that the inference is legitimate if one adds the premise that the truth values of Oscar’s water belief and his twin’s twater belief are relevant for their success. This premise is, however, mistaken. If Oscar’s belief that there is water within one kilometer were wrong, he might be as successful as he actually is. For the liquid which substitutes water might be XYZ (recall that, on our current assumption, XYZ is *not* water!), that is, a liquid which is drinkable, boils at 100 °C etc., in short: the liquid which replaces water might be as life-supporting as water.

⁴ Although the exact wording of this rejoinder cannot be found in Burge’s papers, it is certainly in his spirit (cf. Burge 1979: 86 f. and 92, and Burge 1982a: 99).

⁵ One may object that, apart from H₂O, water contains other chemical constituents, too. So why not also XYZ? This objection can be refuted by substituting the aluminum example for the water example.

Contrast this with Oscar's belief that there is within one kilometer a colourless, transparent, drinkable etc. liquid which boils at 100 °C. The truth value of this belief *is* relevant for Oscar's success: if this belief were wrong, there would be quite probably no equally life-supporting liquid in Oscar's vicinity and he would not therefore be as successful as he actually is.

In sum, Burge's two arguments for the claim that Oscar's twin does not believe that there is water within one kilometer are unconvincing even if one grants him his assumption that XYZ is not water. I now wish to look at two attempts to leap to Burge's defense. It will turn out that the second of these attempts is successful.

3. *Establishing externalism – A first try*

In his essay "A Puzzle about Belief," Saul Kripke formulates the following "strengthened disquotational principle" for beliefs: "A normal English speaker who is not reticent will be disposed to sincere reflective assent to '*p*' if and only if he believes that *p*." (Kripke 1979: 249 and 258).

One may generalize this principle to all languages as follows:

(sDP) A normal person *P* who speaks a language *L* believes that *q* if and only if there is a translation "*r*" of "*q*" into *L* such that *P* would, on reflection, sincerely answer the question whether "*r*" is the case in the affirmative.⁶

But there is no translation of the English expression "water" in the language of Oscar's twin. For the extensions of the English and of the twin-English expression "water" are different. Moreover, we can introduce the term "twater" into English to render the twin-English expression "water" into English, but we cannot introduce an expression into twin-English to render the English expression "water" into twin-English. For if we did this, we would change what we imagine in our thought experiment, that is, we would change the subject matter of our inquiry.

If these considerations are correct, the English sentence "There is water within one kilometer" cannot be translated into twin-English. Thus, there is no translation *t* of this sentence into twin-English such that Oscar's twin would, on reflection, sincerely answer the question whether *t* is the case in the affirmative. Given the strengthened disquotational principle (sDP), it can be concluded that Oscar's twin does not believe that there is water within one kilometer.

Like the previous ones, this argument for externalism has a weakness. It lies in the principle (sDP). What is problematic about this principle is not necessarily that an analysis of beliefs in terms of dispositions to behavior is always misguided.⁷ Rather, the issue with the principle (sDP) is the following: Cases of untranslatability of the kind

⁶ This principle is true only on the assumption that "*q*" and "*r*" do not contain indexical or ambiguous expressions. The same restriction is required for the principle (sDP) formulated below.

⁷ The principle (sDP) does not imply a behavioristic analysis of beliefs since it contains the expressions "on reflection" and "sincerely."

at hand are so rare that in everyday life we do not know how to handle them. Accordingly, the principle (sDP) can hardly provide us with an analysis of our everyday or “folk-psychological” concept of a belief. The following principle is therefore more appropriate:

(sDP') Suppose P is a normal person who speaks a language L . Suppose further that “ r ” is the translation of “ q ” into L . Then the following is the case: P believes that q if and only if it is true that P would, on reflection, sincerely answer the question whether “ r ” is the case in the affirmative.

In contrast to the principle (sDP), the principle (sDP') does not imply anything regarding the question what a person believes if one deals with a case of principled non-expressibility. Accordingly, one cannot argue for externalism as outlined above by relying on the principle (sDP').

4. *The case for externalism*

The case for anti-individualism in the philosophy of mind which am about to present does not invoke the strengthened disquotational principle, as the argument just rejected did, but rather a weaker principle which Kripke calls the “disquotational principle” for beliefs.

