KNOWING WHAT THOUGHT REQUIRES

Part 2: Externalist Dependance Thesis

MARTIN DAVIES

Introduction: Externalism, first-person authority, and armchair knowledge

In the first lecture, I developed an *architecturalist* claim: thought requires a particular kind of internal cognitive machinery. I showed how this claim, when combined with an independently plausible claim about self-knowledge, leads to an instance of *the problem of armchair knowledge*. In this lecture, I show how another instance of the problem of armchair knowledge arises from *externalist* arguments about the contents of mental states.

Externalism about some mental property, M, is the thesis that whether a person (or other physical being) has M depends, not only on conditions inside the person's skin, but also on the person's environment and the way that the person is embedded in that environment. The dependence here is supposed to be conceptual rather than causal; it is the kind of dependence that can be revealed by philosophical theorising. This is an armchair methodology; so, if philosophical theorising yields knowledge, then it is a kind of armchair knowledge. Its status as knowledge does not depend on our conducting any detailed empirical investigation of the world around us. This lecture is about the puzzle that arises when the possibility of armchair knowledge of an *externalist dependence thesis* about mental property M is put together with a thesis of first-person authority for that same mental property.

For the purposes of generating our puzzle, we do not need to be concerned about a precise formulation of the notion of first-person authority. All that we need to suppose is that we each have a distinctively first-personal way of knowing that we ourselves have property M, when we do have it, without needing to conduct any detailed empirical investigation of the environment and our relation to it. This distinctively first-personal knowledge, self-knowledge, is another kind of armchair knowledge.

The puzzle that arises from combining externalism and self-knowledge

is clearly visible when we consider the epistemic status of the premises and conclusion of *modus ponens* inferences of the following kind:

EXT(1)   I am thinking that *p*.

EXT(2)   If I am thinking that *p* then E(me).

Therefore:

EXT(3)   E(me).

('E(*x*)' is some statement about *x*'s environmental embedding.) The occurrence of the particular mental verb 'think' in EXT(1) and EXT(2) is not vital; we could just as well have 'I believe that *p*'. All that is vital is that the mental property of thinking or believing that *p* should meet two conditions. First, it should be a property that is subject to first-person authority. In that case, I can have armchair knowledge of EXT(1). Second, it should be a property for which an externalist dependence thesis:

Necessarily("*x*)(If *x* is thinking that *p* then E(*x*))

holds. In that case, I can have armchair knowledge of EXT(2). Philosophical theorising will underwrite the conditional claim about me in particular by supporting the modal generalisation from which it follows.

Provided that the mental property of thinking or believing that *p* meets these two conditions, I can have armchair knowledge of the premises EXT(1) and EXT(2). In neither case does my knowledge depend for its status as knowledge on an empirical investigation of my environment. Given armchair knowledge of those two premises, it is trivially easy to perform the *modus ponens* inference and so, it seems, to arrive at armchair knowledge of the conclusion EXT(3). But this creates a puzzle since EXT(3) is a statement about my environment and my way of being embedded in it, and we normally expect knowledge about such matters to depend on empirical investigation.

1. The general form of the problem of armchair knowledge

In my view, the puzzle that arises from the combination of externalism and self-knowledge is not an isolated phenomenon. Rather, it is one of a number of puzzles that involve arguments of the more general form (MC):

MC(1)   I have mental property M.

MC(2)   If I have mental property M then I meet condition C.

Therefore:

MC(3)   I meet condition C.

Puzzles that are similar to the one under discussion in this lecture arise from philosophical theorising about the inter-level relationship between descriptions of ourselves as conscious, thinking, language-using persons and descriptions of human animals as information-processing systems (Davies, 2000a). For example, as we saw in the first lecture, there are philosophical theories according to which the cognitive architectural necessary conditions for thought and concept deployment include the truth of the language of thought hypothesis.[20] Other philosophical theories have the consequence that the objectivity of linguistic meaning requires the presence of tacit knowledge of a compositional semantic theory.[21] I know from the armchair that I think many things and deploy many concepts, and I know from the armchair what many hitherto unconsidered sentences of my own language mean. But I do not have armchair knowledge about my own cognitive architecture. I cannot know from the armchair that the language of thought hypothesis is true of me; nor that my language processing system embodies tacit knowledge of a compositional semantic theory. These are matters for empirical investigation.

