

Externalism and Armchair Knowledge*

Martin Davies

[I]f you could know a priori that you are in a given mental state, and your being in that 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.¹

Let us call someone who combines an externalist view of mental content with a doctrine of privileged self-knowledge a *compatibilist*. . . . [I]f compatibilism were true, 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.²

1. Kylie's Puzzle

On summer afternoons in Canberra, the baking sun reflects off Lake Burley Griffin, and the water shimmers. Up behind the university, in the botanical gardens, a cascading stream of water helps to maintain the humidity of the rainforest gully. These are just a couple of Kylie's thoughts on the subject of water, her water thoughts. Amongst Kylie's many other thoughts that involve the concept of water are these: that there is water in the lake, that trees die without water, that water is a liquid and, of course, that water is wet. When Kylie thinks consciously, in a way that occupies her attention, she is able to know what it is that she is thinking.³ This is true for thoughts about water, as for any other thoughts. So when Kylie thinks consciously that water is wet, she knows, even as she thinks, that she is thinking that water is wet.

Kylie is a student of philosophy. She has studied many arguments that purport to show that, in order to think that water shimmers, or cascades, or is wet, in order to think

* Versions of this material were presented at a conference on externalism and self-knowledge held at the University of Bristol in February 1999 and in a symposium with Brian McLaughlin and Brian Loar at the Pacific Division meeting of the American Philosophical Association held in Berkeley in April 1999. I learned much from those occasions and from a seminar with Jessica Brown and Michael McKinsey held in Oxford during the spring of 1999. Comments by Mark Greenberg on what I once thought of as a penultimate version of the paper were extremely useful and a period as a Visiting Fellow in the Research School of Social Sciences, Australian National University provided the opportunity for a complete overhaul of the text. In Canberra, I was helped by discussions with members of the Philosophy Program and other visitors including especially Helen Beebe, Frank Jackson and Michael Martin. Audiences at the University of Western Australia and the University of Melbourne and in the Faculties at ANU helped me to clarify both substance and presentation. I am also grateful to Antonia Barke, Paul Boghossian, Bill Brewer, Kirk Ludwig, Christopher Peacocke, Paul Pietroski, Sarah Sawyer, Stephen Schiffer, Ernest Sosa and Tom Stoneham for comments and conversations. This paper continues some of the themes of and, I hope, improves on 'Externalism, architecturalism, and epistemic warrant' (1998). It certainly inherits some of the debts of that earlier paper, especially a debt to Crispin Wright's British Academy lecture, 'Facts and certainty' (1985). In his 'Cogency and question-begging: Some reflections on McKinsey's paradox and Putnam's proof' (2000), Wright defends a position that is similar to the one that I adopt here though there are differences in detail and in explicit motivation. A systematic comparison of the two approaches must wait for another occasion.

¹ McKinsey (1991), p. 16.

² Boghossian (1997), p. 161.

³ For discussion of the way in which conscious thought occupies attention, and of 'consciously based self-ascriptions', see Peacocke (1998; 1999, Chapter 5).

any water thoughts at all, a thinker must be in some way familiar with water. In Kylie's estimation, it is a complex and delicate question whether these arguments are compelling. But her present judgement is that, in order for someone to think, say, that water is wet, that thinker must be, or at least must have been, in an environment that contains some samples of water.

As she ponders, consciously, the wetness of water, Kylie notes that she is thinking that water is wet. She also judges, on the basis of her philosophical reflections, that if she is thinking that water is wet then she must be, or have been, embedded in an environment that contains samples of water. From these two premises, she draws the obvious inference: she herself is, or has been, in a watery environment. There is nothing counterintuitive or even surprising about this conclusion. She is in Canberra, in the university; in one direction there is the water in the lake, in another direction there is the water in the gully, and there are many other samples of water all around.

There is a puzzle looming here, but just considering the premises and conclusion of Kylie's argument does not bring it out. The argument is obviously valid and she is no less certain about the conclusion than she is about the conjunction of the premises. The puzzle that is looming is not about validity or about certainty but about knowledge. In particular, it is about ways of gaining knowledge.

When Kylie thinks, consciously, that water is wet, she knows that this is what she is thinking; and she knows it in a special first-personal way. This is Kylie's first piece of knowledge. When philosophical theorising goes well, careful evaluation of arguments yields knowledge of true principles. So if Kylie is right in her philosophical reflections on water thoughts, then she has a second piece of knowledge. She knows that if she is thinking that water is wet then she is, or has been, in an environment containing samples of water. It then seems a trivial matter to achieve a third piece of knowledge. For these first two pieces of knowledge appear to offer Kylie a very short route to knowledge that she herself is, or has been, in a watery environment. It is here that the puzzle arises.

What is puzzling is not simply the idea that Kylie knows that there is water in her environment. In the last few days, she has taken showers, drunk lots of mineral water, and walked around the lake and down through the rainforest gully; so of course Kylie knows that there is water in her environment. What is puzzling is, rather, the apparent availability of a particular route to this knowledge. This route goes via Kylie's knowledge about what she is thinking and her knowledge about the conditions that any thinker must meet in order to think water thoughts. The methods that are involved in gaining these two pieces of knowledge are, first, the special first-personal way of knowing what one is thinking and, second, philosophical theorising. But it strikes Kylie as counterintuitive that the use of just these methods should deliver knowledge that there is water in her environment.

One way of dramatising Kylie's puzzle is this. Sitting in her armchair, Kylie may think that water is wet. If she does so, then she can know immediately that this is what she is thinking. Sitting in her armchair, Kylie may consider and evaluate philosophical arguments. If she does this well, then she may end up knowing that if she is thinking that water is wet then she is, or has been, in an environment containing samples of water. Sitting in her armchair, Kylie can draw the obvious conclusion. Yet it is counterintuitive that this exercise in armchair reflection should be a way of knowing that she is, or has been, in a watery environment.

As it happens, Kylie already knows that there is water in her Canberra environment. So, if armchair reflection provides a route to knowledge of this same environmental fact, then what it offers to Kylie is additional warrant for something that she already knows

by the familiar methods of observation and investigation. It offers resources that she might use to confirm that there is water in her environment should a doubt arise. That is already strange enough to be the core of a puzzle. But the puzzle may appear to be even more acute. For if, as things are, armchair reflection can provide this additional epistemic warrant then it seems that it could, in principle, have provided her with a way of learning about her relationship with water *for the first time*.⁴

2. Responding to the Puzzle

Kylie's puzzle is this. An argument that seems palpably valid has two premises:

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.

One premise seems to be knowable in the special first-personal way; the other seems to be knowable by philosophical theorising. But the conclusion is not intuitively something that we would expect to be able to know just by combining self-knowledge with philosophical theorising in this way.

Suppose we allow that self-knowledge is a kind of a priori knowledge and also that philosophical theorising yields a priori knowledge. Then we can say that what is puzzling is that both premises of the argument should be knowable a priori when the conclusion is something that should not, intuitively, be knowable a priori. In the quotations at the beginning of the paper, Michael McKinsey (1991) and Paul Boghossian (1997) pose puzzles like Kylie's in terms of a priori knowledge and, in Section 7, we shall consider those puzzles and Kylie's in the light of a distinction between two notions of a priori knowledge. But we do not need to make use of any particular conception of a priori knowledge in order to introduce Kylie's puzzle.

Someone might respond to the puzzle by recommending that we simply embrace what initially seems so counterintuitive; namely, that self-knowledge and philosophical theorising can together provide a route to knowledge that there is water in the environment. This bold strategy is what Sarah Sawyer proposes: 'introspection becomes a viable method of acquiring knowledge of our environment' (1998, p. 532). Sawyer then adds a consideration that is intended to make this strategy more plausible: '[I]t must be recognised that introspection will yield knowledge *only* of those empirical facts that the subject could already have come to know via empirical means' (ibid.). However, this was already taken into account in Kylie's puzzle. Kylie not only could have come to know that there was water in her environment 'via empirical means'; we assumed that she did already know this by the familiar empirical methods. What is puzzling is that the use of those armchair methods should offer additional warrant or provide confirmation in case of doubt.

⁴ Kylie's puzzle can be dramatised as a problem of armchair knowledge. But her sense of puzzlement is not the product of absolute and non-negotiable convictions about what is, and what is not, knowable from the armchair. She accepts that she may have to allow that more things can be known from the armchair than she had initially expected. But it seems to Kylie that, even against the background of a more generous view of the possibilities of armchair knowledge, it would still be implausible that this *particular* armchair route could lead to this piece of knowledge about her environment.

Bill Brewer (1999) responds to Kylie's puzzle by advancing a claim that is somewhat similar to the consideration that Sawyer puts forward. According to Brewer's account of what is involved in having the concept of water, when Kylie thinks that water is wet it must already be the case that she 'is in a position to express non-inferential knowledge [about water]' (1999, p. 267). What this amounts to is this (ibid.):

[E]ither . . . [she] currently has non-inferential reason of some kind to believe something about [water], or . . . [she] has retained knowledge based upon such a reason in memory.

