

Externalism and memory: don't forget it!

Frank Hofmann

(unpublished manuscript, 1995)

I would like to pose a certain problem about self-knowledge and memory for social externalism. The problem arises from the possibility of slow-switching, and it has the form of a dilemma. The externalist has to choose, as to the content determining external conditions, between initial and recollection circumstances. However, no matter how the externalist chooses, it will turn out that very implausible, even absurd, consequences will follow. Either, memory turns into an absurd faculty, 'pseudo memory'; or, the other horn of the dilemma, discriminating self-knowledge fails. In posing the dilemma, I take up and extend certain arguments given by Paul Boghossian and Peter Ludlow. I hope I can put things together, and add some more aspects, so that social externalism can no longer be saved in any way. Finally, I suggest that the argument not only hits social externalism but also non-social or environmental externalism.

For now, I will restrict the discussion to social externalism, since Ludlow only considers this version of externalism. Only at the end I will extend the argument to non-social, micro-essentialist externalism.

Recently, Peter Ludlow (1995a) has made an attack on Boghossian's argument against externalism (cf. Boghossian 1994). In effect, Ludlow thereby attempts to save social externalism. This he wants to accomplish by having the social externalist make a certain choice. The choice concerns the circumstances which determine the content of a person's, say Peter's, memories. It comes up in the following way.

According to externalism, Ludlow assumes, it is natural to hold that not only the contents of thoughts in general are determined by external conditions, but also, in particular, the contents of our memories. If, as social externalists claim, Peter's mental states get their contents fixed by his social group(s), then his memories should get likewise. For example, at some earlier time, say t_1 , Peter is unaware of the individuating conditions of 'arugula'. He only knows that it is a bitter leafy green vegetable, and, therefore, he defers to his linguistic community at t_1 for those individuating conditions. At some later time, say t_2 , Peter has joined some other linguistic community having different individuating conditions for 'arugula', without his being aware of this slow-switching. At t_2 , Peter recollects some thoughts about arugula. Now, how about the content determining conditions for memory? Is it fixed by the initial circumstances of t_1 , or by those of the later recollection time t_2 ? My claim is that either choice will lead the social externalist into serious trouble.

Ludlow recommends the latter choice. He thinks:

“The consistent social externalist is bound to say that the content of a memory is fixed at the time recollection takes place – for it is the embedding circumstances of that memory which are crucial to the fixing of its content.” (Ludlow 1995a, p. 158)

He seems to have some ‘homogeneity argument’ in mind. If current circumstances fix content for thoughts in general, so they should for memory content likewise. Otherwise, memory-thoughts were somehow ‘frozen out’ and, thus, exempted from the externalist mechanism.

Taking this choice, Ludlow can save some version of self-knowledge (even though only non-discriminatory self-knowledge). For, at any particular time, Peter does have knowledge of his thoughts and memories. He knows what he recollects since the recollected thoughts have the same individuating conditions as all thoughts at recollection time. So Ludlow concludes, against Boghossian,

“Even if I am informed that the content of ‘arugula’ has changed because of changes in my social environment (in effect, that I have become an unwitting victim of slow-switching) it does nothing to undermine my self-knowledge. The changes need not concern me, for the content of my memory is determined by *current* environmental conditions.” (Ludlow 1995a, p. 159, emphasis in original)

However, there is some other serious problem coming up. For, if recollection circumstances determine memory content, then memory turns into an empty, absurd faculty. This is so, since memory can no longer do what it is supposed to do, namely, to recall the very same thoughts one earlier on had entertained. If Peter had come to believe at t1 somehow that arugula only grows in Mediterranean climate, and now, at t2, recalls this thought, then what he recollects will not be the same thought about arugula, but some other thought about, say, tarugula – according to the newly adopted individuating conditions. Peter will no longer be able to have any arugula thoughts which he could express by using the word ‘arugula’, since that word no longer refers to arugula; Peter would have to acquire some new word (or undergo a re-shift) first in order to be able to have thoughts about arugula. Notice that all this happens without Peter having forgotten anything. Furthermore, and equally embarrassing, the truth values of the memories will change. So, for example, Peter’s belief at t1 that arugula only grows in Mediterranean climate will now, when recollected, turn into a falsehood. (I assume that at t1 the belief was true. This I wish to do, since I didn’t know ‘arugula’ when I first read the papers, and so I just make up the story in this way.) For, at time t2 it will be a thought about tarugula which is a vegetable (let’s assume) that grows only in tropical climates. Even worse, if Peter recollects that he had had, at t1, some arugula experience, then also this memory has turned by now into a falsehood, since it has become a memory-thought about tarugula which (let’s assume) Peter never has had any direct encounter with. And all of this has occurred without Peter being able to introspectively become aware of it. Memory, as social externalism will have it, has ‘turned pseudo’. It is no longer a source of knowledge.