Whether we consider these cases of 'down' or this lecture's case of 'out', the general form of the problem of armchair knowledge is this. Armchair philosophical theorising supports an inference from A to B. A can be known from the armchair; B cannot be known without detailed empirical investigation of the world. So a thinker can know that A without detailed empirical investigation of the world, and can likewise know that if A then B.

But it is obvious that B follows; no empirical investigation is needed to see that. So the thinker's knowledge that A and knowledge that if A then B, and the thinker's ability knowledgeably to draw the obvious inference, seem together to provide a route to knowledge that B – and still without any need to rise from the armchair. Yet B was supposed to be something that could only be known by way of detailed empirical investigation of the world. Knowledge that B was supposed to require an investigative, rather than an armchair, methodology. If we are to provide a general solution to the problem of armchair knowledge, then we have to accept that, sometimes, knowing that A and knowing that if A then B, and then knowledgeably drawing the conclusion that B, does not constitute a route to knowledge that B. Sometimes, that is to say, the epistemic warrant or justification that we have for the premises of an argument is not transmitted to the conclusion, even though it is obvious that the premises entail the conclusion.

The motivation for a solution of this kind must wait until the third lecture. In the remainder of this introductory section, I shall make four general comments by way of clarifying both the problem of armchair knowledge and the form of its proposed solution.

1.   The first comment is simply that the issue is not one about validity of arguments. Arguments that seem to give rise to the problem of armchair knowledge, such as arguments of the (MC) form, are palpably valid instances of *modus ponens*. Our problem is not logical but epistemological.

2.   The second general comment is that the issue is not exactly one about closure of knowledge under known entailment. It is true that, if knowledge is not closed under known entailment, then transmission of epistemic warrant must be limited in some way. But the converse is not true in general. In a case where warrant cannot be transmitted from the premises of a valid argument to its conclusion, it might still be that anyone who knows the premises also knows, or can come to know, the conclusion. It might be, for example, that anyone who knows the premises inevitably has an independent warrant

for the conclusion. Or it might be that knowledge of the conclusion does not require a warrant.

3.   The third general comment is that the issue is not one about confidence or subjective probability. To propose a restriction on transmission of epistemic warrant is not to suggest that rational subjects who attach a high degree of confidence to the proposition that A and the proposition that if A then B, should attach a substantially lower degree of confidence to the proposition that B.

There are, of course, cases in which it is rational to be much less confident about the conclusion of a *modus ponens* inference than about either of its premises. The conclusion of a valid argument is no less probable than the conjunction of its premises; but the probability of the conjunction of the premises may be substantially less than the probability of each of the premises taken individually. In the case of *modus ponens* arguments, this can happen if the probability of the conditional premise (if A then B) is high while the conditional probability of B given A is low. In such a case, the conditional is said to be not robust with respect to its antecedent (Jackson, 1987). But the *modus ponens* argument of the (EXT) form is not like this, and we can suppose that if the premises deserve a high degree of confidence individually then they also deserve a high degree of confidence together.

If a thinker is confident that A and confident that if A then B and also confident of their conjunction then she should be no less confident that B. In this sense, believing the premises (together) gives the thinker a reason to believe the conclusion. A thinker who believes that B under these circumstances is not subject to any criticism for doing so; she is not being doxastically reckless or irresponsible. But someone who proposes a restriction on transmission of epistemic warrant says that, even so, epistemically adequate warrants for believing the premises may not themselves add up to an epistemically adequate warrant for believing the conclusion.