Putting the point crudely we can say that, on Brewer's account, if Kylie is in a position to think any water thoughts at all then she already knows some things about water. So, as Brewer continues (ibid.):

[She] is already in a position to arrive at the knowledge that there is (or was) [water] in [her] environment if only [she] turns [her] mind to the matter. Therefore this argument [the argument in Kylie's puzzle] cannot possibly constitute a problematic non-empirical *source* of new empirical knowledge: if its premises are simply true, then the subject already has the wherewithal to arrive at knowledge of its conclusion.

If Brewer is right about this, then it is wrong to cast Kylie's puzzle in terms of armchair reflection providing her with a way of learning about her relationship with water *for the first time*.⁵ But the core of the puzzle remains. For it is strange to allow that Kylie's armchair reflections should provide her with a warrant or justification for believing that there is water in her environment. This is counterintuitive even if we accept that the warrant or justification that would be provided is one for which Kylie would have no need unless doubt arose.

Suppose that a subject believes the premises of a palpably valid argument to be true on the basis of certain considerations and as a result also believes that the conclusion is true. Consider what typically happens if such a subject comes to doubt the truth of the conclusion. Doubting the truth of the conclusion, she also doubts the truth of the conjunction of the premises, and may attach that doubt to one of the premises in particular. But then, by replaying the considerations that led her to accept the premises in the first place, she may be able to rule out certain alternatives to the truth of the premises and so resolve her doubt about them and also her doubt about the conclusion. Typically, the subject's justification for believing the premises offers resources that can be used to resolve a doubt about the truth of the conclusion. But in Kylie's case it is counterintuitive, quite apart from questions of first knowledge, that she could rationally resolve a doubt about the existence of water in her environment by replaying her armchair reflections.⁶

The bullet-biting strategy recommended by Sawyer does not appeal and the additional considerations put forward by her and by Brewer are not adequate to solve the puzzle. Diana Raffman (1998) responds in a different way, charging that arguments that threaten to generate puzzles like Kylie's face a trilemma. Either the truth of the first premise cannot be known in the special first-personal way or the truth of the second premise cannot be known by philosophical theorising alone or else the argument equivocates on what is involved in thinking that water is wet. The worry that the notion of thinking that water is wet might be given a thin construal in the first premise and a thicker construal in

⁵ See again the apparently more acute version of the puzzle mentioned at the end of the last section.

⁶ I am grateful to Michael Martin for stressing the need to connect justification for believing with resolution of doubt.

the second premise is genuine;⁷ but it is open to us simply to stipulate that there is to be no equivocation. Given that stipulation, Raffman's charge becomes the claim that not both of the premises can be known in the prescribed way; that is, in the special first-personal way (first premise) or by philosophical theorising (second premise).

This is, indeed, a very natural response to Kylie's puzzle. The puzzle shows, it may be said, that we cannot consistently combine a certain claim about self-knowledge with a certain claim about philosophical theories relating to water thoughts.⁸ But, in this paper, I shall be arguing that there is another way of responding to Kylie's puzzle. To this end, we can consider a kind of worst-case scenario that might arise with respect to Kylie's specific puzzle or another of the same form. In this scenario, the bullet-biting strategy is inapplicable: it really is problematic to allow that the conclusion of the argument could be known by combining self-knowledge and philosophical theorising. In addition, the premises are both true, and they can be known in the special first-personal way and by philosophical theorising respectively. The premises can be known in those particular armchair ways and, since there is to be no equivocation, the argument is palpably valid. Can we maintain that it *still* does not follow that those armchair methods provide a route to knowledge of the conclusion?

3. Externalism, Self-Knowledge and the Threat of Reductio

The conditional premise, Water(2), that figures in Kylie's puzzle is an instance of a general principle:

(\Box x) (If x is thinking that water is wet then x is (or has been) embedded in an environment that contains samples of water).

I call the claim that this universally quantified conditional holds, not just as things actually are but as a matter of necessity (WaterDep), an *externalist dependence thesis*.

In general, *externalism* about some mental property, M, is the thesis that whether a person 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. Expressed in terms of possible worlds, externalism about M says that there are two possible worlds 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 . This is a thesis about the existence of Twin Earth examples for M.⁹

⁷ According to some theories, the concept of water is descriptive in character so that thinking that water is wet is 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'. (For a defence of description theories, see Jackson (1998). Davies and Humberstone (1980) pointed out that much of what Putnam (1975) says would come out as correct on this construal of natural kind terms.) Given this construal of thinking that water is wet, the force of Raffman's charge is clear. In putting forward Kylie's puzzle about water thoughts, we assume that there is some other way to conceive of thinking that water is wet. Provisionally, we 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 (Evans, 1982; Brown, 1998). This is not to deny that there are chemical kinds that Kylie can think about even though she has had no causal engagement with samples and could not reliably recognise a sample if one were presented. Nor do we deny that there are chemical kinds that Kylie can think about even though there are no samples in her environment (Burge, 1982).

⁸ This, in effect, is the response of McKinsey (1991) and Boghossian (1997); see again the quotations at the beginning of this paper. Wright (2000, Section 6) offers considerations against this response.

⁹ Elsewhere, I have distinguished *modal* externalist theses like this one from *constitutive* modal theses (Davies, 1991, 1992, 1993, 1996, 1998).

If we are given the truth of the externalist dependence thesis, WaterDep, then it is a relatively straightforward matter to construct a Twin Earth example for the property of thinking that water is wet. But it is far from easy to move in the opposite direction from the existence of a Twin Earth example to the truth of any externalist dependence thesis as specific as WaterDep.¹⁰ Kylie's puzzle arises if externalist philosophical theorising leads, not just to Twin Earth externalism about water thoughts, but to an externalist dependence thesis about water thoughts.

3.1 Two problems for externalism and self-knowledge

Kylie's puzzle arises from the combination of two ideas. The first, as we have just noted, is that philosophical theorising leads to externalist dependence theses; in particular, to a dependence thesis about the property of thinking that water is wet. The second idea is that we have a special first-personal way of knowing what we are thinking. When we know in that special way what we are thinking, our knowledge does not depend for its status as knowledge on any empirical investigation of our environment or of our way of being embedded in it.¹¹

Kylie's puzzle points to a seemingly counterintuitive consequence of self-knowledge given an externalist dependence thesis. But there is another, and more familiar, way of developing the idea that there is some tension between the truth of externalist dependence theses and the possibility of self-knowledge. How can I achieve a special kind of knowledge about what I am thinking, knowledge that is not justificatorily based on empirical investigation of my environment, when my thinking what I am in fact thinking depends on my environment? We can call this the *achievement problem* for self-knowledge given externalism to distinguish it from the *consequence problem* for self-knowledge given externalism which is the problem posed by Kylie's puzzle.

Approaches to the achievement problem typically make some use of the fact that, when I think that I am thinking that water is wet, I deploy in thought the very same concepts of water and of being wet that are involved in my thinking that water is wet. So an externalist dependence thesis that is true for my first-order thinking that water is wet will be no less true for my second-order thinking that I am thinking that water is wet. Because the content of my second-order thought embeds the content of my first-order thought, my second-order thinking shares the dependence on the environment that is characteristic of my first-order thinking.¹² Of course, this fact about embedding does not by itself explain how it is that my second-order judgement that I am thinking that water is wet amounts to knowledge. We still need a general account of how authoritative self-knowledge is possible.¹³ But, according to these approaches to the achievement problem, the fact about embedding can be used to show that there is no *special* problem for the

¹⁰ See Gallois and O'Leary-Hawthorne (1996) and compare Wright (2000, Section 3). This view about the difficulty of moving from the existence of Twin Earth examples to a specific externalist dependence thesis seems to stand in contrast with a remark by Boghossian (1997, p. 163): '[A Twin Earth] thought experiment motivates externalism only by motivating a specific form of dependence of mental contents on external facts'.

¹¹ First-personal knowledge is subject to a thesis of *privileged access* along the lines that Brian McLaughlin and Michael Tye formulate in their discussion of these issues (1998, p. 286): 'It is conceptually necessary that if we are able to exercise our actual normal capacity to have beliefs about our occurrent thoughts, then if we are able to occurrently think that p, we are able to know that we are thinking that p without our knowledge being justificatorily based on empirical investigation of our environment.'

¹² Burge (1988), Heil (1988), Peacocke (1999, Chapter 5).

¹³ Burge (1996, 1998), Peacocke (1996, 1998, 1999).

achievement of self-knowledge in the fact that my first-order thinking is subject to an externalist dependence thesis.

