

Inferential Knowledge, Externalism and Self-Knowledge

Ross Gareth Ford

University College London

Submitted for the degree of MPhil in Philosophy

1st April 2004

UMI Number: U602787

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U602787

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract

Privileged self-knowledge says, roughly, that we have non-empirical knowledge of our own thoughts. Externalism about mental content says, roughly, that our mental states are determined at least in part by our environment. It has been alleged that jointly assuming externalism about mental content and privileged self-knowledge are true has the consequence that any subject can have non-empirical knowledge of her own environment and this is intuitively absurd. The thesis investigates in various ways the problem arises and focuses on the following principle:

Knowledge Transmission Principle:

For any subject S and any proposition P and any proposition Q: if S has a certain kind of knowledge that P and S knows that P entails Q, then S knows that Q by way of S's certain kind of knowledge that P and S's knowledge that P entails Q.

After investigating the ways of denying the principle with a view to upholding the compatibility of privileged self-knowledge and externalism about mental content the thesis reaches the following conditional conclusions:

If we allow that the knowledge transmission principle is unrestrictedly true or if we allow for the possibility of so called 'illusions about mental content', then the consequence that a subject can have non-empirical knowledge of her own environment is not absurd.

If on the other hand certain other conditions are in place, then the relevant instance of the transmission principle fails, then we can jointly assume externalism about mental content and privileged self-knowledge are true without them having the absurd consequence that we have non-empirical knowledge of our environment as a result of this joint assumption.

Contents

Acknowledgements	p4
Introduction	p5
Chapter 1	p7
Introduction	p7
§1 Privileged self-knowledge	p8
§2 The relevant instance of externalism about mental content	
§3 outline of the paradox	p13
§4 Transmission and closure	p15
§5 Transmission Principles	p21
§6 Strategies of deriving non-empirical knowledge from externalism and self-knowledge	p24
§7 Summary: the road ahead for transmission failure compatibilists	p29
Chapter 2	p32
Introduction	p32
§1 Two rough arguments for transmission failure	p33
§2 Wright (2000)	p54
§3 Davies (2000)	p60
§4 Summary	p77
Chapter 3	p80
Introduction	p80
§1 Unproblematic non-empirical knowledge that I'm in a water-world	p81
§2 Externalism which allows for content-illusion	p83
§3 Davies: An epistemic principle which denies the possibility of content illusion	p87
§4 Summary	p89
Conclusion	p93
Bibliography	p94

Acknowledgements

I am indebted to Professor Paul F. Snowdon for his generosity of time and help with this thesis during my time as an MPhil student and also for his help with the dissertation I submitted for my BA in Philosophy at University College London (UCL) in 2001-2002 where I first embarked on some of the work presented in this thesis.

I would also like to thank Dr Marcus Giaquinto, Dr Lucy O'Brien and Dr José L. Zalabardo for their comments on various parts of this thesis.

I thank the UCL Friends' Programme for their scholarships in both 2002-2003 and 2003-2004. I would also like to thank UCL Philosophy Department for the Jacobsen Scholarship; and the Newby Trust Ltd. and The Royal Institute of Philosophy for their bursaries in 2003-2004.

Finally, my completing this thesis would not have been possible without the love and support of my parents and grandparents.

Introduction

It is widely held that we have non-empirical knowledge of the contents of our own thoughts (hereafter *self-knowledge*). We can know what we are thinking without examining the world. Externalism about mental content says roughly that the contents of our own thoughts are determined at least in part by our environment. Many philosophers also hold that some form of externalism about mental content is true. It has been claimed that jointly assuming certain sorts of these theses results in us having non-empirical knowledge of our environment and such knowledge is intuitively intolerable. This alleged counter-intuitive result invites the reaction that non-empirical knowledge of our own thoughts is incompatible with externalism about mental content.

In Chapter 1 of this thesis I set out what I take to be the most plausible argument for the conclusion that jointly assuming externalism and self-knowledge results in intuitively intolerable knowledge of our environment. I focus in one specific principle which underlies the transition from jointly assuming self-knowledge and externalism about mental content to the allegedly absurd conclusion that we can have non-empirical knowledge of empirical propositions namely:

Knowledge Transmission Principle:

For any subject S and any proposition P and any proposition Q: if S has a certain kind of knowledge that P and S knows that P entails Q, then S knows that Q by way of S's certain kind of knowledge that P and S's knowledge that P entails Q.

I then isolate a strategy which attempts to deny this principle in order to maintain the joint assumption that externalism about mental content is compatible with non-empirical knowledge of our thoughts. I call this strategy *transmission failure compatibilism*. This strategy is to be distinguished from *bullet biting compatibilism* which maintains the joint assumption of externalism about mental content and self-knowledge by arguing that certain non-empirical knowledge of our environment is not absurd. I set out this transmission failure compatibilism's main strategy and review each step of this strategy throughout the thesis.

In Chapter 2 of the thesis I critically evaluate the transmission failure compatibilist's discussion of the knowledge transmission principle in the case of perceptual knowledge. I then apply the results of my critical discussion in Chapter 2 to the SK paradox. I argue that given certain assumptions the transmission failure compatibilists response to the SK paradox fails. However I take this failure to align on various ways in which bullet biting compatibilism can be argued to be plausible. I also note that a certain variant of transmission failure compatibilism is correct, provided certain epistemic principles are in place.

Thus my overall conclusions are *conditional*. My conclusions are very roughly: if certain conditions obtain, then transmission failure compatibilism fails but this renders bullet biting compatibilism a sound position. If certain other conditions obtain, then a certain variant of transmission failure compatibilism is correct.

Chapter 1

Introduction

One recent discussion in Philosophy concerns the compatibility of privileged first-personal self-knowledge of the contents of our own thoughts (hereafter *self-knowledge*) with certain variants of externalism about mental content or anti-individualism. Self-knowledge says roughly that we can know the contents of our thoughts non-empirically and externalism about mental content says, roughly, that the contents of our thoughts depend at least in part on the external environment. Self-knowledge and externalism about mental content have been alleged to be incompatible when both notions are jointly assumed. This is because when jointly assumed they are alleged to have the consequence that we can know facts about our environment non-empirically and such knowledge is intuitively absurd. In this chapter I want to expound the sort of self-knowledge and externalism about mental content which give rise to the charge that the notions are incompatible. I also want to set out my view on what the epistemological principles are which allow us to make the charge of incompatibility and defend my view against recent objections.

In §1 I set out what I mean by privileged self-knowledge. In §2 I briefly describe the variant of externalism about mental content which I think gives rise to the paradox. In §3 I set out the beginnings of a paradox between externalism and self-knowledge. In §4 I discuss transmission and closure principles about knowledge and warrant in order to aid comprehension of subsequent sections. In §5, again in order to aid comprehension for subsequent sections, I discuss various ways of arguing for the conclusion that arguing for the conclusion that transmission principles fail. In §6 I finish sketching the paradox I introduced in §3 and show that there are at least two different versions of this paradox. I argue in this section that we should focus on only one of these versions. Moreover, I isolate the particular response to the paradox I favour and focus on investigating this response for the rest of the thesis. In §7 I sketch how I shall go about evaluating the response to the paradox I have chosen to focus on.

§1 Privileged self-knowledge

Suppose that you are asked: 'do you think Newcastle is north of London?' You may give one of the following two answers 'yes, I think Newcastle is north of London' or 'no Newcastle is not north of London' depending on your knowledge of the geography of England. Whatever answer you give, there is an intuition that you have not used empirical investigation in order to give any one of these answers. This intuition arises as a result of us being able to, typically, almost immediately give an answer to what is usually considered a simple question. We normally would not have to undertake empirical investigation in order to answer this question because we are able to just report our own thoughts without undertaking such empirical investigation. This intuition does not imply that you can have thoughts about whether Newcastle is north of London without undertaking empirical investigation; indeed empirical investigation of some sort would be a necessary condition of having thoughts of this sort. Rather, what this intuition is saying is that to *access* one's thought that Newcastle is north of London one does not have to undertake empirical investigation.

One might think that a third response (which is not an answer) can be given to the question 'do you think Newcastle is north of London?' The response is: 'I'm not sure whether I think Newcastle is north of London.' However, such a response would be very odd. 'I'm not sure' would be a legitimate response if our question did not ask the subject what he *thought* but rather asked the subject what is true. For example, if we asked the subject 'is Newcastle north of London?', then he could say that he is not sure. But it would seem strange for the subject to claim not to know what he thinks, unless he is suffering from some psychological disorder.

Suppose that you are asked a more difficult question such as: 'do you think that knowledge of non-analytic conceptual truths is a priori?' Again, you could answer 'yes', or 'no', although unlike the last example you are unlikely to answer the question immediately. This is because unless you have been researching the topic of the epistemology of non-analytic conceptual truths, you are unlikely to have a belief state about the epistemology of non-analytic conceptual truths which you can access.

Therefore, in the first instance you are likely to answer 'no' and may spend a substantial amount of time thinking about whether knowledge of non-analytic conceptual truths is a priori before you consider changing or retaining your answer. You may eventually come to answer 'yes' because you have certain beliefs about non-analytic conceptual truths, views about what it is to know something a priori and views about whether the sceptical attacks on all forms of a priori knowledge are successful and so on. Then, you may be able to access these simpler beliefs and 'think out' or deduce your answer to this more difficult question. In this case, your access to the more simpler beliefs still did not require you to undertake empirical investigation. This is because your access to these beliefs just seemed to be 'right there' or available to you as soon as you turned your attention to them. And since these simpler beliefs add up to thought on whether non-analytic conceptual truths can be known a priori, you can, in principle, have non-empirical access to this complex thought.

The important point here is that for 'do you think that P?' type questions you are at least able to settle the question without empirical investigation. And when P is a sufficiently simple proposition for you, you are able to settle the question both without empirical investigation and almost immediately.

Of course one response to the questions 'do you think Newcastle north of London?' and 'do you think that knowledge of non-analytic conceptual truths is a priori?' is to say 'I don't know what you mean'. In this case we would say that you do not possess the relevant concepts to answer these questions, you lack at least one of the concepts of 'Newcastle', 'London' or 'north of' or you lack at least one of the concepts of 'non-analytic conceptual truth', 'knowledge' or 'apriority'.

What do all of the forgoing remarks have to do with self-knowledge? Well, the access you had to certain thoughts which did not require empirical investigation and which, in some cases, was 'right there' when you turned your attention to answering certain questions, is also the same sort of access you have to your thoughts when you are not prompted by the questioning of a third party. You may ask yourself these questions and, typically, you may not even need to ask yourself such questions in order to have access to your own thoughts. So if you think that 'Newcastle is north

of London' and you turn your attention to this thought, then your access to this thought does not require empirical investigation and is 'right there'.

If access to your thought that P does not require empirical investigation for any proposition P, then this puts you in a position to know that you believe or think that P without empirical investigation. Let me elaborate. Suppose that you think that P and P is one of those propositions which is just 'right there' when you turn your attention to it, then, if my forgoing remarks are correct, you have access to your belief that P without having to undertake empirical investigation. If P was a more complex proposition, then access to the thought that P would not be immediate but it would still be non-empirical. Now it seems that once you have access to your thought that P, then you can come to know it in a way which does not require empirical investigation. This is because your recognising that you have the thought that P seems to be all that is required in coming to know that you have the thought that P. This is why I shall call such knowledge *non-empirical knowledge*.

Non-empirical knowledge of our thoughts is *first personal*. By *first-personal* I mean only the agent who has the thought that P can have the non-empirical access to his thought that P and thus come to know that he thinks that P in a way which is non-empirical. If a third party were attempting to come to know whether S thinks that P, then she may look at the agent's behaviour or ask the agent. The methods employed by the third party involve empirical investigation to varying degrees. Thus, this non-empirical knowledge of some of one's thoughts is first-personal and only first-personal.

The way of coming to know our thoughts which we have just considered contrasts with the following way of coming to know about other bodily states. Suppose you are asked 'are you in pain?', in this case it seems to be that your consciousness or attention can alight on various parts of your body and that allows you to determine whether that part of your body is in pain and hence answer the question. Hence the way we come to know our own bodily states seems quite different from how we come to know the contents of our own thoughts. The paradox I shall set out later in this chapter concerns only knowledge of the contents of our own thoughts and, as

we shall, see does not obviously have any implications for knowledge of our own bodily states.

The sort of knowledge this paradox is concerned with is self-knowledge (as defined above). For example, my knowledge that I am thinking that water is wet, which, if the intuitive discussion here is correct, is non-empirical knowledge.

§2 The relevant instance of externalism about mental content

Externalism about mental content is roughly the thesis that the contents of our thoughts are determined, at least in part, by our external environment. I want to introduce a form of externalism about mental content (hereafter *externalism*) which I think is sufficient to generate the incompatibility with self-knowledge. Note that I will not *argue for* this form of externalism. That is to say, I will not produce an argument as to why this form of externalism should be adopted over the many denials of it which are on offer.

Very roughly the form of externalism I want to introduce makes claims of the following form:

If I am thinking that P, then E.

where P is the content of some thought and E is a proposition about the empirical world.

A specific instance of this sort of externalism is:

If I am thinking that water is wet, then E.

where E is some proposition about the empirical world and is related to the content of my thought that water is wet.

What sort of proposition is E taken to be? The answer to this question varies depending on the form of externalism one wants to adopt. E has been taken to be ‘there are samples of water in my environment’¹. E has sometimes assumed to be ‘I (or my speech community) have had such-and-such encounters with water.’² For simplicity I shall use the following version of E for the rest of this thesis:

Water exists and is a natural kind term or members of my linguistic community use the term ‘water’ whether or not water is a natural kind.³

I shall abbreviate this long disjunctive proposition by saying ‘I live in a *water-world*.’

What does this form of externalism amount to? This form of externalism requires that my concept of water vary in counterfactual situations. For example, suppose I live on earth and the stuff I refer to in thought and in speech as ‘water’ is in fact H₂O if water is a natural kind or members of my linguistic community use the term ‘water’ to denote H₂O whether or not water is a natural kind. But suppose instead, of living on earth where ‘water’ corresponds to H₂O, I live on a different world twin-earth identical to earth except that on twin earth there is no H₂O but a chemical XYZ which has all of the macroscopic properties of water. The relevant for of externalism requires that if I am having thoughts about ‘water’, then the stuff I am referring to is XYZ if XYZ is a natural kind or members of my linguistic community use the term ‘water’ to denote XYZ whether or not water is a natural kind. However, this form of externalism requires more than just my concept of water vary in counterfactual situations. It crucially requires that the *existence* of my concept of water depends on either the *existence* of water if water is a natural kind or the *existence* of a speech community that uses the term water whether or not water is a natural kind. In short it depends on the *existence* of a water-world.

¹ Davies (2000)

² Wright (2000), (2001)

³ Brown (1995)

§3 Outline of the paradox

The incompatibility of self-knowledge of the contents of our own thoughts and the specific form of externalism in the last section is typically established as follows. It is argued that the any subject S can run through the following argument (W) without using any empirical resources:

(W)

(W1) I am thinking that water is wet.

(W2) If I am thinking that water is wet, then I'm in a water-world.

Therefore,

(W3) I'm in a water-world.