4.   The fourth general comment is that, in offering putative cases

where warrant is not transmitted from premises to conclusion, we must take care to avoid equivocation on key notions that occur in the proposition that A, as between the first premise and the conditional premise. Thus, for example, we must take care to avoid equivocation on the phrase 'think that *p*' as between EXT(1) and EXT(2). Otherwise we invite a very natural response to the puzzle. Someone may say that it is only in a thin sense of the phrase that I have armchair knowledge that I am thinking that *p*, while it is only in a thick sense of the phrase that philosophical theory can support an externalist dependence thesis.[22]

2. Externalism about content

I said at the beginning of this lecture that externalism is a thesis about dependence: whether a person has mental property M depends, at least in part, on the person's environment and the way that the person is embedded in that environment. It is useful to clarify the relationship between externalist dependence theses such as:

Necessarily("*x*) (If *x* is thinking that *p* then E(*x*))

and what I have elsewhere called constitutive and modal externalist theses.[23]

Constitutive externalism (as it concerns mental property M) says that the fundamental philosophical account of what it is for an individual to have M needs to advert to the individual's physical or social environment. *Modal externalism* says that there are Twin Earth examples for M. That is, according to modal externalism, there are two possible situations w_1 and w_2, differing in environmental conditions, such that an individual, *a*, has M in w_1 but a duplicate individual, *b*, lacks M in w_2. According to modal externalism, having M is not a locally supervenient matter. That is, it does not depend only on what is going on inside the individual's skin.

Suppose that an externalist dependence thesis holds for mental property M:

Necessarily("*x*) (If M(*x*) then E(*x*)).

Then we expect to be able to generate a Twin Earth example for M. To do so, we consider an individual, *a*, who has M in w_1 and therefore meets condition E in w_1 and a duplicate individual, *b*, who fails to meet E in w_2 and therefore lacks M in w_2. Thus, unless it is somehow impossible to have a duplicate of *a* in an environment that fails to meet E, modal (or Twin Earth) externalism goes along with the truth of an externalist dependence thesis. Furthermore, if an externalist dependence thesis for M is supported by philosophical theorising then it seems to be virtually guaranteed that constitutive externalism will also hold for M. So if an externalist dependence thesis holds for mental property M then it is plausible that both constitutive and modal externalism also hold for M.

If M is a constitutively, but not modally, externalist mental property, then we should not expect philosophical theorising to support an externalist dependence thesis for M. But suppose that M is modally externalist. Then it may be tempting to think that it will be a very short step from a Twin Earth example for M to a dependence thesis for M. In a Twin Earth example, *a* has M in w_1 and *b* lacks M in w_2 in virtue of a specific difference in environmental conditions between w_1 and w_2. So, it may be said, whether any individual has M depends on environmental conditions being specifically as they are in w_1 rather than as they are in w_2. But, in fact, it can be quite difficult to derive a specific dependence thesis from a Twin Earth example. To see this, we only have to consider the basic shape of Twin Earth examples that are used to support social externalism (Burge, 1979).

Alf is thinking that arthritis is painful, while TwinAlf is not. Alf has the concept of arthritis, though his understanding of it is incomplete; TwinAlf does not have the concept of arthritis at all. The crucial difference between Alf's social environment and TwinAlf's relates to the linguistic practices of other people. Alf is surrounded by people who use the word 'arthritis' in such-and-such a way while TwinAlf is not. A convincing Twin Earth example

of this kind would establish that having the concept of arthritis or thinking that arthritis is painful is not a locally supervenient matter. The example would establish this by highlighting the difference between meeting and not meeting a social environmental condition. But a Twin Earth example would not motivate the social externalist dependence thesis:

Necessarily("x) (If x is thinking that arthritis is painful then x is surrounded by people who use the word 'arthritis' in such-and-such a way).

For all that the Twin Earth example shows, it may be that particular features of Alf's internal makeup, or particular features of Alf's non-social environment, make Alf peculiarly dependent on the standard-setting role of the linguistic community around him.