An account of how self-knowledge is possible has to show how a second-order judgement that I am thinking that water is wet can be knowledge, when the judgement is made on the basis of my (conscious, first-order) thinking that water is wet. In part, this account will be the same as applies to cases in which no externalist dependence thesis is true for the first-order thinking. The worry is that there is an additional problem to be solved when an externalist dependence thesis does hold. But because of the point about embedding, the philosophical argument that supports the externalist dependence thesis for the first-order thinking will show something else as well. It will show, quite independently of the epistemic status of my second-order judgement, that I can frame that second-order thought only if I am embedded in an environment that contains samples of water. So, at the very starting point for an assessment of the epistemic status of the second-order judgement, it is already guaranteed that I meet the externalist conditions for the first-order thinking.¹⁴ Thus, the truth of an externalist dependence thesis for the first-order thinking poses no special problem for the achievement of self-knowledge.

3.2 *The threat of reductio*

Let us agree that these familiar approaches are successful in dealing with the achievement problem for self-knowledge given externalism. It is essential to that success that, in order to arrive at knowledge that I am thinking that water is wet, I do not need to know anything of externalist philosophical theorising and I do not need to know that the conditions required by an externalist dependence thesis actually obtain.¹⁵ But to the extent that the truth of an externalist dependence thesis is no bar to the achievement of self-knowledge, the consequence problem becomes pressing. When she thinks, consciously, that water is wet, Kylie knows that this is what she is thinking; and the status of her knowledge as knowledge does not depend on any empirical investigation of her environment. Then she engages in some philosophical theorising and comes to know that if she is thinking that water is wet then certain conditions must obtain. She draws the obvious consequence that those conditions do indeed obtain. Yet, as Tyler Burge says, ‘To know that such conditions obtain, one must rely on empirical methods’ (1988, p. 654). This is precisely Kylie’s puzzle.

Paul Boghossian (1997) presents the consequence problem as a *reductio ad absurdum* of the combination of externalism and self-knowledge. (See the second quotation at the beginning of this paper.) In his presentation, externalism is ‘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’ (pp. 161–2). According to Boghossian, one or the other claim has to be rejected.

¹⁴ For a critical assessment of Heil’s (1988) use of the point about embedding, see Brueckner (1990). For further discussion, see Falvey and Owens (1994) and Brueckner (1994).

¹⁵ There is a complication here if we accept Brewer’s (1999) claim that anyone who is in a position to think that water is wet will inevitably know many things about water (see above Section 2). However, even then it remains the case that the self-knowledge is not justificatorily dependent on that empirical knowledge about water.

My aim is to show how both claims can be maintained. Some specific externalist dependence theses are true and can be known by engaging in philosophical theorising. We can know what we are thinking without this knowledge being justificatorily based on investigation of our environment. Yet those armchair reflections that yield knowledge of an externalist dependence thesis and of what I am thinking do not provide a route to knowledge of facts about the world which ‘no one can reasonably believe are knowable a priori’.¹⁶

4. Closure of Knowledge and Transmission of Warrant

Whether it is genuinely possible to respond to Kylie’s puzzle in the way that I am proposing is related to the question whether knowledge is closed under known entailment. If someone knows the premises of an argument and knows that a particular conclusion follows from those premises, does it follow that the person knows, or is at least in a position to know, that conclusion?

Fred Dretske (1970) allows that a father who takes his son to visit the zoo may know that the animal in the pen in front of them is a zebra even without gathering any evidence to rule out the alternative possibility that it is a cleverly disguised mule. Dretske presents the example as a case in which the father knows both:

Zebra(1) The animal in the pen is a zebra.

and:

Zebra(2) If the animal in the pen is a zebra then it is not a cleverly disguised mule.

yet does not know:

Zebra(3) The animal in the pen is not a cleverly disguised mule.

Even more famously, Robert Nozick (1981) presents an analysis of knowledge on which any one of us can know many truths about the external world that are plainly incompatible with a radical sceptical hypothesis, even though we cannot know that the sceptical hypothesis about the external world is itself false.

If knowledge is not closed under known entailment then it can happen that a thinker has justifications or warrants for believing the premises of a valid argument but has no epistemically adequate justification or warrant for believing the conclusion. If that is so then the justifications or warrants for believing the premises do not themselves add up to a justification or warrant for believing the conclusion. In such a case, we shall say that those justifications or warrants are not *transmitted* from premises to conclusion.¹⁷

¹⁶ Boghossian notes the possibility of a response to the threat of reductio along the lines that I propose but is sceptical about its prospects (1997, p. 175): ‘I have to say that I would be very surprised if there turned out to be any such cases [where a priori warrant for the premises of an argument is not transmitted to the entailed conclusion] that survived scrutiny.’ However, I have an ally in Wright (2000) whose response to the threat of reductio is similar to mine.

¹⁷ In an earlier paper (Davies, 1998), I spoke of *transfer* of warrant. Wright’s term ‘transmission’ seems preferable. He says (Wright, 2000) that ‘a particular warrant, w, transmits across a valid argument just in case the argument is cogent when w is the warrant for its premises’. And: ‘A cogent argument is one whereby someone could be moved to rational conviction of the truth of its conclusion.’ As Wright stresses, transmission or non-transmission of warrant is not a property of an argument in itself. It may be that one possible warrant for the premises is transmitted to the conclusion while another possible warrant is not. My only reservation about Wright’s way of introducing the idea of transmission of warrant is that he ties it quite closely to the idea of first knowledge. (See again the discussion of Brewer (1999) in Section 2 above.)

Where knowledge is not closed, warrant is not transmitted. But the converse is not guaranteed to hold. In a particular case where justification or warrant is not transmitted from the premises of a valid argument to its conclusion, it could still be that anyone who knew the premises would also know, or be in a position to know, that the conclusion is true. It might be, for example, that anyone who knew the premises on the basis of a warrant that did not transmit to the conclusion would inevitably have some other warrant for the conclusion. Or, alternatively, it might be that, although warrant does not transmit from premises to conclusion, it is possible to have knowledge of the conclusion without any justification or warrant at all, but with a different kind of entitlement. It may not be obvious how either of these ways of having closure of knowledge without transmission of warrant would apply to Dretske's zebra example.¹⁸ But the transmission question and the closure question are still distinct and there is room for the possibility that a failure of transmission need not be inevitably accompanied by a breach of closure.¹⁹ What is crucial to our strategy for responding to Kylie's puzzle is the idea of there being some limitations on transmission of justification or warrant.

In fact, we have already allowed that Brewer (1999) may be right to claim that if Kylie is in a position to think any water thoughts at all then she is already in a position to know that there is water in her environment. We said that it is still counterintuitive to allow that Kylie's armchair reflections provide her with an additional justification or warrant for believing that there is water in her environment or with resources that she might use to confirm that there is water in her environment should a doubt arise. So we need to provide a motivation for limiting transmission of warrant even in cases where there is no breach of closure.

Our response to Kylie's puzzle is developed in two stages. The aim of the first stage is to uncover a plausible principle for limiting the transmission of observational warrant from premises to conclusion in cases like Dretske's and Nozick's examples. It is also intended that this first principle should apply to Moore's (1959) putative anti-sceptical argument since, intuitively, an everyday observational warrant for believing that here is a hand and here is another does not amount to a justification for denying radical scepticism (Section 6). Moore's argument is naturally described as question begging, and we shall defend 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 (Section 5).²⁰

In Dretske's and Nozick's examples, and in Moore's argument, it is not the case that both premises are known by armchair reflection. So, even if the first stage of our response were to be completed successfully, it would be open to someone to insist that while empirical warrants are not always transmitted, still there can be no failure of transmission for armchair warrants (or for a priori warrants). The aim of the second stage of our response is to transpose the principle for limiting the transmission of observational warrant into a principle for limiting the transmission of the particular kinds of armchair

¹⁸ On Gail Stine's (1976) contextualist account of knowledge, Dretske's zebra example would be a case of closure of knowledge without transmission of warrant. Wright (2000, Section 7) argues that the zebra case and others of Dretske's examples are indeed failures of transmission rather than of closure.

¹⁹ This distinction has often been missed in the literature on epistemic closure but one writer who has consistently emphasised it is Wright. For a clear statement of the distinction see e.g. his (1985, p. 438, note 1).

²⁰ Wright (1985) argues that Moore's putative 'Proof of an external world' founders on a failure of transmission of warrant. Something like the connection that I shall make between begging the question and failure of transmission of warrant is presumably implicit in the title of Wright (2000) though he does not elaborate on the notion of question begging.

justification that figure in Kylie's puzzle (Section 8). However, we also need to examine an important ancestor of Kylie's puzzle that is found in the work of Michael McKinsey (1991), since it is far from obvious that our strategy for responding to Kylie's puzzle would apply to the *reductio* argument that McKinsey presents (Section 7).

