

This is a draft. Citation should be to the official version:

Sawyer, S. 2018. Is there a deductive argument for semantic externalism? Reply to Yli-Vakkuri. *Analysis*. 78 (4): 675-81.

<https://doi.org/10.1093/analys/anx164>

Is there a deductive argument for semantic externalism? Reply to Yli-Vakkuri.

SARAH SAWYER

Juhani Yli-Vakkuri (2017) has recently argued that the Twin Earth thought experiments offered in favour of semantic externalism can be replaced by a straightforward deductive argument from premises widely accepted by both internalists and externalists alike. While acknowledging the role of thought experiments in philosophical theorizing, a deductive argument in favour of externalism from agreed principles would undoubtedly be preferable on the grounds that it would be less vulnerable to competing interpretations and hence would be more likely to settle the debate between internalists and externalists in a decisive manner. As an externalist, I would welcome such a proof. The deductive argument Yli-Vakkuri offers, however, depends on premises which are such that, on standard formulations of internalism, they cannot be satisfied by a single belief simultaneously; it does not therefore, constitute a proof of externalism. The aim of this paper is to explain why.

Yli-Vakkuri takes semantic externalism to be ‘the thesis that the contents of intentional states (such as beliefs) and speech acts (such as

assertions) are not determined by the way those subjects are internally' (2017: 1) and focuses for simplicity on the contents of beliefs. Externalism, thus restricted, is understood as the thesis that the content of a belief is not determined by the way the subject is internally. This is clarified by using the notions of *duplicate*, *correspondence* and *narrowness*. Things that are internally the same are *duplicates*; whenever S and S' are duplicates, each part of S *corresponds to* a part of S' ; and a property P of beliefs is *narrow* iff, necessarily, any corresponding beliefs of duplicate subjects either both have P or both lack P (2017: 3).

Internalism, understood as the claim that belief content is narrow, is then formalized as follows, where the variables range over beliefs, the two-place predicate C expresses the relation of *being corresponding beliefs of duplicate subjects*, and the function symbols c and v express, respectively, *content* and *truth-value* (2017: 5):

$$\text{NARROW}_c: \quad \Box \forall x \forall y (C(x, y) \rightarrow c(x) = c(y))$$

NARROW_c says that, necessarily, any corresponding beliefs of duplicate subjects have the same content, i.e. that content is narrow. (2017: 5)

The deductive argument in favour of externalism makes use of two additional premises. The first premise is formalized as:

$$\text{BROAD}_T: \quad \neg \Box \forall x \forall y (C(x, y) \rightarrow v(x) = v(y))$$

BROAD_T says that it is not necessary that all corresponding beliefs of duplicate subjects have the same truth value. (2017: 5) Truth, according to

Yli-Vakkuri, is a ‘paradigmatic broad semantic property’ (2017: 4), a fact which he says ‘passes without comment’ (2017: 5). The second premise is formalized as:

TRANSPARENCY: $\Box \forall x \ v(x) = v(c(x))$

TRANSPARENCY says that, necessarily, the truth value of a belief is the same as the truth value of its content. (2017: 5) Transparency is taken to be true on the grounds that if it were false, then one of the following would be true: either (i) S ’s belief that p is true but it is not true that p ; or (ii) S ’s belief that p is not true, and p . Since, according to Yli-Vakkuri, neither (i) nor (ii) is possible, TRANSPARENCY is established. (2017: 5)

Yli-Vakkuri’s deductive argument is based on a derivation of the inconsistency of the set {TRANSPARENCY, BROAD_T, NARROW_c} using only standard first-order logic plus **K**, the weakest normal modal logic. Since NARROW_c expresses internalism and the negation of internalism is externalism, this amounts to a derivation of externalism from the conjunction of TRANSPARENCY and BROAD_T. I do not dispute the derivation. However, the claim that this amounts to a proof of externalism rests on Yli-Vakkuri’s further claim that TRANSPARENCY and BROAD_T are accepted by internalists and externalists alike. I will argue, in contrast, that on standard formulations of internalism, a belief is a substitution instance of TRANSPARENCY iff it is not a substitution instance of BROAD_T. That is to say, corresponding beliefs which differ in truth-value are such that their truth-value is not necessarily identical to that of their content; and corresponding beliefs whose truth-value is necessarily identical to that of

their content necessarily have the same truth-value. Specifically, beliefs fall into two distinct categories—indexical beliefs and non-indexical beliefs—and indexical beliefs satisfy $BROAD_T$ but not $TRANSPARENCY$, while non-indexical beliefs satisfy $TRANSPARENCY$ but not $BROAD_T$. The deductive argument, then, rests on a failure to distinguish indexical from non-indexical beliefs and a corresponding ambiguity concerning the notion of a belief's content.