The extension of the expression “water” as used in the language spoken by Oscar’s twin is determined by certain causal chains between him and XYZ. The extension of “water” in this language is accordingly XYZ. If Oscar’s twin assents to the sentence “There is water within one kilometer,” then this sentence is true. Which belief does Oscar’s twin express with his utterance? Surely not the belief that there is water within one kilometer, for this belief is wrong (XYZ is not water). Thus, Oscar’s twin expresses the belief that there is twater within one kilometer.

But could Oscar’s twin not express a wrong belief by, on reflection, sincerely assenting to a true sentence? After all, what does it mean that one expresses a belief by assenting to a sentence?

Instead of attempting to rebut the objection implicit in the first question, I will now answer the second question and then present an improved argument for the claim that Oscar’s twin does not express with his utterance the belief that there is water within one kilometer, but rather the belief that there is twater within one kilometer.

So what does it mean that one expresses a belief by assenting to a sentence? Here Kripke’s disquotational principle comes into the picture. He states it in the following way: “If a normal English speaker, on reflection, sincerely assents to ‘ p ,’ then he believes that p .” (Kripke 1979: 248 f.) Generalized to all languages:

(DP) If a normal speaker S of the language L , on reflection, sincerely assents to the translation of the sentence “ p ” into L , then S believes that p .⁸

⁸ This principle is true only on the assumption that “ p ” and the translation of “ p ” into L do not contain indexical or ambiguous expressions.

The belief expressed by assenting to a sentence is the belief which can be ascribed to the person in question by the principle (DP) and on the basis of the fact that the person assents to this sentence.

The claim to be argued for is:

Oscar's twin does not express the belief that there is water within one kilometer by, on reflection, sincerely assenting to the sentence "There is water within one kilometer."

This claim can now be reformulated as follows:

From the fact that Oscar's twin, on reflection, sincerely assents to the sentence "There is water within one kilometer," further true assumptions and the principle (DP) it cannot be inferred that Oscar's twin believes that there is water within one kilometer.

If one assumes – again following Kripke – that translations from one language into another preserve truth value (cf. Kripke 1979: 250), if one assumes, in addition, that the extension of the twin-English expression "water" is XYZ, whereas XYZ is not water, then one can support the claim in question as follows:

- (1) The only assumption which licenses the inference of the proposition "Oscar's twin believes that there is water within one kilometer" from the principle (DP) and the fact that Oscar's twin assents to the sentence "There is water within one kilometer" is the assumption that the twin-English sentence "There is water within one kilometer" is to be translated with the word-identical English sentence.
- (2) But the assumption that the twin-English sentence "There is water within one kilometer" is to be translated with the word-identical English sentence is incorrect.

For:

- (a) The twin-English sentence is true because the twin-English expression "water" refers to XYZ.
- (b) The English sentence is wrong, for XYZ is not water.
- (c) Translations preserve truth value.

From the premises (a) to (c) one can infer the contention (2). From it and the contention (1) one can finally conclude that the claim to be argued for is true.

It has been shown that Oscar's twin does not express the belief that there is water within one kilometer by, on reflection, sincerely assenting to the sentence "There is water within one kilometer." But which belief does he then express by assenting to this sentence? It will be argued now that Oscar's twin expresses the belief that there is twater within one kilometer.

This claim can be reformulated as follows:

From the fact that Oscar's twin, on reflection, sincerely assents to the sentence "There is water within one kilometer," further true assumptions and the principle (DP) it can be inferred that Oscar's twin believes that there is twater within one kilometer.

This claim is true because the further true assumption of the kind required is the assumption that the twin-English sentence “There is water within one kilometer” is to be rendered into English as “There is twater within one kilometer.”

In brief, Oscar’s twin, by assenting to the sentence “There is water within one kilometer,” does not express a belief about water, but rather one about twater. So far, so good. But does it follow that he does not believe that there is water within one kilometer? No, it does not. For Oscar’s twin could have the water belief as well as the twater belief. To be sure, the water belief would then be in principle inexpressible. In particular, Oscar’s twin would be unable to express it. In other words, it could not be ascribed to him based on the principle (DP).

But is there any reason at all for ascribing to Oscar’s twin not only the twater but also the water belief? Note that, to rescue individualism, one moreover had to ascribe to *Oscar* not only the water, but also the twater belief. Is there – apart from the attempt to stick to individualism – any good reason at all for such a weird duplication of belief ascriptions?

A particularly obstinate individualist could invoke the following principle to defend this duplication:

- (I) If a person *P* believes that ... *A* ..., and if, moreover, *P* associates the concept expressed by “*A*” with the same images and identification procedures as with the concept expressed by “*B*,” then *P* believes that ... *B* ..., too.

