The Problem of Armchair Knowledge[1]

MARTIN DAVIES

1. McKinsey's *reductio* argument: Externalism and self-knowledge

In 'Anti-individualism and privileged access' (1991), Michael McKinsey asks us to consider the following three propositions, where 'E' says that some particular externalist condition for thinking that water is wet is met:[2]

(1)   Oscar knows *a priori* that he is thinking that water is wet.

(2)   The proposition that Oscar is thinking that water is wet conceptually implies E.

(3)   The proposition E cannot be known *a priori*, but only by empirical investigation.

He then argues that (1), (2) and (3) constitute an inconsistent triad as follows (1991, p. 15):

> Suppose (1) that Oscar knows a priori that he is thinking that water is wet. Then by (2), Oscar can simply *deduce* E, using premisses that are knowable a priori, including the premiss that he is thinking that water is wet. Since Oscar can deduce E from premisses that are knowable a priori, Oscar can know E itself a priori. But this contradicts (3), the assumption that E *cannot* be known a priori. Hence (1), (2), and (3) are inconsistent.

McKinsey's conclusion is that 'anti-individualism is inconsistent with privileged access' (ibid.).

In a more recent paper (2002), McKinsey sets out very clearly the principles about privileged access and externalism on which his argument depends. First, (1) is a consequence of a doctrine of privileged access or first-person authority about the contents of our thoughts:

Privileged access to content (PAC)
It is necessarily true that if a person *x* is thinking that *p*, then *x* can in principle know a priori that he himself, or she herself, is thinking that *p*.

Second, if we take E to be an externalist condition in the sense that it 'asserts or implies the existence of contingent objects of some sort external relative to Oscar' then (2) is a consequence of a doctrine of semantic externalism applied to the predicate 'is thinking that water is wet':

Semantic externalism (SE)
Many *de dicto*-structured predicates of the form 'is thinking that *p*' express properties that are wide, in the sense that possession of such a property by an agent logically or conceptually implies the existence of contingent objects external to that agent.

---

[2] Proposition (2) is actually numbered (2b).

If it is correct that what can be deduced from premises that are knowable a priori can itself be known a priori then (1) and (2) are jointly inconsistent with (3).[3] More generally, if no proposition that asserts or implies the existence of contingent external objects can be known a priori then no pair of propositions like (1) and (2) can be true together. But if (PAC) and (SE) are both correct then *some* such pairs must be true. So (PAC) and (SE) cannot both be true: anti-individualism, as rendered by (SE), is inconsistent with privileged access, formulated as (PAC). If 'is thinking that *p*' expresses a 'wide' property then I cannot know with first-person authority that it is true of me.

Here, then, is the *reductio* (1991, 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.

McKinsey's *reductio* argument about externalism and self-knowledge can be adapted to provide the first instance of the epistemological problem with which I am concerned in the present paper.[4] For my own expository purposes, it is useful to separate Oscar's palpably valid argument for E from the epistemological commentary that generates the puzzle. In order to make the problem vivid we can, in Oscar's argument, substitute a specific claim about the environment for the placeholder 'E'. In order to avoid detailed consideration of different notions of a priori knowledge we can, in the epistemological commentary, make use of the intuitive notion of knowledge that is available from the armchair.

1.1 Externalism and a first instance of the problem of armchair knowledge

Consider the argument (WATER):

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 plausible that my first-personal knowledge that I am thinking and what I am thinking does not depend for its status as knowledge on my conducting any detailed empirical investigation either of the information processing going on inside my head or of the physical and social environment in which I am situated. I am able to know from the armchair that I am a thinking being and that I think many particular things. So I can have armchair knowledge of the first premise WATER(1). But if philosophical arguments yield knowledge then there is more that I can know from the armchair. If externalism is

---

[3] In the more recent paper, McKinsey (2002) points out that his argument for the inconsistency of the triad (1), (2) and (3) depends only on a closure principle about a priori knowability, which he calls the closure of a priority under logical implication (CA): Necessarily, for any person *x*, and any propositions *p* and *q*, if *x* can know a priori that *p*, and *p* logically implies *q*, then *x* can know a priori that *q*. See also his paper in this volume.

[4] See also Brown, 1995; Boghossian, 1997.

correct then I can know, not only that I have thoughts with certain particular contents, but also that having those thoughts imposes requirements on my environment. In particular, we suppose that externalist philosophical theory motivates the *externalist dependence thesis*:

> **WaterDep**   Necessarily($\forall x$) (If x is thinking that water is wet then x is (or has been) embedded in an environment that contains samples of water).

So, philosophical theorising yields armchair knowledge of the conditional premise, WATER(2).[5]

Both the premises WATER(1) and WATER(2) can be known from the armchair and it does not require any empirical investigation to see that the conclusion WATER(3) follows. But it is overwhelmingly plausible that some empirical investigation is required if I am to settle the question whether or not I am embedded in an environment that contains samples of water. I cannot, without empirical investigation, come to know that the answer to this question is that my environment does indeed contain samples of water. So, while WATER(1) and WATER(2) can be known from the armchair, WATER(3) seems to fall outside the scope of armchair knowledge. Externalist philosophical theory, when taken together with a plausible claim about self-knowledge, gives rise to an instance of what I call *the problem of armchair knowledge*.

2.   Wright on Moore: Limitations on the transmission of evidential support

In his British Academy Lecture, 'Facts and certainty', Crispin Wright reflects on the intuitive inadequacy of Moore's (1959) anti-sceptical argument (MOORE), which we can represent as follows:

> MOORE(1)   Here is one hand, and here is another.
>
> MOORE(2)   If here is one hand and here is another, then an external world exists.
>
> Therefore:
>
> MOORE(3)   An external world exists.

Moore's experience provides good but defeasible evidence for MOORE(1). But the question is whether this evidential support is transmitted to MOORE(3) across the *modus ponens* inference in which the elementary piece of conceptual analysis, MOORE(2), figures as the conditional premise.

2.1 A pattern for non-transmission

Wright (1985, 435–6) asks us to consider three examples in which the question of transmission of evidential support can arise:

> (A) The transmission of support from: Five hours ago Jones swallowed twenty deadly nightshade berries,
>
> to: Jones has absorbed into his system a fatal quantity of belladonna,
>
> and thence to: Jones will shortly die.
>
> (B) The transmission of support from: Jones has just written an 'X' on that piece of paper,

---

[5] Henceforth, I omit the parenthetical 'or have been'.

to: Jones has just voted,

and thence to: An election is taking place.

(C) The transmission of support from: Jones has kicked the ball between the two white posts,

to: Jones has scored a goal,

and thence to: A game of football is taking place.

In examples (B) and (C), but not in (A), Wright says, 'the evidential support afforded by the first line for the second is itself conditional on the *prior* reasonableness of accepting the third line. . . . [K]nowledge of the first does not begin to provide support for the second unless it is *antecedently* reasonable to accept the third' (p. 436). Moore's mistake then is to suppose that the structure of evidential support in (MOORE) is like that in example (A) when it is really like that in (B) and (C) (p. 437):

> Once the hypothesis is seriously entertained that it is as likely as not, for all I know, that there is no material world as ordinarily conceived, my experience will lose all tendency to corroborate the particular propositions about the material world which I normally take to be certain.

If (MOORE) provides an example of non-transmission of evidential support across a palpably valid *modus ponens* inference, then it seems that other cases discussed by Wittgenstein in *On Certainty* (1969) provide examples as well. Consider *OC*, 208–11:

> 208.   I have a telephone conversation with New York. My friend tells me that his young trees have buds of such and such a kind. I am now convinced that his tree is . . . . Am I also convinced that the earth exists?
>
> 209.   The existence of the earth is rather part of the whole picture which forms the starting-point of belief for me.
>
> 210.   Does my telephone call to New York strengthen my conviction that the earth exists?
>
>   Much seems to be fixed, and it is removed from the traffic. It is so to speak shunted onto an unused siding.
>
> 211.   Now it gives our way of looking at things, and our researches, their form. Perhaps it was once disputed. But perhaps, for unthinkable ages, it has belonged to the *scaffolding* of our thoughts. . . .

The argument that we need to consider here is (TREE):

> TREE(1)   My friend in New York has a . . . tree in his garden.
>
> TREE(2)   If my friend in New York has a . . . tree in his garden, then the earth exists.
>
> Therefore:
>
> TREE(3)   The earth exists.