If Alf's internal constitution and his non-social environment were the same, but the surrounding community did not play that standard-setting role, then he would not count as having the concept of arthritis. In the Twin Earth example, TwinAlf, who shares both internal constitution and non-social environment with Alf, does indeed lack that concept. But it hardly follows that any subject whose linguistic community fails to play the standard-setting role for the concept of arthritis must lack that concept. It clearly does not follow, for example, that someone internally unlike Alf, or someone in a different non-social environment from Alf's, would inevitably be without the concept of arthritis unless he enjoyed the support of a standard-setting linguistic community. Nor does it follow from the Twin Earth example that *I* am dependent on *my* linguistic community for the concept of arthritis in the way that Alf is dependent on his.

Let us now consider the basic shape of Twin Earth examples that are used to support externalism about thoughts involving natural kind concepts (Putnam, 1975). Oscar is thinking that water is wet, while TwinOscar is not. Oscar has the concept of water; TwinOscar lacks that concept. The crucial difference between Oscar's situation and TwinOscar's is that the stuff in Oscar's environment (water aka H_2O) is different from the stuff in TwinOscar's

environment (twater aka XYZ). A convincing Twin Earth example of this kind would establish that having the concept of water or thinking that water is wet is not a locally supervenient matter. It would do so by highlighting the difference between meeting and not meeting a physical environmental condition, namely, being surrounded by samples of water. But it would be wrong to think that a Twin Earth example would work by motivating a specific externalist dependence thesis such as:

Necessarily("x) (If x is thinking that water is wet then x is surrounded by samples of water).

For all that the example of the Oscar/TwinOscar duplicate pair shows, there could be someone else, also in TwinOscar's waterless environment but not a duplicate of Oscar, who did have the concept of water. For perhaps there is something about Oscar's internal makeup that makes his having the concept of water peculiarly dependent on his environmental relations. All the more so, there could be someone in an environment quite different from Oscar's or TwinOscar's but still with twater and no water, who had the concept of water. There could be, for all that the Twin Earth example shows.

The point of these last few paragraphs has been that it is not generally possible to read off an externalist dependence thesis from the difference between worlds w_1 and w_2 in a Twin Earth example. A Twin Earth thought experiment provides a counterexample to a claim of local supervenience and so licenses an existentially quantified statement about worlds and individuals. But an externalist dependence claim involves universal quantification over individuals and universal quantification over worlds. So it is unsurprising that there should be a gap between Twin Earth examples and externalist dependence theses.

3. Externalism and self-knowledge

We have seen that it is far from straightforward to motivate a specific externalist dependence thesis even when modal externalism is true. So con-

sider a philosopher who holds that modal externalism is true for the mental properties of thinking thought contents that involve natural kind concepts. This externalist philosopher may still deny that externalist dependence theses about natural kind thoughts lead to the problem of armchair knowledge. Indeed, even a philosopher who accepts both modal and constitutive externalism for a wide range of mental properties could maintain that there are no true externalist dependence theses at all that give rise to the problem of armchair knowledge. I cannot show that such a position is incorrect. But my own view is that problematic externalist dependence theses cannot be avoided and that, in any case, the problem of armchair knowledge arises in many other areas of philosophical theorising. So I want to pursue the question of how we might respond to the problem of armchair knowledge as it arises in the case of externalism and self-knowledge.

3.1 Externalism and natural kind concepts

Let us suppose that it is possible to motivate some such dependence thesis as:

WaterDep Necessarily("*x*) (If *x* is thinking that water is wet then *x* is (or has been) embedded in such-and-such ways in an environment that contains samples of water).

Given that assumption, we can consider the following argument in which the conditional premise follows from WaterDep:

WATER(1) I am thinking that water is wet.

WATER(2) If I am thinking that water is wet then I am (or have been) embedded in an environment that contains samples of water.