5. Begging the Question and Transmission of Warrant

Suppose that someone suspects that the animal in the pen is a cleverly disguised mule and that a speaker attempts to convince this suspicious hearer of the truth of Zebra(3) by advancing the argument in Dretske's example. Even in the absence of any further details about the case, there is a quite strong intuition that the speaker is begging the question. If we imagine a speaker who advances Moore's argument in an attempt to convince a sceptic of the existence of the external world then the intuition of question begging is even stronger. However, in order to know what to make of these intuitions, we need an account of what it is to beg the question. Once we have an account, we can ask how this dialectical phenomenon is related to questions about transmission of epistemic warrant in the thinking of a single subject.²¹

5.1 *Begging the question*

In an important discussion of begging the question on which we shall rely heavily, Frank Jackson (1987, p. 107) argues that: 'The utility of valid argumentation . . . in convincing audiences of conclusions lies in the evidence implicitly offered for borrowing by the presentation and selection of premises.' By advancing an argument, rather than flatly asserting the conclusion, the speaker invites the hearer to borrow his evidence; and by his choice of premises the speaker provides some information about what the nature and structure of this evidence is.

Because advancing an argument with a particular choice of premises indicates what kind of evidence is available for borrowing, it is possible for a speaker to engage in what Jackson calls 'misleading advertising' (1987, p. 108). But in order to isolate 'begging the question proper', Jackson asks us to consider examples in which an argument is advanced and the advertising of evidence for borrowing is not misleading. What is important for begging the question is the particular way in which evidence that is available for borrowing supports the premises of the argument. The issues are particularly clear if set out in terms of a Bayesian account of the confirmation of hypotheses by evidence.

So long as we ignore the existence of background assumptions, we can say that evidence *E* supports hypothesis *H* if the probability of *H* given *E*, $\text{Pr}(H/E)$, is greater than the prior probability of *H*, $\text{Pr}(H)$. Once background assumptions are in play, *E* supports *H* relative to background assumptions *B* if the probability of *H* given *E* plus the background assumptions *B*, $\text{Pr}(H/(E\&B))$, is greater than the prior probability of *H* given *B* alone, $\text{Pr}(H/B)$. It can happen that *E* supports *H* relative to background assumptions *B*₁ but not relative to background assumptions *B*₂. So, in particular, it may happen that *E* supports *H* relative to a speaker's background assumptions but not relative to the background assumptions of a hearer.

Suppose, then, that a speaker propounds an argument to a hearer who needs to be convinced of the argument's conclusion *C*. Suppose, too, that *A* is among the argument's premises. By advancing the argument, the speaker advertises that he has evidence of a certain kind for *A*. We are supposing that this advertisement is not misleading. So the

²¹ I am grateful to Helen Beebe for raising the question about the relationship between begging the question and transmission of warrant, and for helping me to get clear about my answer to it.

evidence offered for borrowing is available and does indeed support *A relative to the speaker's background assumptions*. But it does not follow that the evidence supports *A* relative to the hearer's background assumptions. In fact, the hearer's doubt about *C* may virtually guarantee that the speaker's offered evidence does not support *A* relative to the hearer's background assumptions and, in that case, advancing the argument will be ineffectual. It will not provide the hearer with any grounds for accepting the conclusion *C* because the offered evidence will not provide the hearer with any grounds for accepting the premise *A*.²²

In summary, then, it may be that (1987, p. 111):

[A] given argument to a given conclusion is such that anyone – or anyone sane – who doubted the conclusion would have background beliefs relative to which the evidence for the premises would be no evidence.

Of this kind of case, Jackson says (p. 112): 'Such an argument could be of no use in convincing doubters, and is properly said to beg the question.'

5.2 Justification and the resolution of doubt

Now that we have an account of begging the question, we can turn to the relation between this dialectical phenomenon and the transmission of epistemic warrant from premises to conclusion. To this end, we shall borrow and adapt an example from Crispin Wright.²³ It has the evidential structure of begging the question proper, but now within the thought of a single subject. Acceptance of the background assumptions that are required if the subject's own evidence for the premises is to be evidence cannot be rationally combined with doubt about the conclusion.

At the MCG,²⁴ I see what seems to be an Australian Rules football match in progress, and one of the players kicks the ball between the tall central uprights. The crowd roars, the goal umpire makes a sign with his hands and then waves two flags. Ordinarily, this would count as overwhelming evidence that a goal had been scored. So I would have a warrant for believing the first premise of the following obviously valid argument:

Goal(1) A goal has been scored.

Goal(2) If a goal has been scored then a football match is in progress.

Therefore:

Goal(3) A football match is in progress.

But there are imaginable background beliefs relative to which the apparent evidence for that premise would be no evidence.

If I believe that what I am watching is the seventeenth take for a scene in a movie, then the ball sailing between the uprights, the roaring crowd and the two flags do nothing to confirm the hypothesis that a goal has genuinely been scored. Even if I believe that it is an open question whether there is a football match or a movie scene or some other

²² In the simplest case, the conclusion *C* is entailed by the premise *A* and *C* is itself one of the speaker's background assumptions *B*. Since this is not to be a case of misleading advertising, the offered evidence, *E*, supports the premise *A* relative to the speaker's background assumptions; so $\Pr(A/E\&B) > \Pr(A/B)$. Also, since *A* entails *C*, $\Pr(C/E\&B) \geq \Pr(A/E\&B)$. But it does not follow that the speaker's evidence, *E*, provides support for the conclusion *C* even relative to the speaker's own background assumptions. For those assumptions include *C* and so $\Pr(C/E\&B) = \Pr(C/B) = \Pr(C/C) = 1$.

²³ See Wright (1985, p. 436) and, in much more detail, Wright (2000, Section 2).

²⁴ The Melbourne Cricket Ground is used for Australian Rules football several times each weekend during the winter.

elaborate pretence before my eyes then, for me, the evidence does not support the claim that a goal has really been scored. Relative to these background beliefs, the trajectory of the ball and the behaviour of the players, the umpires and the crowd do not amount to evidence that supports Goal(1).

However, the mere possibility of such background beliefs does not show that advancing the argument would be begging the question as Jackson defines that notion. For it is not true that anyone sane who doubted the conclusion, Goal(3), would have background beliefs relative to which the offered evidence for the premise Goal(1) would be no evidence.²⁵ In order to generate a clear case of begging the question we can make one of two changes to the example. We can modify the setting of the example so that anyone sane who doubted Goal(3) in that setting would believe that they were watching a movie scene. We could do this by making the movie scene possibility especially salient. Alternatively, and more simply, we can adjust the argument itself so that the conclusion actually speaks of the movie scene possibility:

Goal(1) A goal has been scored.

Goal(2') If a goal has been scored then a football match is in progress and this is not just a movie scene.

Therefore:

Goal(3') This is not just a movie scene.

Advancing this argument and advertising the availability of the evidence of the ball, the crowd and the flags in support of the premise Goal(1) could be of no use in convincing a hearer who doubted Goal(3'). It really would be begging the question. And if I were to doubt whether this was a real match or just a movie scene then reviewing that evidence in an attempt to resolve my own doubt would be no less futile.

Beginning in a state of doubt and appreciating the futility of relying on the evidence of ball, crowd and flags for a resolution of that doubt, I might seek out an answer to the question whether it is a football match or a movie scene that I am watching. Suppose that I gather independent evidence that it is, after all, a genuine match. Given this background information, I do have a justification or warrant for believing that a goal has been scored; and intuitively this warrant is provided by the evidence of ball, crowd and flags. But, Crispin Wright (2000, p. 142) says, '[I]t would be absurd to regard that warrant as transmissible across the entailment to [Goal(3)]. You don't get any *additional* reason for thinking that a game is in progress by having the warrant for [Goal(1)]'.

It seems to me that Wright is correct to say that there is a failure of transmission of warrant here although it is clearer if we consider the argument in which the conclusion explicitly mentions the movie scene possibility. There is a background assumption that is required if my evidence for Goal(1) is to count as evidence. But acceptance of this background assumption cannot be rationally combined with doubt about Goal(3'). So the particular evidential justification that I have for believing Goal(1) does not provide me with a resource that can be used to settle a doubt about Goal(3'). If I came to doubt the independent evidence that I had gathered and so started to wonder whether, after all, I

²⁵ Someone arriving at the MCG at 2.15pm on a winter Saturday might doubt that a football match is in progress because he believes that the day's fixture is an evening rather than an afternoon match. Or perhaps he believes that it is an afternoon match but that there are still several minutes to go before the start. Entering the stand with this belief he sees one of the players kick the ball between the goalposts and he hears the crowd roar. He realises that a goal has been scored and that the match is already in progress. (Afternoon matches start at 2.10pm.)

was just watching a movie scene, the evidence of ball, crowd and flags would be of no help to me in resolving that doubt.