What Wittgenstein's remarks seem to suggest is that the evidential support for TREE(1) that is provided by my telephone conversation with my friend in New York is not transmitted to TREE(3).

2.2 Epistemic achievement and entitlement

Towards the end of 'Facts and certainty' (pp. 470–1), Wright considers the possibility that there are propositions (including some of Wittgenstein's 'hinge' propositions) that lie outside the domain of *cognitive achievement*. Evidential support or epistemic warrant would not be transmitted to such propositions just because, lying outside the domain of cognitive or epistemic achievement, they are also 'outside the domain of what may be known, reasonably believed, or doubted'. But although these propositions would not be known in the sense that involves epistemic achievement, they would still be known in a more inclusive sense. Thus, *OC*, 357–9:

> 357.   One might say: "'I know' expresses *comfortable* certainty, not the certainty that is still struggling."
>
> 358.   Now I would like to regard this certainty, not as something akin to hastiness or superficiality, but as a form of life. . . .
>
> 359.   But that means I want to conceive it as something that lies beyond being justified; as it were, as something animal.

Wright actually explores the idea that these propositions lie outside the domain of cognitive or epistemic achievement because they lie outside the domain of truth-evaluability or are *not fact-stating*. But it seems that the structure of Wright's proposal as involving less inclusive and more inclusive notions of knowledge might be retained even if we do not go so far as to deny the fact-stating status of the propositions to which only the more inclusive notion ('comfortable certainty') applies. We might distinguish between a stricter notion of knowledge that is an *achievement* and a more inclusive notion that embraces assumptions that we are epistemically *entitled* to make.

Knowledge may be an achievement in the sense that it requires that a question-settling justification or warrant be provided for believing the known proposition. A rational thinker who is engaged in an epistemic project may regard the question whether q is true as being open *pro tem*, and he may seek to bring to bear considerations that settle the question. Such a thinker might achieve knowledge that q by, for example, gathering evidential support for q and against alternatives, or by following through an a priori argument in favour of q, or by assembling considerations in favour of some premise, p, from which q palpably follows.

A fact-stating assumption may be one that we are epistemically entitled to make in the context of a particular epistemic project in the sense that a rational thinker is entitled to rely on the assumption in the conduct of that project. The project may lead to knowledge even though it involves taking the assumption for granted. No evidential support or other question-settling warrant for the assumption needs to be provided within that project. In this rough and intuitive characterisation of epistemic entitlement, the notion of making an assumption should be construed in a thin way so as to include the case where it simply does not occur to a thinker to doubt that something is the case. Being epistemically entitled to make an assumption thus includes being epistemically entitled to ignore, or not to bother about, certain possibilities.[6] So suppose that a thinker sets out to settle the question whether q is true and that the thinker is entitled to ignore

---

[6] Burge writes (1993, pp. 458–9): 'We are entitled to rely, other things equal, on perception, memory, deductive and inductive reasoning . . . Philosophers may articulate these entitlements. But being entitled does not require being able to justify reliance on these resources, or even to conceive such a justification.'

certain possible ways in which *q* might be false. Then the thinker's project may yield knowledge that *q* even though the positive considerations that the thinker assembles within that project do not rule out those particular alternatives to *q*.

In some contexts, I may be entitled to the assumption that a football match is taking place; I may be entitled to ignore the possibility, for example, that I am watching a rehearsal on a movie set. Against the background of the assumption that it is a football match and not a movie rehearsal that I am watching, the perceptual evidence of Jones kicking the ball between the two white posts counts very strongly in favour of the proposition that Jones has scored a goal and against many alternative possibilities. By watching the trajectory of the ball, and perhaps by observing also the behaviour of the referee and the crowd, I can come to know that Jones has scored a goal. I have an epistemically adequate question-settling justification for that belief.

From the proposition that Jones has scored a goal it surely follows that a game of football is taking place. So if I believe that Jones has scored a goal and I appreciate the entailment then I should also believe that a game of football is taking place. If I appreciate the entailment then, since I am justified in believing that Jones has scored a goal, I am also justified in believing that a game of football is taking place. But I cannot take the question-settling justification for the first belief that is provided by watching the trajectory of the ball and augment it by recognition of the entailment so as to provide myself with a question-settling justification for the second belief. Even if I am poised to make the inference from the premise that Jones has scored a goal to the conclusion that a game of football is taking place, the perceptual evidence of Jones kicking the ball between the two white posts is of no use to me in the project of rationally settling the question whether a game of football is indeed taking place. If I begin by regarding the question as open *pro tempore* – if, for example, I regard the possibility that I am watching a rehearsal on a movie set as a live option – then I cannot take the perceptual evidence as counting in favour of the premise. For the perceptual evidence supports the premise only against the background of the assumption that it is a football match and not a movie rehearsal that I am watching.

In summary, because I am entitled to the background assumption, I do have an epistemically adequate question-settling justification for believing the premise. That is my epistemic achievement. But, even given my appreciation of the entailment, I cannot redeploy that justification for believing the premise as a question-settling justification for believing the conclusion.[7] This, I think, is the lesson of Wright's example (C), and much the same could be said of his example (B) about voting. And I agree with Wright that the structure of evidential support in Moore's argument is relevantly similar to the structure in examples (B) and (C), even though the nature of my entitlement to the background

---

[7] I hope that the terminology of epistemic 'achievement' may provide a helpful contrast with epistemic entitlement. But I do not want to suggest that regarding a question as open and then closing it is the only kind of epistemic achievement. Sometimes, following through an argument does not put us in a position to provide a question-settling justification for believing the conclusion, but does serve to make plain that we are rationally committed to believing the conclusion. Further reflection on the structure of evidential support may reveal the role that the conclusion plays as a background assumption in epistemic projects and we may be able to show that we are epistemically entitled to make that assumption. Coming to see all this would be an epistemic achievement, but not the epistemic achievement of providing a question-settling justification.

assumption that there is an external world is surely different from the nature of my entitlement to the background assumption that I am watching a football match.

These notions of entitlement and background assumptions did not, however, figure explicitly in my first attempt (1998) to use Wright's ideas about non-transmission of epistemic warrant as a way of avoiding McKinsey's *reductio*.

2.3 Early versions of the limitation principles

Confronted by McKinsey's *reductio* argument, and with Wright's discussion of Moore in mind, I proposed a principle that would limit the transmission of epistemic warrant from the premises to the conclusions of even palpably valid inferences:[8]

> **First Limitation Principle (early version):**
> Epistemic warrant cannot be transmitted from the premises of a valid argument to its conclusion if, for one of the premises, the truth of the conclusion is a precondition of our warrant for that premise counting as a warrant.

This principle appears to have the consequence that epistemic warrant cannot be transmitted from the premises to the conclusion of Moore's argument. It also seems to account for the non-transmission of evidential support in Wright's examples (B) and (C) and in Wittgenstein's example (TREE). But, in this initial formulation, the principle is problematic in a number of respects. It makes use of the unexplained notion of a precondition. If this notion is interpreted simply as a necessary condition, then the principle is certainly open to counterexamples.

Yet more pressing than these worries about the principle is the fact that it is not at all clear how it applies to the example (WATER) that is motivated by McKinsey's *reductio* argument. The First Limitation Principle is modelled on Wright's account of cases in which *evidential* support is not transmitted. But our knowledge of WATER(1) and WATER(2) is not based on evidence; it is armchair knowledge.

What seems to be needed to block the unwanted transmission of armchair warrant from WATER(1) and WATER(2) to WATER(3) is not the First Limitation Principle, but something like:[9]

> **Second Limitation Principle (early version):**
> Epistemic warrant cannot be transmitted from the premises of a valid argument to its conclusion if, for one of the premises, the truth of the conclusion is a precondition of the knower even being able to believe that premise.

According to externalist philosophical theory, my being embedded in an environment that contains water is a necessary condition for my believing or even thinking *that water is wet*. It is also a necessary condition for my thinking any other thought in which the concept of water is deployed; in particular, for my thinking *that I am thinking that water*

---

[8] The actual formulation in Davies, 1998, p. 351, is: Epistemic warrant cannot be transferred from A to B, even given an a priori known entailment from A to B, if the truth of B is a precondition of our warrant for A counting as a warrant.