Let us examine indexical beliefs first. Suppose S sincerely utters the sentence 'That is an apple', thereby expressing an indexical (demonstrative) belief. We are assuming internalism, and hence are committed to $NARROW_c$. As such, we are committed to the claim that necessarily, any corresponding beliefs of duplicate subjects have the same content. Nonetheless, and in accordance with $BROAD_T$, corresponding indexical beliefs of duplicate subjects may well differ in truth-value: we need only suppose, for example, that S is pointing at an apple (making her belief true) while S' is pointing at a pear (making her belief false). How are we to reconcile the claim that corresponding indexical beliefs of duplicate subjects necessarily have the same content while not necessarily having the same truth-value? Clearly the content of such beliefs cannot be 'object-dependent', since this would violate $NARROW_c$.¹ Rather, the content of the beliefs of S and S' must be understood as object-independent. And we find just such a view articulated by Burge (1977) and then defended by Segal (1989).

According to Burge and Segal, a sentence containing an indexical term expresses a belief the content of which is akin to an open sentence rather

¹ For object-dependent thought theories see Evans (1982) and McDowell (1977, 1984).

than to a closed sentence. That is, the content of an indexical belief is specified as a predicative, propositional fragment rather than a complete proposition. An indexical belief understood in this way requires that its content be applied in a context in order for it to be truth-evaluable. This means the same content can be applied across a range of different contexts with different resulting truth-values. The content of corresponding indexical beliefs of duplicate subjects can thus be understood as identical precisely because the duplicates can be understood as applying the same predicative, propositional fragment in different contexts.² The difference in truth-value between corresponding indexical beliefs of duplicate subjects is explained by the difference in contextual application and is not due to a difference in their content. In the example above, the apple and the pear determine that the truth-values of the corresponding beliefs of *S* and *S'* differ, but the contents of the corresponding beliefs may nonetheless be identical—thought by *S* in the context of the apple, and thought by *S'* in the context of the pear. The view involves a distinction between an indexical belief, which is truth-evaluable, and its content, which is not. An indexical belief is truth-evaluable because it involves a contextual application of a content which is in itself neither true nor false. It is this distinction between a belief and its content that entails the falsehood of TRANSPARENCY for indexical beliefs. We can now see that rejecting TRANSPARENCY does not imply, as Yli-Vakkuri maintains,

² The view is consistent with the propositional fragment being either broad or narrow, which explains why it is consistent for the view to be held by both Burge, a paradigmatic externalist, and Segal, a paradigmatic internalist. See Sawyer (2011). The important point here is that the nature of indexical beliefs as akin to open sentences is consistent with NARROW_c and BROAD_T but inconsistent with TRANSPARENCY.

either (i) *S*'s belief that *p* is true but it is not true that *p*, or (ii) *S*'s belief that *p* is not true, and *p*. Rather, (i) and (ii) fail to capture the correct logical form of indexical beliefs by representing the content of all beliefs as complete propositions. The content of an indexical belief is, in contrast, an incomplete propositional fragment that acquires a truth-value only in the context of application.

Farkas (2008) and Pitt (2013) provide an alternative internalist account of indexical beliefs. Given the internalist commitment to NARROW_c, they also maintain that a difference in truth-value between corresponding indexical beliefs is due to a difference in contextual application rather than due to a difference in content.³ But there are two striking differences. First, neither Farkas nor Pitt talk of the content of an indexical belief as a predicative, propositional fragment; and second, they do not draw a distinction between an indexical belief and its content. But there are corollaries of each of these claims in their account. In fact, it would not be implausible to think of the views as notational variants. With regard to the first difference, Pitt says that 'indexical contents are *thin*' (2013: 61, original emphasis), by which he means that they are not truth-evaluable in their own right but only relative to a context of application. This, then, plays the same explanatory role as Burge and Segal's claim that indexical contents are

³ Both Farkas and Pitt take their view of indexical beliefs to be part of a wider, internalist understanding of context-dependent beliefs. Context-dependent but non-indexical beliefs include, on this view, descriptive beliefs, such as the belief that the President of the United States is a criminal (Pitt, 2013: 53), and general beliefs, such as the belief that all dogs are quadrupeds (Pitt, 2013: 53). For present purposes I restrict my discussion to their account of indexical beliefs.