So, rejecting Boghossian’s principle

(1) If S forgets nothing, then what S knows at t1, S knows at t2.

amounts to a pretty strange sort of memory. Note that Ludlow does agree that Peter has nothing newly learned nor forgotten anything about arugula. Nevertheless, his earlier knowledge is unnoticedly and (up to some further change) in principle lost. And, as Ludlow has pointed out (Ludlow 1995b), slow-switching, within the externalist conception, is not just a logical possibility, but can occur rather frequently, at least for social externalism. So it is fair to say that memory has a *systematic* malfunction built into which disqualifies it as a source of knowledge.

To illustrate the difficulty further, it seems to me that the externalist conception is also not compatible with a certain view of concepts. That is, let us think of thoughts as the exercising of (conceptual) abilities. So assume

(TC) Having a thought is the employment of concepts; since concepts are capacities, thought is the exercising of (conceptual) capacities.

Then, according to the externalist, the content of my memory-thoughts have also switched (and therefore I misrepresent my own earlier thoughts). So my concepts, my conceptual abilities have changed. However, not only have I not (subjectively) realized that they have changed, but also they have changed without there (objectively) being anything that as a cause has brought about some causal *effect* in me, some change in my thought instrument, my brainpower. Nothing has had any effect on me. Isn't it weird to assume that my conceptual capacities have changed without there being any such effects? Especially if one holds that

(EDC) My capacities are a product of my interaction with my environment, are a common or final effect of certain causes. That is – in causal terms – what it means that I have acquired them, that I have learned to think.

(EDC), at the same time, is the proper sense in which 'causes determine meaning' holds. 'Determine' has to be read as 'bring about as an effect'. Content is determined by permanent effects.

Now, consider the other horn of the dilemma. Assume that the *initial* circumstances fix the content of memories. This is the alternative which Boghossian's intuition leads him to take (cf. Boghossian 1994, p. 39). Then, indeed, we avoid the absurdity just laid out; we have true memory which really 'freezes thoughts'. However, in this case we get into exactly the conflict with self-knowledge which Ludlow wanted to avoid by choosing the other alternative. Namely, as for example Boghossian has pointed out, self-knowledge will no longer be epistemically transparent. That is, in some thoughts 'arugula' will have arugula as its reference, and in others it will have tarugula, but Peter will by no means be able to tell from inside. There will be no discriminating self-knowledge – which is the form of self-knowledge we normally think we have (cf. Boghossian 1994, p. 35). True memory is saved only at the price of sacrificing self-knowledge.

An illustration of the difficulty would be the following case. Assume Peter says 'I recollect that I thought that arugula grows in Mediterranean climate' and, at the same time, he comes to think that somebody else, say Sabine, has now the same thought he earlier on had, which he expresses by saying 'Look, Sabine also thinks that arugula grows in Mediterranean climate'. Then, the content of the 'object' thought he self-

ascribes is different from the one he ascribes to Sabine. Thus, Peter's interpretative capacity becomes split. Peter does not realize that he ascribes two different contents.

One could maybe think of saving self-knowledge by introducing the idea that memory-thoughts get a 'special marker'. All memory-thoughts are thus (introspectively) distinguished from non-memory-thoughts, and Peter therefore can know apriori that his memory of the arugula encounter is not one of tarugula. However, multiple iteration of slow-switching will undermine that possibility. For after another slow-switching to the 'third earth', Peter will refer by 'arugula' to tharugula, and ... on 'n-th earth' to n-arugula. So he would have to have (n-1) different memory markers in order to be able to (introspectively) distinguish between (n-1) different meanings of 'arugula' corresponding to (n-1) different 'earths'. Otherwise, authority of content ascription would also fail for memories stemming from different 'earths'. But that sounds not very plausible. For, do I know when (on which 'earths') I first learned that Cicero was a Roman orator? There certainly are no such memory markers. Memory-thoughts are not that different from newly acquired thoughts.

Finally, I believe that argument extends also to other forms of externalism, non-social or micro-essentialist externalism. According to Putnam, for example, reference is fixed by the microstructure of indexically identified samples. On earth 'water' refers to H₂O, on twin-earth to XYZ. The argument given above also affects this version of externalism, since it did not depend on the specific mechanism by which content is externally determined. It does not matter if content is externally fixed by one's social group(s), or by the microstructure of the stuff in one's environment. All that counts is the possibility of there being switches in the content determining external conditions. As soon as such switches are considered possible, the dilemma shows up. Either Peter's memories turn into falsehoods (by becoming XYZ-thoughts), or he does not have discriminating self-knowledge. So non-social externalism is no better off than social externalism.

References:

- Boghossian, P. (1994), "The transparency of mental content", *Philosophical Perspectives* 8, 33-50.
- Ludlow, P. (1995a), "Social externalism, self-knowledge, and memory", *Analysis* 55:3, 157-159.
- Ludlow, P. (1995b), "Externalism, self-knowledge, and the prevalence of slow-switching", *Analysis* 55:1, 45-49.