5.3 *Transmission of warrant in the absence of doubt*

On a Saturday during the football season there is no serious prospect that the MCG would be used as a film set. As I sit in the stand, it does not occur to me to doubt that it is a genuine match that I am watching. Against the background of my assumption (my not doubting) that a football match is in progress, the trajectory of the ball and the behaviour of the players, the umpires and the crowd add up to a very good justification or warrant for believing that a goal has been scored. In addition, a small amount of conceptual analysis provides me with a compelling justification for believing the conditional premise, Goal(2'), since it is only in the context of a genuine match that a goal can be scored. Does my observation of the ball, the crowd and the flags, combined with my elementary piece of conceptual analysis, add up to a justification or warrant for believing that a football match, rather than a movie scene, is in progress?

It was in these terms that we introduced the idea of transmission of epistemic warrant (Section 4). But someone might suggest that, if the idea of transmission is understood in these terms, then warrant is indeed transmitted in the case that we have just described. Since I believe the premises of a palpably valid argument, I am surely justified in believing the conclusion as well; indeed, it would be irrational for me not to do so. This suggestion about transmission of epistemic warrant has to be rejected. The crucial question is not whether I am right to believe that a football match is in progress, nor even whether I have some epistemic warrant for believing it. The question is whether the epistemic warrants that I have for believing the premises add up to an epistemically adequate warrant for the conclusion.

Against the background of the assumption that a genuine Australian Rules football match is in progress, the evidence of ball, crowd and flags counts in favour of the hypothesis that a goal has been scored and against a host of alternative hypotheses. For example, the evidence counts against the hypothesis that only a behind has been scored,²⁶ and against the hypothesis that the ball has been kicked out of bounds. In short, the evidence rules out various ways in which the hypothesis that a goal has been scored could have been false, and it is for this reason that the evidence provides a resource for resolving doubt. It is also by ruling out alternatives that the evidence confers knowledge. This is how evidence constitutes an *epistemic* warrant. But the evidence of ball, crowd and flags does not count in favour of the hypothesis that a football match rather than a movie scene is in progress and against alternative hypotheses. The evidence, even taken together with the considerations that support the conditional premise, does nothing to rule out the most obviously salient alternative hypothesis, namely, that it is a movie scene that I am watching. That evidence would be of no help in resolving doubt and it does not confer knowledge. My epistemic warrants for the premises do not add up to an epistemic warrant for the conclusion.

In fact, there are two ways of conceiving my justificatory situation when, on a winter Saturday at the MCG, the movie scene possibility is utterly ignored. On one account, what I see of ball, crowd and flags constitutes my justification for believing that a goal has been scored. The background assumption that a football match, rather than a movie scene, is in progress is one that I am entitled to make, without justification, in the circumstances

²⁶ A behind is scored if the ball is kicked between the shorter outer uprights but not between the tall central uprights.

of being at the MCG on a Saturday afternoon during the football season. It is a substantive philosophical question what the nature of this entitlement would be. But one defining mark of the entitlement is that the dependence of my evidential justification on that unjustified assumption does not prevent my knowing that a goal has been scored. On the other account, my justification for believing that a goal has been scored properly consists in what I see of ball, crowd and flags *together with* some justification for the background assumption that it is a football match, and not a movie scene, that is in progress. On both accounts, the evidence of ball, crowd and flags contributes to an epistemic warrant for Goal(1) but contributes nothing to an epistemic warrant for Goal(3').

6. Moore's Argument and a Limitation Principle

Failure of transmission of epistemic warrant is the analogue, within the thought of a single subject, of the dialectical phenomenon of begging the question.²⁷ We said at the end of Section 4 that Moore's argument:

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.

is naturally described as question begging and, unsurprisingly, Jackson classifies it as a case of question begging proper. Anyone sane who doubted the conclusion, Moore(3), would have background beliefs relative to which the perceptual evidence for Moore(1) that is offered for borrowing would be no evidence.²⁸ Acceptance of the background assumptions that are required if the offered evidence for Moore(1) is to be evidence cannot be rationally combined with doubt about Moore(3). So, if I were beset by sceptical doubt, then attempting to resolve that doubt by reviewing the perceptual evidence for Moore(1) would be a futile exercise.

In ordinary circumstances, it does not occur to me to doubt that there is an external world. Against the background of my assumption (my not doubting) that there is an external world, my perceptual experience adds up to a very good justification for believing Moore(1). An analysis of the concepts of hand and of external world provides me with a justification for believing the second premise, Moore(2). But the obvious parallel with the case of a goal being scored shows that my observation of two hands, combined with my modest piece of conceptual analysis, does not add up to an epistemic warrant for the conclusion that an external world exists.

As we continue with the parallel, there seem to be two ways of conceiving the everyday justificatory situation. On one account, the background assumption that there is an external world is one that I am entitled to make without justification.²⁹ The role of this

²⁷ In a case of begging the question, the hearer is provided with no grounds for accepting either the premise or the conclusion. Helen Beebe has raised the point that this makes it difficult to see how there can be an analogy with failure of transmission of warrant since, from the hearer's point of view, there is no warrant to transmit. The response to this worry is that we need to focus on the situation of the speaker, rather than the hearer, in a case of begging the question.

²⁸ See Jackson (1987, p. 113). Commenting on Moore's argument, Wright says (1985, 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.'

²⁹ In essence, this is the proposal that Wright (1985) canvasses though it is there set in the context of the suggestion that propositions like Moore(3) are not 'fact-stating'.

unjustified assumption in the justification of everyday beliefs like the belief that here is a hand and here is another does not prevent those beliefs from being knowledge.³⁰ On the other account, a proper justification for my belief that here is a hand and here is another would have to include a justification of the background assumption that there is an external world. The decision between these accounts is not a trivial matter; large issues in epistemology turn on it.³¹ But both accounts agree that the epistemic warrant for Moore(1) that is furnished by perceptual experience is not transmitted to Moore(3).

6.1 Completing the first stage of the plan

The aim of the first stage of the two-stage plan announced at the end of Section 4 was to uncover a principle for limiting the transmission of observational warrant from premises to conclusion in cases such as Moore's argument as well as Dretske's and Nozick's examples. In an earlier paper (1998), I suggested a limitation on transmission of epistemic warrant along the following lines:

First Limitation Principle (first 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.

The motivation for this principle was straightforward. Our reflections on examples such as Moore's argument are supposed to be captured by the idea that the truth of Moore(3) is a 'precondition' for our observational warrant for Moore(1) to count as a warrant. But this initial formulation of the principle is problematic, not least because it involves the unexplained notion of a precondition.³² If the notion is interpreted so that a precondition is simply a necessary condition then the First Limitation Principle is certainly open to counterexamples. For example, Bill Brewer (this volume; see also 1999) points out that, on that interpretation of the notion of precondition, the principle has the consequence that warrant can never be transmitted to a necessary truth – a result that is disastrous for logic and mathematics. In addition, on that interpretation the principle has the somewhat counterintuitive consequence that warrant could never be transmitted to the proposition that I have a warrant for something or that someone has a warrant for something.

Our discussion of begging the question suggests a better way of codifying the idea behind the First Limitation Principle:

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.

Where the original version speaks (in effect) of the conclusion C's being a necessary condition for the warrant to count as a warrant, the revised version says, instead, that doubt about C cannot be rationally combined with acceptance of certain background

³⁰ It is consistent with the second way of conceiving the everyday justificatory situation to say that the background assumption is itself known without justification. In that case, we would have closure of knowledge without transmission of warrant.

³¹ Because it is not clear where a justification of the background assumption could come from, the second way of conceiving the everyday justificatory situation tends in the direction of scepticism.

³² This worry about the notion of precondition is explicit in Davies (1998, p. 352) and is echoed by Wright (2000).

assumptions. There are two features of this revised version that promise to make it less open to counterexamples than the original version. One is the introduction of the notion of *rational combination*; the other is the introduction of an explicit mention of *background assumptions*.³³

First, the original version's use of the notion of a necessary condition has the result that if the principle blocks transmission of warrant to C then it also blocks transmission of warrant to any proposition that is entailed by C and, in particular, to any logical truth. But the notion of rational combination is to be interpreted in such a way that it will often be possible for a thinker rationally to combine acceptance of one proposition, A, with doubt about another proposition, B, even though A entails B. We shall say that a thinker *cannot* rationally combine acceptance of A with doubt about B only when doubt about B *immediately* constitutes a reason for not accepting A. With the notion of rational combination interpreted in this way, it may be that acceptance of A cannot be combined with doubt about B, but can be rationally combined with doubt about some proposition that is entailed by B. It may also be that acceptance of various propositions can be rationally combined with doubt about a proposition that is, in fact, necessarily true, or even a priori necessarily true.³⁴

A version of the principle incorporating just this first improvement would be as follows:

Epistemic warrant cannot be transmitted from the premises of a valid argument to its conclusion if, for one of the premises, doubt about the truth of the conclusion cannot be rationally combined with acceptance that the warrant for that premise counts as a warrant.

