

Singapore Management University Institutional Knowledge at Singapore Management University

Research Collection School of Social Sciences
(Open Access)

School of Social Sciences

1-2002

Externalism and Knowledge of Comparative Content

Yoo Guan Tan

Singapore Management University, ygtan@smu.edu.sg

Follow this and additional works at: http://ink.library.smu.edu.sg/soass_research

 Part of the [Social and Behavioral Sciences Commons](#)

Citation

Tan, Yoo Guan, "Externalism and Knowledge of Comparative Content" (2002). *Research Collection School of Social Sciences (Open Access)*. Paper 5.

http://ink.library.smu.edu.sg/soass_research/5

Available at: http://ink.library.smu.edu.sg/soass_research/5

This Working Paper is brought to you for free and open access by the School of Social Sciences at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection School of Social Sciences (Open Access) by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email libIR@smu.edu.sg.

Externalism and Knowledge of Comparative Content

Yoo Guan Tan

July 2002

Paper No. 5-2002

Externalism and Knowledge of Comparative Content

Y. G. Tan

Concepts are the constituents of thoughts, which in turn, are the contents of propositional attitudes. They are also what the predicates of our language express. According to a tradition going back to Plato, questions about comparative content – questions of the form “Is concept *F* the same as concept *G*?” – are purely about relations of ideas, and so are answerable *a priori*. This does not mean that no experience at all is necessary to answer such questions, for experience may be needed to grasp their content. Call a piece of information about *F*s *extraneous* if it is not required to obtain a proper understanding of the concept *F*. Then what the traditional assumption says is that no extraneous information is necessary to answer a question about comparative content. Henceforth, I shall refer to this assumption about concepts as the *transparency thesis*, or **T** for short.

In recent years, an account of mental content known as externalism has been gaining increasing acceptance among philosophers. According to this account, the contents of a person’s propositional attitudes – beliefs, desires, hopes, – are partly determined by factors in his social and physical environment, of which he may have no knowledge. A quick survey of the current literature indicates that there is a strong convergence of opinion, among philosophers on both sides of the externalist-internalist divide, regarding the connection between externalism and **T**. As Falvey and Owens put it, “There is no question but that externalism is incompatible with introspective knowledge of comparative content. ... Given externalism, there is simply no reason to suppose that a subject should be capable of making correct judgments of comparative content if he lacks the relevant empirical

knowledge.” (p. 110-111 [5]) Hilary Putnam, Colin McGinn, Tyler Burge and others have made similar claims.

It is not my aim in this paper to defend **T**. My concern is only to consider and answer the question “Is externalism compatible with **T**?” I believe that the matter is not as clear-cut as it may seem to those philosophers who answer the question negatively. Not every form of externalism, I shall argue, is incompatible with **T**. In particular, the externalism that is warranted by the standard Twin-Earth arguments is neutral with respect to **T**. I argue for this claim in section (I) and (II). If I am right, then Putnam’s argument in **The Meaning of “Meaning”** against what he calls the traditional theory of meaning – the theory that meanings are in the head – is inconclusive. For **T** provides one good sense in which concepts (and hence meanings) are in the head. I pick up this point in section (III).

I

It is possible to distinguish two types of argument for content externalism. The first type of argument depends essentially on the assumption that a person can have a propositional attitude even if he lacks a complete grasp of some concept in its content. Suppose A, whose use of the word ‘arthritis’ has so far been unexceptional, now comes to affirm “My arthritis has spread to my thigh.” Since arthritis is specifically an inflammation of joints, this shows that A’s understanding of the word ‘arthritis’ is defective. Nevertheless, this need not prevent us from attributing to A the (false) belief that he has arthritis in his thigh. Now imagine a counterfactual situation in which A’s physical and mental histories (non-intensionally described) are exactly the same but the word ‘arthritis’ is

used in the linguistic community for various forms of rheumatoid ailments including the one in A's thigh. If this situation were actual, A would not have any beliefs whose contents involved the concept *arthritis*. He would not, for instance, have the belief that he has arthritis in his thigh; in its place would be a belief that could be true. It follows that the contents of a person's propositional attitudes are not independent of the linguistic and social conventions that exist within his community. For convenience, I shall refer to this form of externalism as *externalism1*.