It is argued that S can know (W1) non-empirically (as we saw in §1). And S does not require empirical knowledge that (W2) in order to derive (W3) from his knowledge that (W1); I shall return to this claim in §6. Hence, S can know (W3) non-empirically. But our intuitions are that S does not know non-empirically that (W3) since non-empirical knowledge of (W3) is absurd.

Spelled out in more detail the reductio says that the following is possible for any subject S:

(1) S knows non-empirically that (W1).

and

(2) (W2).

Therefore,

(3) S knows non-empirically that (W3).

But,

(4) It is not the case that S knows non-empirically that (W3).

Therefore,

(5) S knows non-empirically that (W3) and it is not the case that S knows non-empirically that (W3).

For brevity I shall refer to (1) to (5) as the *SK paradox*.

How does the proponent of the SK paradox derive (3) from (1) and (2)? There are broadly two views on how this can be done. In order to explain these views I need to introduce so called *transmission* and *closure* principles about knowledge and warrant. These can be quite tricky notions and so to aid comprehension I shall discuss them at length in §4 and §5. I can then return to the question of how (3) is derived from (1) and (2) in §6.

§4 Transmission and closure

It might be thought compelling that any subject can gain knowledge by thinking through a valid argument which has true premises. More specifically, it might be thought to be true that if a subject, S, knows the premises of an argument, recognises that the argument is valid, then S can gain knowledge of the argument's conclusion from his recognition of the validity of the argument and his knowledge of the argument's premises. However, consider the following argument from Dretske.⁴

(Z)

(Z1) That animal is a Zebra.

(Z2) If that animal is a Zebra, then that animal is not a cleverly disguised mule.

Therefore,

(Z3) That animal is not a cleverly disguised mule.⁵

There is a strong intuition that if S knows the premises of (Z), S does not come to know the conclusion of (Z) by way of the kind of knowledge he has of the premises of (Z) and his recognition that (Z) is valid. Why is this intuition so strong? Well, S typically knows (Z1) on the basis of his perceptual experiences, usually his visual experiences, and this somehow furnishes S with the knowledge that (Z1). A small amount of reflection allows S to know (Z2). But does S's perceptual knowledge that

⁴ Dretske (1970)

⁵ I'm taking cleverly disguised here to mean 'visually cleverly disguised as a zebra'

(Z1) and S's reflective knowledge that (Z2), *give* S knowledge that (Z3)? If the animal was a mule cleverly disguised as a zebra, wouldn't S be having just the same visual experiences he would have if the animal really was a zebra? Has S checked that the animal is not a mule cleverly disguised as a zebra? Has S checked the zookeeper isn't conspiring to hoax his visitors? Typically S will not have checked these things, so S does not know them. And even if S had checked these things, then he may know (Z3) but he wouldn't know (Z3) *by way of his perceptual knowledge of (Z1) and his reflective knowledge of (Z2) and recognising (Z) is valid*. Rather, S may know (Z3) in another way for example by way of the testimony of a trusted source, induction, memory or some other means.

Some other arguments which are alleged to fall foul of the intuition that we can know the conclusion of a valid argument by way of knowing the argument's premises and recognising the argument is valid are listed below.

(M)

(M1) That is a hand.

(M2) If that is a hand then, the external world.

Therefore,

(M3) The external world exists.

(RED)

(RED1) The wall is red.

(RED2) If the wall is red, then it is not the case that the wall is white cleverly lit by red.

Therefore,

(RED3) It is not the case that the wall is white cleverly lit by red lighting.

Again there is a strong intuition that if I know (M1) by way of my typical visual experiences of my hand, know (M2) by reflection, I do not *thereby* know (M3). The same observations follow *mutatis mutandis* for the (RED) argument.

Arguments like the (Z), (M) and (RED) contrast with the following argument (F):

(F)

(F1) The animal in the garbage is a fox.

(F2) If the animal in the garbage is a fox, then it is not a cat.

Therefore,

(F3) The animal in the garbage is not a cat.⁶

In the case of (F) my visual experiences of a fox furnish me with knowledge that (F1). I also know (F2) by reflection. And our intuition is that I can come to have a warrant for believing (Z3) by way of my visual knowledge that (F1) and my knowledge that (Z2).

In normal cases, like (F), we expect our typical perceptual knowledge that the major premise is true and our knowledge of the entailment to give us knowledge that the conclusion is true. We expect, in other words, the following knowledge transmission principle to hold true:

Knowledge Transmission Principle.

For any subject S and any proposition P and any proposition Q: if S has a certain kind of knowledge that P and S knows that P entails Q, then S knows that Q by way of S's certain kind of knowledge that P and S's knowledge that P entails Q.⁷

When the knowledge transmission principle holds true I shall say the *S's knowledge that P transmits across the known entailment*. When there is a counterexample to the knowledge transmission principle I shall say *S's knowledge that P fails to transmit across the known entailment* or when the context is clear I shall say there is a *knowledge transmission failure*.

⁶ Brown (2004), p58

⁷ I've restricted the principle to a simple entailment since for my purposes this simple formulation is the only formulation I need. For simplicity, I am also waiving difficulties from counter-examples which say the principle fails because S does not believe the conclusion of every valid argument whose premises S knows.

Note that the knowledge transmission principle is knowledge specific, that is, certain kinds of knowledge that P may fail to transmit across the known entailment while other kinds of knowledge do not suffer from a failure. That is, it is possible there can be one way of knowing that P that fails to transmit across the known entailment and there can be a distinct way of knowing that P which does in fact transmit across the known entailment.

For example, our intuition is that our typical perceptual knowledge that (Z1) failed to transmit across the known entailment (Z2). But suppose I have a special kind of perceptual knowledge that (Z1) which is different from my typical perceptual knowledge of a zebra. Suppose that I have knowledge that (Z1) by way of my having a visual experience of a zebra *together with* an auditory experience of the zebra making a braying noise which no other animal can make. It seems here that we may say this special kind of perceptual knowledge transmits across the known entailment (Z2) whereas my typical perceptual knowledge that (Z1) does not transmit across the same known entailment.⁸

How do issues concerning transmission of knowledge relate to issues concerning closure principles about knowledge? To see how the issues relate consider the (Z) argument again. Our intuition was that I usually know that (Z1) (that animal is a zebra) as a result of my typical perceptual experiences of a zebra and that I know (Z2) (if that animal is a Zebra, then that animal is not a cleverly disguised mule), by reflection. But our intuition was that my typical perceptual knowledge that (Z1) did not transmit across the known entailment (Z2). But we also allowed I could know that (Z3) in another way, by induction or memory. I could even ask the zookeeper if the animal is a cleverly disguised mule and then investigate whether the zookeeper is a trustworthy source of information, and so on in order to gain knowledge that (Z3).

⁸ This example is in Pryor (ms) p5 n5

Consider a typical closure principle.

Knowledge Closure Principle

For any subject S any proposition P and any proposition Q: if S knows that P and S knows that P entails Q, then S knows that Q.

All the knowledge closure principle states is an acceptable combination of knowledge ascriptions for S, it remains silent about the way in which S has obtained his knowledge that Q.

More explicitly the knowledge closure principle says:

For any subject S any proposition P and any proposition Q: if S knows that P and S knows that P entails Q, then S knows that Q *in any way whatsoever*.

Hence our intuition is that arguments such as (Z), (M) and (RED) are counter-examples to the knowledge transmission principle, when we have typical perceptual knowledge that their major premise are true and know their minor premises, but such arguments may not be counter-examples to the knowledge closure principle. On the other hand for arguments such as (F) we would say that since (F) satisfies the knowledge transmission principle when we have typical perceptual knowledge that (F1), then (F) satisfies the knowledge closure principle when we have typical perceptual knowledge that (F1) and knowledge that (F2).

More generally the knowledge closure principle is a necessary condition for the knowledge transmission principle and the knowledge transmission principle is a sufficient condition for the knowledge closure principle.

The discussion so far has used the term 'knowledge' and has asked whether S knows the conclusion of a valid argument by way of the kind of knowledge S has for the premises of this valid argument. However, recent discussion of this topic is usually conducted using the term *warrant*. It will be easier for me to discuss this area if I

follow the convention set by recent literature. In order to do this I need to set out what I mean by the term *warrant*. The use of the term 'warrant' may have some motivation since some philosophers would deny that we *know* propositions that (M3) the external world exists and perhaps deny that we *know* many other ordinary propositions. However these philosophers may grant that we have a warrant or justification for believing ordinary propositions and perhaps even propositions like (M3).

When I say S's *warrant* for believing a certain proposition is w, I will mean S's justification for believing a certain proposition is that w. I'm being deliberately vague about the term warrant for the moment because firstly, the recent literature on transmission failure is also vague about the term and secondly, the subsequent discussion of transmission failure should give a more precise picture of what is meant by warrant (see §5 for this). I allow that warrants are defeasible in the sense that a warrant for P does not entail that P is true.

For example, on some views my visual experiences of a zebra in a cage is sufficient to give me a warrant for the proposition that animal is a zebra and I could label this warrant say w_1 . Similarly, if I'm told by testimony from a trustworthy source that there is a zebra in the cage, then on some views, this gives me warrant for the proposition there is a zebra in the cage. But this testimonial warrant is a different sort of warrant which is afforded by my visual experiences of the zebra in the cage, so this warrant other can be labelled anything other than w_1 , lets call it w_2 . If on some views my auditory experiences alone of the noises made by a zebra are sufficient to give me warrant for the proposition there is a zebra in the cage, then we might want to call this warrant w_3 . Similarly, if the combination of both my visual and auditory experiences of a zebra are sufficient to give me warrant for the proposition there is a zebra in the cage, then we might want to call this warrant w_4 . Note w_4 is distinct from w_1 and w_3 .

Furthermore, when I say that someone's *warrant for believing* that p is that w, or that I will not only mean that the person has a certain justification for believing the

proposition, but that the person *does* believe the proposition, and that the person believes the proposition in some sense on the basis of the relevant justification.⁹

I shall use the term *warrant to believe* to mean that the person has a certain justification for believing the proposition *without* actually believing the proposition.¹⁰

The zebra example can now be framed using the term ‘warrant for believing’. Suppose my perceptual experience of a zebra in a cage gives me a warrant for believing that the animal in the cage is a zebra. Let’s call this warrant for believing *w*. Again, a little reflection allows me to come to know that (Z2) if the animal in the cage is a zebra, then it is not a cleverly disguised mule. The question now is: do I acquire a warrant for believing that the animal in the cage is not a cleverly disguised mule by way of my warrant *w* for believing (Z1) and my knowledge that (Z2)? Again, the intuition is that I do not. Consider:

Warrant Transmission Principle.

For any subject *S* and any proposition *P* and any proposition *Q*: if *S* has a warrant, *w*, for believing *P* and *S* knows *P* entails *Q*, then *S* has a warrant for believing *Q* by way of his warrant *w* for believing *P* and knowledge that *P* entails *Q*.

When the warrant transmission principle holds true I shall say *S*’s warrant *w* transmits across the known entailment. When there is a counterexample to the transmission principle I shall say *S*’s warrant *w* fails to transmit across the known entailment or when the context is clear I shall say there is a *transmission failure*.

Our intuition is that my typical perceptual warrant, *w*, for believing (Z1) fails to transmit across the known entailment (Z2).

Note that the transmission principle is *warrant specific*, that is it allows that some warrants a subject has for believing *P* can fail to transmit across the known entailment while other warrants for believing *P* do not suffer from a failure. That is,

⁹ I got this definition from Silins (ms)

¹⁰ I got this definition from Silins (ms)

it is possible there can be a warrant for believing, w , that fails to transmit across the known entailment and there can be a distinct warrant for believing, w^* , which does in fact transmit across the known entailment.

For example, our intuition is that our typical visual warrant, w , for believing (Z1) failed to transmit across the known entailment (Z2). But, again, suppose I have a warrant, w^* distinct from w , for believing (Z1). Suppose also that w^* consists of my having a visual experience of a zebra together with an auditory experience of the zebra making a braying noise which no other animal can make. It seems here that we may say w^* transmits across the known entailment (Z2).

What is the relation between the warrant transmission principle and the knowledge transmission principle? This is where things get tricky. Philosophers use warrant transmission principles to talk about knowledge transmission principles. But it is not usually not very clear about what justifies their doing this. In the section below I shall discuss using warrant transmission principles to talk about knowledge transmission principles.

§5 Transmission Principles

The some philosophers take the relation between warrant and knowledge to be that roughly:

For any subject S : if S has a certain kind of knowledge that P , then S has a certain kind of warrant for believing P corresponding to this kind of knowledge.

For example, if S knows non-empirically that he is thinking that P , then S has a non-empirical warrant for believing that he is thinking that P .

It is as a result of the relation between warrant and knowledge above that many philosophers take the knowledge transmission principle to entail the warrant transmission principle when the S 's kind of knowledge that P is related to S 's kind of warrant that P . That is to say they take:

Knowledge Transmission Principle.

For any subject S and any proposition P and any proposition Q: if S has a kind of knowledge that P and S knows that P entails Q, then S knows that Q by way of S's certain kind of knowledge that P and S's knowledge that P entails Q.

to entail:

Warrant Transmission Principle.

For any subject S and any proposition P and any proposition Q: if S has a warrant, w , for believing P and S knows P entails Q, then S has a warrant for believing Q by way of his warrant w for believing P and knowledge that P entails Q.

when the kind of knowledge that P (S has) in the knowledge transmission principle entails S has the warrant, w , for believing P.

The philosopher's who take this entailment to be true, may also argue that a failure of the warrant transmission principle entails a failure of the knowledge transmission principle for the appropriate kind of knowledge related to the warrant w . Consider Jessica Brown recommending just such a move:

“In principle, we have a counterexample to the transmission of knowledge whenever any component of knowledge fails to transmit across a valid argument” (2003, p118)¹¹.

What is the problem of arguing this in this way for the transmission of knowledge?

The problem is that the following argument is invalid:

- (1) X is a necessary condition for knowledge that P.
- (2) X does not transmit across a known entailment.

Therefore,

- (3) Knowledge does not transmit across a known entailment.

¹¹ Davies (2000) tacitly makes this move see p393

Why is (1)-(3) invalid? It is because it is possible that a certain of knowledge can fail to have a property (in this case transmission across the known entailment) which is consistent with knowledge itself having this property. It may be that several necessary conditions on knowledge interact which leave the result that a certain kind of knowledge transmits across a known entailment yet a necessary condition on knowledge fails to transmit across an entailment.¹²

The better way for arguing for failure of the knowledge transmission principle would have been to argue as follows: Assume that for any subject S: S has a certain kind of knowledge that P and S knows that (if P, then Q). And then show how S does not obtain a warrant, *w*, for believing Q by way of this certain kind of knowledge that P and knowledge that (if P, then Q). And since no such warrant *w* can be obtained S does not know that Q by way of this certain kind of knowledge that P and knowledge that (if P, then Q).

However, as we have seen above many discussions of the failure of the knowledge transmission principle argue that such a principle fails because the corresponding warrant transmission principle fails. This is not ideal, many philosophers I discuss later will argue for the failure of the knowledge transmission principle in this way. I therefore want to tentatively suggest how we can work around this invalid way of arguing for the failure of the knowledge transmission principle.

One way to work around this invalid argument would be to treat the term warrant for believing in the warrant transmission principle as a *knowledge sufficient warrant*.

w is a knowledge sufficient warrant for P iff (if P is true and S has warrant *w* for believing P, then S knows that P).