But it is plausible that one cannot rationally combine acceptance that a particular warrant counts as a warrant for a certain premise with doubt about whether one has, or whether anyone has, a warrant for anything. So this version would still have the consequence that epistemic warrant could never be transmitted to the proposition that I have a warrant for something or that someone has a warrant for something. To the extent that this is genuinely a counterintuitive consequence, the problem lies in the fact that the half-improved version of the principle does not yet say anything about background assumptions.

The second attractive feature of the revised version that is motivated by our discussion of begging the question is that it is explicit that there is a role for background assumptions to play. It may be that evidence E supports a hypothesis H only relative to certain background assumptions. But we do not need to include amongst those assumptions the very proposition that E supports H. It may be that my justification for believing the premise of an argument is constituted by certain evidence and that my justification counts as a justification only against the background of certain assumptions. But it does not follow that the proposition that my evidence counts as a warrant must

³³ Wright's proposed account (2000, Section 2) of examples in which there is a failure of transmission for what he calls 'information-dependent' warrants already incorporates, in effect, at least the second of these two features. In fact, the importance of the role of background assumptions should have been clear from Wright (1985; see e.g. p. 436) but it was not reflected in the original version of the First Limitation Principle (Davies, 1998, p. 351.)

³⁴ This still leaves it open that there may be a proposition, B, such that doubt about B cannot be rationally combined with acceptance of any propositions. The revised version of the First Limitation Principle, as presently formulated, has the consequence that epistemic warrant cannot be transmitted to such a proposition. Similarly, any argument advanced in support of such a proposition would be *vacuously* question begging. We shall not pursue here the question whether the vacuous case should be ruled out.

itself be included amongst those background assumptions. Nor does it follow that acceptance of those background assumptions cannot be rationally combined with doubt about whether anyone has a warrant for anything.

There can be no doubt that even this revised version of the First Limitation Principle will face counterexamples and will require further modifications. This is not the place to pursue the matter at any length, but we can very briefly indicate two ways in which the principle could be improved further.

First, although the principle talks about background assumptions there is still some unclarity about the way in which these assumptions have to be accepted. At a couple of points we have spoken of not doubting as an alternative to explicitly assuming and, in fact, absence of doubt seems to be the way in which background assumptions usually figure in our justificatory practices. Further work on the First Limitation Principle should lead to a clearer specification of the role of background assumptions in the justification of beliefs.

Second, it would be natural to interpret the principle as indicating that there is some unique set of background assumptions relative to which a warrant for some proposition A counts as a warrant. But this indication of uniqueness does not seem quite right. We begin, perhaps, with a default set of assumptions; we do not doubt that certain possible circumstances do not obtain. If any one of these circumstances were to obtain then our putative warrant for A would not, after all, be a warrant. Also, if we were to doubt that these circumstances do not obtain – if we were seriously to consider that these circumstances might obtain – then we would no longer be justified in believing A. These circumstances are defeaters of our warrant. But the defeaters may themselves be defeated and warrant may be restored. So our warrant may count as a warrant relative to a set of background assumptions that is different from the default set. Further work should incorporate the notion of defeaters and defeaters of defeaters (and so on) into the First Limitation Principle.³⁵

Even without these further improvements, the principle seems to give a reasonable account of the failure of transmission of observational warrant in Moore's argument, in Dretske's and Nozick's examples, and in the case of my seeing a goal being scored. To the extent that this principle enjoys a measure of plausibility, we can allow that the first stage of the two-stage plan that was announced at the end of Section 4 is complete.

6.2 Problems for the second stage of the plan

The second stage of the plan was to transpose the principle that limits the transmission of observational warrant into a principle for limiting the transmission of the particular kinds of armchair justification that figure in Kylie's puzzle. But this now looks problematic. The motivation for the First Limitation Principle comes from examples in which entertaining a doubt undermines or defeats a warrant that is constituted by observational evidence. The warrant is undermined because acceptance of certain (default) background assumptions cannot be rationally combined with the doubt. But it is far from clear how this idea could be applied in the case of Kylie's puzzle.

The first problem is that neither self-knowledge nor the kind of knowledge that comes from philosophical theorising seems to be based on observational evidence. In response to

³⁵ I am grateful to Michael Martin for discussion of the issues raised in these last two paragraphs. The indication of uniqueness that is present in the First Limitation Principle (revised version) is not a feature of Jackson's (1987) account of begging the question, so it is not obvious that the envisaged refinements would involve any departure from the idea of a close connection between begging the question and failures of transmission of warrant.

this first problem it might be said that the notion of observational evidence figures in the examples that motivate the revised version of the First Limitation Principle but does not figure in the principle itself. All that is really required for a successful transposition is that there should be background assumptions for self-knowledge or for philosophical reasoning. But although this response is correct, it highlights two further problems.

The second problem is that it is not at all obvious what the (default) background assumptions for self-knowledge or for philosophical theorising are. The third problem is that it is even less obvious that acceptance of those assumptions, whatever they may be, cannot be rationally combined with doubt about whether I am embedded in an environment that contains samples of water.³⁶

However, despite these problems for the second stage of the announced plan, I shall, in the final section of the paper, propose a principle that limits the transmission of armchair (or a priori) warrant. Before that, we need to consider an important ancestor of Kylie's puzzle.

7. McKinsey's Original Recipe

Michael McKinsey (1991) presented a *reductio* argument that has been widely regarded as establishing a model for puzzles like Kylie's. Brian McLaughlin and Michael Tye speak of 'McKinsey's recipe for trying to show that [a] version of externalism is incompatible with privileged access' and describe the recipe as follows (1998, p. 290):

[F]ind some E such that (1) E cannot be known a priori . . . yet (2) the version of externalism implies that it is a conceptual truth that if one is thinking that p, then E. According to McKinsey, if such an E can be found, it can be successfully argued that *either* one cannot know a priori that one is currently thinking that p *or* it is not a conceptual truth that if one is thinking that p, then E, from which it may be inferred that the version of externalism is incompatible with privileged access.

They continue (*ibid.*): 'McKinsey's recipe is, we believe, a perfectly fine one.'

In the passage quoted at the beginning of this paper, McKinsey does not offer any explicit defence of the truth of the conditional statement about a priori knowledge (1991, p. 16): '[I]f you could know a priori that you are in a given mental state, and your being in that state conceptually or logically implies the existence of external objects, then you could know a priori that the external world exists.' He merely asserts (*ibid.*): 'It's just that simple.' So from that passage alone it is not clear whether McKinsey has in mind what we have called the achievement problem for self-knowledge given externalism, or the consequence problem, or some other problem. But elsewhere in his paper he spells out the argument for *reductio* in terms of a thinker being able to deduce E from premises that are knowable a priori, and so being able to know E a priori as well. It is for this reason that I regard McKinsey's *reductio* argument as being an ancestor of Kylie's puzzle.

If we are to carry through the second stage of our announced plan for responding to Kylie's puzzle, then we must defend the idea that even armchair or a priori warrants may depend on empirical background assumptions. But at just this point a distinctive feature of McKinsey's original presentation of his recipe becomes important. McKinsey (1991, p. 9) says that a priori knowledge is 'knowledge obtained independently of empirical

³⁶ Because of these striking differences between Moore's argument and Kylie's puzzle, some philosophers have proposed that an argument rather like the one in Kylie's puzzle can be used effectively against a sceptic about the external world even though Moore's argument cannot. See, for example, Warfield (1998).

investigation', but he 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*.

Let us say that *strict a priori* knowledge is knowledge that is not justificatorily based on empirical investigation and *also involves no empirical assumptions*.³⁷ The lesson that McKinsey draws from his reductio argument is this. If thinking that *p* is conceptually or logically dependent on environmental conditions (about which I could not have strict a priori knowledge) then I cannot have strict a priori knowledge that I am thinking that *p*. So we cannot know about our own externalist mental states without either launching an empirical investigation or making some assumptions about the external physical world.

We have already noted that there are two ways of conceiving the justificatory situation when background assumptions are in play. On one account, what we ordinarily take to be the justification really does constitute the justification for believing the proposition in question, and we are entitled to make the background assumptions without justification. On the other account, the justification for believing the proposition in question properly consists of what we would ordinarily take to be the justification together with a justification for the background assumptions. There are also two ways of conceiving the significance of the lesson that McKinsey draws from his reductio argument.

On the first account, the status of my self-knowledge as knowledge would not depend on my having a justification for the empirical assumptions that I made. Instead, I would have an entitlement to make some empirical assumptions without justification. On this account, the lesson that McKinsey draws about strict a priori knowledge would be consistent with our knowledge about our own externalist mental states being a priori in the weaker sense of being not justificatorily based on empirical investigation.

On the second account, the status of my self-knowledge as knowledge would depend in part on my having a justification for any empirical assumption that I made and this justification would presumably have to involve empirical investigation. Thus, any dependence on empirical background assumptions would already be enough to ensure that self-knowledge does not even meet the standard of being not justificatorily based on empirical investigation. So, on this account, there is no significant difference between a priori knowledge as initially defined and strict a priori knowledge.