[9] The actual formulation in Davies, 1998, p. 353, is: Epistemic warrant cannot be transferred from A to B, even given an a priori known entailment from A to B, if the truth of B is a precondition of the knower even being able to believe the proposition A.

is wet. So the truth of WATER(3) is a necessary condition of my even being able to think WATER(1), and this triggers the Second Limitation Principle.[10]

The early version of the Second Limitation Principle has the desired result; but in other respects it is far from satisfactory. The worry is not, primarily, that the principle is open to counterexamples, but rather that no independent motivation for the principle has been provided.[11] In short, the early version of the Second Limitation Principle appears to be completely *ad hoc*. One of my aims in what follows is to provide a proper motivation for limitation principles that account for the non-transmission of epistemic warrant in Wright's examples, including particularly (MOORE), and in McKinsey's example (WATER).

3. Aunty's argument: A second instance of the problem of armchair knowledge

In 'Aunty's own argument for the language of thought' (1992),[12] I put forward an argument for the language of thought hypothesis. The argument is relatively non-empirical in character and it proceeds in two main steps. The first step makes use of neo-Fregean resources. Thinking involves the deployment of concepts and having concepts involves commitments to certain patterns of inference. In particular, conceptualised thought involves performing certain inferences *in virtue of their form* and this is then glossed in terms of tacit knowledge of the corresponding inferential rule. The second step makes use of a quite general connection between tacit knowledge of rules and syntactically structured representations.[13]

3.1 Eliminativism and an intuition of non-negotiability

Aunty's argument supports a conditional: *If* we are thinking beings *then* the language of thought (LOT) hypothesis is true of us; that is, we are LOT beings. Although the argument is relatively non-empirical in character, the question whether we really are LOT beings is a substantive empirical one and answering it requires detailed empirical investigation.[14] It seems reasonable to allow that it is epistemically possible (whether or

---

[10] We assume that the unexplained notion of a precondition is to be interpreted simply as a necessary condition.

[11] This is not quite true. It was suggested that we should want to block the transmission of warrant in certain putative anti-sceptical arguments even if the truth of the sceptical hypothesis would render one of the premises unthinkable rather than just robbing it of its warrant (Davies, 1998, p. 353).

[12] See also Davies, 1991. The Aunty in question is Jerry Fodor's. He represents her as a conservative figure who is more likely to favour connectionism than to accept that there are good reasons to adopt the language of thought hypothesis (Fodor, 1987, p. 139): 'It turns out that dear Aunty is, of all things, a New Connectionist Groupie.' As I envisage her, she has some sympathy for the views of the later Wittgenstein but is fundamentally a neo-Fregean. I claim that the neo-Fregean framework offers Aunty the resources to construct her own argument for the claim that conceptualised thought requires the truth of the language of thought hypothesis.

[13] A background assumption for the whole argument is that personal-level events of conscious judgement and thought are underpinned by occurrences of physical configurations belonging to kinds that figure in the science of information-processing psychology. These physical configurations can be assigned the contents of the thoughts that they underpin. They are 'proposition-sized' bearers of causal powers. This assumption is what Fodor (1985, 1987) calls *intentional realism* and it is close to the assumption of *propositional modularity* (Ramsey, Stich and Garon, 1990). In my view, we are committed to this assumption by some of our everyday practices of mental talk and explanation, but I shall not spell out the nature of this commitment here.

[14] This is so even given the assumption that intentional realism is true of us.

not it is likely) that we may turn out not to be LOT beings. But then Aunty's argument would support an eliminativist *modus tollens*. From the premise that we are not LOT beings we would be able to conclude that we are not thinking beings.[15]

Imagine, for a moment, that empirical evidence decisively supported the thesis that we are not LOT beings. It seems that, in those circumstances, we would face a stark choice between two alternatives. On the one hand, we could perform the *modus tollens* inference and cease to regard each other and ourselves as thinking beings. On the other hand, we could conclude that there is something wrong with Aunty's argument. But the first alternative seems rationally to require that we abandon our familiar descriptions of ourselves and others as believing and wanting things, as hoping and fearing things, as engaging in reasoning and planning; and there are powerful intuitions proclaiming that this option is not genuinely available to us. Our everyday engagement in folk psychological practice seems to be philosophically non-negotiable. So, we would be driven to the second alternative. If we found ourselves to be in a disobliging world then we would be bound to reject Aunty's argument. We would have to conclude that the philosophical theories that support the argument are in some way flawed.

It may well seem to you that, if this is how things would be in a disobliging world, then we should already conclude now that Aunty's argument is the product of flawed philosophical theories. But in my view we can respect the intuition of non-negotiability even while embracing the philosophical theories that support Aunty's argument. We can accept that those philosophical theories provide the best elaboration and precisification of our current conception of a thinking being and that Aunty's argument correctly draws out a necessary condition for falling under that conception. But we can also allow that it is part of our current conception that we ourselves are thinking beings: being one of us is a sufficient condition for falling under that conception. Suppose that these claims about a necessary condition and a sufficient condition for falling under our current conception of a thinking being are both correct. It follows that if we are not LOT beings then our current conception dictates both that we are and that we are not thinking beings. In a disobliging world, our current conception of a thinking being would be of no use to us, since it would dictate contradictory answers to the question whether we are thinking beings.

If we turn out not to be LOT beings then we must negotiate our way to a revised conception of what it is to be a thinking being.[16] This conceptual negotiation would proceed under two constraints. The revised conception should be one under which we fall; so it should not involve a commitment to the truth of the LOT hypothesis.[17] And the revised conception should rationally sustain as much as possible of our folk psychological practice.[18] It is by acknowledging this pair of constraints on the process of

---

[15] Compare what Ramsey, Stich and Garon write (1990, p. 500): '*If* connectionist hypotheses [of a particular sort] turn out to be right, so too will eliminativism about propositional attitudes.'

[16] The process of revision will be informed by the particular ways in which the world turns out to be disobliging.

[17] On the assumption that the philosophical theories supporting Aunty's argument do provide the best elaboration and precisification of our current conception, we need to revise that conception in order to avoid a commitment to the truth of the LOT hypothesis.

[18] We would not abandon the idea that we engage in deductive inference; but we would, presumably, adjust our conception of what is involved in accepting or performing an inference in virtue of its form.

revision that we honour the intuition of non-negotiability concerning our engagement in folk psychological practice.

In response to the worry about eliminativism, what is being proposed is that the concept of a thinking being has at least two components. There is an exemplar component that specifies sufficient conditions: we, at least, are thinking beings. And there is a more theoretical component which, according to Aunty's argument, imposes a necessary condition: thinking beings are LOT beings. There is no logical guarantee that the items that meet the sufficient conditions also meet the necessary conditions and in a disobliging world the two components lead to contradictory verdicts on cases.

The worry about eliminativism does not, in the end, constitute an objection to Aunty's argument. The importance of the worry is, rather, that it prompts us to uncover a particular structure in our conception of a thinking being. The real problem for Aunty's argument is that it gives rise to a second instance of the problem of armchair knowledge.

3.2 Aunty's argument and armchair knowledge

Suppose that the LOT hypothesis is, in fact, true and that the concept of a thinking being is in good order. It seems that, by relying on my grasp of the exemplar component of the concept of a thinking being, I can know that I am a thinking being. In fact, it seems that I have more than one way of knowing this. Since at least some thinking is conscious, first-personal awareness of my own conscious mental states also assures me that I am a thinking being. Either way, provided that the LOT hypothesis is in fact true, this knowledge seems to be available to me ahead of any empirical investigation of the information-processing mechanisms inside my head.

By relying on my grasp of the theoretical component of the concept of a thinking being, engaging in some inferences to the best philosophical explanation, and following through Aunty's argument I can, if the argument is a good one, come to know that a thinking being must be an LOT being. I know that if I am a thinking being then I am an LOT being.

Without conducting any detailed empirical investigation, I can have two pieces of knowledge that provide the premises for a simple *modus ponens* inference:

LOT(1)   I am a thinking being.

LOT(2)   If I am a thinking being then I am an LOT being.

Therefore:

LOT(3)   I am an LOT being.

But it is highly plausible that settling the question whether the LOT hypothesis is true will be the result of experiments, computational modelling and, more generally, detailed comparison of the successes and failures of competing research programmes. So Aunty's argument gives rise to a second instance of the problem of armchair knowledge. For, if the argument is a good one, then both LOT(1) and LOT(2) can be known from the armchair; yet knowledge of LOT(3) requires an investigative methodology rather than an armchair methodology.