Hence, if P is in fact true and S has warrant *w* for believing that P, then S knows that P. When I use the term *warrant for believing P* from now on I shall mean a *knowledge sufficient warrant for P*.

¹² This point is made against closure principles by Warfield (2004). But, I think, Warfield's point applies to transmission principles too.

When warrant and knowledge are related in this way it is hard to see that there will be cases where warrant fails to transmit across a known entailment and yet knowledge does transmit across the same entailment. This is because it is difficult to see how the necessary conditions of truth and knowledge sufficient warrant would interact to secure a contrary result.

§6 Strategies of deriving non-empirical knowledge from externalism and self-knowledge

We can now return to the SK paradox and examine how (3) can be derived from (1) and (2). Recall the paradox was as follows for any subject S who considers the (W) argument.

- (1) S knows non-empirically that (W1) I am thinking that water is wet.
- (2) (W2) If I am thinking that water is wet, then I'm in a water-world: (i.e. if (W1), then (W3)).

Therefore,

- (3) S knows non-empirically that (W3) I'm in a water-world.

(3) then conjoins with (4) below:

- (4) It is not the case S knows non-empirically that (W3) I'm in a water-world.

In order to derive the absurd conclusion (5):

- (5) S knows non-empirically that (W3) and it is not the case that S knows non-empirically that (W3).

One proposal about how (3) is validly derived from (1) and (2) is proposed by McKinsey (2003). McKinsey holds that (1), (2) and the *McKinsey closure principle* below are sufficient to establish (3):

McKinsey Closure Principle:

Necessarily, for any subject S: *if* S can know non-empirically that (W1) and in fact (if (W1), then (W3)), *then* S can know non-empirically that (W3).¹³

The McKinsey closure principle is a consequence of a more general closure principle McKinsey holds true which he labels (CA):

Necessarily, for any person x and any propositions p and q , if x can know non-empirically that p and [in fact] p logically implies q , then x can know non-empirically that q .¹⁴

I shall call this strategy the *McKinsey proposal*.

A second proposal says that (3) follows from (1) and (2) by assuming (i) that S knows that (W2) a priori and (ii) holding that S's non-empirical knowledge that (W1) transmits across the a priori known entailment (W2). I shall call this the *transmission proposal*. Note that the transmission principle requires the SK paradox to have an extra premise, not only must (W2) in fact be true but S has to know (W2) a priori. Note also that *if* S's non-empirical knowledge that (W1) fails transmit across the a priori known entailment (W2), then we cannot derive (3) S knows non-empirically that (W3) I'm in a water-world.

The McKinsey proposal threatens the transmission proposal in the following way. The McKinsey proposal has the consequence that regardless of whether the transmission proposal is true, (3) can be derived from (1) and (2).

McKinsey's argument that his proposal has this threat is as follows. Consider the following entailment:

¹³ The SK paradox changes slightly when (CA) is used since the relevant 'can' operators need to be inserted in the other premises. I'm waiving this difficulty here.

¹⁴ McKinsey (2003), I've changed his use of 'a priori' to 'non-empirical' but this won't matter for our purposes here.

(G) Laura is thinking that George is cute logically implies George exists.¹⁵

According to McKinsey, if we can know our own thoughts non-empirically, then antecedent of (G) can be known non-empirically, by Laura. Hence by (CA) Laura can know George exists non-empirically. But Laura knowing George exists non-empirically is intuitively absurd, hence Laura cannot know George exists non-empirically. Hence only, (G), (CA) and allowing that Laura can have non-empirical knowledge of the antecedent of (G) are required to generate a paradox. Also note the following:

“[t]he *source* of one’s warrant for believing that George exists could not be that one has correctly deduced this conclusion from the relational premise in question, since one would not be warranted in believing the relational premise in the first place, unless one were *already* warranted in believing that George exists. Thus, in the very cases to which my argument for incompatibilism most clearly applies, warrant does not transmit from the cognitive premise to the externalist consequence. But contrary to what both Davies and Wright contend, this does not show that there is anything wrong with my argument. For the argument does not assume that warrant is always transmitted from the relevant cognitive premises to the externalist conclusions.” (2003, p102-3).

So, according to McKinsey, Laura’s knowledge that George exists cannot have been obtained by deducing it from the entailment (G). This is because, according to McKinsey, Laura must already know George exists in order to think that George is cute. Hence according to McKinsey the relevant application of transmission of knowledge fails in this case: Laura cannot gain (non-empirical) knowledge George exists by way of her non-empirical knowledge that she is thinking George is cute even if she could have a priori knowledge of (G).

Of course we cannot know George exists non-empirically, so if McKinsey is correct and Laura must already know George exist in order to know that she is thinking George is cute, then Laura’s knowledge of this thought will be empirical in some sense.

¹⁵ Note that George is a man who does in fact exist and whom Laura is looking at.

McKinsey's position is neatly summarised in the following quote (McKinsey uses a priori where I have used non-empirical):

“When my argument for incompatibilism is applied to a particular relational cognitive premise like (3), we assume for *reductio* that (3) is knowable a priori (by Laura). This assumption, as it turns out, is contrary to fact. But if (3) *were* knowable a priori, then CA generates the absurd consequence that Laura could know a priori that George exists. And surely, that is the correct result. For if (3) *were* knowable a priori, then it would be knowable without empirical investigation. Hence any assumption on warrant for which knowledge of (3) is based would itself have to be knowable without empirical investigation. For otherwise, knowledge of (3) would, contrary to our assumption for *reductio*, be based in part on empirical investigation after all, since it would be based in part on whatever empirical investigation is required to provide warrant for the relevant empirical assumption. In this particular case, knowledge of (3) clearly depends on the agent's having warrant for the assumption that George exists. Hence one could not have a priori knowledge of (3), unless one also had a priori knowledge that George exists. Here then is a clear case in which warrant fails to transmit from a premise to a deductive consequence of that premise, even though my closure principle for apriority CA yields the right result. And in general we may conclude that failure of warrant to transmit to a given conclusion is *not* a good reason to suspect that closure of apriority fails in the same case. Failure of warrant transmission is perfectly consistent with successful closure of apriority.” (2003, pp102-3)

Now one can see how McKinsey would apply his strategy to the (W) argument. He would make the following claims:

- (a) S cannot know (W1) I am thinking that water is wet unless S already knows that (W3) I am in a water world.
- (b) Hence, from (a), when we assume (1) for *reductio* we can use (2) and (CA) to generate the absurd consequence that S can know that (W3) a priori.
- (c) Hence, from (a), we cannot know that (W3) by way of knowledge that (W1). Therefore, the relevant instance of transmission which would result in S knowing that (W3) non-empirically by way of non empirical knowledge that (W1) and a priori knowledge that (W2) fails.

I think there are two problems with the above claims. The first problem is with the truth of claim (a). The second problem is with the plausibility of the McKinsey's closure principle (CA). The first problem is that, so far, in the discussion of self-knowledge it has not been obvious that in order to know that we think certain thoughts we need to know that the environment is in such-and-such a way (e.g.

George exists in this environment or the environment is a water-world). Sure, externalism used in the SK paradox holds, roughly, that *if in fact* a subject is thinking or thinks certain thoughts, then the environment is such and such a way. For example, the relevant instance of externalism discussed in the SK paradox says *if in fact* S thinks that water is wet, then S is in a water world. But neither externalism nor the discussion of self-knowledge here obviously have the consequence that in order to *know that* (W1), S *already* has to *know* that (W3). Therefore, it is reasonable to suggest that McKinsey's claim (a) makes certain controversial demands on knowledge of our own thoughts, controversial demands which do not obvious follow from our pre-theoretical discussion of self-knowledge earlier and the truth of the relevant instance of externalism. It is therefore sensible to work with an SK paradox that treats claim (a) above as false, the transmission strategy allows us to do this.

The second problem with the McKinsey strategy is that (CA) on some readings invites counter-examples. Suppose we the phrase 'S can' as 'S has the psychological propensity to'. On this reading (CA) commits us to saying if S has the psychological propensity to know that all of Peano's axioms are true non-empirically, then S has the psychological propensity to know certain complex truths of arithmetic non-empirically. But this is surely false certain subjects simply do not have the psychological propensity to know the complex truths of arithmetic. Many other examples could be formulated along these lines. By contrast the transmission strategy does not use a dubious closure principle in its statement of the self-knowledge paradox.¹⁶

Suppose instead we treat the phrase 'S can' in (CA) as 'S is at least in a position to'. There are, however, again counter-examples to (CA) when 'can' is read in this way. For example, suppose P logically entail a proposition Q and suppose Q is a truth which cannot be known (perhaps Q is a necessary truth which we cannot comprehend). Then even if we grant S can know P non-empirically and P logically

¹⁶ Another counter-example could be constructed if Q is a truth so horrendous that it will kill any subject as soon as she entertains it. Here the antecedent of (CA) is satisfied but the consequent will not be since any subject who entertains Q will be dead.

entails Q, S is not in a position to know that Q in any way whatsoever including a non-empirical way.

As a result of the problems associated with (CA) it seems better to work with the transmission proposal.

§7 Summary: the road ahead for transmission failure compatibilists

It seems that the most plausible statement of the SK paradox is given by what has been called the transmission strategy. The statement of the SK paradox given by the transmission strategy is the following could be true of any subject S:

- (1) S knows non-empirically that (W1) I am thinking that water is wet.
- (2) (W2) If I am thinking that water is wet, then I'm in a water-world: (i.e. if (W1), then (W3)).
- (2a) S knows a priori that (W2).

Therefore,

- (3) S knows non-empirically that (W3) I'm in a water-world.

(3) then conjoins with (4) below:

- (4) It is not the case S knows non-empirically that (W3) I'm in a water-world.

In order to derive the absurd conclusion (5):

- (5) S knows non-empirically that (W3) and it is not the case that S knows non-empirically that (W3).

(From now on I will use 'SK paradox' to denote the transmission proposal's statement of the SK paradox).

The transmission strategy method of deriving (3) from (1) and (2a) is to say that S's non-empirical knowledge that (W1) transmits across the a priori known entailment (W2). Hence S acquires non-empirical knowledge that (W3) by way of his non-empirical knowledge that (W1) and a priori knowledge that (W2).

Let's call a *compatibilist* someone who wants to hold that (1) and (2) are true but (5) is false. Compatibilists come in two varieties. *Bullet biting compatibilists* deny (4) so they can allow that (3) follows from (1) and (2a). *Transmission failure compatibilists* deny that (3) follows from (1) and (2a)

The transmission failure compatibilist claims that even if (1) and (2a) are true (3) does not follow from (1) and (2a). In order claim this the transmission failure compatibilist needs to claim that S cannot acquire non-empirical knowledge that (W3) by way of his non-empirical knowledge that (W1) and a priori knowledge that (W2). That is to say, the transmission failure compatibilist needs to deny that S's non-empirical knowledge that (W1) transmits across the a priori known entailment (W2).

What is the transmission failure compatibilist saying? The transmission failure compatibilist's position is that *even if* we know the contents of our thoughts in a non-empirical way and we know that the relevant externalist premise is true a priori, *then* we not, as a result, know empirical propositions in a non-empirical way. The transmission failure compatibilist can remain silent over the questions of whether:

- (i) we can know what we are thinking non-empirically
- (ii) over whether the relevant instance of externalism is in fact true
- (iii) if the relevant instance of externalism is true, we can know that it is true a priori.

For the rest of this thesis I want to investigate the compatibilist's strategy of denying that (3) follows from (1) and (2a). In so doing I too will remain silent over the question of whether each of (i), (ii) and (iii) are true.

How does the compatibilist argue that that S's non-empirical knowledge that (W1) transmits across the a priori known entailment (W2)? Described in broad terms the compatibilist makes two steps here. His first step is to argue that certain kinds of perceptual knowledge fail to transmit across known entailments and he'll argue that such cases exhibit a certain kind of epistemic structure. His next step is to say that this epistemic structure is present in the case of arguments involved in the SK paradox. More specifically he'll say such an epistemic structure is present in the case

of arguments like (W) when we have non-empirical knowledge of (W)'s major premise and a priori knowledge of (W)'s minor premise. Once he's secured this second step the compatibilist concludes that non-empirical knowledge that (W1) does not transmit across the a priori known entailment (W2) and hence (3) cannot be derived from (1) and (2).

So to summarise the main steps the compatibilist makes are:

Step 1: A certain epistemic structure is present in the cases where we intuitively think a certain kind of perceptual knowledge fails to transmit across a known entailment.

Step 2: The epistemic structure identified in Step 1 is present in the case where we have non-empirical knowledge that (W1) and a priori knowledge that (W2).

Although as we have noted earlier those who argue for failure of knowledge transmission usually do so by arguing that the appropriate instance of warrant transmission fails. As we noted pursuing such a strategy is not ideal but we can work around it by treating the relevant warrant in the warrant transmission principle as a knowledge-sufficient warrant. Hence the analogous steps that transmission failure compatibilist makes are:

Step 1': Step 1: A certain epistemic structure is present in the cases where we intuitively think a certain kind of perceptual warrant fails to transmit across a known entailment.

Step 2': The epistemic structure identified in Step 1' is present in the case where we have non-empirical knowledge that (W1) and a priori knowledge that (W2).

Step 2' allows the transmission failure compatibilist to conclude S's non-empirical warrant that (W1) fails to transmit across S's a priori knowledge that (W2). Hence, according to the compatibilist, S cannot gain a non-empirical knowledge sufficient warrant (since this is how I am defining warrant) for the consequent of (W2) (which is (W3)). Therefore, according to the compatibilist, S does not know (W3) non-empirically.

Chapter 2

Introduction

In this section I want to discuss the first step which the transmission failure compatibilist makes. The transmission failure compatibilist argues that certain perceptual warrants fail to transmit across known entailments when a certain epistemic structure is present.

Two transmission failure compatibilists who make at least this first move are Davies¹⁷ and Wright¹⁸. However, I will not immediately start discussing the conditions under which Davies and Wright hold that a certain perceptual warrant transmits across a known entailment in this chapter. Before I start discussing Davies and Wright I want to set out what I take to be two arguments which are at least a good rough approximation of how one could argue for the result that certain perceptual warrants fail to transmit across a known entailment when a certain epistemic structure is in place. The first of these arguments is connected up with some very specific theories of how our beliefs are warranted. The second of these arguments is more general in the sense that it remains neutral over precisely how our beliefs are warranted. Then with these two arguments set out, I shall sketch a number of possible responses one can make to them in order to deny or explain away our intuition that a certain perceptual warrant fails to transmit across a known entailment. I shall do this in §1. The picture I develop in §1 will be useful for situating the arguments of Davies and Wright and the responses to them in the latter sections. In §2 I turn to discussing Wright's argument that certain warrants fail to transmit across known entailments when a certain epistemic structure is present. I shall argue that the first two arguments I sketch in §1 is a good rough guide to Wright's position. After this I note some recent responses to Wright's argument and focus in depth on one of them. In §3 I discuss Davies' argument that certain warrants fail to transmit across known entailments. I argue that Davies' argument is a good rough approximation to the second argument I identify in §1. I then argue that at least two of the recent criticisms of Davies rest on a misunderstanding of his

¹⁷ (2000)

argument. I then advance my own third criticism of Davies and tentatively suggest an amendment Davies can use to sustain his position. §4 I summarise the results of this Chapter.