It is clear from a later paper (McKinsey, forthcoming) that McKinsey himself favours the second account of the justificatory situation when background assumptions are in play. He explicitly argues that there is no relevant difference between 'knowledge that has certain empirical presuppositions' and 'knowledge that is directly based on empirical investigation'. This, of course, is why, in the original paper, he moves without comment from the apparently less strict notion of a priori knowledge to the strict one.

We cannot expect there to be any limitation on the transmission of strict a priori warrant in McKinsey's sense. So suppose, for a moment, that I could know strictly a priori that I am thinking that water is wet, and could know strictly a priori that if I am thinking that water is wet then environmental condition *E* holds. We have no reason to deny that the strict a priori warrant for those premises could be transmitted to the conclusion that condition *E* holds. But, as McKinsey points out and as we must surely

³⁷ It is not clear that there could be any motivation for ruling out empirical assumptions about the external physical world while allowing empirical assumptions about, for example, the workings of my brain.

agree, it is absurd to suppose I could know strictly a priori that E holds. McKinsey's reductio argument is compelling when it is cast in terms of strict a priori knowledge as he defines that notion. On that construal, McKinsey's recipe is indeed 'a perfectly fine one'.

Someone who accepts the a priori truth of externalist dependence theses and who wants to maintain the idea of authoritative self-knowledge must insist on a distinction between strict a priori knowledge and knowledge that is a priori in the weaker sense of being not justificatorily based on empirical investigation. As we have seen, making this distinction involves allowing for the idea of epistemic entitlement without justification.

With the distinction in place, it is possible to respond to McKinsey's reductio argument by allowing that self-knowledge is not strictly a priori but maintaining that it is, nevertheless, not justificatorily based on empirical investigation.³⁸ In fact, for reasons that have nothing to do with externalism, it is implausible that our knowledge of our own mental states is strictly a priori in McKinsey's sense. For example, Christopher Peacocke considers a range of ways in which the causal processes that underlie our normal ability to make knowledgeable self-ascriptions of mental states can go wrong. Some of these ways are associated with symptoms of schizophrenia; others are more science-fictional. The upshot is that there can be dissociations between a conscious thought's being initiated by a subject and a conscious thought's being experienced as self-initiated. Peacocke concludes (1999, p. 244–5):

There is then a reliance, in everyday conscious self-ascriptions, on these dissociations not actually obtaining. The ordinary self-ascriber is entitled to such presuppositions. . . . [A]lthough there are many deep respects in which self-knowledge cannot be assimilated to perceptual knowledge, there is, even in consciously based self-ascription, reliance on a network of causal relations whose obtaining is by no means necessary.

But even if we can offer an adequate response to McKinsey's reductio argument as he originally presented it, the argument remains important because its pattern can be repeated for a more relaxed notion of a priori knowledge than the one that McKinsey himself favours. This is what Boghossian's reductio argument shows³⁹ and, of course, Kylie's puzzle can be cast in terms of a priori knowledge as well. Kylie knows a priori (in a way that is not justificatorily based on empirical investigation) that she is thinking that water is wet. She also knows a priori that if she is thinking that water is wet then she is embedded in an environment that contains samples of water. But it is counterintuitive that the use of just these a priori methods should provide Kylie with knowledge that there is water in her environment.

8. A Second Limitation Principle and the Resolution of Kylie's Puzzle

At the end of Section 6, we noted some problems for the idea of adapting the First Limitation Principle to block the transmission of epistemic warrant in Kylie's puzzle. After our consideration of McKinsey's original recipe in Section 7, we know that there must be background assumptions for self-knowledge that we are entitled to make without justification. But this does not yet tell us how to move from the First Limitation Principle to something that will resolve Kylie's puzzle.

³⁸ It would seem that something like this thought is pivotal to Wright's (2000, Section 7) argument that McKinsey's reductio argument falls foul of a failure of transmission of warrant. However the details of Wright's diagnosis are, at least on the surface, quite different from the resolution that I propose. I hope to return in future work to a comparison and evaluation of the two accounts.

³⁹ See also Brown (1995).

In the earlier paper (Davies, 1998) that I have already mentioned, I proposed a second principle that was intended to apply to puzzle cases like Kylie's:

Second Limitation Principle (first 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.

The idea behind this principle was clear. In the case of Moore's argument or Dretske's puzzle, if the conclusion were false then our putative warrant for the first premise would not, after all, be a warrant, though that premise would (let us allow for the purposes of this discussion) still be something that we could entertain. In the case of Kylie's puzzle, if the conclusion were false then, according to the externalist philosophical theorising that supports the second premise, I could not think that water is wet. What is more, I could not think any thought in which the concept of water is deployed and so, in particular, I could not even think that I am thinking that water is wet.⁴⁰ So, if the conclusion were false then I could not believe or even entertain the first premise. The Second Limitation Principle simply makes this relationship between conclusion and premise the triggering condition for failure of transmission of warrant and so is guaranteed to have the desired result. Unfortunately, however, the Second Limitation Principle, in that early version, appears both desperately ad hoc and open to counterexamples.

The Second Limitation Principle makes use of the notion of a precondition and so faces all the problems associated with that notion. In addition, the principle has the somewhat counterintuitive consequence that warrant could never be transmitted to the proposition that I am able to believe something or that someone is able to believe something.⁴¹ These worries about counterexamples are serious. But what is more important is to provide a principle that not only offers a resolution of Kylie's puzzle but also draws on some of the motivation for the revised version of the First Limitation Principle, so avoiding the charge of being ad hoc. If we could achieve that much then discussion of counterexamples could be left for another occasion.

But we face serious problems. It is not at all obvious what the background assumptions for self-knowledge or for philosophical theorising are, and it is even less obvious that acceptance of those assumptions cannot be rationally combined with doubt about whether I am embedded in an environment that contains samples of water. These problems are addressed in the next two subsections.

8.1 Towards a motivated principle

As we saw in Section 7, Peacocke argues that in the case of self-knowledge, as in the case of perceptual knowledge, there is 'reliance on a network of causal relations whose obtaining is by no means necessary' (1999, p. 245). But I want to focus on an even more basic background assumption, not only for self-knowledge, but for any kind of justification or warrant.⁴² For any proposition, A, if there were no such proposition as A, if there were no such thing to think, to entertain, to believe, to doubt, to confirm or to disconfirm as the proposition A, then there could be no question of anything justifying

⁴⁰ Here we make use of 'the fact about embedding' that was important for work on the achievement problem for self-knowledge given externalism.

⁴¹ This counterintuitive consequence is explicit in Davies (1998, p. 354) and is also mentioned by Wright (2000).

⁴² Even when we have improved limitation principles available (see the next subsection) it does not appear to be possible to resolve Kylie's puzzle by appealing only to the background assumptions that Peacocke mentions.

my believing A. So a putative warrant for A counts as a warrant only against the background of the assumption that there is such a proposition as the proposition A.

If we incorporate just this idea into the Second Limitation Principle then the result is as follows:

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 there being any such proposition for the knower to think as that premise.

But now, in order to move the principle closer to the revised version of the First Limitation Principle and closer to the notion of begging the question, we need to make a second change. Instead of talking about preconditions or necessary conditions the principle should speak of rationally combining doubt about the conclusion with acceptance of what we have now identified as a background assumption:

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.

The principle that we have arrived at seems to be well motivated rather than ad hoc. In fact, there is only one difference between this principle and the revised version of the First Limitation Principle. Where the first principle speaks of the warrant for a premise counting as a warrant only against the background of certain assumptions, the principle that we have just formulated speaks of the specific assumption that there is such a proposition as that premise. We might, perhaps, allow that calling into question the assumption that there is such a proposition as the premise to be believed is simply a very radical way of undermining our justification for believing that premise. If we were to allow that then the principle that we have arrived at could be regarded as a consequence of the revised version of the First Limitation Principle.

Unfortunately, however, the principle that we have arrived at is not adequate to resolve Kylie's puzzle. The reason is that it is clearly possible rationally to combine doubt about the truth of the conclusion (Water(3)) that there are samples of water in my environment with acceptance that there is such a thing for me to think as the premise (Water(1)) that I am thinking that water is wet. Many philosophers who do not endorse externalist arguments would be prepared to combine that doubt with that acceptance. Certainly, doubt about Water(3) does not immediately constitute a reason for not accepting that there is such a thing for me to think as Water(1).