This principle allows the individualist to ascribe not only the twater but also the water belief to Oscar’s twin. For the latter associates the same images and identification procedures with the twater concept as with the water concept.

However, the principle (I) has – apart from the fact that it seems to be entirely *ad hoc* and only postulated to rescue individualism – one crucial disadvantage. It implies that one must ascribe duplicated beliefs to the protagonists of Putnam’s elm-beech example, too (cf. Putnam 1975: 226 f.).

Like most of us, I cannot discriminate between elms and beeches. Moreover, I associate the same images with elms as with beeches. If asked what an elm is, I answer: some leaf tree. And I answer precisely the same if asked what a beech is. Nonetheless, I refer to elms when using the word “elm,” and to beeches when using the word “beech.” The two expressions differ in extension. How can this be if I am associating the same descriptions and images with both expressions? Putnam’s answer is: The difference in extension is due to the fact that my usage of the word “elm” and my usage of the word “beech” are linked by causal chains to different kinds of trees, viz., elms and beeches, respectively.

External causal chains fix the reference of words, not internal states such as images, identification procedures or dispositions.

Proponents of the principle (I) must ascribe to me – as to most of their fellow beings – an elm belief if and only if they ascribe to me a beech belief. For I associate the same images and identification procedures with the two notions. The absurdity of the principle (I) is much more evident in this example than in the water-twater example. Why?

Well, regarding the just mentioned example, one can argue that the expression “twater” is a philosophical term of art and that accordingly the duplicated ascription of water and twater beliefs is not at odds with our ordinary practice of ascribing and denying beliefs. One cannot argue in this way regarding the elm-beech example. The words “elm” and “beech” are not philosophical terms of art but words used in everyday life. The claim that one must ascribe to someone an elm belief if and only if one must ascribe the corresponding beech belief clashes with our ordinary practices of ascribing and denying beliefs in the most obvious way. And this cannot be explained away by drawing attention to the alleged fact that the expressions in question are mere technical terms.

To make it conspicuous that there is a clash of the kind described, let us look at an example of the elm-beech variety in more detail. If I assent to the sentence “This is an elm,” I believe that this is an elm. Yet, I do not believe, in addition, that this is a beech. This is because I know that elms and beeches are different kinds of trees and I therefore believe that this is not a beech tree.

There is of course a subtle shift here from “I believe that this is *not* a beech” to “I do *not* believe that this is a beech.” This shift is not illegitimate, though. There is simply no reason to suppose that I have contradictory beliefs in this case. Transitions such as the one under consideration are routine in everyday life. Of course, there are far-fetched examples in which the transition rule implicit in this routine breaks down. Kripke’s puzzling Pierre is such an example (cf. Kripke 1979: 254–259). But this does not refute the view that there are plenty of cases in everyday life where this rule can be applied – and *is* applied – without danger.

But is the duplicated ascription of beliefs, quite apart from the fact that it is at odds with our everyday practice of belief ascription, maybe philosophically required? Why not argue that the belief that there is a beech in front of me is the same belief as the belief that there is an elm in front of me?

An argument for this identity thesis could run as follows: The extension of the word “beech,” in my idiolect, is the same as the extension of the word “elm,” in my idiolect. Therefore, the meaning of both words, in my idiolect, is identical, too, for the non-extensional components of their meanings, in my idiolect, are the same anyway. Thus, the belief that there is an elm tree in front of me is identical to the belief that there is a beech tree in front of me.

To this, I reply that the extension of the expression “beech”, in my idiolect, is not at all identical to that of the expression “elm,” in my idiolect. Suppose that, as a matter of fact, there is a beech – not an

elm – in front of me. And I have assented to the sentence “There is an elm in front of me.” Suppose further my interlocutor – of whom I know that he is an expert about elm trees and beech trees who furthermore does not wish to mislead me – corrects me with the words “No. This is a beech and not an elm.” Under these conditions, *I will accept the correction*. This implies that the sentence “There is an elm in front of me,” to which I have assented before, is wrong. The extension of the expression “elm,” in my idiolect, is accordingly not identical to the extension of the expression “beech,” in my idiolect.

A further manoeuvre of the proponent of individualism could be to reformulate the principle (I) in the following way:

(I') If a person *P* believes that ... *A* ..., and if, moreover, *all members of P's language community* associate the concept expressed by “*A*” with the same images and identification procedures as with the concept expressed by “*B*,” then *P* believes that ... *B* ..., too.