§1 Two rough arguments for transmission failure

In this section my goal is to set out two rough and ready arguments to show that a certain epistemic structure must be in place for certain perceptual warrants fail to transmit across known entailments. I shall conclude that there are two such arguments. The first argument (Argument 1) requires that we subscribe to a specific theory of warrant whereas the second argument whereas the second argument (Argument 2) remains neutral over the theory of warrant assumed. I shall set out these arguments in §1.4. In order to do this I need introduce the notion of an *undercutting defeater* in §1.1, a *problematically warranted argument* in §1.2 and set out some theories of how our beliefs warranted in §1.3 before actually setting out the two arguments in §1.4. In §1.5 I review some response that can be made to the arguments in §1.4, they reveal how some responses to the first argument don't tell against the second argument and that our intuitions in the transmission failure cases trade on some general principles about the nature of warrant.

§1.1 Undercutting defeaters

A *prima facie warrant* for believing a proposition P is a proposition P is a warrant for believing P which may or may not have the status of a warrant.

When a *prima facie warrant*, *w*, in fact has the status as a warrant, I shall say *w* is a *genuine warrant*. When a *prima facie warrant*, *w*, in fact has the status as a warrant, I shall say *w* is *not a genuine warrant*

For example, my perceptual experiences of a zebra give me a *prima facie warrant* for, *w*, believing (Z1) that animal is a zebra. But my warrant *w* can be defeated in a number of ways. For example, I could be told by a trustworthy source 'a fox is always kept in the pen' where you say you see a zebra. Or a trustworthy source

¹⁸ (2000), (2001) and (2003)

could tell me 'I slipped hallucinogenic drugs in your coffee before we went to walk around the zoo today'. In the first example, the warranting status of, *w*, remains but there may be good reasons (a trustworthy source telling me things to the contrary) which count against my prima facie warrant, *w*, for believing (*Z1*). In the second example, it seems as though my believing that I am under the influence of a hallucinogenic drug *removed* the warranting status of my prima facie warrant *w*. I shall be concerned with the latter type of defeaters in this section.¹⁹

For any *rational* subject, *S*: a proposition φ is an *undercutting defeater* *S*'s prima facie warrant, *w*, for believing *P* iff *S*'s belief that φ removes *w*'s warranting status as a warrant for believing *P*.

The idea of an undercutting defeater is in standard circumstances *S*'s prima facie warrant, *w*, for believing *P* would be a warrant for believing *P* but in the circumstances that *S* believes φ , *S*'s belief that φ would rationally prevent *S* from taking *w* as a warrant for believing *P*.²⁰

For example, suppose I have a (prima facie) warrant, *w*, for believing that that is an orange on the basis of my visual experience of an orange and suppose it is true that I'm having a hallucination of an orange when in fact there is no orange there. Now in such a case, if I become aware of the proposition 'I'm having a hallucination of an orange when in fact there is no orange there', then it removes the capacity of my visual experiences to provide a warrant for believing that that is an orange. So I'm having a hallucination of an orange when in fact there is no orange there is an undercutting defeater for my prima facie warrant *w*. What about making the following modification to this example: it is false that I'm having a hallucination of an orange when in fact there is no orange there but I nonetheless believe this falsehood. In the modified case is the status of my prima facie warrant, *w*, removed? I think the intuition here is 'yes' provided that we are dealing with rational subjects. If I believed that I'm having a hallucination of an orange when in fact there is no orange there, then it would not be rational of me to regard my visual experiences of an orange as giving me a warrant for believing that that is an orange.

¹⁹ These distinctions between kinds of defeaters may bear some relation to Type-I and Type-II defeaters in Pollock (1974), p41 ff.

²⁰ It's difficult to say what would happen in a case where the subject is not rational

Here's another example. Suppose I have the typical prima facie visual warrant, w , for believing $(Z1)$ that animal is a zebra and suppose I believe not- $(Z3)$ is true, that is, I believe that that animal is a mule cleverly disguised as a zebra. In this case if I am behaving rationally I would regard my visual experience of a zebra as not giving me a warrant for $(Z1)$. Hence, not- $(Z3)$ is an undercutting defeater for my (prima facie) visual warrant, w , for believing $(Z1)$ since if I believed not- $(Z3)$, the warranting status of my typical prima facie visual warrant w would be removed.

Note that not- $(Z3)$ is not the only undercutting defeater for my typical perceptual warrant, w , for believing $(Z1)$; I list some propositions which are undercutting defeaters for w below:

(N1) My perceptual faculties are working correctly.

(N3) I am not a brain in a vat having the perceptual experience that that animal is a zebra.

(N3) I have not recently had hallucinogenic drugs slipped in the coffee he drank just before having his perceptual experience of the zebra
and so on to (Nn) .

Note also that the notion of an undercutting defeater is as I have sketched it is just a relation between two propositions and a certain warrant for believing one of those propositions and hence an undercutting defeater makes no strong demands on what it is for a belief to be warranted. That is to say, the notion of an undercutting defeater alone does not say in order for S 's prima facie warrant, w , for believing P to be a warrant for believing P , S is required to have a warrant to (or for) believing all undercutting defeaters for w are false. Certain theories of how our beliefs are warranted may secure such a requirement. There is, however, one weak demand the notion of an undercutting defeater makes on a theory of warrant: in order for a prima facie warrant, w , for believing P to have the status of a warrant for a rational subject, S , S cannot have beliefs that any undercutting defeaters for w are true.

Consider a slightly different case where I have a different prima facie warrant, w^* , for believing $(Z1)$. Suppose that w^* consists in both my having a visual experience of a zebra and an auditory experience of that zebra making a braying noise which no

other animal can make. Is not-(Z3) an undercutting defeater for my warrant, w^* , for believing (Z1)? Would my belief that the animal is a cleverly disguised mule, remove the warranting status of my prima facie warrant w^* ? I want to suggest that such a belief would *not* remove the warranting status of a prima facie warrant w^* . Provided I recognise that the braying noise I am aware of can only be made by a zebra, then the rational thing for me to do would be to discard my belief that the animal is a cleverly disguised mule.²¹ This is because, should I recognise that any animal except for a zebra, including a cleverly disguised mule, cannot make such a braying noise, the rational thing for me to do would be to stop believing that the animal is a cleverly disguised mule. Let us bear this case in mind when we revisit our intuitions about perceptual warrant transmission in the next subsection.

§1.2 Problematically warranted arguments

Cast your mind back to the following results in Chapter 1:

(a) One result was that I could have a typical perceptual warrant, w , for believing (Z1) that animal is a zebra gained from a visual experience of a zebra. And our intuition was that w does not transmit across the known entailment (Z2) if that animal is a zebra, then it is not a cleverly disguised mule.

(b) A second result was that I could have an extraordinary perceptual warrant, w^* , for believing (W1) gained from having both the visual experience of a zebra and the zebra making a braying noise which no other animal can make. And our intuition was that when I had the warrant, w^* , for believing (Z1) it did transmit across the known entailment (Z2).

As we saw in the last subsection not-(Z3) (i.e. the animal a cleverly disguised mule) was an undercutting defeater for my warrant w above but not for my warrant w^* . These results may indicate that there is a connection between the failure of a certain warrant, w , to transmit across a known entailment when the consequent of that entailment says that the undercutting defeater of that warrant is false.

²¹ Remember, I'm taking 'cleverly disguised mule' to mean 'a mule cleverly disguised to *look* like a zebra'

For example, (Z3) says that the undercutting defeater for my warrant, w , for believing (Z1) is false. Now our intuition was that w failed to transmit across the known entailment (Z2). On the other hand, (Z3) does not say anything about whether the undercutting defeaters for w^* are false and our intuition was w^* transmitted across the known entailment (Z2). I want to suggest that this feature goes some way to explaining why we have the intuitions we do about whether certain warrants transmit across known entailments. In order to develop this suggestion I want to introduce the notion of a *problematically warranted argument*.

An argument Φ where S has a warrant, w , for believing the major premise of Φ is a *problematically warranted argument* iff

- (i) Φ is of the form [A, (if A, then B) therefore B]²²
- (ii) S has a warrant, w , for believing A
- (iii) S knows that (if A, then B)
- (iv) not-B is an undercutting defeater for S's warrant, w , for believing A.

Note that whether a certain argument is problematically warranted depends on both the premises of the argument itself and the type of warrant S has for believing the major premise of the argument (premise A). In order to see this consider the next few examples.

Example 1:

Recall, the (Z) argument. The argument is of the form [A, (if A, then B) therefore B] so it satisfies condition (i) above. My typical visual experiences of a zebra providing me with a warrant, w , for believing (Z1) satisfying condition (ii) above. It is uncontroversial that I know that (Z2) satisfying condition (iii) above. And we have already noted not-(Z3) is an undercutting defeater for my warrant, w , for believing (Z1) satisfying condition (iv) above. In this case we would say that the (Z) argument

²² This notion could be extended for arguments of other forms too but for my purposes I won't need to do that here.

where I have the warrant, w , for believing (Z1) argument is a problematically warranted argument.

Example 2:

Now suppose instead of w , my warrant for believing (Z1) is the warrant w^* as described above which is furnished both by my visual experience of a zebra and my auditory experience of the zebra making a distinctive braying noise. Conditions (i), (ii) and (iii) are satisfied as before. However, condition (iv) is not satisfied: as we have seen not-(Z3) is not an undercutting defeater for my warrant for believing (Z1). Hence, the (Z) argument where I have the warrant, w^* (rather than w), for believing (Z1) is a not a problematically warranted argument. So here the argument remains the same as in example 1 but the warrant for believing its major premise has been changed and this changes the argument's status as a problematically warranted argument.

Example 3:

Suppose now that I have the warrant w for believing (Z1) (as defined in example 1).

But in this case I consider the following argument:

(Z')

(Z1) That animal is a zebra.

(Z'2) If that animal is a zebra, then it is not a mule.

Therefore,

(Z'3) That animal is not a mule.

It is uncontroversial that I know that (Z'2). Again conditions (i), (ii) and (iii) above are satisfied. Is condition (iv) above satisfied? Is not-(Z'3) is an undercutting defeater for my warrant, w , for believing (Z1)? The answer here is 'no' since if I believed the animal was a mule and I acquired a typical visual warrant for believing (Z1), then it would be rational for me to stop believing that the animal is a mule. Hence, here my warrant, w , for believing (Z1) has remained the same as in example 1 but changing the relevant premise of the argument (the B proposition in the

template), results in the argument (Z') where I have warrant, w, for believing (Z1) not being a problematically warranted argument. The same explanation here follows *mutatis mutandis* for the argument (F) when we have a typical perceptual warrant for (F)'s major premise.

I think having an argument with a certain warrant for its major premise which is problematically warranted does not *in itself* entail that such a problematically warranted argument is a case where this warrant fails to transmit across the minor premise of such an argument. I think certain other assumptions about accounts of how our beliefs are warranted and what it is for something to count as a warrant may in fact secure this result. In order to reach this conclusion I shall discuss accounts of warrant in the next sub-section.

§1.3 Accounts of warrant

Using the notion introduced above I want to introduce several *theories of warrant*, theories of the necessary conditions that need to obtain in order for us to have a warrant for believing a proposition. These various theories of warrant will be helpful when it comes to discerning the epistemic structure which obtains in the cases where we have an intuition that there is a transmission failure. The accounts of warrant set out here are modifications of the accounts of warrant set out by Pryor (ms) but with some important differences. Pryor uses the term 'rests on' or 'rests in part on' whereas I do not. Also my notion of a warrant for believing is different from Pryor's. However even with these similarities I think the theories of warrant are roughly the same as the accounts of warrant used in Pryor (ms).

In order to discuss the accounts of how our beliefs are warranted below I need introduce the notion of one warrant being independent from another. The intuitive idea is that:

For any subject S: S's warrant, w^* , for believing (or to believe) P is independent of S's warrant, w^+ , to believe (or for believing) iff S did not use her warrant w^+ in obtaining her warrant w^* .

The *conservative account of warrant* says for any subject S:

S's prima facie warrant, w , for believing P is a warrant *only if* S has a warrant (or warrants), independent of w , to believe that each of the undercutting defeaters for w are false.

So for example, suppose I have the typical prima facie perceptual warrant w for believing (Z1). Then on the conservative account of warrant in order for w to be I warrant I need to have a warrant, independent of w , (or warrants independent of w) to believe that each undercutting defeater for w is false. That is to say, I need to have a warrant (or warrants), independent of w , to believe each proposition from (N1) to (Nn) below is false:

(N1) My perceptual faculties are not working correctly.

(N3) I am a brain in a vat having the perceptual experience that that animal is a zebra.

(N3) I have recently had hallucinogenic drugs slipped in the coffee he drank just before having his perceptual experience of the zebra

.

.

.

(Nn)

where (N1) to (Nn) range over all of the undercutting defeaters for my warrant, w , for believing (Z1).

In order to set out the weak conservative account (to be defined below) I need to introduce the notion of an *entitlement to believe* a proposition.

Any subject, S, has an entitlement to believe P iff S has no sufficient reason to believe that P is false and S has some sort of justification to believe that P needn't have been acquired or earned.

On this view, 'entitlement' can be seen as having a more primitive epistemic status than 'warrant' whereas warrant is usually (though not always) acquired through evidence or arguments or operation of some reliable faculty perception or reasoning, and so on. Entitlement, on the other hand, need not be acquired or earned in this way. Moreover, an entitlement could be an a priori notion or it may not be. We may have an entitlement for P for pragmatic reasons or for other reasons. Moreover, if S has an entitlement to believe P, S need not at this moment (or indeed ever) in fact believe P or think that P. Also, entitlement can be defeated: if S comes to possess sufficient reason that P is false. Again this is a vague notion of entitlement but it will do for the purposes of fleshing out this account. The key thing here is that entitlement in this sense is not merely lacking evidence. Moreover warrant need not be something a subject possesses just by virtue of being a rational subject.

The weak conservative account of warrant says for any subject S:

S's prima facie warrant, w , for believing P is a warrant *only if* S has an entitlement (or entitlements), independent of w , to believe that each of the undercutting defeaters for w are false.

For example, on the weak conservative account of warrant my typical prima facie perceptual warrant, w , for believing (Z1) is a warrant only if I have entitlement (or entitlements) to believe each of the propositions (N1) to (Nn) below are false:

(N1) My perceptual faculties are not working correctly.

(N3) I am a brain in a vat having the perceptual experience that that animal is a zebra.

(N3) I have recently had hallucinogenic drugs slipped in the coffee he drank just before having his perceptual experience of the zebra

.
.
.

(Nn)

where (N1) to (Nn) range over all of the undercutting defeaters for my warrant, w , for believing (Z1).

I shall use the term *conservative accounts of warrant* to refer to both the conservative account of warrant and weak conservative account of warrant.

What about the rival accounts of warrant to the conservative account of warrant?
There are broadly two accounts which rival the conservative account of warrant.

The liberal and intermediate accounts of warrant deny the main thesis of the conservative accounts of warrant that:

For any subject, S, S's prima facie warrant, *w*, for believing P is a warrant *only if* S has either an entitlement (or entitlements) or a warrant (or warrants), independent of *w*, to believe that each of the undercutting defeaters for *w* are false.