The early and unsatisfactory version of the Second Limitation Principle that was mentioned towards the end of the previous section does at least have the advantage of providing a way out of this instance of the problem of armchair knowledge. If the argument that supports LOT(2) is correct then the truth of the conclusion, LOT(3), is a

necessary condition for my being a thinking being, for my being able to think anything at all, and so for my being able to think or believe the premise LOT(1). As I go on to offer more adequately motivated limitation principles, my aim is that they should account for the non-transmission of epistemic warrant in (LOT), as well as in (MOORE) and (WATER).

4. Interim report: In the armchair, down and out

In my view, being a thinking person depends on being embodied and embedded in the right way. I call the claim about embodiment, that thought requires a particular kind of internal cognitive machinery, an *architecturalist* claim. The claim about being embedded, that there are requirements that our environment must meet if we are to have thoughts with certain contents, is an *externalist* claim. Both claims are supported by philosophical arguments of a relatively a priori kind – arguments advanced from the armchair.

My concern in this paper is with the epistemological problem that these arguments pose. For both architecturalist and externalist arguments generate instances of the problem of armchair knowledge. When the arguments are combined with a claim about self-knowledge they seem to yield deeply implausible consequences about what it is possible to know from the armchair. Given the plausibility of the claim about knowledge of our own thoughts, the problem of armchair knowledge is naturally regarded as casting doubt on the arguments that generate it. The moral that many will draw is that armchair philosophical theorising cannot take us from everyday folk-psychological claims about our thoughts and their contents either *down*, to substantive claims about the cognitive machinery that underpins our thinking, or *out*, to substantive claims about the world that our thoughts concern. But I shall be taking a different approach.

In my view, philosophical theorising, conducted in the armchair, can indeed support both conditional claims that link the personal level of folk psychology with the subpersonal level of information-processing mechanisms and conditional claims that link mind and world. In the armchair, we can proceed both down and out, to know what thought requires. But I also want to maintain the plausible claim about first-personal knowledge of our thoughts and their contents.

In the armchair, I can know what thought requires. In the armchair, I can know about my thoughts and their contents. But I cannot, purely by armchair reflection, settle the question whether the conditions that thought requires are conditions that actually obtain. In general, from the facts that I can have armchair knowledge of a conditional (if A then B), and that I can have armchair knowledge of the antecedent of the conditional (A), it does not follow that I can gain armchair knowledge of the consequent of the conditional (B). In my view, then, the solution to the problem of armchair knowledge lies in limitations on our ability to achieve knowledge by inference from things that we already know. Sometimes, the epistemic warrant or justification that we have for believing the premises of an argument is not transmitted to the conclusion of the argument, even though the argument is palpably valid. Sometimes (in the terminology of section 2.2), even given my appreciation of the validity of the argument, I cannot redeploy the justification for believing the premises as a question-settling justification for believing the conclusion.

Placing limitations on the transmission of epistemic warrant from premises to conclusion in palpably valid arguments may strike you as an extreme measure.

Knowledge by inference is surely a vital component in our epistemic practices. So it may seem much more promising to reject the externalist and architecturalist arguments that generate instances of the problem of armchair knowledge. In my view, Wright's treatment of Moore's anti-sceptical argument furnishes considerations that count against a blanket rejection of the idea of limiting the transmission of epistemic warrant. But there may still be a concern about the apparently *ad hoc* step from the First Limitation Principle, which emerged fairly naturally from what Wright said, to the Second Limitation Principle, which is needed to deal with (WATER) and (LOT).

In the remainder of this paper, I shall try to motivate my approach in two ways.[19] First, I shall show that instances of the problem of armchair knowledge, or closely related problems about transmission of epistemic warrant, are relatively widespread. It would not be right to suppose that the problem is only generated by a couple of idiosyncratic and easily rejected philosophical arguments. Second, I shall show that the proposed limitations on transmission of warrant are far from being *ad hoc.* Failure of transmission of epistemic warrant is the analogue, within the thought of a single subject, of the dialectical phenomenon of begging the question.[20]

5. Problems about transmission of epistemic warrant: Six examples

So far, we have considered two examples of the problem of armchair knowledge and one closely related problem about transmission of epistemic warrant in a putative anti-sceptical argument.

> Example 1 (WATER): Environmental requirements for thought
>
> Example 2 (MOORE): Moore's anti-sceptical argument
>
> Example 3 (LOT): Subpersonal-level requirements for thought

In this section, I shall add three further examples.

Example 4: Indexical thoughts

The instance of the problem of armchair knowledge that results from externalist philosophical theorising about thoughts involving natural kind concepts (Example 1) clearly belongs in a larger category. There are, for example, other varieties of externalism, including the externalism about so-called 'object-dependent thoughts' that is familiar from the work of Gareth Evans and John McDowell.[21] More generally, these externalist examples belong with other substantive requirements for thought that issue from philosophical theories about thought content such as teleological theories of content.[22] It is not difficult to see how an instance of the problem of armchair knowledge could be generated from the claim that, in order to be a thinker, a being must have such-

---

[19] I shall not attempt to set my approach against the background of a general epistemology. For some of the issues that would need to be addressed, see Jessica Brown's paper in this volume.

[20] I am not alone in proposing a connection with begging the question here. See the title of Wright, 2000; and see McLaughlin, 2000, pp. 104–5. James Pryor (forthcoming) says: 'This notion of transmission-failure is basically a new piece of terminology for talking about an old phenomenon: the phenomenon of *begging the question.*' But although this looks like a point of agreement, Pryor actually disagrees with the approach that Wright and I take because he does not regard begging the question as a dialectical phenomenon.

[21] Evans, 1982; McDowell, 1984, 1986.

[22] See, for example, Millikan, 1989. See also McLaughlin's discussion (2000, pp. 107–9) of teleological theories such as Dretske's (1995, chapter 5) and of Davidson's, 1987, example of Swampman.

and-such a kind of selectional history and must not have come into existence just a few minutes ago. Here, I borrow material from Evans (1982) to provide an externalist example that involves the indexical concept 'here'.

First, according to Evans, being able to think about a particular place is not a trivial matter:[23]

> We are prepared to suppose that there is a determinate thought here – that the subject has a definite place in mind – because we know that subjects do have a capacity to select one position in egocentric space, and to maintain a stable dispositional connection with it. . . . If the subject . . . does know which place his thought concerns . . . this will be manifestable only in manifestations of that stable dispositional connection . . .

What this suggests is that someone who is unable, for a while, either to maintain a stable dispositional connection with a position or to keep track of his movement through space is likewise unable, for that while, to have (determinate) indexical thoughts about places.

Second, Evans presents a vivid example of a thinker who fails to keep track of his movement through space:[24]

> A person might lie in bed in hospital thinking repeatedly 'How hot it was here yesterday' – supposing himself to be stationary in the dark. But his bed might be very well oiled, and be pulled by strings, so that every time he has what he takes to be the same thought, he is in fact thinking of a different place, and having a different thought.

As Evans describes the case, this thinker has several instantaneous thoughts about different places. But we can adapt the example by imagining that the person thinks, slowly, carefully, not wanting to knock anything over in the dark, 'There's a bottle of whisky just here'. In general, it is plausible that a thinker who essays a 'here'-thought, but who is moving through space even as he thinks, fails to think any determinate thought at all. If the thought that he essays as he moves several yards is, 'There's a bottle of whisky just here', then there is no place such that the correctness of the putative thought would turn on whether there is a bottle of whisky at that place. The subject has no determinate place in mind.

Suppose now that I am stationary in bed in the dark, thinking, 'There's a bottle of whisky just here' – a thought that is true if there is indeed a bottle of whisky located at a particular position just next to the bed. Suppose, too, that it is correct, as a matter of philosophical theory, that someone who neither maintains a stable dispositional connection with a position nor keeps track of his movement through space is unable to have indexical thoughts about places. And now consider the following argument:

BED(1)   I am thinking that there's a bottle of whisky just here.

BED(2)   If I am thinking that there's a bottle of whisky just here then I am stationary.

Therefore:

BED(3)   I am stationary.

---

[23] Evans, 1982, p. 161.
[24] Ibid., p. 201.

By the assumption of first-person authority, I can know BED(1) from the armchair. If I follow through the philosophical theorising that is indicated in the previous paragraph, then I can also have armchair knowledge of BED(2). But it is highly implausible that I can settle the question whether I am stationary, rather than being moved silently along a darkened hospital corridor, just by giving thought. The conclusion, BED(3) seems to fall outside the scope of armchair knowledge.