But we can accept the conclusion of this argument without having to abandon **T**. Let *arthritis** be the concept that is expressed by 'arthritis' in the counterfactual situation. Then surely no experience other than what is required for a proper understanding of the concepts *arthritis* and *arthritis** is necessary for knowing that they are different concepts. Thus, if someone unfamiliar with the word 'arthritis' were given two explanations of its meaning, one in accordance with the concept *arthritis* and the other in accordance with the concept *arthritis**, he would know that they cannot both be correct. For the argument to entail the falsity of **T**, it must imply that some of our concepts have identity conditions that are not fully determined by our linguistic practices. But the present argument does not have this implication. Therefore, we may safely conclude that the externalism1 is not incompatible with **T**.

However, to be fair to Falvey and Owens, **T** is not the thesis that they have in mind when they claim that "externalism is incompatible with *a priori* knowledge of comparative content." The target of their arguments is a stronger version of **T** which I shall designate **T'**:

(**T'**) With respect to any two of his thoughts or beliefs, an individual can know authoritatively and directly (that is, without relying on inferences from his observed environment) whether or not they have the same content. (p. 109-110 [5])

According to Falvey and Owens, **T'** is false if externalism₁ is true. The following is a variation of one of their arguments. Suppose there is a community L* which actually uses 'arthritis' to mean *arthritis**, and suppose A is a frequent visitor to L*. Then, according to externalism₁, the content of some of A's propositional attitudes will change when he moves between L* and his own community L. When A is in his own community L, some of his thoughts will contain the concept *arthritis* but none will contain the concept *arthritis**. When he is in L*, the situation will be reversed. Now suppose A is asked, shortly after his return from L*, whether, when he utters the sentence "Arthritis is ___" he expresses the same thought that he expressed when he uttered the same sentence last week. Presumably, A will answer yes. But if he was in L* last week, then he will be mistaken. And clearly, nothing that is accessible to A through introspection alone will reveal his mistake to him. Therefore **T'** is false.

Although this argument seems persuasive, it is nevertheless fallacious. Let us assume that A will answer the question affirmatively if he relies solely on introspective evidence. Let us also grant that A will be mistaken if he answers affirmatively. None of this, however, is enough to yield the conclusion of the argument. The crucial question we need to ask is whether A's mistake is one that would be excluded by **T'**. I shall argue that A's affirmative response cannot have the meaning that the argument reads into it. Consequently A's case is not a counterexample to **T'**.

Given A's answer, we may infer that A believes that when he utters the sentence "Arthritis is ___", the thought that he expresses with it is the same as the thought he expressed last week with the same sentence. Let p be the thought A now expresses and q be the thought he expressed last week. Since p and q are different thoughts, A's belief is false. However, for A's case to be a counterexample to **T'**, A's belief must rest upon a judgment about the content of p and q . For otherwise, there will be no mistake about comparative content to hold up against **T'**.

It is essential to note that not every judgment of the form "X and Y are (are not) the same thought" is a judgment of comparative content in the sense used in the various transparency theses. If I affirm, on good authority, that the thought expressed by a certain Finnish sentence is the same as the thought expressed by a certain Polish sentence, I have made a judgment concerning two thoughts. But since I do not understand either of the sentences, my judgment is not a judgment of comparative content in the sense required by **T** and **T'**. In this sort of judgments, the Xs and Ys must themselves be thoughts, nor just names or descriptions of thoughts. Thus for A's case to be a genuine counterexample to **T'**, A's mistake must rest upon a judgment in which both p and q figure as parts of the content.