The *liberal account of warrant* says that for any subject S:

S's prima facie warrant, *w*, for believing P is a warrant only if S lacks evidence that each of the undercutting defeaters for *w* are true.

For example, on the liberal account of warrant my typical prima facie perceptual warrant for (Z1) is a warrant provided I lack evidence that each of the propositions (N1) to (Nn) are true:

(N1) My perceptual faculties are not working correctly.

(N3) I am a brain in a vat having the perceptual experience that that animal is a zebra.

(N3) I have recently had hallucinogenic drugs slipped in the coffee he drank just before having his perceptual experience of the zebra

.
. .
.

(Nn)

where (N1) to (Nn) range over all of the undercutting defeaters for my warrant, *w*, for believing (Z1).

The *intermediate account of warrant* says that for any subject S:

S's prima facie warrant, w , for believing P is a warrant only if S lacks evidence that each of the undercutting defeaters for w are true and in fact the undercutting defeaters for w are false.

For example, on the intermediate account of warrant my typical prima facie perceptual warrant for (Z1) is a warrant provided I lack evidence that each of the propositions (N1) to (Nn) are true and in fact the propositions (N1) to (Nn) are false:

(N1) My perceptual faculties are not working correctly.

(N3) I am a brain in a vat having the perceptual experience that that animal is a zebra.

(N3) I have recently had hallucinogenic drugs slipped in the coffee he drank just before having his perceptual experience of the zebra

.
. .
.

(Nn)

where (N1) to (Nn) range over all of the undercutting defeaters for my warrant, w , for believing (Z1).

One of the crucial differences between the conservative accounts of warrant and the liberal and intermediate accounts are that they in order for a subject's typical prima facie warrant for believing a proposition do not require that the subject bother about propositions which are undercutting defeaters for w in the sense that the subject need not have beliefs about such defeaters or have any positive warrant or entitlement to believe that such defeaters are false.

Note that Pryor has commented that that a reliabilist, who roughly says our belief that P is warranted iff it is produced and sustained by a reliable process, may fall into

the intermediate category. Whereas, internalists, those who say, roughly, our beliefs are warranted depending on what it is like from the subject's perspective, will take a liberal or conservative line.²³ I'm not going to make a commitment as to what, if any, theories of warrant above reliabilists or internalists would subscribe to but it may be that Pryor's remarks carry over for my accounts of warrant above too.

²³ Pryor (ms) p9.

§1.4 Transmission Failure and theories of warrant

Having set out what I mean by the various accounts of warrant above and what a problematically warranted argument is, I now want to suggest can use these notions together with certain other assumptions about what constitute a warrant to give an argument for the conclusion that certain warrants fail to transmit across known entailments.

Before I give these arguments it is worth thinking about our intuitions in the transmission failure cases again. Our intuition is that my warrant for believing (Z1) (that animal is a zebra) fails to transmit across the known entailment (Z2) if that animal is a zebra, then it is not a cleverly disguised mule. We have this intuition because, we think my typical perceptual warrant for (Z1) *in some sense* ‘presupposes’ or ‘assumes’ (Z3) in order for my typical perceptual warrant for (Z1) to count as a warrant. The arguments below will sharpen this extremely rough intuition.

§1.4.a Argument 1:

Argument 1 starts off from the plausible claim that there are problematically warranted arguments:

Problematically warranted argument premise.

(PWA)

Φ is a problematically warranted argument (hence Φ is of the form [A, (if A, then B), therefore B]) where S has a warrant, w, for believing A.

The argument then claims (PWA) entails the following claim:

Principle of independent entitlement claim:

(PIE)

In order to have a warrant, w , for believing A in the problematically warranted argument Φ , S is required to have a warrant or entitlement for B which is independent of w .

The argument then deploys the premise below:

Additional warrant claim:

(AWC):

If S 's warrant for believing A in the problematically warranted argument Φ requires S to have a warrant or entitlement for B which is independent of w , *then* S cannot obtain any kind of new warrant for B by way of his warrant, w , for believing A and his knowledge that (if A , then B).

An alternative way of putting (AWC) is:

If S 's warrant for believing A in the problematically warranted argument Φ requires S to have a warrant or entitlement for B which is independent of w , *then* S 's warrant, w , for believing A fails to transmit across the known entailment (if A , then B).

By modus ponens on (AWC) and (PIE) we reach the conclusion that any problematically warranted argument where S has warrant, w , for believing the argument's major premise is a transmission failure.

In brief argument 1 is: Assume (PWA). (PIE) follows from (PWA). Assume (AWC). Then when combined with the additional premise (AWC), (PIE) results in the conclusion that any problematically warranted argument where S has warrant, w , for believing the argument's major premise is a transmission failure. So Argument 1 S 's warrant, w , for believing A 'presupposes' or 'assumes' B in order for the warrant w to count as a warrant.

The conservative accounts of warrant hold that (PIE) *always* follows from (PWA) for the following reasons. They hold that in order to have a warrant, w , for believing A , S needs entitlement or warrant, independent of w , to believe that each undercutting defeater for w is false. Since (PWA) stipulates Φ is a problematically argument, not- B is an undercutting defeater for w . Hence, on the conservative accounts of warrant, S requires warrant or entitlement, independent of w , to believe B (i.e. to believe not- B is false) in order for w to count as a warrant for believing A .

In contrast to the conservative accounts of warrant, the liberal and intermediate accounts of warrant do not hold that (PIE) *always* follows from (PWA). Rather, the intermediate or liberal accounts of warrant will say (PIE) follows from (PWA) depending on what kind of warrant the warrant w is. For the liberal account of warrant the bare minimum requirement for a prima facie warrant, w , for believing P to count as a warrant S needs to lack evidence that undercutting defeaters for w are true. Hence, for the liberal when only just this bare minimum requirement is met in the case of the problematically warranted argument, S only lacks evidence that not- B is true. S does not have warrant or entitlement to believe B which is independent of w in this case. Hence, for a liberal account of warrant, when only its bare minimum requirement on warrant is met, (PIE) does not follow from (PWA).

However, if the liberal account of warrant is correct, one can go beyond its bare minimum requirements in order to have a warrant for a certain proposition. For example, rather than just lacking evidence that not- B is false, in the last example, S could go out and acquire positive evidence that B is true and hence obtain a warrant, independent of w , to believe B (or even for believing). In this modified example S does in fact have a warrant, independent of w , to believe B . So in this modified example, the liberal would say that (PIE) does follow from (PWA). So the key conclusion here is that the liberal account of warrant does not hold that (PIE) always follow from (PWA). Similarly the intermediate account of warrant would say (PIE) does not always follow from (PWA).

§1.4.b Argument 2:

In contrast to argument 1, argument 2 does not require us to subscribe to any particular theory of warrant. In order to set out argument 2, I need to set out what I mean by an *alternative* to a proposition.

An *alternative* to a proposition θ is a proposition π , distinct from θ , which entails that θ is false.

For example, some of the alternatives to (Z3) that animal is not a cleverly disguised mule could be:

The animal is a cleverly disguised mule and the British Library is located in Spain.

The animal is a cleverly disguised mule and the zoo where this mule is located is run by a dishonest zookeeper who is trying to hoax visitors.

Note that the alternatives to (Z3) needn't themselves be true, they simply have to entail (Z3) is false.

Argument 2 again starts off with the premise (PWA).

Problematically warranted argument premise.

(PWA)

Φ is a problematically warranted argument of the form [A, (if A, then B), therefore B] where S has a warrant, w, for believing A.

It then claims (PWA) entails the following claim:

The *failing to rule out alternatives claim*.

(FROA)

S's prima facie warrant for believing B by way of his warrant, w, for believing A and his knowledge that (if A, then B) cannot rule out any alternatives to B.

Argument 2 then adds a second premise called the *ruling out alternatives claim*.

(ROA)

For any proposition P: a prima facie warrant *w* is a warrant for believing P is a genuine warrant only if *w* rules out some of the alternatives to P.

The consequence of (FROA) and (ROA) is that S's prima facie warrant for believing B by way of his warrant, w, for believing A and his knowledge that (if A, then B) is not a genuine warrant. Hence, S does not obtain a warrant for believing B by way of his warrant, w, for believing A and his knowledge that (if A, then B). That is to say S's warrant w fails to transmit across the known entailment (if A, then B).

(PFW) and (ROA) have the consequence that any problematically warranted argument is a failure of S's warrant, w, for believing A to transmit across the known entailment (if A, then B) since cannot obtain a genuine warrant for B by way of his warrant w for believing A and his knowledge of the entailment.

Note that Argument 2 is much more of a threat than Argument 1 because it does not require us to subscribe to any particular account of warrant in order to reach the conclusion that problematically warranted arguments are transmission failures.

§1.5 Responses to the arguments

In this section I want to look at how one can respond to the arguments for transmission failure in the last subsection. This will help us situate the responses to Davies and Wright later on in the chapter.

§1.5.a Responses to Argument 1:

The most obvious response to argument 1 is to advocate a liberal or intermediate account of warrant. This response allows that (PIE) does not always follow from (PWA) depending on what warrant one has for the major premise A of the problematically warranted argument. However, this response is quite weak, in that, even when it doesn't allow us to derive (PIE) from (PWA), it still faces a threat from Argument 2. Why does this response face a threat from argument 2? It does so because Argument 2 remains neutral over what account of warrant is the correct account.

A second response to Argument 1 is to deny the (AWC) premise.

Additional warrant claim:

(AWC):

If S's warrant, w, for believing A in the problematically warranted argument Φ requires S to have a warrant or entitlement for B which is independent of w, then S cannot obtain any kind of new warrant for B by way of his warrant, w, for believing A and his knowledge that (if A, then B).

This response says that even if it is granted S's warrant, w, for believing A in the problematically warranted argument Φ requires S to have a warrant or entitlement for B which is independent of w, S still in fact can obtain some kind of a new warrant for B by way of his warrant, w, for believing A and his knowledge that (if A, then B).

Let's see what this response amounts to in the case of the (Z) argument. Suppose I have a prima facie warrant, w, for believing (Z1) on the basis of my typical visual experiences of a zebra. Suppose further that I have warrant or entitlement, independent of w, to believe that each undercutting defeater for w is false. Hence, on any account of warrant my prima facie warrant w is in fact a genuine warrant. Now suppose that I deny (AWC) and I say

'I've acquired a new warrant, let's call this w+, for believing (Z3) by way of my warrant, w, for believing (Z1).'

It seems that I would be open to the following sorts of retorts:

(i) 'Your warrant w+ for believing (Z3) is no genuine warrant at all, it does not give you an additional *reason* to believe (Z3) in the sense that it does not allow you to rule out any alternatives to (Z3).'

(ii) 'If you are rational and you doubted (Z3) w+ would not resolve your doubt.'

(iii) 'If you are rational and doubted (Z3), then you would revisit the (Z) argument. This would lead you to doubt the conjunction of the premises. You would grant that we know (Z2) and hence doubt (Z1). Then when you examine the prima facie warrant you have for (Z1) your prima facie warrant, w, for believing and note that it does nothing to resolve your doubt that the animal is a cleverly disguised mule. Hence your warrant w+ for believing (Z3) will do nothing to resolve your doubt about (Z3) since it is obtained in part by way of your warrant w.'

However, these retorts seem to be trading on certain principles about warrant:

Warrant and rational doubt principle.

(WRD)

For any prima facie warrant w for a proposition P: if w is a genuine warrant for believing (or to believe) P, then w is able to resolve the doubt of a rational subject about P.

And the following premise from Argument 2. The *ruling out alternatives claim*.

(ROA)

For any warrant prima facie warrant w for a proposition P: if w is a genuine warrant for believing (or to believe) P, then w is able to rule out some alternatives to P.

I'm not going to take a position on whether the above principles of warrant are true. Indeed the truth of the above principles has recently been a matter of controversy²⁴. All I want to point out is that if I were to reject the above principles about warrant, then I could say that my warrant w^+ for believing (Z3) is a genuine warrant. Hence, if I were to take w^+ as a genuine warrant I would have to concede that it is of extremely limited strength in that it could not persuade any rational doubter about (Z3) (including myself), nor could it rule out any alternatives to (Z3).

If we deny (WRD) and (FROA) then it seems to bring out a distinction between two types of transmission principles.

Strong warrant transmission principle.

For any subject S and any proposition P and any proposition Q: if S has a warrant, w , for believing P and S knows P entails Q, then S has a *strong* warrant for believing Q by way of his warrant w for believing P and knowledge that P entails Q.

²⁴ See Boghossian's 'principle of the universal accessibility of reasons' (2000, p253). See also reliabilist justifications of induction.

Where a *strong warrant* for believing a proposition ψ is a warrant which resolves rational doubt about ψ and rules out some alternatives to ψ .

Weak warrant transmission principle:

For any subject S and any proposition P and any proposition Q: if S has a warrant, w , for believing P and S knows P entails Q, then S has a *weak warrant* for believing Q by way of his warrant w for believing P and knowledge that P entails Q.

Where a *weak warrant* for believing a proposition ψ is a warrant which does not resolve rational doubt about ψ or does not rule out any alternatives to ψ .

While the strong warrant transmission principle is false for problematically warranted arguments, the weak warrant transmission principle is true.

§1.5.b Response to Argument 2:

The response to Argument 2 is more or less along the same lines as the last response to Argument 1. The typical response is to deny the (ROA) premise.

(ROA)

For any proposition P: a *prima facie warrant* w is a warrant for believing P is a genuine warrant only if w rules out some of the alternatives to P.

Again I do not want to endorse the denial of (ROA) here. However, if I denied (ROA), then I could say the following about the zebra case just considered:

'My warrant w^+ for believing (Z3) is a genuine warrant but it is of extremely limited strength in that it cannot rule out any alternatives to (Z3).'

Hence we deny that the relevant strong transmission principle is true in this case and allowing that the relevant instance of the weak transmission principle is true in this case.

In this section that follows I want to suggest that there is strong textual evidence to show that Wright is committed to something along the lines of Argument 1 and Davies is committed to something along the lines of Argument 2. Hence, the response to the arguments we have considered in this section will help us to situate the type criticisms many commentators raise against Davies and Wright.

§2 Wright (2000)

§2.1 Exposition of Wright

To what extent can Wright (2000) be held to subscribe to Argument 1? There is no direct evidence that Wright would subscribe to (PWA). This is where we need to examine Wright's (2000) template for transmission failure.

On Wright's (2000) template²⁵ (what I shall call the *A-B-C template*) an argument of the form [A, (A, then B), therefore B] fails to be a case where a certain warrant [w] for A transmits across the known entailment if A, then B when there is some proposition C such that:

- (i) A entails B.
- (ii) My warrant [w] for A consists in my being in a state which is subjectively indistinguishable from a state in which the relevant C would be true.
- (iii) C is incompatible with A.
- (iv) C would be true, if B were false.

According to Wright the argument (Z) meets the template with A=(Z1) that animal is a zebra B=(Z3) that animal is not a cleverly disguised mule and C=not-(Z3) that animal is a cleverly disguised mule. The argument obviously meets conditions (i), (ii)

²⁵ Wright (2000), p155

and (iii) of the template. And condition (ii) is satisfied S could be in a state subjectively indistinguishable from his actual state in which C the animal is a cleverly disguised mule and not a zebra when he has a typical perceptual warrant for A.