Example 5: Colour concepts

In 'Naming the colours', David Lewis begins from the thought that our folk theory of colours contains principles linking colours and colour experiences, such as:[25]

> When a red thing is before someone's eyes, it typically causes in him an experience of redness.

If our concepts of colours and of colour experiences are concepts of properties of objects and of inner states that are implicitly defined by our folk theory, then conceptual analysis is liable to lead us to such 'definitions' as these:[26]

> D1 *Red* is the surface property of things which typically causes experience of red in people who have such things before their eyes.

> D2 *Experience of red* is the inner state of people which is the typical effect of having red things before the eyes.

The problem with D1 and D2 is that what they say, while true, does not distinguish the pair <red, experience of red> from other similar pairs such as <green, experience of green>. A further chapter must be added to the folk theory of colour in order to individuate specific colours, and Lewis suggests that this further chapter can come in different versions, each specifying relatively parochial examples that serve well enough the needs of some subcommunity of the population. Thus, amongst followers of Australian Rules football, it will suffice to say 'that red is the colour of the diagonal stripe on an Essendon Football Club jumper'.[27]

 With this much by way of background, we can consider the following *modus ponens* inference:

> RED(1)   This [pointing at the diagonal stripe on an Essendon jumper] is red.

> RED(2)   If this is red then there is a type of colour experience and a type of inner state that is typically caused in people who have this before their eyes.

> Therefore:

> RED(3)   There is a type of colour experience and a type of inner state that is typically caused in people who have this before their eyes.

By relying on my mastery of the exemplar component of the concept of red (the parochial exemplar component that applies to my group), I can know that this Essendon stripe is red. Indeed, I have more than one way of knowing this since I can often know what colour something is just by looking at it. Having seen many Essendon jumpers, I can recognise this item as being the colour of the Essendon diagonal stripe. Either way,

---

[25] Lewis, 1997, p. 326.

[26] Ibid., p. 327.

[27] Ibid., p. 335. In American English, 'jersey' is more natural than 'jumper' for the item of clothing worn by football players.

knowledge of RED(1) is available to me ahead of any investigation of other people's colour experiences or inner states. By relying on my grasp of the theoretical component of the concept of red (including the principles D1 and D2), I can know that if something is red then there is a type of colour experience and a type of inner state that is typically caused in people who have that thing before their eyes. So I can know RED(2). But it is implausible that, without rising from the armchair save perhaps to look at an Essendon football jumper, I can know the conclusion RED(3).

At the beginning of 'Naming the colours', Lewis says that 'it is a Moorean fact that there are colours rightly so-called'.[28] This remark suggests that certain claims about colours and colour experiences have the status of presuppositions or unquestioned background assumptions in our everyday use of colour concepts to classify the things that we see. It also suggests that these claims, like Moore's conclusion, cannot have epistemic warrant transmitted to them from premises that acquire their warrant in our everyday epistemic projects. That is just what I shall be claiming.

Example 6: Meaning and tacit knowledge

The third new example to be introduced in this section concerns the meaning of sentences that are never used. Ordinary speakers of English are credited with speaking a language in which sentences that no one ever gets around to using nevertheless have determinate meanings. But what facts about ordinary speakers and their language use could make it correct for us to describe them in this way? This is the problem of *meaning without use.*

A number of philosophers of language, including Brian Loar (1981) and Stephen Schiffer (1993),[29] have argued persuasively that this problem cannot be solved without appeal to the structure of the mechanisms of language processing in speakers' heads. I myself would argue specifically that our assignments of meaning without use are correct only if speakers have subpersonal-level tacit knowledge of a compositional semantic theory for their language.[30]

Suppose, for a moment, that Loar, Schiffer and I are right about this. Then the *modus ponens* inference to be considered is as follows:

> MEANING(1) Sentence *s* means that *p* in my language and would do so whether or not I ever used it.
>
> MEANING(2) If sentence *s* means that *p* in my language and would do so whether or not I ever used it then I have tacit knowledge of a compositional semantic theory for my language.
>
> Therefore:
>
> MEANING(3) I have tacit knowledge of a compositional semantic theory for my language.

---

[28] Ibid., p. 323: 'It won't do to say that colours do not exist; or that we are unable to detect them; or that they never are properties of material things; or that they go away when things are unilluminated or unobserved; or that they change with every change in the illumination, or with every change in an observer's visual capacities; or that the same surface of the same thing has different colours for different observers. Compromise on these points, and it becomes doubtful whether the so-called colours posited in your theory are rightly so-called.'

[29] Schiffer, 1993, is responding to Lewis, 1992.

[30] See Davies, 2000b.

Suppose that *s* is a hitherto unused and unconsidered sentence built from words and constructions that occur in other sentences that I have used. When I hear or consider the sentence *s* for the first time, I am able to assign it a meaning, say the meaning that *p*. I may know that this is what *s* means. I may know that this is what *s* does and did and would mean whether or not I used it. Furthermore, I may know this without engaging in any empirical investigation of my language-processing system. So I have armchair knowledge of the first premise. Then, if the development of the arguments advanced by Loar and Schiffer is correct, I also have armchair knowledge of the conditional premise. But the conclusion, which follows so obviously from these premises, concerns the structure of the language-processing system and surely I cannot gain knowledge about this cognitive structure without a substantial programme of empirical research. Armchair methodologies suffice for knowledge of the premises, but knowledge of the conclusion requires an investigative methodology.

6. Limitation principles and begging the question

In the previous section, I tried to show that instances of the problem of armchair knowledge, or closely related problems about transmission of epistemic warrant, are relatively widespread. My aim in this section is to motivate limitation principles on transmission of epistemic warrant by making use of the idea that failure of transmission of epistemic warrant is the analogue, within the thought of a single subject, of the dialectical phenomenon of begging the question.

6.1 Moore's anti-sceptical argument and a revised limitation principle

It is often said that Moore's argument begs the question against the sceptic, but what we need is an explicit account of what makes an argument question-begging and for this I rely on Frank Jackson.[31] He says that an argument begs the question when 'anyone – or anyone sane – who doubted the conclusion would have background beliefs relative to which the evidence for the premises would be no evidence'.[32]

According to Jackson's view of what is achieved by advancing an argument for a conclusion, the speaker invites the hearer to borrow evidence, or other considerations, in favour of the premises of the argument. By her choice of premises she provides an indication as to what kinds of considerations these are. Typically, evidence counts in favour of a proposition only relative to particular background assumptions and often the relevant background assumptions are shared between speaker and hearer. But when background assumptions are not shared, it is possible that the considerations that count in favour of the premises relative to the speaker's background assumptions do not count in favour of the premises relative to the hearer's background assumptions. Suppose that a speaker sets out to convince a doubting hearer of the truth of some conclusion. The speaker begs the question against the hearer if the hearer's doubt rationally requires him to adopt background assumptions relative to which the considerations that are supposed to support the speaker's premises no longer provide that support. A question-begging argument 'could be of no use in convincing doubters'.[33]

---

[31] Jackson, 1987.
[32] Jackson, 1987, p. 111.
[33] Ibid., p. 112.

Convincing a doubter and settling a question both involve ruling out various ways in which a proposition could have been false. In the case of a speaker who is trying to convince a doubting hearer, the speaker's evidence for her premises rules out various ways in which those premises could have been false, ways that are left open by the speaker's background assumptions. The hearer who doubts the conclusion of the argument may have background assumptions that leave a wider range of possibilities open, and the speaker's evidence for the premises may not rule all those possibilities out. Indeed, the speaker's evidence may leave untouched the ways in which, according to the hearer, the conclusion could be false.

In a similar way, a thinker who has question-settling justifications for believing the premises of an argument is able to rule out various ways in which those premises could have been false. These are ways that are left open by background assumptions that the thinker is, in that context, epistemically entitled to make. But it does not follow, even given the thinker's appreciation of the validity of the argument, that the thinker can redeploy his justifications for believing the premises so as to provide himself with a question-settling justification for believing the conclusion. For it may be that in regarding the question of the truth of the conclusion as open *pro tem*, the thinker regards as live options certain possibilities that he was entitled to ignore when only the premises were under consideration. So the considerations that furnished epistemically adequate question-settling justifications for believing the premises may be inadequate to settle the question of the truth of the conclusion.