In comparison with the original principle, the principle (I') has, for the individualist, one crucial merit: it does not have the unwelcome consequence that one must ascribe to me – as well as to all others who are in the same situation as the protagonists of Putnam's elm-beech example – the elm belief as well as the beech belief. This is because not all members of my language community associate the same images and identification procedures with the elm concept and the beech concept. There are beech and elm experts in my language community who, in contrast to me and most of us, can discriminate elm trees from beech trees.

Nonetheless, the advocate of individualism cannot embrace the principle (I'). For, in the first place, like the original principle, this principle is completely *ad hoc* and implies a weird duplication of belief ascriptions. In the second place, this principle does not justify ascribing to Oscar's twin the twater belief as well as the water belief. For not all members of his language community associate the same images and identification procedures with the twater and the water notion. In the language community of Oscar's twin, there are experts who, in contrast to Oscar's twin, know about the chemical constitution of XYZ and whose twater concept is accordingly associated with different identification procedures than their water concept (in case they have one).

Finally, a proponent of individualism could modify the principle (I) as follows:

(I'') If a person *P* believes that ... *A* ..., and if, moreover, *P* associates the concept expressed by “*A*” with the same *functional role* as with the concept expressed by “*B*,” then *P* believes that ... *B* ..., too.⁹

Like the principle (I'), the principle (I'') does not have the unwelcome consequence that one must ascribe to me – as well as to all others who are in the same situation as the protagonists of Putnam's elm-

⁹ The functional role of a concept, as it is understood here, is determined by the concept's relations to other concepts, sensory “inputs” and behavioral “outputs.”

beech example – the elm belief as well as the beech belief. For I do not associate with the concept expressed by the word “elm” the same functional role as with the concept expressed by the word “beech,” because I believe that elms and beeches are different kind of trees.

Nonetheless, an adherent of individualism cannot rely on the principle (I'') to ascribe to Oscar's twin the water belief. The reason for this is the following: Since there is no translation of the English sentence “... water ...” into twin-English, twin-Oscar's thought that ... water ... is for him in principle inexpressible. His thought that ... twater ..., on the contrary, is for him expressible (by the twin-English sentence “... water ...”). Therefore, twin-Oscar's thought that ... twater ... does not bear the same relation to behavioral “output” as his thought that ... water In brief, he does not associate the same functional role with his water concept and with his twater concept. Thus, a proponent of individualism cannot use the principle (I'') to ascribe to Oscar's twin the water belief, in addition to the twater belief.

Before I come to a conclusion, I wish to address two rather fundamental objections. The first objection is that my argument for externalism is based on a fallacy: From the thesis that there is *no* reason (based on the principle (I), (I') or (I''), respectively, or any other principle) to *ascribe* to Oscar the twater belief it does not follow that there is *a* reason to *deny* that Oscar has the twater belief (cf. Kemmerling 2017: 285).

To rebut this challenge, I wish to give two responses. My *first response* is based on a principle of economy which can be formulated in the following way: If two doxastic descriptions (for example, “Oscar has the water belief, but not the twater belief” vs. “Oscar has the water belief as well as the twater belief”) are available, choose *ceteris paribus* the description which ascribes to the person in question fewer belief properties. My argument against the principle (I) and related principles is supposed to show that the *ceteris-paribus* condition is satisfied: there is no reason to ascribe to Oscar the twater belief. By the economy principle just stated, one can infer that one must deny that Oscar has the twater belief.

The issue becomes trickier if one allows the following three doxastic descriptions:

- (1) Oscar has the water belief but not the twater belief.
- (2) Oscar has the water belief as well as the twater belief.
- (3) Oscar has the water belief. The claim that he has the twater belief lacks a truth value.

It is now not evident any longer which of these descriptions is the most economical description of Oscar. Is it (1) or (3)?

But even if this question cannot be answered, the doctrine of individualism is refuted by the considerations advanced above. For, according to this doctrine, the same belief sentences are true regarding dop-

pelgangers such as Oscar and his twin.¹⁰ This is, however, not the case, for “Oscar’s twin believes that there is twater within one kilometer” is true, whereas “Oscar believes that there is twater within one kilometer” is wrong (option (1)) or lacks a truth value (option (3)).¹¹

My *second response* to the above objection consists in turning my considerations regarding the principle (I) against individualism in still another way than that described so far:

Individualism is mistaken because, first, its proponents must endorse the principle (I) (they identify concepts with images, identification procedures etc.) and, second, this principle is at odds with the way we would describe the protagonists in the elm-beech example. If I assent to the sentence “This is an elm tree,” I believe that this is an elm tree. But I do not believe, in addition, that this is a beech tree. For I know that a beech tree and an elm tree are different kinds of trees and I do not have contradictory beliefs regarding this issue.