Now it might be argued that A's belief must involve precisely such a judgment because he can only arrive at the belief by thinking as follows:

- (a1) p is the thought I express now with S (where S = "Arthritis is ___").
- (a2) q was the thought I expressed last week with S.
- (a3) p and q have the same content.
- (a4) So the thought I express now with S is the same as the thought I expressed last week with S.

The judgment that p and q have the same content is thus part of the justification of A's mistaken belief.

But this contention is unsustainable, first, because there are other ways for A to arrive at his belief, and second and more importantly, because the line of reasoning it ascribes to A is one which, in the context of the externalist framework that informs the discussion, A cannot pursue.

To take the second point first, it should be noted that the embedding context of A's false belief is his home community L. By the argument's own assumptions, the concept *arthritis** is one which does not exist in L. In particular, it is a concept which A does not have when he is in L. Since q is a thought that embodies the concept *arthritis**, q is not a thought that can be ascribed to A when he is in L. But if A cannot have this thought, then neither can he have any thought that has q as part of its content. It follows that A can neither have the thought expressed by (a2), nor the thought expressed by (a3), in the reasoning above. But if these thoughts are unavailable to A, then so is the reasoning that the argument attributes to him. The reasoning embodied in (a1) to (a4) cannot therefore be the basis of A's false belief.

We have shown that A's false belief cannot consist in the judgment that p and q have the same content. For in order for A to have this thought, he must be able to think its two constituent thoughts in the same context, but by the argument's own premise, this is something that he cannot do. But if A cannot make this judgment, then, as we have noted earlier, his mistake will have no bearing on the truth or falsity of **T'**. To make the point another way, all that **T'** requires of A is that if he could entertain the question whether p and q are the same thought, then he would be able to know a

priori that they are not the same. Admittedly, the consequent of this conditional is not true of A. But neither is the antecedent. Consequently, the case described in the argument is not a genuine counterexample to **T'**.

Let us now return to the first point and see how we can give a different explanation of A's mistake. This will help us to get clear on the content of A's false belief. We have already ruled out one line of justification for A's belief on the ground that it attributes to him a thought which, by hypothesis, he cannot have. This effectively limits the justification that A could have for belief to the following line of reasoning:

- (b1) p is the thought I express now with S
- (b2) p was also the thought I expressed last week with S
- (b3) So the thought I express now with S is the same as the thought I expressed last week with S.

On this theory, A's mistake is to be found in (b2), the belief that p was the thought he expressed last week with the sentence "Arthritis was ____". Since q was the thought that was expressed last week and q is a different thought from p , his belief is false. Also, it is clear that nothing that is available to A through introspection alone will reveal this mistake to him. Hence, this account is consistent with both of our initial concessions to the argument, first, that A will answer affirmatively if he relies on introspective evidence alone, and second, that he will be mistaken if he answers affirmatively. However, since no mistake of comparative judgment has been made by A on this account, we may conclude, as before, that the case is not a counterexample to **T'**.

The case described by Falvey and Owens therefore fails to be counterexample to **T'** because A cannot have a thought which contains both *arthritis* and *arthritis** as constituents. But what if someone who is capable of having such a thought (call him B) is substituted for A in the argument? What effect will this have on the argument? First, if B can have thoughts about arthritis when in L*, and about arthritis* when in L, then there is nothing to prevent him from using the sentence “Arthritis is ____” to express the same thought in both L and L*. Therefore it cannot be presumed that B will be mistaken if he believes that the thought he now express is the same as the thought he expressed last week. Second, if B can have a thought which contains both *arthritis* and *arthritis** as constituent, then it would be question-begging to insist that B is incapable of distinguishing the one concept from the other. But if B is capable of making this distinction, then it cannot be presumed that if he were to make a mistake in answering the question, he would not be able to uncover the mistake through introspection alone.

Thus we may conclude that the argument given by Falvey and Owens is powerless against **T'**. For either the subject is in a position to make the required comparative judgment or he isn't. If he is able to make the judgment, then the argument's premises will be undermined. But if he is not able to do so, then his case will have no relevance to **T'**. Either way, the argument collapses.