8.2 *The final step*

There is one more step that we need to take if we are to resolve Kylie's puzzle. We need to take full account of the fact that the argument that gives rise to Kylie's puzzle has more than one premise. Return, for a moment, to the dialectical phenomenon of begging the question. When a speaker advances an argument with several premises, he implicitly offers the hearer evidence for, or more generally considerations in favour of, each of the premises. It seems to be a legitimate extension of what Jackson actually says (1987, p. 111) to suggest the following. Advancing a multi-premise argument will be begging the question if anyone sane who doubted the conclusion *and accepted the evidence offered in support of all but one of the premises* would have background beliefs relative to which the evidence offered in support of *the remaining premise* would be no evidence. For that

would be enough to guarantee that the argument (1987, p. 112) ‘could be of no use in convincing doubters’.

It might be that (as in the original definition) anyone sane who doubted the conclusion would, whether or not they accepted the evidence offered in support of the other premises, have background beliefs relative to which the evidence offered in support of one of the premises would be no evidence. In such a case, the question begging nature of the argument is likely to be obvious. But suppose that what is offered in support of one of the premises is a complex and subtle piece of philosophical theory. Then it might not be obvious what assumptions could sanely be combined with acceptance of that philosophical theory and doubt about the conclusion of the argument.

The revised version of the First Limitation Principle was motivated by the idea that failure of transmission of warrant is the analogue, within the thought of a single subject, of begging the question. So we should expect the First Limitation Principle to incorporate the extension to multi-premise question begging thus:

First Limitation Principle (multi-premise 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.

Then we can incorporate the same extension into the improvement on the Second Limitation Principle that we have already arrived at:

Second Limitation Principle (multi-premise 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 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.

The second of these principles is far from being ad hoc since it is certainly closely related to the first and can perhaps even be regarded as a consequence of it. Furthermore, it promises to block transmission of warrant from the premises to the conclusion of the argument in Kylie’s puzzle.

The puzzle arises because of the threat that Kylie’s warrant for Water(1) and her warrant for Water(2) should jointly constitute a warrant for Water(3). Her warrant for Water(2) is a piece of philosophical theory that shows, in effect, that if there is to be such a thing for me to think as that water is wet then there must be samples of water in my environment. That same theory would equally have the consequence that if there is to be such a thing for me to think as that I think that water is wet then, likewise, there must be samples of water in my environment. So acceptance of the assumption that there is such a proposition for me to think as Water(1) and acceptance of the warrant for Water(2) cannot be rationally combined with doubt about the truth of Water(3).⁴³ The

⁴³ We have already agreed that doubt about the truth of Water(3) can be rationally combined with acceptance of the assumption that there is such a proposition for me to think as Water(1). But there are still two ways in which the Second Limitation Principle (multi-premise version) could be triggered. It could be that doubt about the truth of Water(3) cannot be rationally combined with acceptance of the warrant for the conditional premise Water (2). Or it could be that, while that combination is rationally possible, that pair of doubt and acceptance cannot rationally be further combined with acceptance that there is such a

multi-premise version of the Second Limitation Principle is triggered and our resolution of Kylie's puzzle is, at last, complete.

proposition for me to think as Water(1). Further consideration needs to be given to the first of these possibilities.

References

- Boghossian, P.A. 1997: What the externalist can know a priori. *Proceedings of the Aristotelian Society*, 97, pp. 161–75. Reprinted in C. Wright, B.C. Smith and C. Macdonald (eds), *Knowing Our Own Minds*. Oxford: Oxford University Press, pp. 271–84.
- Brewer, B. 1999: *Perception and Reason*. Oxford: Oxford University Press.
- Brown, J. 1995: The incompatibility of anti-individualism and privileged access. *Analysis*, 55, pp. 149–56.
- Brown, J. 1998: Natural kind terms and recognitional capacities. *Mind*, 107, pp. 275–303.
- Brueckner, A. 1990: Scepticism about knowledge of content. *Mind*, 99, pp. 447–51.
- Brueckner, A. 1994: Knowledge of content and knowledge of the world. *Philosophical Review*, 103, pp. 327–43.
- Burge, T. 1979: Individualism and the mental. In P.A. French, T.E. Uehling and H.K. Wettstein (eds), *Midwest Studies in Philosophy, Volume 4: Studies in Metaphysics*. Minneapolis: University of Minnesota Press, pp. 73–121.
- Burge, T. 1982: Other bodies. In A. Woodfield, (ed.), *Thought and Object: Essays on Intentionality*. Oxford: Oxford University Press, pp. 97–120.
- Burge, T. 1988: Individualism and self-knowledge. *Journal of Philosophy*, 85, pp. 649–63.
- Burge, T. 1996: Our entitlement to self-knowledge. *Proceedings of the Aristotelian Society*, 96, pp. 91–116.
- Burge, T. 1998: Reason and the first person. In C. Wright, B.C. Smith and C. Macdonald (eds), *Knowing Our Own Minds: Essays on Self-Knowledge*. Oxford: Oxford University Press, pp. 243–70.
- Davies, M. 1991: Individualism and perceptual content. *Mind*, 100, pp. 461–84.
- Davies, M. 1992: Perceptual content and local supervenience. *Proceedings of the Aristotelian Society*, 92, pp. 21–45.
- Davies, M. 1993: Aims and claims of externalist arguments. In E. Villanueva (ed.) *Philosophical Issues Volume 4: Naturalism and Normativity*. Atascadero, CA.: Ridgeview Publishing Company, pp. 227–49.
- Davies, M. 1996: Externalism and experience. In A. Clark, J. Ezquerro and J.M. Larrazabal (eds.), *Philosophy and Cognitive Science: Categories, Consciousness and Reasoning*. Dordrecht: Kluwer Academic Publishers, pp. 1–33. Reprinted in N. Block, O. Flanagan and G. Güzeldere (eds.), *The Nature of Consciousness: Philosophical Debates*. Cambridge, MA.: MIT Press, 1997, pp. 309–28.
- 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, pp. 321–61.
- Davies, M. and Humberstone, L. 1980: Two notions of necessity. *Philosophical Studies*, 38, pp. 1–30.
- Dretske, F. 1970: Epistemic operators. *Journal of Philosophy*, 67, pp. 1007–23.
- Evans, G. 1982: *The Varieties of Reference*. Oxford: Oxford University Press.
- Falvey, K. and Owens, J. 1994: Externalism, self-knowledge, and skepticism. *Philosophical Review*, 103, pp. 107–37.
- Gallois, A. and O’Leary-Hawthorne, J. 1996: Externalism and scepticism. *Philosophical Studies*, 81, pp. 1–16.
- Heil, J. 1988: Privileged access. *Mind*, 97, pp. 238–51.
- Jackson, F.C. 1987: *Conditionals*. Oxford: Blackwell Publishers.

- Jackson, F.C. 1998: Reference and description revisited. In J.E. Tomberlin (ed.), *Philosophical Perspectives 12: Language, Mind, and Ontology*. Oxford: Blackwell Publishers, pp. 201–18.
- McKinsey, M. 1991: Anti-Individualism and privileged access. *Analysis*, 51, pp. 9–16.
- McKinsey, M. forthcoming: Forms of externalism and privileged access.
- McLaughlin, B.P. and Tye, M. 1998: Externalism, Twin Earth, and self-knowledge. In C. Wright, B.C. Smith and C. Macdonald (eds), *Knowing Our Own Minds*. Oxford: Oxford University Press, pp. 285–320.
- Moore, G.E. 1959: Proof of an external world. In *Philosophical Papers*. London: Allen and Unwin, pp. 127–50.
- Nozick, R. 1981: *Philosophical Explanations*. Oxford: Oxford University Press.
- Peacocke, C. 1996: Entitlement, self-knowledge and conceptual redeployment. *Proceedings of the Aristotelian Society*, 96, pp. 117–58.
- Peacocke, C. 1998: Conscious attitudes, attention, and self-knowledge. In C. Wright, B.C. Smith and C. Macdonald (eds), *Knowing Our Own Minds*. Oxford: Oxford University Press, pp. 63–98.
- Peacocke, C. 1999: *Being Known*. Oxford: Oxford University Press.
- Putnam, H. 1975: The meaning of ‘meaning’. In *Mind, Language and Reality*. Cambridge: Cambridge University Press, pp. 215–71.
- Raffman, D. 1998: First-person authority and the internal reality of beliefs. In C. Wright, B.C. Smith and C. Macdonald (eds), *Knowing Our Own Minds*. Oxford: Oxford University Press, pp. 363–9.
- Sawyer, S. 1998: Privileged access to the world. *Australasian Journal of Philosophy*, 76, pp. 523–33.
- Stine, G.C. 1976: Skepticism, relevant alternatives, and deductive closure. *Philosophical Studies*, 29, pp. 249–61.
- Warfield, T.A. 1998: A priori knowledge of the world: Knowing the world by knowing our minds. *Philosophical Studies*, 92, pp. 127–47.
- Wright, C. 1985: Facts and certainty. *Proceedings of the British Academy*, 71, pp. 429–72.
- Wright, C. 2000: Cogency and question-begging: Some reflections on McKinsey’s paradox and Putnam’s proof. To appear in *Philosophical Issues*, 10, pp. 140–63.