Therefore:

WATER(3) I am (or have been) embedded in an environment that contains samples of water.

It is worth noting two points about this example.

First, the conditional premise, WATER(2), and the conclusion, WATER(3), speak of an environment that contains samples of water rather than of an environment that contains samples of H_2O. The difference does not matter for the truth of the conditional premise; but it is important for the premise's epistemic status. It would be implausible that armchair philosophical theorising could deliver knowledge of a conditional premise that mentioned water in the antecedent and H_2O in the consequent, without drawing on empirical investigations to support the identity claim that water is H_2O. Speaking of samples of water rather than samples of H_2O protects the status of the conditional premise as a piece of armchair knowledge. But this does not make the conclusion something that should, intuitively, be knowable from the armchair. I can and do know that my environment contains samples of water; but that knowledge is achieved by empirical investigation.

The second point to note is that thinking that water is wet is not to be equated with thinking something along the lines of: 'the chemical kind that exists in my actual environment and which falls from clouds, flows in rivers, is drinkable, colourless, odourless etc. is wet'. This is important if we are to avoid a charge of equivocation. (Recall the fourth general comment in Section 1 above.)

It is plausible that if I were thinking a thought with the content 'the chemical kind that exists in my actual environment and which . . . etc.' then I could know in the special first-personal way that I was doing so. So the first premise, so interpreted, would be a piece of armchair knowledge. But, on this interpretation, the conditional premise is not something knowable from the philosopher's armchair; indeed, it need not even be true. Thinking a thought with that content does not require the thinker to be in an environment that contains samples of water. For a thinker can deploy in thought the description 'the chemical kind that exists in my actual environment and which . . . etc.' even though no chemical kind fits that description.

We could patch the conditional premise to make it plausibly true and knowable by way of philosophical theorising if we added a conjunct to the antecedent of the conditional and had it say:

If I am thinking that the chemical kind that exists in my actual environment and which . . . etc. is wet *and some chemical kind fits the description that I deploy in thought* then I am embedded in such-and-such ways in an environment that contains samples of water.

But then, to preserve the validity of the argument, we would need to patch the first premise as well and have it say:

I am thinking that the chemical kind that exists in my actual environment and which . . . etc. is wet *and some chemical kind fits the description that I deploy in thought*.

Of course, with the second conjunct added, this is no longer something that can be known in the special first-personal way.

In short, if thinking that water is wet were conceived as deploying a definite description in thought, then we would be open to a charge of equivocation between the first premise and the antecedent of the conditional premise. Without the equivocation, it would not be plausible that both premises could be known from the armchair. So the argument (WATER) would fail as an example of the problem of armchair knowledge. If we are to use (WATER) as an example then we must not regard the thought that water is wet as involving a definite description ('the chemical kind . . . etc.'). It is far from obvious what is the right way to conceive of thoughts about natural kinds, but provisionally we can suppose that the thought that water is wet is about water in somewhat the same way that so-called recognition-based thoughts are about their objects.[24] Against this presumed background, we can summarise our epistemological commentary on (WATER) as follows.

I can know the contents of my own thoughts in the special first-personal way; so I can have armchair knowledge of the first premise, WATER(1).

Also, we are supposing, externalist philosophical theorising yields armchair knowledge of the conditional premise, WATER(2). But, while the conclusion, WATER(3), is something that might well be known without any great difficulty, it seems to fall outside the scope of armchair knowledge; some empirical investigation is required.

3.2 Two problems for first-person authority given externalism

The idea that there is some tension between externalism about content and first-person authority has been developed in more than one way. The problem posed by arguments like (WATER), or more generally by arguments of the (EXT) form, might be called the *consequence problem* for first-person authority given externalism. It should be distinguished from a different problem, which we can call the *achievement problem* for first-person authority given externalism. How can I achieve an especially authoritative kind of knowledge about my own mental states, given that my being in those mental states depends on my environmental relational properties? For I am not, in general, especially authoritative about such properties.