Without doubt, it is a merit of this argument for externalism that one does not have to invoke economy considerations to deny that I have the beech belief (but only the assumption that I do not contradict myself regarding the issue in question). Nor is there a reason to claim that the sentence “I believe that this is a beech tree” lacks a truth value, for it is evidently wrong.

Brian Loar has argued that the same content clause can be used to ascribe beliefs with different contents (cf. Loar 1988: 102).¹² To establish this thesis, he contrives a variant of Kripke’s well-known thought experiment of puzzling Pierre. Let us first look at Kripke’s original story (cf. Kripke 1979: 254–259). Suppose Pierre, a normal French speaker, assents to the sentence “Londres est jolie.” Assuming the disquotational principle, we can ascribe to him the belief that London is pretty. Imagine, now, that Pierre moves to an unattractive part of London. Nobody there knows any French. He manages to pick up English without using a translation manual. Yet, he does not realize that the names “London” and “Londres” refer to the same town. He assents to “London is not pretty,” thereby expressing the belief that London is not pretty.

Since Pierre apparently does not have contradictory beliefs, it can be inferred (based on his English assertion) that Pierre does not believe that London is pretty. This is puzzling, though, because his French assertion allows one to attribute to Pierre the belief that London is pretty.

Let us now turn to Loar’s modification of Kripke’s original story (cf. Loar 1988: 10). Loar’s Pierre does not move to an unattractive part of London, but rather right to the city center. He is rather impressed by

¹⁰ The subject term of the belief sentences can, of course, differ.

¹¹ If option (3) cannot be ruled out, externalism can only be supported by the water-twater example if this doctrine is understood as being equivalent to *anti-individualism* (viz., the negation of individualism).

¹² By “content” I mean what Loar calls “psychological content.”

it and therefore assents to the sentence "London is pretty." The belief sentence "Pierre believes that London is pretty" is accordingly true not only because of his French assertion "Londres est jolie," but also because of his English assertion "London is pretty." In Loar's view, the "French" belief sentence and the "English" belief sentence ascribe different contents although they contain the same content clause.

Loar suggests that, just as the same content clause can be used to ascribe different contents, different content clauses may ascribe beliefs with the same content (cf. Loar 1988: 105 f.) In his view, the ascriptions "Oscar believes that there is water within one kilometer" and "Oscar's twin believes that there is twater within one kilometer" are of this kind: they contain different content clauses but ascribe the same content. It is for this reason that, according to Loar, Burge's case for anti-individualism fails.

This critique of Burge's argument can easily be dismissed. For Burge does not argue that externalism is true because the two belief ascriptions mentioned above ascribe different contents. As should be clear from section 1, he derives externalism from the claim that Oscar believes that there is water within one kilometer, whereas his twin does *not* believe that there is water within one kilometer.

Moreover, it should be apparent by now that this way to argue for externalism is to be preferred to the argument for externalism which Loar attributes to Burge. For even if Oscar's twin believes that there is twater within one kilometer, he may also believe that there is water within one kilometer.

But if Loar is right in claiming that the same content clause can be used to ascribe different contents, one may wonder whether a negation of the belief ascription "*P* believes that *q*" must necessarily deny that *P* has a belief with the same content as the belief ascribed to *P* by "*P* believes that *q*." The "English" negation "Pierre does not believe that London is pretty"¹³ may possibly not deny of Pierre a belief with the same content as the belief attributed to Pierre by the "French" ascription "Pierre believes that London is pretty."¹⁴

In the same vein, one may suggest that "Oscar's twin does not believe that there is water within one kilometer" does not deny of Oscar's twin a belief with the same content as the belief ascribed to Oscar by "Oscar believes that there is water within one kilometer." The upshot of these considerations is that Oscar and his twin could have beliefs with the same contents even though everything said in this essay on behalf of externalism is correct.

In response, let me consider, first, the original Pierre example in more detail. The "English" negation "Pierre does not believe that London is pretty" is derived from Pierre's assertion "London is not pretty,"

¹³ I call this negation "English" because it is based on the English assertion "London is not pretty."