Now does this template commit Wright to (PWA)? I think it does. Suppose we take any problematically warranted argument [ψ , (if ψ , then θ), therefore θ] where S has warrant, w for believing ψ . If we put $C = \text{not-}\theta$, then condition (iv) is satisfied. If $\text{not-}\theta$, then $\text{not-}\psi$ since by assumption if ψ , then θ . Hence, for $A = \psi$ and $C = \text{not-}\theta$ condition (iii) is satisfied. Is condition (ii) satisfied $A = \psi$ and $C = \text{not-}\theta$? I think 'yes' since because by assumption $\text{not-}\theta$ is an undercutting defeater for S's warrant, w , for believing ψ , then if it were true that $\text{not-}\theta$, S would still be in the same state he would be in which he has warrant, w for believing ψ . So Wright is committed to some premise similar to (PWA) but his template may commit him to something much stronger than this premise as we shall see in §2.1.

Given condition (ii) why do we claim S has a warrant for believing A, rather than just saying S has a warrant for the more tentative disjunction (A or C)? For example, why do we claim S has a warrant for believing that animal is a zebra as a result of his visual typical experiences of a zebra rather than claiming S has a warrant for believing that animal is a zebra or that animal is a cleverly disguised mule? Here is where Wright's commitment to one of the conservative accounts of warrant becomes relevant (and hence his commitment to the (PIE) claim). On Wright's view in order to have a warrant for believing A, we require warrant or entitlement to believe C is false (which is independent of our warrant for believing A). Hence, according to Wright in order to have a warrant for believing (Z1) we require a warrant or entitlement to believe $\text{not-}(Z3)$ which is independent of our warrant for believing (Z1).

There is some textual evidence to suggest Wright would subscribe to (PIE).

“So it would seem I must have an appreciable entitlement to affirm B already, independent of the recognition of its entailment by A” (2000, p155)

There is also evidence that Wright subscribes to something like (AWC).

“While you have no doubt quite properly taken it for granted the conditions were generally suitable for the acquisition of reliable information by purely perceptual means, it would be absurd to pretend you had gained a reason for thinking so – at least in the specific respects that you didn’t have to reckon with disguised mules or deceptive lighting – just by dint of the fact those specific possibilities are logically excluded by the beliefs which, courtesy of your background assumption, you have now confirmed” (2000, p154)

§2.2 Responses to Wright

There are a number of criticisms that can be made about Wright’s template. The first criticism is that Wright is committed to one of the conservative accounts of warrant and many philosophers would reject this account of warrant and this account of warrant has a number of puzzling or undesirable consequences.²⁶ I shall not review this response because even if it is correct such a response would still be threatened by Argument 2. The other sort of criticism is more restricted, it remains silent over whether there are any problems with the conservative accounts of warrant which Wright subscribes; instead this criticism examines weaknesses in Wright’s A-B-C template above and proposes amendments to the template. I’m going to focus solely on the last criticism. This is because this criticism may be important when we go on to look at Wright’s response to the SK paradox in Chapter 3.

Wright’s (2000) template has been recently criticised. Consider the following argument:

(F)

(F1) The animal in the garbage is a fox.

(F2) If the animal in the garbage is a fox, then the animal in the garbage is not a cat.

Therefore,

(F3) The animal in the garbage is not a cat.

²⁶ Pryor (ms) and Brown (2003) and Brown (forthcoming) offer this line of thought.

Here we have $A=(F1)$ $B=(F3)$ and $C=\text{not}(F3)$. Conditions (i), (ii) and (iii) are obviously satisfied. Condition (ii) is also met in certain cases. This case is when my warrant for w is given to me on my visual or perceptual experience of a fox (as it usually is), then (ii) is also met: there is a case where the warrant w is subjectively indistinguishable from a state in which C would be true namely a cat cleverly disguised as a fox. Hence, the warrant w in these usual circumstances would fail to transmit across the known entailment (F2), according to this template.²⁷

More generally, for any $C=\text{not}A$, conditions (iii) and (iv) hold for any B entailed by A . Moreover, condition (ii) holds in this case provided my warrant w for A consists in my being in a state which is subjectively indistinguishable from a state in which $\text{not}A$. Hence the A - B - C templates charges a large number of perfectly good arguments with being transmission failures.²⁸

This suggests that any argument with a perceptual warrant w for the relevant A will fit the template and hence be a failure of the perceptual warrant to transmit across the known entailment if A , then B . This is surely too strong, intuitively we would hold that a warrant w does transmit across known entailment in the case of arguments like (F). Two amendments to Wright's template have been offered the first is by Wright himself and the second is an amendment made by Brown on Wright's behalf.

Wright (2003) has suggested the following amendment to his template in 2000. The amendment is for (iii) to be reformulated as:

(iii*) C is incompatible with the reliable operation of the cognitive capacities involved in generating the warrant for A .

"This amendment allows the perceptual warrant for (F1) to be transmitted across the known entailment (F2) since neither the supposition the animal is a fox nor the supposition the animal is a cat is incompatible with the reliable operation of my cognitive faculties" (Brown 2004, p60).

²⁷ Brown (2004) advances this specific example

²⁸ This is McLaughlin (2003)'s general point.

Unfortunately, it may allow a perceptual warrant w for believing (Z1) to be transmitted across the known entailment (Z2) and this is something we intuitively do not want and a result which is incompatible with the standard view. Why may such a warrant w transmit across the known entailment (Z2)? According to Brown it depends on how we individuate cognitive processes and individuate reliability.

“The hypothesis that that animal is a cleverly disguised mule is not incompatible with the general reliability of perception, or with the general reliability of the process of forming beliefs about an animal’s type on the basis of appearance. Of course, whether we type the relevant belief forming process as perception, or as the process of forming beliefs about an animal’s type on the basis of its appearance, this process would not be reliable in forming some particular beliefs, those about the type of animal in the zoo cage.” (Brown 2004, p61)

Brown has suggested on Wright’s behalf perhaps (iii*) could be amended to (iii**) in order for the perceptual warrant w to transmit across the (F) argument and not across the (Z) argument:

(iii**) If C obtains, then the cognitive capacities involved in generating the warrant for A would not be reliable in this particular use.

Other than her suggestion that Wright should replace (iii) with (iii**), Brown has also suggested another way in which Wright’s template can be amended to sustain the standard view. Brown’s suggestion is to leave (i), (ii) and (iii) the same and replace (ii) with (ii*):

(ii*) If C were true, one would be in a state which is subjectively indistinguishable from the state which warrants A.

According to Brown:

“This revised condition neatly blocks McLaughlin’s objection that, on Wright’s template, warrant fails to transmit across every valid argument from A to B where premise A has a non-inferential, non-logically conclusive warrant. Suppose with McLaughlin that $C = \text{not-}A$. Conditions (iii) and (iv) both hold for any B entailed by A. But it does not hold that condition (ii*) holds. Suppose my non-inferential warrant for A is not logically conclusive, so I could be in a subjectively indistinguishable state even if A were false. It follows if C (=not-A) were true, then I could be in a state subjectively

indistinguishable from that which warrants A, but it does not follow that I would be.” (Brown 2004, p59).

Brown notes this feature allows our typical perceptual warrant to transmit across the (F) argument and not transmit across the (Z) argument which is just the result the template needs.

I think Brown’s suggestion of replacing (ii) with (ii*) is on the right track to amending the template. My only worry would be whether the revised template could handle a case like this:

(F’)

(F1) The animal in the garbage is a fox.

(F2’) If the animal in the garbage is a fox, then (the animal in the garbage is not a dog and that animal in the garbage is not a cat cleverly disguised as a fox).

Therefore,

(F3’) That animal in the garbage is not a dog and that animal in the garbage is not a cat cleverly disguised as a fox.

Consider A=(F1) B=(F’3) and C=not-(F3’).

Condition (i) of the modified template is satisfied. Condition (iv) is satisfied since if (F’3) were false, not-(F’3) would be true. Not-(F’3) is incompatible with (F1) since if the animal in the garbage is a dog or a cat cleverly disguised as a fox, then that animal is not a fox. What does condition (ii*) say? Condition (ii*) is unsatisfied since if C were true (the animal in the garbage is a dog or a cat cleverly disguised as a fox, then would be in a state subjectively distinguishable from the state which warrants (F1) when the animal is a dog.

But intuitively we would see (F’) as a failure of a typical perceptual warrant for (F1) to transmit across the known entailment (F’2). So more work may need to be done to Wright’s template to rule out cases like the (F’) argument.

§3 Davies (2000)

I want to suggest that Davies (2000) gives an argument along the lines of Argument 2 when arguing that certain arguments are transmission failures. Davies argues that transmission of warrant across known entailment is subject to the following limitation principle (L):

(L): “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 premiss 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.” (p402).

Now at the moment (L) does not look to be obviously related to Argument 2. In order to see why it is related, we need to look into the arguments which Davies uses to establish (L). These arguments seem to be establishing that certain arguments *beg the question in Jackson’s sense* and certain arguments *fail to rule out alternatives to their conclusion*.

Before I examine these arguments I want to show why Davies, unlike Wright, does not need to subscribe to one of the particular accounts of warrant above in order to argue that certain arguments are failures of transmission. That is to say, Davies can hold either one of the conservative accounts of warrant or a liberal account of warrant or an intermediate account of warrant in order to argue that certain arguments are failures of transmission. In order to see this note in the passage below how Davies seems to allow for two uses ‘background of certain assumptions’ and does not commit himself to either one of them:

In this passage Davies is discussing the (M) argument (Davies uses ‘justification where I use the term warrant):

(M1) Here is a hand.

(M2) If this is a hand, then the external world exists.

Therefore,

(M3) The external world exists.

“On one account, the background assumption that [(M3)] is one which I am entitled to make without justification. The role of this unjustified background assumption in the justification of everyday beliefs like [(M1)] does not prevent those beliefs from being knowledge. On the other account, a proper justification for my belief that [(M1)] would have to include justification of [(M3)].” (p401).

So on the first use of ‘background assumptions’ Davies is referring to the conservative or weak conservative account warrant: in order for S to have a warrant for believing (M1), S requires independent warrant or entitlement to believe (M3).

On the second use of the first passage, Davies holds a liberal or intermediate account of warrant: we just do not need to bother about propositions such as (M3) until evidence comes our way that propositions such as (M3) are false. On this liberal or intermediate view we merely need to lack evidence that certain propositions such as (M3) are false.

If this reading of Davies above is correct then Pryor (ms, p9-10) is incorrect to label Davies as holding conservative account of warrant, since the text clearly suggests above that Pryor is not committed to a specific account of warrant.

I’ll now go on to consider Davies’ discussion of *begging the question in Jackson’s sense* and *ruling out relevant alternatives* in order to reach the conclusion that a certain epistemic structure obtains when we have the intuition that a certain perceptual warrant fails to transmit across a known entailment. In particular I will emphasise that Davies must *combine* both of these discussions to reach this conclusion. Many commentators tend to think Davies uses the discussions of begging the question in Jackson’s sense and ruling out relevant alternatives *separately* to discern the epistemic structure which obtains in cases where certain perceptual warrants fail to transmit across known entailments.²⁹

§3.1.a Davies on begging-the-question in Jackson’s sense

Davies, following Jackson, claims an argument *begs the question* when the evidence for one of the argument’s premises is *no evidence* relative to the background beliefs of the

²⁹ Brown (forthcoming, Ch 7) may have read Davies in this way.

audience who are listening to that argument. For example, consider the argument (Z) again:

(Z)

(Z1) That animal is a Zebra.

(Z2) If that animal is a Zebra, then that animal is not a cleverly disguised mule.

Therefore,

(Z3) That animal is not a cleverly disguised mule.

Suppose that the audience who are being advanced (Z) have background beliefs that not-(Z3). Perhaps they think that the zoo is involved in a mass hoax of the public in the location where (Z) is being advanced or perhaps they believe not-(Z3) for other reasons or for no reason at all. In such a case, whatever the audience think of the evidence for (Z2), even if the audience take themselves to know (Z2), the perceptual evidence offered to them by the speakers assertion of (Z1) is completely compatible with both (Z1) and not-(Z1). Hence the typical perceptual evidence being advanced to the audience is no evidence relative to the audiences background belief that not-(Z3).

Thus if an audience has background beliefs such that they would deny that the evidence for the premises is any evidence, then this a case of question-begging *for this particular audience*. The argument need not be question begging for the speaker, he may have background beliefs such that his evidence for one of the premises does count as evidence. Similarly, an argument may be question begging for one audience but not another. It may be that a different audience has background beliefs relative to which the evidence borrowed from the speaker is evidence.

So far question begging *for certain audiences* has been described. For Davies an argument is *question-begging proper* if and only if

“An 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 will be no evidence.” (p111, Jackson.).

Thus, the audience used to assess whether an argument is question-begging proper is anyone – or anyone sane – *who doubts the conclusion*. So for example, anyone or anyone sane who doubts (Z3) will take the typical perceptual evidence for (Z1) as no evidence relative to this background belief that not-(Z3), according to Davies, hence (Z) is question-begging proper.

Note that in the giving this description of a question-begging proper argument I have used the (Z) argument where the person advancing the argument has the typical visual warrant, w , for believing (Z1). As we noted earlier (Z) where S has the typical visual warrant, w , for believing (Z1) is a problematically warranted argument. So could it be that when an argument is question-begging proper, it is in some way connected to what I have called a problematically warranted argument? I think 'yes'. What is the precise connection between an argument which begs the question proper and a problematically warranted argument?

Recall, an argument Φ where S has a warrant, w , for believing A is a *problematically warranted argument* iff

- (i) Φ is of the form [A, (if A, then B) therefore B]
- (ii) S has a warrant, w , for believing A
- (iii) S knows that (if A, then B)
- (iv) not-B is an undercutting defeater for S's warrant, w , for believing A

Recall the definition of an undercutting defeater:

For any rational subject, S: a proposition φ is an *undercutting defeater* S's prima facie warrant, w , for believing P iff S's belief that φ removes w 's warranting status as a warrant for believing P.

Now if I believe not-B, then my warrant prima facie warrant, w , for believing A will not be a warrant at all. Hence when I believe not-B, the argument Φ where I have the warrant, w for believing A is not a problematically warranted argument since I do not have the warrant, w , for believing A (i.e. condition (ii) above is not satisfied). Suppose we take doubt about θ as a belief that not- θ , then when an argument is

question begging proper, anyone sane who doubts the conclusion of the relevant argument, will take it that certain prima facie warrants for at least one of the arguments premises is no warrant relative to their doubt. In the absence of doubt, a sane or rational person considering a question begging proper argument may take it that her prima facie warrant for believing the at least one of the premises of the argument is a genuine warrant. This seems to indicate that if Ψ is a question begging proper argument where a subject does *not* have doubt about Ψ 's conclusion, then Ψ is a problematically warranted argument.

So the main point here is that in establishing that certain arguments are question begging proper, I think Davies has effectively found another way of individuating problematically warranted arguments. Hence, Davies subscribes to something like (PWA).