Approaches to the achievement problem typically make some use of the fact that the content of my second-order thought that I am thinking that water is wet, for example, embeds the content of my first-order thought that water is wet. So the content of the second-order thought is dependent on the environment in just the same way as the content of the first-order thought.[25] This does not yet explain how it is that my second-order judgement amounts to knowledge. We still need a general account of how authoritative self-knowledge is possible.[26] But, according to these approaches to the achievement problem, the fact about embedding can be used to show that no special or additional problem for the achievement of self-knowledge flows from the fact that the content of a first-order thought is, in part, externalistically determined.

An account of how self-knowledge is possible has to show how a sec-

ond-order judgement that I am thinking that water is wet, made on the basis of my (first-order) thinking that water is wet, can be knowledge. If I am to make that second-order judgement – indeed, if I am even to frame that second-order thought – then I must meet the requirements for thinking that thought. Since the second-order thought embeds the first-order thought that water is wet, the requirements for thinking the second-order thought include the requirements for thinking the first-order thought. Thus, quite independently of the epistemic status of the second-order judgement, I can make that judgement only if I meet the requirements for thinking the first-order thought. At the very starting point for an enquiry into the epistemic status of the second-order judgement, it is already guaranteed that I meet the externalist conditions for thinking the first-order thought. So, externalism poses no special problem for the achievement of self-knowledge.

In his seminal contribution to this topic, Tyler Burge says (1988, pp. 653–4):

Among the conditions that determine the contents of first-order empirical thoughts are some that can be known only by empirical means. To think of something as water, for example, one must be in some causal relation to water – or at least in some causal relation to other particular substances that enable one to theorize accurately about water. . . . To know that such conditions obtain, one must rely on empirical methods. To know that water exists, or that what one is touching is water, one cannot circumvent empirical procedures. But to *think* that water is a liquid, one need not *know* the complex conditions that must obtain if one is to think that thought.

Let us agree with this. In order to know that I am thinking that water is a liquid, or that water is wet, I do not need to know anything of externalist philosophical theory, and I do not need to know that the conditions required by that theory actually obtain. But to the extent that the truth of externalist dependence theses is no bar to the achievement of self-knowledge, the consequence problem for first-person authority given externalism becomes press-

ing. Without empirical investigation, I know that I am thinking that water is a liquid. If I learn something about externalist philosophical theory then I can also know that if I am thinking that water is wet then certain conditions must obtain. I can draw the obvious consequence that those conditions do indeed obtain. Yet, as Burge says: 'To know that such conditions obtain, one must rely on empirical methods'. This is the consequence problem for first-person authority given externalism.

Paul Boghossian presents this problem as a *reductio ad absurdum* of the combination of externalism about mental content and privileged self-knowledge. Externalism is here 'the view that what concepts our thoughts involve may depend not only on facts that are internal to us, but on facts about our environment' (1997, p. 161). The claim about self-knowledge is 'that we are able to know, without the benefit of empirical investigation, what our thoughts are in our own case' (ibid.). The conclusion for *reductio* is that, if both these claims were true, then 'we would be in a position to know certain facts about the world a priori, facts that no one can reasonably believe are knowable a priori' (ibid.). According to Boghossian, either the externalist claim or the claim about self-knowledge must be rejected. My aim is to show how both claims can be maintained.

4. McKinsey's *reductio* and two notions of a priori knowledge

In an earlier and influential paper, Michael McKinsey presented a somewhat similar *reductio* (1981, p. 16):

[I]f you could know a priori that you are in a given mental state, and your being in that mental state conceptually or logically implies the existence of external objects, then you could know a priori that the external world exists. Since you obviously *can't* know a priori that the external world exists, you also can't know a priori that you are in the mental state in question.

But McKinsey's presentation has a number of distinctive features of which one is particularly important for our purposes.