All this is consistent with saying that the thinker who has justifications for believing the premises of an argument is also justified in believing the conclusion. Indeed, it is consistent with saying that the speaker is epistemically entitled to believe the conclusion. The point about non-transmission of epistemic warrant is not that the thinker should believe the premises but not believe the conclusion. It is not that the thinker's beliefs in the premises are epistemically in good order while his belief in the conclusion would be epistemically out of order. It is that the thinker cannot take the question-settling justifications for believing the premises and augment them by recognition of the validity of the argument so as to provide himself with a question-settling justification for believing the conclusion.

The reason for this non-transmission of question-settling warrant is that the thinker's operative considerations amount to epistemically adequate justifications for believing the premises only against the background of certain assumptions that the thinker is entitled to make. Simply regarding a question – here, the question of the truth of the conclusion – as open *pro tem* does not rob the thinker of that entitlement. But it may be that a doubt about the truth of the conclusion would rationally require the thinker to adopt different background assumptions relative to which the operative considerations would no longer amount to epistemically adequate justifications for believing the premises. The proposal is that, in such a case, the thinker cannot consistently make use of the original background assumptions within the context of an epistemic project that begins with the thinker regarding the question of the truth of the conclusion as open.

The analogy between convincing a doubter and providing an epistemically adequate question-settling justification for believing thus motivates the following principle about transmission of epistemic warrant:

First Limitation Principle (revised version)
Epistemic warrant cannot be transmitted from the premises of a valid argument to its conclusion if, for one of the premises, the warrant for that premise counts as a warrant only against the background of certain assumptions and acceptance of those assumptions cannot be rationally combined with doubt about the truth of the conclusion.

In order to apply this principle to any particular argument, we need to identify assumptions such that, for one of the premises, it is only against the background of those assumptions that the operative considerations amount to an epistemically adequate question-settling warrant for that premise. Then we need to show that acceptance of those assumptions cannot be combined with doubt about the truth of the conclusion. Wright's diagnosis of the failure of transmission of evidential support from the premises to the conclusion of Moore's argument seems to fit this pattern.[34]

6.2 Subpersonal requirements for thought and two generalised limitation principles

It is not so clear, however, that this revised version of the First Limitation Principle explains the failure of transmission of warrant from LOT(1) and LOT(2) to LOT(3).[35] The epistemic warrant for LOT(1) is constituted either by grasp of the exemplar component of the concept of a thinking being or else by awareness of my own conscious mental states. But in neither case is there an obvious candidate for the role of background assumption without which the epistemic warrant would not count as a warrant.

There is, however, a very basic assumption that lies in the background of any epistemic project; namely, the assumption that there is such a proposition as the proposition for which one is attempting to provide evidence, justification or warrant. The notion of a proposition that figures in this assumption is not to be construed in a metaphysically committed way. If a thinker is attempting to provide a warrant for believing A then the basic background assumption is simply that there is such a thing to think as A. If there were no such thing to think as A then there could be no question of anything constituting an epistemically adequate warrant for believing A. So we can make explicit a second principle that is arguably a consequence of the first:

Second Limitation Principle (revised version)
Epistemic warrant cannot be transmitted from the premises of a valid argument to its conclusion if, for one of the premises, acceptance of the assumption that there is such a proposition for the knower to think as that premise cannot be rationally combined with doubt about the truth of the conclusion.

One way in which the assumption that figures in this principle could turn out to be false would be if one of the purported conceptual constituents in the premise were revealed to be internally incoherent, dictating contradictory answers to the question whether some particular item falls under the concept. To that extent, the principle holds some promise of providing a solution to the instance of the problem of armchair

---

[34] See the discussion in Davies 1998, 2000a; and cf. Wright, 1985.
[35] In the case of the conditional premise LOT(2), it might reasonably be said that it is only against the background of the assumption of intentional realism that the premise is supported by the neo-Fregean philosophical theory on which Aunty's argument draws. But it is surely not true that acceptance of this assumption cannot be rationally combined with doubt about the conclusion LOT(3). Many philosophers sanely believe that intentional realism is true but the language of thought hypothesis is false.

knowledge that is posed by Aunty's argument. For the worry about eliminativism prompted us to uncover a particular structure in our conception of a thinking being. But, on the other hand, it is clear that acceptance of the assumption that there is such a thing to think as that I am a thinking being – and, in particular, acceptance of the assumption that the concept of a thinking being is in good order – can be rationally combined with doubt about the truth of the LOT hypothesis. It is only the acceptance of Aunty's argument that generates rational tension between acceptance of the background assumption and doubt about the conclusion.

It is clear what kind of modification of the principle is required if it is to provide a solution to the problem of armchair knowledge that arises from Aunty's argument and the required modification is not merely opportunistic or *ad hoc*. To see this, we need to return to begging the question and focus on the fact that arguments may have several premises. Suppose that a speaker advances a multi-premise argument in an attempt to convince a hearer who doubts that argument's conclusion. The speaker offers various considerations for borrowing; they are considerations that count in favour of the premises relative to the speaker's background assumptions. If the hearer's doubt by itself rationally requires him to adopt background assumptions relative to which one of the speaker's premises is no longer supported by the considerations that she offers for borrowing then the speaker begs the question against the hearer. That is the kind of case that Jackson describes.

But there is a more complicated scenario in which it is no less true that the argument, as advanced by the speaker, will be of no use in convincing the doubting hearer. If the hearer is to be convinced then he must accept the considerations that the speaker offers in support of her premises. In addition, he must not differ from the speaker in his background assumptions in such a way that the premises are no longer supported by those considerations. Suppose that the hearer's doubt about the conclusion, when put together with acceptance of the considerations that the speaker offers in support of some of the premises, rationally requires him to adopt background assumptions relative to which some other one of the premises is no longer supported by the considerations offered in support of it. That is enough to ensure that the argument, as advanced by the speaker, will be of no use in convincing the hearer. So, if failure of transmission of epistemic warrant is the analogue, within the thought of a single subject, of the dialectical phenomenon of begging the question then we should expect the following pair of limitation principles, of which the second is arguably a consequence of the first:[36]

> **First Limitation Principle (generalised version)**
> Epistemic warrant cannot be transmitted from the premises of a valid argument to its conclusion if, for one of the premises, the warrant for that premise counts as a warrant only against the background of certain assumptions and acceptance (i) of those assumptions and (ii) of the warrants for the other premises cannot be rationally combined with doubt about the truth of the conclusion.

> **Second Limitation Principle (generalised version)**
> Epistemic warrant cannot be transmitted from the premises of a valid argument to its conclusion if, for one of the premises, acceptance (i) of the assumption that

---

[36] These generalised versions of the two principles are essentially the same as the 'multi-premise' versions of Davies, 2000a, p. 412.

there is such a proposition for the knower to think as that premise and (ii) of the warrants for the other premises cannot be rationally combined with doubt about the truth of the conclusion.

It is this last principle that provides a solution to the instance of the problem of armchair knowledge that is posed by Aunty's argument.

7. Applying the limitation principles

In this section, I shall show how the generalised versions of the two limitation principles account for the failure of transmission of epistemic warrant from premises to conclusion in five of our six examples.

Example 2 (MOORE): Moore's anti-sceptical argument

We have seen that the First Limitation Principle in either its early version (section 2.3) or its revised version (section 6.1) accounts for the non-transmission of epistemic warrant in Moore's argument. The same goes, of course, for the generalised version of the First Limitation Principle.

Example 3 (LOT): Subpersonal requirements for thought

Suppose that a thinker accepts that there is such a thing to think as the premise LOT(1), that he himself is a thinking being. Suppose, in particular, that he accepts that there is no internal incoherence, no source of contradictions, in the concept of a thinking being. In that case, the thinker must accept the assumption that the items, such as himself, that meet the sufficient condition for falling under the concept also meet the necessary condition. Acceptance of that assumption does not, by itself, rationally preclude doubt about whether he himself is an LOT being. But suppose, in addition, that the thinker accepts the epistemic warrants for the premises LOT(1) and LOT(2).

The epistemic warrant for believing LOT(1) is provided either by the exemplar component of the concept of a thinking being or else by his awareness of his own conscious mental states. But it is the warrant for believing the conditional premise LOT(2) that figures crucially in the solution to the problem of armchair knowledge. That warrant is provided by a battery of philosophical theory and by Aunty's argument. Acceptance of the assumption that the items that meet the sufficient conditions for falling under the concept of a thinking being also meet the necessary conditions *and of the warrant for LOT(2)* cannot be rationally combined with doubt about whether the thinker himself is an LOT being. So the generalised version of the Second Limitation Principle is triggered and epistemic warrant cannot be transmitted from LOT(1) and LOT(2) to the conclusion LOT(3).