In response to our criticism, our opponents might try to modify their argument as follows. Suppose that A's community L has a word, say 'tarthritis', which expresses the concept *arthritis**. In L, then, 'arthritis' stands for arthritis and “tarthritis” stands for tarthritis (=arthritis*). However A is unaware of this difference and uses the two words interchangeably. For

instance, just as he would affirm “My arthritis has spread to my thigh”, he would also affirm “My tarthritis has spread to my thigh”, and conversely. Now suppose we ask A whether the thought he associates with “Arthritis is ___” (S) is the same as the thought he associates with “Tarthritis is ___” (S[^]). Presumably he will answer yes. But since S and S[^] do not express the same thought, A will be mistaken. Furthermore, no amount of introspection by A will reveal the mistake to him. Therefore **T'** is false.

But this argument too is inconclusive. In what follows, let *p* be the thought expressed by S, and *q* the thought expressed by S[^], in the language of L. As before, let us assume that A will answer the question affirmatively. From this, we may infer that A believes that the thought he associates with S is the same as the one he associates with S[^]. But in order for the case to be a counterexample to **T'**, A's belief must be false and it must be based upon a judgment of comparative content. However, neither of these requirements is clearly met. To see that this is so, suppose that A has arrived at his belief by reasoning as follows:

- (c1) *q* is the thought I associate with S.
- (c2) *q* is also the thought I associate with S[^].
- (c3) So the thought I associate with S is the same as the thought I associate with S[^].

Three points may be made about this piece of reasoning. First, none of the thoughts involved in it is a judgment of comparative content in the required sense. Second, one cannot deny the reasoning to A on the ground that the concept *arthritis** does not exist in L, for by assumption it does. Nor can one deny it to A on the ground that S and S[^] do have the same meaning in the language of L, for neither (c1) nor (c2) is a claim about what S or S[^]

mean. Rather they are about what A means by these sentences, and clearly the two kinds of claims are independent of each other. It is therefore perfectly possible for both (c1) and (c2) to be true of A. Third, if the reasoning embodied in (c1)-(c3) is sound, then A's belief, as given by his reply, will be true, not false. Since A's belief need not be false in the situation described in the argument, this case, like the previous one, is not a counterexample to **T'**.

Thus far, we have seen no decisive reason for thinking that there is any conflict between externalism¹ and **T'**, let alone the weaker **T**. But there is another form of externalism, one which is driven by a different set of considerations from those which motivated externalism¹. Perhaps, a better case against **T** can be built upon this form of externalism.

II

The first type of argument for externalism is based on the assumption that a subject can have a thought despite lacking a complete understanding of some concept in its content. This assumption is not essential to a second type of argument for externalism, of which the most famous example is undoubtedly Putnam's Twin Earth argument. To briefly summarize, suppose that somewhere in the galaxy is a Twin-Earth that is exactly like Earth except in one respect. The one difference is that the liquid which people on Twin-Earth call 'water' is not H₂O but a different liquid whose chemical formula may be abbreviated as XZY. Despite the difference in their microstructure, H₂O and XZY are indistinguishable in terms of their observable properties. Let A be a scientifically ill-informed English speaker with a faultless understanding of the word 'water', and let B be his replica on Twin-Earth. By hypothesis, A and B have exactly same physical and mental

histories non-intensionally described. Yet when A utters the sentence “Water is wet” and B does the same, they express different beliefs. For A’s belief is true just in case H₂O is wet, whereas B’s belief is true just in case XYZ is wet. It follows that even when a thought is fully grasped by a subject, its identity might not be completely determined by the subject’s intrinsic properties, but might be dependent upon factors which are external to the subject – in this case, the subject’s physical environment. I shall refer to this form of externalism as *externalism2*.