McKinsey says that a priori knowledge is 'knowledge obtained independently of empirical investigation' (1981, p. 9). This sounds just like Boghossian, for whom the notion of a priori knowledge in play is simply that of knowledge that is available 'without the benefit of empirical investigation'. But McKinsey makes a significant addition when he talks about the character of self-knowledge (ibid; emphasis added):

[W]e can in principle find out about these states in ourselves 'just by thinking', without launching an empirical investigation *or making any assumptions about the external physical world*.

The lesson that McKinsey draws from his *reductio* is that we do not have this special kind of knowledge of our own mental states in cases where the truth of an ascription of the mental state depends on the existence of objects external to the subject of the ascription. If the lesson is unpacked in line with McKinsey's characterisation of self-knowledge then it becomes this: We cannot know about our own externalist mental states without launching an empirical investigation or at least making some assumptions about the external physical world.

It is not obvious that this is something with which Burge would disagree. In a sentence just before the passage that I quoted towards the end of the last section, he says (1988, p. 653): 'It is uncontroversial that the conditions for thinking a certain thought must be presupposed in the thinking.' And at the end of that quoted passage, when he says, 'But to *think* that water is a liquid, one need not *know* the complex conditions that must obtain if one is to think that thought', he immediately adds: 'Such conditions need only be presupposed.' On Burge's account, in thinking that water is wet, or in thinking that I am thinking that water is wet, I presuppose or assume that the conditions necessary for me to think that thought do obtain. So (just as McKinsey would insist), my knowledge that I am thinking that water is wet is not knowledge that I can have 'without making any assumptions about the external physical world'.[27]

The notion of a priori knowledge that McKinsey uses is very strict. Suppose that, in that very strict sense, I could know a priori that I am thinking that water is wet, and could also know a priori that if I am thinking that water is wet then environmental condition E holds. Then McKinsey assumes, and I accept, that the strict a priori warrant would be transmitted from the premises of the *modus ponens* inference to the conclusion, that it, to the proposition that condition E holds.[28] But, as McKinsey points out and as we must surely agree, it is absurd to suppose that E could be known a priori. So it cannot be that I can both know what I am thinking and know the truth of an externalist dependence thesis a priori, in that strict sense.

But an externalist who accepts a thesis of first-person authority may respond to McKinsey's r*eductio* argument by conceding one point while insisting on another. First, an externalist who accepts a thesis of first-person authority may concede that self-knowledge is not a priori in the strict sense. This is what the quotation from Burge (1988) suggests and in any case, quite apart from externalism, it is plausible that self-knowledge is not a priori in the strict sense; there seem to be empirical assumptions or presuppositions in the background.[29] But, unless background assumptions always stand in need of justification, this concession is consistent with the idea that self-knowledge is a priori in the weaker sense of not being justificatorily based on em-

pirical investigation.[30] So, second, an externalist who accepts a thesis of first-person authority may insist that both self-knowledge and knowledge of externalist dependence theses are a priori in that weaker sense.

Part of the importance of McKinsey's *reductio* argument is that, although McKinsey himself focuses on a priori knowledge in the strict sense, the pattern of the argument can be repeated for the weaker notion of a priori knowledge.[31] Suppose that we take a priori knowledge to be knowledge that does not rest on empirical investigation but may still depend on empirical background assumptions (such as the assumption that 'the conditions for thinking a certain thought' are met). This makes it much more plausible that externalist dependence theses can be known a priori and much more likely that first-person authority involves a priori self-knowledge. But it still seems absurd that I could know a priori, even in this weaker sense, that I am (or have been) embedded in an environment that contains samples of water.

In the end, it does not matter whether we use the term 'a priori knowledge' at all. Whatever label we use for knowledge that is not justificatorily based on empirical investigation, externalism and self-knowledge together give rise to a problem, an instance of the problem of armchair knowledge. In the third lecture, I offer a solution to the problem, making use of the idea of limitations on the transmission of epistemic warrant.