Example 6 (RED): Colour concepts

The problem about transmission of epistemic warrant that is presented by Lewis's account of colour concepts has a solution that is very similar to the solution to the problem presented by Aunty's argument. For the concept of red, like the concept of a thinking being, has an exemplar-based sufficient-conditions component and a theory-based necessary-conditions component.

According to the (parochial) exemplar component of the concept of red, being the colour of the Essendon stripe is sufficient for being red: Essendon stripes (at least) are red

things. From the theoretical component we can derive a necessary condition for being red: If something is red then there is a type of colour experience and a type of inner state that is typically caused in people who have that thing before their eyes. But there is no logical guarantee that there is a single type of colour experience and a single type of inner state that is typically produced in people by the diagonal stripe on an Essendon jumper. If there is not, then the two components of the concept yield contradictory pronouncements. If the world turns out to be disobliging in this respect then our current colour concepts will be of no use to us and we must negotiate our way to revised, presumably relativised, colour concepts.

Acceptance of the assumption that there is such a thing to think as the premise RED(1) involves acceptance of the assumption that the items that meet the sufficient conditions for falling under the concept of red also meet the necessary conditions. But acceptance of this *and of the warrant for RED(2)* cannot be rationally combined with doubt about the truth of RED(3). So the generalised version of the Second Limitation Principle is again triggered and epistemic warrant cannot be transmitted from RED(1) and RED(2) to the conclusion RED(3).

Example 1 (WATER): Environmental requirements for thought

We can also confirm that the generalised version of the Second Limitation Principle provides a solution to the instance of the problem of armchair knowledge that arises from externalism and self-knowledge.[37]

The warrant for the conditional premise, WATER(2), is a piece of philosophical theory that supports both the thesis:

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

and the thesis:

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

The theory supports the first thesis because it supports the claim:

> Necessarily(If there is such a thing for me to think as *that water is wet* then I am (or have been) embedded in an environment that contains samples of water).

Equally, it supports the claim:

> Necessarily(If there is such a thing for me to think as *that I am thinking that water is wet* then I am (or have been) embedded in an environment that contains samples of water).

So acceptance (i) of the assumption that there is such a thing for me to think as WATER(1) and (ii) of the warrant for WATER(2) cannot be rationally combined with doubt about the truth of WATER(3). According to the generalised version of the Second Limitation Principle, then, epistemic warrant cannot be transmitted from the premises WATER(1) and WATER(2) to the conclusion WATER(3).

---

[37] The revised version of the principle is not adequate to this task. There is no immediately obvious incompatibility between, on the one hand, acceptance of the assumption that there is such a thing for me to think as that I am thinking that water is wet and, on the other hand, doubt as to whether I am (or have been) embedded in an environment that contains samples of water. It is only in the context of a philosophical theory of externalism that there is a tension between this acceptance and this doubt.

Example 4 (BED): Indexical thoughts

The solution to the instance of the problem of armchair knowledge that is presented by indexical thoughts follows the contours of Example 1 (WATER). The warrant for the conditional premise:

> BED(2)   If I am thinking that there's a bottle of whisky just here then I am stationary.

is a piece of philosophical theory that also supports the conditional:

> If I am thinking that I am thinking that there's a bottle of whisky just here then I am stationary.

The theory supports these conditional theses because it also supports the claims:

> If there is such a thing for me to think as that there's a bottle of whisky just here then I am stationary.

> If there is such a thing for me to think as that there's a bottle of whisky just here then I am stationary.

Thus, acceptance (i) of the assumption that there is such a thing for me to think as BED(1) and (ii) of the warrant for BED(2) cannot be rationally combined with doubt about the truth of BED(3), and this again triggers the generalised version of the Second Limitation Principle.

8. Limitation principles and the objectivity of meaning

In Section 5, I provided six examples to substantiate the claim that problems about transmission of epistemic warrant are relatively widespread. In Section 6, I argued that limitation principles on transmission of epistemic warrant can be motivated by an analogy between providing a warrant and convincing a doubter. Failure of transmission is the analogue of begging the question. In Section 7, I showed how five of the six problems (three instances of the problem of armchair knowledge and two closely related problems) can be solved by appeal to the generalised versions of the First and Second Limitation Principles. It remains to say something about the final example.

Example 6 (MEANING): Meaning and tacit knowledge.

According to the philosophical theory that supports the conditional premise, MEANING(2), if sentences that are never used or even considered are to have determinate meanings then the language user must have tacit knowledge of a compositional semantic theory. If a speaker had only phrasebook knowledge of the meanings of a finite set of sentences then there would be no basis for crediting her with speaking a language in which sentences outside that set had determinate meanings.[38] In the absence of tacit knowledge of a compositional semantic theory, the application of the concept of meaning to an unused sentence, *s*, would be indeterminate. Any specific judgement about the meaning of *s* in this speaker's language would be incorrect. But it does not appear to follow from this philosophical theory about the objectivity of meaning that if MEANING(3) were false then there would be no such thing for me to think as MEANING(1). Rather, if MEANING(3) were false because I did not have tacit knowledge of a compositional semantic theory then MEANING(1) would be thinkable

---

[38] Schiffer, 1993; Davies, 2000b.

but false. So it is not very plausible that the Second Limitation Principle will be applicable to this example.

The solution to the instance of the problem of armchair knowledge that is posed by the argument about meaning and tacit knowledge must lie with the First Limitation Principle. What we need to show is that the warrant for MEANING(1) counts as a warrant only against the background of certain assumptions and that acceptance of those assumptions cannot be combined with doubt about the truth of MEANING(3) – or at least that acceptance of those assumptions *together with the warrant for MEANING(2)* cannot be combined with doubt about MEANING(3).[39] A fully satisfying account of the issues surrounding the warrant for MEANING(1) would require nothing less than an adequate epistemology of understanding. But perhaps it is sufficient for present purposes to suggest that one route to knowledge of meaning is, under appropriate conditions, to take an impression of meaning at face value.

Suppose, for a moment, that the philosophical theory about the objectivity of meaning is correct and that things are as that theory says they need to be. In particular, suppose that I have tacit knowledge of a compositional semantic theory and this tacit knowledge underwrites the meanings of sentences that belong to my language even though I never get around to using them. Meaning in my language is not constituted by my having an impression of meaning, both because unconsidered sentences have meanings and because impressions of meaning can, in principle, be misleading or illusory.[40]

Suppose that s is a hitherto unused and unconsidered sentence built from words and constructions that occur in other sentences that I have used. And suppose that, in virtue of my having tacit knowledge of compositional meaning rules for those words and constructions, s determinately means that p in my language. If I now hear or consider s for the first time (hearing it in reality or in my mind's ear, as it were) then I may hear it *as* meaning that p and, taking that impression of meaning at face value, I may judge that s does mean that p. The suggestion is that, under appropriate conditions, this judgement amounts to knowledge.

We do not have to be in the grip of a purely reliabilist epistemology to find it plausible that one condition that is necessary, if this judgement is to be knowledge, is that the same states of tacit knowledge that contribute to the constitution of s as meaning that p should figure in the causal explanation of s's being heard as meaning that p. If taking an impression of meaning at face value is to be a route to knowledge, then the mechanisms that generate the impression of meaning should be mechanisms that reliably track the truth about meaning. It would be too restrictive to insist that every knower should be able to conceptualise this requirement and explicitly assume that it holds. Language users with no conception of mechanisms that embody tacit knowledge of semantic rules, or even with no conception of mechanisms that generate impressions of meaning, can surely come to know what sentences mean by taking impressions of meaning at face value. But, on the other hand, if a language user has the conceptual sophistication to consider this

---

[39] This is not, strictly speaking, the only way in which the First Limitation Principle could be triggered. But it is the most promising way.