What implication does externalism2 have for **T**? Among philosophers who have considered this question, an overwhelming majority seem to think that we must abandon **T** if we accept externalism2. For convenience, I shall refer to this prevailing view as the *incompatibility thesis*. The central argument for the incompatibility thesis is the slow switching argument. I shall not, however, discuss this argument here as it is similar in structure to the earlier argument by Falvey and Owens and therefore open to the same objections. To establish the incompatibility thesis, an argument would have to show that if externalism2 were true, then no one, no matter how well informed, could know *a priori* that *water* and *twater* (the concept expressed by ‘water’ on Twin Earth) were different concepts. Such a conclusion is clearly beyond the reach of standard versions of the switching argument.

The incompatibility thesis is unwarranted. In the context of this discussion, unless it can be shown that externalism2 entails that

(O) No one can know *a priori* that *water* and *twater* are different concepts

there will be no case for the incompatibility thesis. Since externalism2 is entailed by the Twin Earth argument, any proposition that is derivable from

externalism₂ must be derivable from that argument. But **O** cannot be derived from the premises of the Twin Earth argument. Hence **O** is not entailed by externalism₂ and the incompatibility thesis is unfounded.

To see that the Twin Earth argument does not entail **O**, we need first to state the argument. The basic argument for externalism₂ is as follows:

- (A1) Concept determines extension.
- (A2) The extension of ‘water’ is different on Earth and Twin Earth.
- (A3) So *water* and *twater* are different concepts.
- (A4) Twins (such as A and B) share all their intrinsic properties.
- (A5) So the concept *water* is not fixed by the intrinsic properties of a competent speaker.

If **O** is a consequence of the Twin Earth argument, then it must be derivable from the set {(A1), (A2), (A4)} (plus various common background assumptions). In what follows, let C be an ideal interpreter who has a perfect grasp of the concepts *water* and *twater* but no knowledge of the relevant chemistry, and let

$p = \textit{water}$ and \textit{twater} are different concepts.

$q = \textit{twater}$ is not water.

From (A1) it follows that q entails p . If we assume that C knows this entailment and knowledge is closed under known entailment, we can deduce, by the following steps, that if C knows that q , then C knows that p :

- (B1) If [C knows that q and C knows that q entails p], then C knows that p .

(B2) C knows that q entails p .

(B3) So if C knows that q , then C knows that p .

And from (B3) we may derived

(B4) If C can know *a priori* that q , then C can know *a priori* that p .

But there is no valid inference from (B4) to the proposition that C cannot know *a priori* that p . Thus **O** is not derivable from (A1). Nor is **O** derivable from the other two premises. This needs no argument in the case of (A2), since the premise makes no mention of concepts. This same applies to (A4). To get **O**, (A4) must be combined with the assumption that only concepts that are intrinsically determined can satisfy **T**. But such a move would clearly be question begging, as the assumption is just a re-statement of the incompatibility thesis. Since **O** cannot be derived from the set of premises which yield externalism₂, it cannot be a consequence of externalism₂. The incompatibility thesis is therefore unwarranted.

Why then have philosophers like Putnam and McGinn been so ready to accept the incompatibility thesis? I suspect that part of the explanation has to do with their belief that the only type of theory of meaning that could square with the results of the Twin Earth argument is a multiple component theory. On Putnam's own theory, the meaning of 'water' is envisaged as an ordered-pair consisting of (a) the stereotype - a qualitative description of a normal sample of water, and (b) its extension. If one extends this account of meaning to concepts, then it's fairly easy to construct a valid argument for **O**. Thus let the concept *water* be represented by $\langle s_1, e_1 \rangle$ and the concept *twater* by $\langle s_2, e_2 \rangle$, where s_1 and s_2 are the respective stereotypes, and e_1

and e_2 , the respective extensions. To establish **O**, a preliminary step is to establish that if p then q :

(C1) $water = \langle s_1, e_1 \rangle$ and $twater = \langle s_2, e_2 \rangle$.