§3.1.b Davies on arguments failing to rule out alternatives to their conclusion

Why couldn't Davies just stop after establishing arguments where we have certain prima facie warrants for believing their conclusion are question-begging-proper arguments? If Davies stops at this point he is faced with at least two problems. First, a question begging proper argument where any sane person has a doubt about the conclusion of the argument will not be a case of transmission failure. Why not? Since the sane or rational person doubts the conclusion he would normally take at least one of the typical prima facie warrants for the premises of A to not be a genuine warrant for A. Hence, one of the necessary conditions for transmission failure cases is absent, namely the condition that a person have a typical warrant the major premise of the argument. The second problem can be seen if we consider a case of a question-begging proper argument when a sane person does *not* doubt the argument's conclusion. Davies, so far presented, has remained silent about whether such a case is a case of transmission failure. The second point can be put another way, provided a subject S does not doubt the conclusion of a question begging proper argument, then as we saw above such an argument can be considered to be a problematically warranted argument. Hence Davies has established something roughly like the (PWA) premise. But Argument 2 requires other premises in addition to (PWA) in order to charge problematically warranted arguments as being

transmission failures. I am going to suggest that it is Davies' discussion of certain arguments failing to rule out alternatives to their conclusions which establishes the premises needed to supplement (PWA) in Argument 2.

What are Davies' remarks on certain arguments failing to rule out alternatives to their conclusions? Davies focuses on cases of question begging proper arguments where we have a (rational or sane) subject, S, who does not doubt the conclusion the question-begging proper argument. Hence we are considering the following sort of case:

- (a) We have an argument of the form [A, (if A, then B) therefore B].
- (b) S has a typical warrant, w, for believing A..
- (c) S knows that (if A, then B).
- (d) S does not doubt that B.
- (e) If S doubted B, then S would not consider himself to have the warrant, w, for believing A.

The question is: does S acquire a warrant for believing B, by way of his warrant w for believing A and his knowledge that (if A, then B)? In order to investigate this question, let's say S has a prima facie warrant w+ for believing B by way of his warrant w for believing A and his knowledge that (if A, then B). The question is: is w+ a genuine warrant? Davies' answer to this question is negative.

I think Davies is committed to the, second premise of Argument 2, the *ruling out alternatives claim*.

(ROA)

A prima facie warrant *w* is a warrant for believing P is a genuine warrant only if *w* rules out some of the alternatives to P.

Consider Davies' remarks on the matter below:

Davies is referring to the following argument (G) in the passage:

(G1) A goal has just been scored.

(G2) If a goal has just been scored, this is not a take in a movie scene.

Therefore,

(G3) This is not a take in a movie scene.

“Against the background 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 against a host of other alternative hypotheses. For example, the evidence counts against the hypothesis that...the ball has been kicked out of bounds. In short, evidence rules out the 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 evidence confers knowledge. This is how evidence constitutes epistemic warrant.* But the evidence of ball, crowds 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 considerations which support the conditional premiss, *does nothing to rule out the most obviously salient alternative hypothesis,* namely that it is a movie scene that I am watching. The evidence would be of no help in resolving my doubt and it does confer knowledge. My epistemic warrants for the premises do not add up to an epistemic warrant for the conclusion.” (2000, p399-400, my italics).

The first italicised part of the passage above indicates why I think Davies holds (ROA). The second italicised part of the passage provides further evidence that Davies holds (ROA) since I take it to be saying that the evidence for (G1) does not count against *any* alternative to (G3). It has recently been claimed that Davies is committed to something stronger than (ROA). The stronger claim that Davies is alleged to be committed to is:

Salient alternatives claim:

(SA)

A prima facie warrant w is a warrant for P is a genuine warrant for P only if w rules out relevant or salient alternatives to P.³⁰

³⁰ “Davies claims that warrant fails to transmit from the premises to the conclusion of the goal argument since warrant for the premises fails to rule out the salient alternatives to the conclusion.” (Brown, forthcoming p262)

Now the third italicised part of the passage above may be thought to support the attribution of (SA) to Davies. However, given the first two italicised passages I simply take the third italicised passage to be pointing out since the evidence for (G3) does not rule out any alternative to (G3), then such evidence will not rule out the most salient alternative to (G3).

How does the above passage bear on whether the prima facie warrant w^+ for B? Since one of the background assumptions that is required for the warrant, w , for A, to count as a warrant is the background assumption that B, then, according to Davies, S's prima facie warrant w^+ for believing B will do nothing to rule out any alternatives to B. Hence, w^+ is not a genuine warrant. That is to say Davies is committed to the third and final premise of Argument 2:

The failing to rule out alternatives claim.

(FROA)

S's prima facie warrant for believing B by way of his warrant, w , for believing A and his knowledge that (if A, then B) cannot rule out any alternatives to B.

For concreteness let's consider what Davies would say about the (Z) argument where S has the typical visual warrant, w , for (Z1) that animal is a zebra. w is a genuine warrant for Davies since w rules out various alternatives to (Z1) (i.e. some or at least one alternative). For example, w rules out the alternatives that that animal is a mouse, is a dog, is a parrot, etc. However, w only does this against a background of certain assumptions and one of those assumptions is that (Z3) that animal is not a cleverly disguised mule is true. Hence w cannot rule out not-(Z3). Suppose we call the prima facie warrant for believing (Z3) obtained from S's warrant w and knowledge that (Z2) if that animal is a zebra w_1 . The question is does w_1 rule out any alternative to (Z3)? The answer to this question, according to Davies, would be 'no' since in order to obtain the warrant w_1 S has been assuming in some sense that the alternatives to (Z3) are false (or put another way S has been assuming, in some sense, (Z3) to be true). Therefore, since w_1 does not rule out any alternatives to

(Z3), w1 is not a genuine warrant. So here we have a case of S's warrant, w, for believing (Z1) failing to transmit across the known entailment (Z2).

This exposition of Davies shows that he is roughly committed to Argument 2. Davies then codifies this argument in his limitation principle (L)

§3.2 Responses to Davies (2002)

There are a number of criticisms that can be raised against Davies. The first criticism is that Davies' discussion of warrants ruling out alternatives has controversial consequences. The second criticism claims that question-begging proper arguments do not suggest such arguments are transmission failures. I'm going to argue that these first two criticisms do not have much force since they are either based on misattributing views to Davies or that they fail to recognise how Davies' discussions of begging the question proper and ruling out alternatives work together. A third criticism of my own concerns an ambiguity in the doubt clause of the limitation principle. As we have seen I think there is a problem taking both of these arguments separately on my view they work together to explain why certain arguments fail to transmit across known entailments.

§3.2.a Response 1: Davies' discussion of warrants ruling out alternatives has controversial consequences

This main proponent of this criticism is Brown (forthcoming)³¹. Brown argues that Davies' discussion of warrants ruling out alternatives has at least two undesirable consequences: (a) it rules out our having a warrant for believing propositions obtained from false premises and (b) it rules out our having a warrant for believing false propositions.³²

What is Brown's argument for the undesirable consequence (a)?

³¹ pp259-269

³² A third criticism is that Davies' discussion of ruling out alternatives yields counter-intuitive results on both internalist and externalist accounts of warrant. I don't have space to discuss this here but this line of thought too may involve attributing (SA) to Davies.

Brown asks us to consider the following argument:

(McE)

(McE1) I have just seen McEnroe win this year's Wimbledon final.

(McE2) If I have just seen McEnroe win this year's Wimbledon final, then McEnroe is this year's Wimbledon champion.

Therefore,

(McE3) McEnroe is this year's Wimbledon champion.

Brown then sets the details of this case up as follows. (McE1) is false since the subject considering the (MCE) argument, Julie, is watching a repeat of last year's Wimbledon final where McEnroe beat Connors. The BBC is showing a repeat of last year's final rather than showing this year's final live due to technical details. It is uncontroversial that Julie knows (McE2). And (McE3) also happens to be true since McEnroe has in fact just beaten Connors in this year's Wimbledon final and hence is this year's Wimbledon Champion.

If we allow that Julie has a warrant for believing (McE1) we would intuitively want to allow that this warrant transmits across Julie's knowledge of the entailment (McE2). However, if we attribute the salient alternatives claim (SA) to Davies, he cannot support this intuition:

(SA)

A prima facie warrant w is a warrant for P is a genuine warrant for P only if w rules out relevant or salient alternatives to P.

Davies cannot support our intuitions here when he is attributed (SA). Why not? Let w_1 be the prima facie warrant Julie obtains from her warrant for believing (McE1) and her knowledge that (McE2). Why can't Davies hold w_1 is a genuine warrant if he holds (SA)? One relevant alternative to (McE3) is that McEnroe is not this year's Wimbledon champion and the broadcast is a replay. Julie's warrant, w , does not rule at least this relevant alternative out. Hence Julie's warrant w_1 for believing (McE3)

does not rule at least this relevant alternative out. Therefore, if (SA) is true, w1 is not a genuine warrant for believing (McE3).

However, as I have argued earlier, there is not conclusive textual evidence to attribute (SA) to Davies. Indeed the text supports Davies subscribing to a much weaker ruling out alternatives claim:

(ROA)

A prima facie warrant w is a warrant for believing P is a genuine warrant only if w rules out some of the alternatives to P.

Given (ROA) Julie's (defeasible) warrant for believing (McE1) does in fact rule out some alternatives to (McE1) such as the match has not been rained off and nobody has yet won, the match has not been postponed due to a bomb scare, Becker got to this year's final and beat Connors to win. Sure, (McE1) is false and hence Julie cannot know (McE1) but Julie does have a warrant for believing (McE1), given (ROA). Does Julie's prima facie warrant, w1, for believing (McE3) rule out some alternatives to (McE3)? The answer is 'yes' since w1 is obtained partly by way of Julie's warrant for believing (McE1), her warrant w1 rules out the alternatives that no one is this year's champion due to the match being postponed, or that Becker is this year's champion. Hence, w1 would be a genuine warrant for Davies on my reading of Davies which attributes (ROA) to Davies rather than (SA).

Brown's second charge is that Davies discussion of warrant and ruling out alternatives leads to the counter intuitive consequence (b) that one cannot have a warrant for believing a false proposition. The charge proceeds as follows. (McE1) is false. A relevant alternative to (McE1) is that Julie has not seen this year's final and the broadcast is a repeat. But Julie's warrant for believing (McE1) cannot rule out this relevant alternative. Hence, if (SA) is attributed to Davies, Davies cannot say Julie has a warrant for believing (McE1).

The basic response to Brown's second charge from Davies is to make the same move as above. Deny that Davies is committed to (SA) and instead hold that Davies is committed to (ROA) which allows Davies to hold that false beliefs are warranted.

§3.2.b Response 2: question-begging proper arguments do not suggest such arguments are transmission failures

Beebe (2001) (Beebe has also been seconded by Brown (forthcoming)) has claimed that if an argument begs the question proper, then such an argument is not a case of transmission failure.

Beebe's criticism is based on the contention that there are two ways in which an argument can fail to convince: either the argument is a failure of transmission of a certain warrant to transmit across our knowledge that the premises of the argument entail the conclusion or a valid argument can fail to convince by being incapable of persuading someone who doubts the conclusion.

Beebe then claims that Davies has shown that the problematic arguments like (Z) and (M) fail to convince in the second sense of being incapable of persuading someone who doubts their conclusion but this does not allow Davies to conclude that the arguments fail to convince in the sense that such arguments are transmission failures. Therefore, according to Beebe, the fact that problematic arguments like (Z) and (M) fail to convince in the second sense does not lend any plausibility to the claim that in some cases warrant is not transmitted from the premises of a valid argument when we know the argument's premises entail the argument's conclusion.

Beebe argues for this claim by noting that if a valid argument begs the question proper for a subject, then on Davies' view, the subject will not take his alleged evidence for the one of premises to be any evidence at all given his background beliefs. For example, a subject who doubts that that animal is a cleverly disguised mule will not take any perceptual experiences to be evidence for (Z1) that animal is a zebra and hence will not take himself to have warrant for believing (Z1) on the basis these perceptual experiences. But then in this case, goes Beebe's response, we do not have a case of transmission failure since one of the necessary conditions for a

transmission failure does not obtain – the necessary condition being that the subject have warrant for the major premise of the problematic arguments we have been considering. Hence, even if an argument begs the question proper, it is not a case of transmission failure according to Beebee. This is nicely put in the following quotation:

“We can diagnose an argument as question-begging in a way which simply does not touch on the question of warrant can be transferred from its premises to its conclusion. Just because an argument fails to convince sceptics, nothing at all follows about whether or not anyone who does regard the premises as warranted is entitled to transfer warrant to [its conclusion].” (2001, p359).

I think Beebee’s response crucially misses how Davies’ discussion of question-begging proper arguments and ruling out alternatives work together. Davies could make the following reply to Beebee’s argument. Davies can grant that when an argument begs the question proper anyone sane who doubts the conclusion will not take themselves to have warrant for believing all of the premises of the argument and hence a necessary condition for the argument to be a case of transmission failure for any sane person *who doubts the conclusion of that argument*. However, what Davies will say is that a question-begging proper argument *for any sane person who does not doubt the conclusion of that argument* is a transmission failure. Recall, that my reading of Davies’ argument above is, roughly, that question-begging-proper arguments where subjects who do *not* doubt the conclusion of that argument, fail to produce a genuine warrant for believing their conclusion, if we take a genuine warrant for a proposition to be a warrant which rules out some alternatives to that proposition.

One might think Beebee has anticipated this reply from Davies in the following passage:

“One might try to claim at this point that although there are indeed two senses in which an argument can fail to convince, arguments which fail to convince in a sense they beg the question just are arguments which fail to transmit warrant. In the absence of any reason to suppose that the two senses come apart, we can legitimately claim that they are co-extensive, and hence the failure of an argument to convince sceptics can be used as indirect support for the limitation principle.

However, I think it can be shown that the two senses do come apart with the following argument (A):

A1. *The Australian* says the Crows won the Grand Final

A2. If *The Australian* says the Crows won the Grand Final, then the Crows won the Grand Final

A3 [Therefore] The Crows won the Grand Final.

This, of course, is a perfectly straightforward argument, which might easily and legitimately, convince a normal person that the Crows won the Grand Final. But if we try to run (A) in the face of die-hard sceptics, the sceptics will of course doubt the conclusion, since if they doubt the existence of the world they will likewise doubt the existence of Australian rules football teams. And (A) cannot convince the sceptics into believing (A3) because, as with any argument from perceptual premises, their reasons for doubting (A3) will lead them to deny the implicitly offered perceptual evidence for (A1) is evidence: they will deny that perceptual newspaper-headline experience gives one any reason whatever to suppose that either newspapers or the events they mention really exist." (p359).

The basic line of thought from Beebe is that certain arguments fail to convince sceptics about the external world, such as argument (A). But (A) is intuitively an argument which is not a transmission failure. Hence, if an argument fails to convince a sceptic about the external world who doubts its conclusion, then such an argument is not a transmission failure for non-doubters and many other kinds of doubters about the conclusion.

Again Davies could make the following reply to this challenge from Beebe. There is a distinction between an argument which begs the question proper and an argument which begs the question against the sceptic about the external world. The former sense of begging the question is when anyone sane who doubts the conclusion of the argument regards the prima facie evidence offered for the premises of the argument as no evidence relative to his doubt. The latter case of begging the question is when a sceptic about the external world doubts the conclusion of the argument regards the prima facie evidence for the premises of the premises of the argument as no evidence relative to his doubt. Note that for some arguments the two sense of question begging are equivalent. For example consider the argument (M):

(M)

(M1) Here is a hand.

(M2) If this is a hand, then the external world exists.