¹⁴ This ascription is called "French" because it is based on the French assertion "Londres est jolie."

the principle of disquotation and the claim that Pierre does not have contradictory beliefs. In this way, it is linked to his usage of the name "London." On the contrary, the "French" ascription "Pierre believes that London is pretty" is linked to Pierre's usage of the name "Londres." A case may be made that, due to these linkages to different usages of the names "London" and "Londres," respectively, the "English" negation and the "French" ascription deny or ascribe, respectively, beliefs with different contents.

However, my negation "Oscar's twin does not believe that there is water within one kilometer" is not linked to utterances of Oscar's twin in the way the "English" negation and the "French" ascription are linked to Pierre's utterances. My negation is rather based, first, on the claim that, given that Oscar's twin has the twater belief, there is no reason to ascribe to him the corresponding water belief and, second, on certain economy principles.

Because my negation "Oscar's twin does not believe that there is water within one kilometer" is not linked to particular utterances of Oscar's twin, one cannot pick out one particular content among those described by its content clause, arguing that this content is denied by it. Rather, any content described by its content clause is denied by it. This is a fortiori true for the content ascribed to Oscar by "Oscar believes that there is water within one kilometer."

Let me be clear. I do not intend to advocate Loar's view that one content clause can be linked to a variety of contents. I only argue that *if* this view is correct, my negation "Oscar's twin does not believe that there is water within one kilometer" denies, for each of them, that Oscar's twin has it.¹⁵

5. Conclusion

The argument for externalism advanced in this essay makes a series of assumptions which are not expressly supported here. Among them are the following: (i) the claim that XYZ is not water, (ii) the causal

¹⁵ A proponent of Loar's view may rejoin that, in asserting "Oscar's twin does not believe that there is water within one kilometer," one denies only that Oscar's twin has a certain "social" content ("social" contents are individuated by content clauses; cf. Loar 1988: 110). I do not think that this rejoinder is particularly promising because it postulates a sharp contrast between two kinds of belief negations: some deny only "social" contents, others (such as the "English" negation "Pierre does not believe that London is pretty") also "psychological" contents. There is no indication whatsoever in everyday discourse that such a sharp division exists.

It may be objected that it is also implausible to hold that some belief negations (such as the "English" negation "Pierre does not believe that London is pretty") deny particular contents, whereas others (such as my negation "Oscar's twin does not believe that there is water within one kilometer") deny any content described by the content clause. If one agrees with this, one is driven to the conclusion that there is no convincing theory of belief negations within Loar's framework at all. So much the worse for the Loar-style objection to the argument for externalism put forward in this article.

theory of reference regarding the word “water,” (iii) the disquotational principle for beliefs, and (iv) the assumption that translations from one language into another preserve truth value. Here I have done little more than to profess that I think these assumptions are unproblematic. What is established in this article, is, however, that anyone who accepts them must embrace externalism regarding beliefs.

References

- Bach, K. 1994. *Thought and Reference*. 2. rev. ed. Oxford and New York: Oxford University Press.
- Burge, T. 1979. “Individualism and the Mental.” *Midwest Studies in Philosophy* 4: 73–121.
- Burge, T. 1982a. “Other Bodies.” In A. Woodfield (ed.). *Thought and Object. Essays on Intentionality*. Oxford and New York: Clarendon Press, 97–120.
- Burge, T. 1982b. “Two Thought Experiments Reviewed.” *Notre Dame Journal of Formal Logic* 23: 284–293.
- Burge, T. 1986. “Intellectual Norms and Foundations of Mind.” *The Journal of Philosophy* 83: 697–720.
- Crane, T. 1991. “All the Difference in the World.” *The Philosophical Quarterly* 41: 1–25.
- Kemmerling, A. 2017. *Glauben. Essay über einen Begriff*. Frankfurt am Main: Vittorio Klostermann.
- Kripke, S. A. 1979. “A Puzzle about Belief.” In A. Margalit (ed.). *Meaning and Use. Papers Presented at the 2. Jerusalem Philosophical Encounter, April 1976*. Dordrecht and Boston: D. Reidel Publishing Company, 239–283.
- Loar, B. 1988. “Social Content and Psychological Content.” In R. H. Grimm and D. D. Merrill (eds.). *Contents of Thought*. Tucson: The University of Arizona Press, 99–110.
- Putnam, H. 1975. “The Meaning of ‘Meaning’.” In H. Putnam. *Mind, Language and Reality. Philosophical Papers, Volume 2*. Cambridge, New York and Melbourne: Cambridge University Press, 215–271.