(C2) $s_1 = s_2$

(C3) So $water = twater$ if and only if $e_1 = e_2$.

(C4) So if p then q

The next step is to assume that C knows (C4) and that his knowledge is closed under known implication. Then, by a similar argument to (B1)-(B4), we get the result that

(C5) If C can know *a priori* that p , then C can know *a priori* that q .

From (C5), it is a sure step to **O** by modus ponens, given the undeniable fact that C cannot know *a priori* that q .

Although this line of thought may explain why so many philosophers subscribe to the incompatibility thesis, it doesn't justify the thesis because of the unsupported assumption that the two-component theory of meaning comes with externalism². From our earlier discussion, it should be clear that such an assumption is unsustainable. For what the Twin Earth argument shows is no more than this: that the meaning of a NK term is not determined by the intrinsic properties of a person who fully understands the term. From this, it logically follows that the concept that is associated with such a term in the language must have an external determinant. But the proposition that such a concept has an external determinant is weaker than the proposition that the external determinant of such a concept is the extension, which in turn is weaker, being less specific, than the proposition

that the extension is a component of the concept. Thus, we see that the two-component theory is not a consequence of externalism².

To reinforce the conclusion that we have just reached, I will argue that there is an alternative account of NK terms which is consistent with externalism², but which does not imply that terms which are associated with different concepts must have different extensions. This account is suggested by the following remark of Putnam: “[M]y ways of recognizing water (my ‘operational definition’, so to speak) . . . , like the ostensive one, is simply a way of pointing out a standard – pointing out the stuff in the actual world such that for x to be water, in any world, is for x to bear the relation $same_L$ to normal members of the class of entities that satisfies the operational definition. ‘Water’ on Twin Earth is not water, even if it satisfies the operational definition, because it doesn’t bear $same_L$ to the *local* stuff that satisfies the operational definition, and local stuff that satisfies the operational definition but has a different microstructure different from the rest of the local stuff that satisfies the operational definition isn’t water either, because it doesn’t bear $same_L$ to the *normal* examples of the local ‘water’.” (p. 232 [3]) These remarks open up an account of NK concepts which is altogether different from the two-component theory.

The meaning of ‘water’ is a concept – a thing that determines an extension without the aid of contexts. The meaning of an indexical expression, on the other hand, is a function from context to extension. The word ‘water’ cannot therefore be regarded as an abbreviation of a complex indexical expression such as ‘the local stuff that is transparent, etc.’ But although an indexical word type is not coupled with any concept, its tokens can be used to express different concepts in different contexts. This is essentially Frege’s view of indexicals. “With words like ‘here’ and ‘there’,”

Frege observes, “the merely wording, as it is given in writing, is not the complete expression of the thought, but the knowledge of certain accompanying conditions of utterance, which are used as means of expressing the thought, are needed for its correct apprehension.” (p 94 [7]) Thus the same sentence containing the word ‘here’ will express different thoughts when uttered at different places, of which some may be true, others false. What is expressed by the word ‘here’ in these various utterances must, accordingly, be different. (For a forceful defence of Frege’s theory, see Gareth Evans’s **Understanding Demonstratives**).

We are now in a position to state *the indexical theory* of NK terms. On this theory, our concept of water would be a condition that could only be expressed by a clause containing an indexical expression, such as the following:

(**W**) $(\forall L)$ (“water” is true of L iff L bears the same liquid relation as the local stuff which is transparent, colourless, odourless, etc.)

When we use **W** to fix the extension of “water”, we express a condition K1 that is satisfied by a liquid if and only if the liquid has the chemical structure of water. Thus K1 applies to water but not to twater. Note that this fact is independent of contexts. Any expression that has K1 as its associated concept will be true of water but not of twater, regardless of where the expression is uttered. K1 is therefore not a function from contexts to reference, but a complete concept. In contrast, when **W** is used by Twin Earth speakers to fix the extension of their word ‘water’, what they express is a condition K2 that is satisfied by twater but not water. It follows that K1 is not the same concept as K2. A plausible thought is that K1 is the concept *water* and K2 is the concept *twater*. Thus, like the two-component theory, this

theory has the welcomed consequence that *water* and *twater* are different concepts.