[40] The problem of meaning without use goes along with a problem of *meaning despite use* (Davies, 2000b). Some examples of sentences that are typically used to communicate something other than what they mean (such as, 'No head injury is too trivial to be ignored'; Bennett, 1976, p. 17) may provide examples of meaning illusions.

requirement and actually doubts that it holds then this seems to rule out the possibility of gaining knowledge of meaning simply by taking impressions of meaning at face value.[41]

When, as in this case, there is a logical gap between having an impression and that impression's being veridical, one is justified in taking the impression at face value only against the background of an assumption (a not calling into question) that certain reliabilist conditions related to the production of that impression are met. The impression furnishes an epistemic warrant for the judgement that things are as they seem to be only against the background of that assumption. The assumption against the background of which an impression of meaning furnishes a warrant when it is simply taken at face value may not be very specific; it may speak of reliability in general rather than of mechanisms that embody tacit knowledge in particular. But, given the philosophical theory that provides the warrant for MEANING(2), a general assumption of reliability can be elaborated into the particular assumption about impressions of meaning being generated by mechanisms that embody tacit knowledge of semantic rules. So, it is not possible rationally to combine acceptance of (i) the assumption of reliability against the background of which the warrant for MEANING(1) counts as a warrant and (ii) the philosophical theory that provides the warrant for MEANING(2) with doubt about the truth of MEANING(3). This is what we needed to show in order to trigger the generalised version of the First Limitation Principle.

If I were to doubt that I have tacit knowledge of a compositional semantic theory for my language then I could not resolve that doubt by reviewing the considerations that would ordinarily count in favour of MEANING(1) and MEANING(2). For, in the presence of that doubt, and given the considerations in favour of MEANING(2), the consideration that would ordinarily count in favour of MEANING(1) would no longer justify that belief. Analogously, if *you* were to doubt that I have tacit knowledge of a compositional semantic theory for my language then I would be begging the question against you if I tried to convince you by offering those considerations for borrowing.

In ordinary circumstances, it does not occur to me to doubt that the reliabilist conditions for gaining knowledge by taking an impression of meaning at face value are met. Against the background of that assumption (that not calling into question), the impression of meaning provides knowledge that s means that p by ruling out various relevant alternatives to MEANING(1), such as that s means that q or that s means that r.[42] But, even taken together with the philosophical theory that supports MEANING(2), the impression that s means that p does nothing to rule out the most obviously salient alternative to MEANING(3), namely that I do not have tacit knowledge of a compositional semantic theory and that my impressions of objective meaning are illusory. My epistemic warrants for the two premises of the *modus ponens* inference do not add up to an epistemic warrant for the conclusion. Warrant is not transmitted from premises to conclusion.

---

[41] Peacocke, 1999, p. 51, says that thinkers sometimes operate in 'the mode of taking the deliverances of a given informational system . . . at face value' and continues: 'It is in the nature of such modes of operation that they have both an objective and a subjective dimension involving reliability.' What I say in this paragraph closely follows Peacocke, ibid., pp. 51–2. Also, see Davies, 2000b, for an argument in favour of a requirement of awareness of linguistic structure.

[42] I take it that I am entitled to this assumption (this not calling into question) so that what is achieved is knowledge *simpliciter* and not just knowledge relative to that assumption.

Conclusion

I began (Sections 1–3) with the instance of the problem of armchair knowledge that arises from McKinsey's *reductio* argument, a closely related problem about transmission of epistemic warrant in Moore's anti-sceptical argument, and a second instance of the problem of armchair knowledge that arises from Aunty's argument for the language of thought. Wright's discussion of Moore's argument provides support for the general idea of limitations on the transmission of epistemic warrant, but my early proposals for limitation principles do not provide a satisfactory resolution of the problems generated by (WATER) and (LOT). In the second half of the paper, I have tried to improve on that situation.

I have shown (Section 5) that instances of the problem of armchair knowledge, or closely related problems about transmission of epistemic warrant, are relatively widespread. It would not be right to suppose that they arise only from a couple of idiosyncratic philosophical arguments. I have then motivated some principled limitations on transmission of epistemic warrant (Section 6) and shown how these provide solutions to three instances of the problem of armchair knowledge and two closely related problems (Section 7). In the final section, I have considered one instance of the problem at greater length. There are many difficult questions concerning the epistemology of understanding. But I am reasonably confident that even this last instance of the problem of armchair knowledge can be solved in a well motivated way. Being in the armchair, down and out, still seems like an attractive philosophical position.

Bibliography

Boghossian, P.A. 1997: What the externalist can know a priori. *Proceedings of the Aristotelian Society*, 97, 161–75.

Brown, J. 1995: The incompatibility of anti-individualism and privileged access. *Analysis*, 55, 149–56.

Burge, T. 1993: Content preservation. *Philosophical Review*, 102, 457–88.

Davidson, D. 1987: Knowing one's own mind. *Proceedings and Addresses of the American Philosophical Association*, 60, 441–58.

Davies, M. 1991: Concepts, connectionism, and the language of thought. In W. Ramsey, S. Stich and D. Rumelhart (eds), *Philosophy and Connectionist Theory*. Hillsdale, NJ: Lawrence Erlbaum Associates, 229–57.

Davies, M. 1992: Aunty's own argument for the language of thought. In J. Ezquerro and J.M. Larrazabal (eds), *Cognition, Semantics and Philosophy: Proceedings of the First International Colloquium on Cognitive Science*. Dordrecht: Kluwer Academic Publishers, 235–71.

Davies, M. 1998: Externalism, architecturalism, and epistemic warrant. In C. Wright, B.C. Smith and C. Macdonald (eds), *Knowing Our Own Minds*. Oxford: Oxford University Press, 321–61.

Davies, M. 2000a: Externalism and armchair knowledge. In P. Boghossian and C. Peacocke (eds), *New Essays on the A Priori*. Oxford: Oxford University Press, 384–414.

Davies, M. 2000b: Persons and their underpinnings. *Philosophical Explorations*, 3, 43–62.

Dretske, F. 1995: *Naturalizing the Mind*. Cambridge, MA: MIT Press.

Evans, G. 1981: Semantic theory and tacit knowledge. In S. Holtzman and C. Leich (eds), *Wittgenstein: To Follow a Rule*. London: Routledge and Kegan Paul, 118–37. Reprinted in *Collected Papers*. Oxford: Oxford University Press, 1985.

Evans, G. 1982: *The Varieties of Reference*. Oxford: Oxford University Press.

Fodor, J. 1985: Fodor's guide to mental representation. *Mind*, 94, 55–97. Reprinted in *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press, 1990.

Fodor, J. 1987: *Psychosemantics*. Cambridge, MA: MIT Press.

Jackson, F.C. 1987: *Conditionals*. Oxford: Blackwell Publishers.

Lewis, D. 1992: Meaning without use: Reply to Hawthorne. *Australasian Journal of Philosophy*, 70, 106–10.

Lewis, D. 1997: Naming the colours. *Australasian Journal of Philosophy*, 75, 323–42.

Loar, B. 1981: *Mind and Meaning*. Cambridge: Cambridge University Press.

McDowell, J. 1984: De re senses. *Philosophical Quarterly*, 34, 283–94.

McDowell, J. 1986: Singular thought and the extent of inner space'. In P. Pettit and J. McDowell (eds), *Subject, Thought, and Context*. Oxford: Oxford University Press, 137–68.

McKinsey, M. 1991: Anti-individualism and privileged access. *Analysis*, 51, 9–16.

McKinsey, M. 2002: Forms of externalism and privileged access. To appear in *Philosophical Perspectives*, 16.

McLaughlin, B.P. 2000: Self-knowledge, externalism and skepticism. *Proceedings of the Aristotelian Society Supplementary Volume* 74, 93–117.

Millikan, R.G. 1989: Biosemantics. *Journal of Philosophy*, 86, 281–97. Reprinted in *White Queen Psychology and Other Essays for Alice.* Cambridge, MA: MIT Press, 1993, 83–101.

Moore, G.E. 1959: Proof of an external world. In *Philosophical Papers*. London: Allen and Unwin, 127–50.

Peacocke, C. 1999: *Being Known*. Oxford: Oxford University Press.

Pryor, J. forthcoming: Is Moore's argument an example of transmission-failure?

Ramsey, W., Stich, S. and Garon, J. 1990: Connectionism, eliminativism and the future of folk psychology. In J.E. Tomberlin (ed.), *Philosophical Perspectives 4: Action Theory and Philosophy of Mind*. Atascadero, CA: Ridgeview Publishing Company, 499–533.

Schiffer, S. 1993: Actual-language relations. In J.E. Tomberlin (ed.), *Philosophical Perspectives, 7: Language and Logic.* Atascadero, CA: Ridgeview Publishing Company, 231–58.

Wittgenstein 1969: *On Certainty*. Oxford: Basil Blackwell.

Wright, C. 1985: Facts and certainty. *Proceedings of the British Academy*, 71, 429–72.

Wright, C. 2000: Cogency and question-begging: Some reflections on McKinsey's paradox and Putnam's proof. *Philosophical Issues*, 10, 140–63.