predicative, propositional fragments. With regard to the second difference, Farkas and Pitt agree with the fundamental claim that indexical content requires a contextual application in order to be truth-evaluable. Whether an indexical belief is then understood as the indexical content itself or the truth-evaluable result of a contextual application of the indexical content seems to be a matter of terminological preference. Farkas and Pitt opt for the former; Burge and Segal opt for the latter. The important point for present purposes is that both views entail the falsehood of TRANSPARENCY. This is because both views reject the claim, implicit in the statement of TRANSPARENCY, that the content of a belief is a complete, truth-evaluable proposition. The content of an indexical belief is, on neither account, a complete, truth-evaluable proposition.

Let us now examine non-indexical beliefs. Non-indexical beliefs are beliefs which have, as their content, complete propositions. For such beliefs, TRANSPARENCY is clearly true: the truth value of a non-indexical belief is the same as the truth value of its content. Rejecting TRANSPARENCY for non-indexical beliefs would, I agree, lead to the kinds of absurdities Yli-Vakkuri mentions.⁴ Again, for the sake of the argument, we are assuming internalism and hence are committed to NARROW_c. As such, we are committed to the claim that necessarily, any corresponding non-indexical beliefs of duplicate subjects have the same content. But given

⁴ The specific examples Yli-Vakkuri offers to illustrate this point involve singular uses of names. Since I take names to be predicates with a demonstrative (and hence indexical) element in their singular use, I do not think the examples work on the grounds that they express indexical beliefs with contents which are incomplete propositional fragments. See Burge (1973). See also Sawyer (2010). Nonetheless, I agree with the claim.

TRANSPARENCY, this means that necessarily, any corresponding non-indexical beliefs of duplicate subjects have the same truth-value. This amounts to a rejection of BROAD_T. One might wonder how this is possible. Yli-Vakkuri reflects a common intuition when he says that truth is a paradigmatic broad semantic property and that it requires no further comment. But the intuition that truth is a broad semantic property is grounded in the idea that the truth or falsity of a belief depends on facts beyond the believer. This, I take it, is true. It does not follow from this, however, that corresponding beliefs of duplicate subjects might have different truth values. To see this, we need simply look at the internalist view defended by Segal (2000), which, I will argue, implies that corresponding non-indexical beliefs of duplicate subjects necessarily have the same truth-value. If this is right, then there is a standard internalist view of non-indexical beliefs according to which TRANSPARENCY is true and BROAD_T is false.

Consider the example in Putnam (1975) of the duplicate subjects Oscar, on Earth, and Toscar, on Twin Earth. According to the externalist, Oscar refers to water (H₂O) by his use of the term 'water', while Toscar refers to twin water (XYZ) by his use of the term 'water'. As such, they express beliefs with different contents when they utter the same-sounding sentence 'water always contains oxygen'.⁵ The content of Oscar's belief is *water always contains oxygen*, while the content of Toscar's belief is *twater always contains oxygen*. And it is clear that Oscar's belief is true while Toscar's

⁵ We are to assume that neither Oscar nor Toscar knows the compositional structure of the watery stuff in their environments, but we can suppose for the purposes of this example that they have learnt about oxygen and compositional structure through other examples, and have come to form the corresponding beliefs by means of conjecture.

belief is false. This illustrates the truth of BROAD_T on the externalist understanding of the content of non-indexical beliefs. According to the internalist theory proposed by Segal, in contrast, Oscar and Toscar each express beliefs with the same content because their corresponding beliefs have the same (narrow) cognitive role. The term ‘water’ as used by Oscar and Toscar has an extension which includes water, twin water, and anything else that is (roughly speaking) indistinguishable from their point of view. Let us use the term ‘water*’ to refer to all such substances. This allows us to capture the content of the corresponding beliefs of Oscar and Toscar. The content of their beliefs is simply *water* always contains oxygen*. And now we can see why it is impossible for the truth-value of Oscar’s belief to differ from the truth-value of Toscar’s belief. The belief that water* always contains oxygen is false no matter who utters it, because water* does not always contain oxygen—twin water provides a counterexample. It is worth noting that there is a perfectly good sense in which the truth-value of the belief that water* always contains oxygen depends on facts about the world beyond Oscar and Toscar. This respects the intuition that truth is a paradigmatic broad semantic property. But nonetheless, Segal’s view implies that beliefs of duplicate subjects will never differ in truth-value. BROAD_T is, on this internalist understanding of the content of non-indexical beliefs, false.⁶

⁶ Farkas (2013: 174) suggests that utterances of sentences containing the term ‘water’ might express the same content on Earth and Twin Earth despite a difference in truth-conditions. This would constitute an extension of her account of indexical beliefs mentioned above. However, she does not endorse the view and does not specify what the content of such beliefs would be. As such, I merely note that this provides a potential internalist alternative to Segal’s view of non-indexical beliefs.