However, there are two crucial differences between the two theories. The first is that the indexical theory, unlike the two-component theory, does not carry the assumption that C could not know that *water* and *twater* were different concepts without knowing that they had different extensions. To see this, consider a parallel case: two utterances of “The local water is hard”, one made in London and the other in Reading. In each context, “the local water” means: *the total sample of water around here*. Hence, the two utterances do not express the same thought, for one might be true and the other false. Given what the term “the local water” means in each context, its extension will not be the same in the two utterances. But it is nevertheless possible for the water in one total sample to stand in the “same liquid” relation to the water in the other total sample. Hence, the fact that a different thought is expressed by each of the two utterances is compatible with the supposition that the local water in London and the local water in Reading are of the same kind of stuff.

The same consideration can be made (with the necessary enlargement of the scope of “local”) with respect to two utterances of **W**, one on Earth and the other on Twin Earth. In each case, the reference of “local” is part of the expression of the condition that is expressed by the utterance. Since the reference of “local” is different in each utterance, the condition expressed by **W** on Earth is different from the condition expressed by **W** on Twin Earth. However, given the form of **W**, the extensions of “water” and “twater” will not be particular samples of stuff, but kinds of stuff, so that a liquid can be water (or twater) even if it is not located on Earth (or Twin Earth). But the crucial point is this: on this account, even if the extension of “water” and

“twater” were the same, the meaning of the two words would still be different. The difference between *water* and *twater* therefore need not be tied to any fact about the real nature of twater and water. But if the difference in the two concepts may be independent of their extensions, then there is no conclusive ground for the claim that C could not know that *twater* and *water* are different concepts without knowing that twater is not water.

The second difference between the two accounts of meaning is this. On the two-component theory, the concept linked to the word “water” is not fully known to a competent speaker, for a component of the concept – the extension – is not part of this knowledge. On the indexical theory, the concept *water* - the condition expressed by tokens of an indexical sentence like **W** on Earth – is fully grasped by a competent speaker. This has an important consequence for our present concern. Let R1 and R2 be the references of “local” in an utterance of **W** on Earth and an utterance of **W** on Twin Earth, respectively. Now it is true that C would not know that *water* and *twater* were different concepts if he did not know that R1 was not the same as R2, and C could not know that R1 was different from R2 without empirical investigation. But it does not follow from this that C could not know *a priori* that *water* and *twater* were different concepts. This is because, on this account, no one could grasp the concepts *water* and *twater* unless he knew that R1 was not the same as R2. Hence although this piece of knowledge is *a posteriori*, it is not extraneous to C’s understanding of the question “Is *water* the same concept as *twater*?” We see once again that externalism₂ does not support **O**.

I would like to end this section by showing how C can know *a priori* that *water* and *twater* are different concepts. This will strengthen our contention that **O** is logically independent of externalism₂. One way in

which C can arrive at the knowledge that *water* and *twater* are distinct concepts is by reasoning along the lines of the following argument:

- (D1) *Water* and *twater* are the same concept just in case the ‘water’ on Earth is synonymous with “water” on Twin Earth.
- (D2) If ‘water’ on Earth is synonymous with ‘water’ on Twin Earth, then it will be an analytic truth that *twater* is water.
- (D3) If it is an analytic truth that *twater* is water, then it will be knowable *a priori* that *twater* is water.
- (D4) It is not knowable *a priori* that *twater* is water.
- (D5) Hence ‘water’ on Earth is not synonymous with ‘water’ on Twin-Earth.
- (D6) Hence *water* and *twater* are different concepts.

If this argument is sound, then C knows that *p* (D6), since he has arrived at it by correct reasoning. But since none of the premises requires C to know that “water” has a different extension on Earth and Twin Earth, it follows that C could know that *p* without knowing that *q*. Also if the argument’s premises are all knowable *a priori*, then C’s knowledge of *p* will be *a priori*. I believe that both of these requirements are met by the argument (D1-D6). Moreover, none of its premises are in obvious conflict with externalism₂. This shows, once again, that externalism₂ does not entail the claim that C cannot know *a priori* that *water* and *twater* are different concepts.

To sum up, the incompatibility thesis is warranted only if the claim that no one can know *a priori* that *water* and *twater* are different concepts is entailed by externalism₂. But externalism₂ does not have this consequence. There might be versions of externalism, such as one which incorporates a two component conception of meaning, which are incompatible with **T**, but

these are not obtainable from the existing arguments for externalism. Therefore we may conclude that the incompatibility thesis, though widely believed, is groundless.

III

In **The Meaning of “Meaning”**, Putnam maintains that the traditional theory of meaning – “the theory that (1) words have ‘intensions’ which are something like concepts associated with words by speakers, and that (2) intension determines extension – cannot be true of natural-kind words like ‘water.’” (p. 234 [3]) I believe that Putnam did not succeed in establishing this claim in his paper. To support my case, I rely on the results of sections (I) and (II).

The basic argument for Putnam’s claim can be put in the form of a dilemma:

- (A1) Either the concept *water* is “in the head” or it is not.
- (A2) If it is “in the head”, then it cannot determine extension.
- (A3) If it is not “in the head”, then it cannot be what a C knows when he understands the word ‘water’.
- (A4) So either the concept *water* cannot determine extension or it cannot be what C knows when he understands the word ‘water’.

There are several interpretations of the term “in the head’ which are consistent with Putnam’s intentions in **The Meaning of “Meaning”** and elsewhere. For present purposes, we need only consider the following two:

- (a) An item is in the head just in case it is fully determined by the subject's internal properties.
- (b) An item is in the head just in case it is transparent (in the sense required by **T**).

From our earlier discussion, it should be clear that (a) and (b) are not logically equivalent. Let's see how the argument fair under each interpretation.

If we interpret "in the head" in accordance with (a), then (A2) is true (assuming externalism). But (A3) is not true, for from the assumption that *water* is not wholly determined by the intrinsic properties of a speaker, it does not follow that *water* is not completely grasped by such a speaker when he understands the word 'water'. To motivate this result, one will have to assume that the chemical facts about water are somehow involved in the concept *water*, but this assumption has been shown to be gratuitous. On the other hand, if we interpret "in the head" in accordance with (b), then (A3) is true (assuming **T**). But (A2) is now false, for as we have shown, *water* need not fail to be transparent if satisfies externalism₂, and if *water* satisfies externalism₂, then it does determine extension. Hence, it appears that whichever interpretation is put on the expression "in the head", Putnam's argument against the traditional theory of meaning is unsound.

Singapore Management University
469 Bukit Timah Road
Singapore 259756
Email: ygtan@smu.edu.sg

References

- (1) Tyler Burge, "Individualism and the Mental", *Midwest Studies in Philosophy* Vol IV, Minneapolis, P. French *et al.* (eds), MN: University of Minnesota Press.
- (2) John McDowell and Philip Pettit (eds), *Subject, Thought and Context*, (Oxford University Press, 1986)
- (3) Hilary Putnam, "The Meaning of 'Meaning'", in *Mind, Language and Reality*, (Cambridge University Press, 1984)
- (4) Hilary Putnam, *Representation and Reality*, (MIT Press 1991)
- (5) Kevin Falvey & Joseph Owens, "Externalism, Self-Knowledge, and Scepticism", *Philosophical Review*, volume 103, number 1, pp. 107-137 (January 1994).
- (6) Colin McGinn, *Mental Content*, (Blackwell 1989).
- (7) G. Frege, "The Thought" in *Philosophical Logic*, P. F. Strawson (ed), (Oxford University Press 1967)