Now consider the example in Burge (1979) of Alf and counterfactual Alf, each of whom utters the same-sounding sentence ‘Arthritis can spread to the thigh’. In Alf’s community the term ‘arthritis’ is defined as an ailment specifically confined to the joints, whereas in counterfactual Alf’s community, the term ‘arthritis’ has a broader definition and applies in addition to similar ailments outside the joints. According to the externalist, Alf expresses the belief that arthritis can spread to the thigh, a belief which is false, while counterfactual Alf expresses the distinct belief that tharthritis can spread to the thigh, a belief which is true. The intuition depends on individuating the content of a belief by appeal to the broad property of its being embedded in a particular socio-linguistic community, and once again it illustrates the truth of BROAD_T on the externalist understanding of the content of non-indexical beliefs. But Segal rejects this interpretation. His commitment to NARROW_c leads him to attribute Alf and counterfactual Alf beliefs with the same content on the grounds that the beliefs have the same (narrow) cognitive role. According to Segal, both Alf and counterfactual Alf believe that tharthritis can spread to the thigh, and both beliefs are true. More generally, the truth-value of corresponding non-indexical beliefs of duplicate subjects will, on Segal’s internalist theory, necessarily have the same truth-value. This is because corresponding non-indexical beliefs of duplicate subjects have the same content, and the content, being a complete proposition, has its truth-conditions essentially. BROAD_T is false.

In conclusion, while the set {TRANSPARENCY, BROAD_T, NARROW_c} is indeed inconsistent, there is a natural internalist understanding of indexical beliefs according to which they satisfy BROAD_T but not TRANSPARENCY, and a natural internalist understanding of non-

indexical beliefs according to which they satisfy TRANSPARENCY but not BROAD_T. On neither view need we give up NARROW_c. Whether we choose to accept such internalist positions, or whether we choose instead to embrace externalism, will, I maintain, be determined by independent factors such as intuition, thought experiment and perceived explanatory gain. What is clear is that we do not yet have a proof of externalism.⁷

The University of Sussex
Brighton, BN1 9QN, UK
s.a.sawyer@sussex.ac.uk

References

- Burge, T. 1973. Reference and proper names. *The Journal of Philosophy*. 70: 425-39.
- Burge, T. 1977. Belief de re. *The Journal of Philosophy*. 74: 338-62.
- Burge, T. 1979. Individualism and the mental. In P. French, T. Uehling & H. Wettstein (eds). *Studies in Epistemology, Midwest Studies in Philosophy* 4: 73-121. Minneapolis: University of Minnesota Press.
- Evans, G. 1982. *The Varieties of Reference*. J. McDowell (ed.) Oxford: Oxford University Press.
- Farkas, K. 2008. *The Subject's Point of View*. Oxford: Oxford University Press.

⁷ Thanks to an anonymous referee and to Michael Morris for comments. Thanks also to Corine Besson for comments and helpful discussion.

- McDowell, J. 1977. On the sense and reference of a proper name. *Mind*. 86: 159-85.
- McDowell, J. 1984. De re senses. In C. Wright (ed.) *Frege: Tradition and Influence*. Oxford: Basil Blackwell.
- Pitt, D. 2013. Indexical thought. In U. Kriegel (ed.) *Phenomenal Intentionality*. New York: Oxford University Press: 49-70.
- Putnam, H. 1975. The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science* 7: 131-91.
- Sawyer, S. 2010. The modified predicate theory of proper names. In S. Sawyer (ed.) *New Waves in Philosophy of Language*. New York: Palgrave MacMillan Press: 206-25.
- Sawyer, S. 2011. Internalism and externalism in mind. In J. Garvey (ed.) *The Continuum Companion to the Philosophy of Mind*. London: Continuum International Publishing Group: 133-50.
- Segal, G. 1989. The return of the individual. *Mind*. 98: 39-57.
- Segal, G. 2000. *A Slim Book about Narrow Content*. Cambridge, Mass.: MIT Press.
- Yli-Vakkuri, J. 2017. Semantic externalism without thought experiments. *Analysis* 77 <https://doi.org/10.1093/analys/anx127>