Therefore,

(M3) The external world exists.

Here when *anyone* sane or rational doubts (M3), they will effectively be sceptics about the external world provided they are able to recognise the consequences of doubting (M3); and hence they will regard the typical prima facie evidence for (M1) as no evidence relative to their doubt. Similarly sceptic as about the external world will of course doubt (M3) and regard the typical prima facie evidence for (M1) as no evidence relative to their doubt.

Now let's consider argument (A) to which Beebe refers. Let's grant Beebe has established that (A) begs the question for sceptics about the external world. Does (A) *beg the question proper*? Does anyone sane who doubts (A3) regard the typical prima facie perceptual evidence for (A1) as no evidence relative to their doubt? I think the answer is 'no'. Sane people can doubt (A3) for a *variety of reasons* they may think the Crows are going through a bad spell at the moment and have no chance of winning the final, or they may think that the final was postponed because of bad weather and so on (let's call such doubts *mundane doubts about (A3)*). It is true that some sane people will doubt (A3) because they are sceptics about the external world and we have already granted (A) begs the question against such sceptics. The question is does (A) beg the question against anyone who doubts the conclusion? The answer is negative. The typical prima facie perceptual evidence for (A1) does count as evidence for at least some mundane doubters about (A3). For example, suppose I think the final has been rained off due to bad weather (and hence doubt (A3)), then I pick up the Australian newspaper and read that the Crows have won the final; it seems that the evidence from the Australian newspaper does count as evidence relative to this sort of doubt. Hence, for some sane people who doubt (A3) the prima facie perceptual evidence for (A1) will be no evidence relative to their doubt. But for at least some sane doubters about (A3), that is, at least some mundane doubters about (A3). This is sufficient to establish that (A) does not beg the question proper.

More generally for any argument so long as there is *one way* in which a sane or rational person can doubt the conclusion can still regard the evidence offered for the premises as evidence the argument does not beg the question proper. Note in the

case of a doubt about (M3) the external world exists it seemed that the only way to rationally doubt (M3) was to effectively be a sceptic about the external world.

§3.2.c Response 3: ways of doubting

I want to argue here that as it stands (L) is too strong and is open to counter-examples. Before I give this argument, one might wonder why I am bothering to raise such a problem given Davies says:

“There can be no doubt that even this revised version of [(L)] will face counter-examples and will require further modifications.” (p404).

I am raising the particular counter-example that follows for two reasons. First, the counter-example I shall raise will show that the limitation principle does not quite properly codify Davies’ argument that question-begging proper arguments are transmission failures (when considered by subjects who do not doubt the conclusion of such arguments). Secondly, the counter-example I raise here, as we shall see in Chapter 3, may have a bearing on Davies’ verdict on the SK paradox..

Recall that Davies’ limitation principle (L) says: “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 premiss 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.” (2000, p402).

Here’s the counter example. Consider the argument (F):

(F)

(F1) The animal in the garbage is a fox

(F2) If the animal in the garbage is a fox, then it is not a car.

Therefore,

(F3) The animal in the garbage is not a cat.

(F) is intuitively an argument where a typical visual warrant, w, for (F1) transmits across the known entailment (F2).

Suppose I have a typical visual warrant, *w*, for believing (F1). Hence, my warrant *w* is made against at least the background example that the animal in the garbage is not a cat cleverly disguised as a fox. But suppose that I doubt (F3) in one specific way: I believe that the animal in the garbage is a cat cleverly disguised as a fox. Can my specific doubt here be rationally combined with the background assumptions required for *w* to be a warrant for believing (F1)? If I doubt that (F3) in this specific way then it would not be rational for me to make the background assumption that the animal in the garbage is a cat cleverly disguised as a fox. Hence, according to Davies' limitation principle (L) the argument (F) is a failure of my warrant, *w*, for believing (F1) to transmit across the known entailment (F2).

This counter-example reveals that (L) codifies arguments which are not question-begging proper arguments as transmission failures. This is because (F) is not a question-begging proper argument as shown above. Recall, one requirement on non-question-begging proper arguments was that there is at least one way in which a sane or rational person who doubts the conclusion of the argument who regards the typical prima facie perceptual evidence for the premises as evidence relative to his doubt. (L), as it stands, captures arguments which meet such requirements and the intuition is that this is incorrect.

A tentative suggestion of how to modify the limitation principle is as follows:

(L-modified) 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 premiss counts as a warrant only against the background of certain assumptions and acceptance of those assumptions cannot be rationally combined with *any kind of* doubt about the truth of the conclusion.

The italicised words indicate where this modified limitation principle differs from (L).

(L-modified) allows that the argument (F) where one has a typical perceptual warrant for believing the (F1) to transmit across the known entailment (F2) for the following reasons. While doubting (F3) on the grounds that the animal in the garbage is a cat

cleverly disguised as a fox, it is not the only kind of doubt I can have about (F3). I may doubt (F3) on the basis of some inductive evidence that no animals other than cats hang around my garbage and cats are the only animals in the area. I could doubt (F3) as a result of a testimony from a trustworthy source and so on. It seems these latter ways of doubting are rationally combinable with my background assumption that the animal in the garbage is not a cat cleverly disguised as a fox.

§4 Summary

In this Chapter I have distinguished two rough arguments (Argument 1 and Argument 2) which can be used to support our intuitions that many typical perceptual warrants fail to transmit across known entailments. As a result of considering the responses that could be made to these arguments I revealed that the arguments were committed to the following general principles about warrant.

Warrant and rational doubt principle:

(WRD)

For any prima facie warrant w for a proposition P: if w is a genuine warrant for believing (or to believe) P, then w is able to resolve the doubt of a rational subject about P.

And the following premise from Argument 2. *The failing to rule out alternatives claim.*

(FROA)

For any warrant prima facie warrant w for a proposition P: if w is a genuine warrant for believing (or to believe) P, then w is able to rule out some alternatives to P.

If we deny these principles then we can allow that our typical perceptual warrants for the major premise of problematically warranted arguments fail to transmit across known entailments. But we have to concede that such a warrant is very limited in strength and we won't get any stronger warrants for the conclusion of

problematically warranted arguments as a result of our typical perceptual warrant for the argument's major premise and conclusion. That is to say, we allow there is weak warrant transmission in the case of problematically warranted arguments:

Weak warrant transmission principle:

For any subject S and any proposition P and any proposition Q: if S has a warrant, *w*, for believing P and S knows P entails Q, then S has a *weak* warrant for believing Q by way of his warrant *w* for believing P and knowledge that P entails Q.

Where a weak warrant is a warrant for believing a proposition ψ is a warrant which does not resolve rational doubt about ψ and does not rule out any alternatives to ψ .

We deny that there is strong transmission:

Strong warrant transmission principle.

For any subject S and any proposition P and any proposition Q: if S has a warrant, *w*, for believing P and S knows P entails Q, then S has a *strong* warrant for believing Q by way of his warrant *w* for believing P and knowledge that P entails Q.

Where a strong warrant is a warrant for believing a proposition ψ is a warrant which resolves rational doubt about ψ or rules out some alternatives to ψ .

Secondly, I've argued that that the two arguments I've distinguished(Argument 1 and Argument 2) are good approximations of Wright and Davies' templates as to why certain arguments are transmission failures. I also observed that Argument 2 was the stronger argument in the sense that it remained neutral over the theory of warrant required to support its conclusion. Hence, this makes Davies' template for transmission failure of certain perceptual warrants a much stronger template than Wright's, in the sense that one cannot reject it by subscribing to an alternative account of warrant.

Thirdly, I briefly reviewed the charge that Wright's template for transmission failure is far too strong and suggested a new template amended by Brown on Wright's behalf is on the right track but still needs to exclude some further problem cases. I also argued that Davies has the resources to reply to at least two of the current criticisms to his template for transmission failure (the responses from Brown (forthcoming) and Beebee). But I then suggested a different problem for Davies' template and tentatively suggested an amendment to resolve this problem.

I shall now apply the results in this chapter when I consider the transmission failure compatibilist's second step in resolving the SK paradox in Chapter 3.

Chapter 3

Introduction

Recall that the overall transmission compatibilist strategy is:

Step 1': A certain epistemic structure is present in the cases where we intuitively think a certain kind of perceptual warrant fails to transmit across a known entailment.

Step 2': The epistemic structure identified in Step 1' is present in the case where we have non-empirical knowledge that (W1) (I am thinking water is wet) and a priori knowledge that (W2) (if I am thinking that water is wet, then I am in a water-world).

Step 2' allows the transmission failure compatibilist to conclude S's non-empirical warrant that (W1) fails to transmit across S's a priori knowledge that (W2). Hence, according to the compatibilist, S cannot gain a non-empirical knowledge sufficient warrant for the consequent of (W2) (which is (W3)). Therefore, according to the compatibilist, S does not know (W3) non-empirically resolving the SK paradox.

In this chapter I want to review both step 2' of the compatibilist solution in light of the results established in Chapter 2.

First, in §1, I look at ways in which the compatibilist can be simply granted step 2' and examine the various replies one could make to stop the compatibilist using step 2' to establish that non-empirical warrant for (W1) fails to transmit across the a priori known entailment (Z2). I shall argue that these replies indicate a different way we can allow that we have both non-empirical warrant that (W1) and a priori knowledge that (W2).

Secondly, in §2, I examine step 2' itself given the possibility of so called 'illusions of content'. I examine whether step 2' is correct given that I upheld the result that the Wright template was far too strong in the last chapter. Here I conclude that the transmission failure compatibilist solution as stated in Chapter 1 fails but this failure

shows why bullet-biting compatibilism succeeds. I shall argue that this result concurs with Wright's own change of view in his papers in 2000 and 2001.

Thirdly, §3, in I continue with my examination of step 2' and note examine Davies' argument for the conclusion that the (W) is a transmission failure. I attribute this conflicting result to the fact that Davies, unlike Wright, does not allow for the possibility of *an illusion of content*.

§1 Unproblematic non-empirical knowledge that I'm in a water-world

One of the main conclusions from Chapter 2 was that we could deny that arguments like (Z) where I have a typical perceptual warrant for believing (Z1) were failures of transmission provided if we denied certain principles about warrant. More generally Chapter 2 showed my problematically warranted argument Φ where w is a warrant for believing Φ 's major premise is a failure of w to transmit across Φ 's minor premise if we denied the certain principles about warrant. These principles of warrant were:

Warrant and rational doubt principle:

(WRD)

For any prima facie warrant w for a proposition P: if w is a genuine warrant for believing (or to believe) P, then w is able to resolve the doubt of a rational subject about P.

And the following premise from Argument 2. The *ruling out alternatives claim*:

(ROA)

For any warrant prima facie warrant w for a proposition P: if w is a genuine warrant for believing (or to believe) P, then w is able to rule out some alternatives to P.

Denying these principles allowed us to say that the relevant instance of principle below is true:

Weak warrant transmission principle.

For any subject S and any proposition P and any proposition Q: if S has a warrant, *w*, for believing P and S knows P entails Q, then S has a *weak* warrant for believing Q by way of his warrant *w* for believing P and knowledge that P entails Q.

Where a weak warrant is a warrant for believing a proposition ψ is a warrant which does not resolve rational doubt about ψ and does not rule out any alternatives to ψ .

I am going to grant for the sake of argument that the compatibilist has established step 2'. That is to say I'm going to grant that the transmission failure compatibilist has established that our non-empirical warrant for believing (W1) and a priori knowledge that (W2) exhibit the same sort of epistemic structure as the cases where perceptual warrant fails to transmit. On either of my rough arguments in Chapter 2 this means granting the transmission-failure compatibilist that the (W) argument where S has a non-empirical warrant for (W1) and a priori knowledge that (W2) is a problematically warranted argument. More specifically, S has a non empirical warrant, *w*, for believing (W1), S knows a priori that (W2) and not-(W3) (it is not the case that I'm in a water world) is an undercutting defeater for S's warrant *w*. Let's note immediately that, on this analysis, S cannot doubt that he is in a water-world or believe that he is not in a water-world since this would defeat S's warrant, *w*, for believing (W1).

So what does it amount to say that S can have a weak warrant for (W3) by way of his non-empirical warrant, *w*, for believing (W1) and his a priori knowledge that (W2)? It would mean that S has a non-empirical warrant for believing (W3) which cannot rule out any alternatives to (Z3) and cannot persuade a rational doubter that (W3). Hence if (W3) is in fact true, and S has this non-empirical warrant for believing (W3), then S knows (W3) non-empirically (given the warrants for believing have been defined as knowledge-sufficient warrants).

Is this result that S knows non-empirically that (W3) a win for the proponent of the SK paradox over the transmission failure compatibilist? This result obviously conflicts with the result the transmission failure compatibilist wants to establish. However, it does not look like this is a win for the proponent of the SK paradox. This is because the non-empirical knowledge that (W3) which S possesses is harmless. It is harmless in the sense that it cannot persuade a rational doubter that S is in a water world and it cannot rule out any alternatives to S being in a water-world. This allows us to explain away the intuition that premise (4) of the SK paradox is true:

(4) S does not know that (W3) non-empirically.

S cannot be moved to use run through the (W) argument in his head in order to resolve a doubt about (W)'s conclusion or rule out any alternatives to (W)'s conclusion. So it looks like this is a win for the bullet-biting compatibilist. Of course this result is *conditional* on the principles about warrant (WRD) and (FROA) both being false and if these principles were true there would be no such thing as a weak non-empirical warrant for (W3).

§2 Externalism which allows for content-illusion

In 2000 Wright argued that when S has a non-empirical warrant, w, for believing (W1) and S knows a priori that (W2)

There was an proposition C which fitted his A-B-C template below:

An argument of the form [A, (A, then B), therefore B] fails to be a case where a certain warrant [w] for A transmits across the known entailment if A, then B when there is some proposition C such that:

- (v) A entails B
- (vi) My warrant [w] for A consists in my being in a state which is subjectively indistinguishable from a state in which the relevant C would be true
- (vii) C is incompatible with A

(viii) C would be true, if B were false

For C = “the seeming thought which I attempt to token by “I believe water is wet” is content-defective owing the reference failure of the purported natural kind term “water” in my language.” (Wright 2000, p156).

C is a proposition which captures the possibility of an *illusion of content* where both my encounters and my speech communities encounters with water are taken to be illusory. I don't have the space to discuss whether such an illusion of content is a genuine possibility given the relevant instance of externalism. But I will in this section look at the results for the SK paradox if it is a genuine possibility.

According to Wright, we then make an a priori presumption or have an a priori entitlement or have an a priori warrant to believe that not-C which is independent of our non-empirical warrant for (W1).

Wright in 2000 claimed each of the four conditions of the A-B-C template are met for the (W) argument.

Condition (i) is obviously met given we are taking (W2) to be true.

Condition (ii) is met since “My warrant for (W1) consists in my being in a state which is subjectively indistinguishable from a state in which the relevant proposition C, that the seeming-thought which I attempt to express by “I believe water is wet” is content defective owing the reference failure of the purported natural kind term “water” in my language, would be true.

Condition (iii) is obviously satisfied C is incompatible with (W1).

According to Wright (*in 2000*) condition (iv) is satisfied C would be true if B (W3) were false. But since the Wright has changed his view:

“But there is a misgiving about it which needs to be confronted. It concerns the fourth condition of the disjunctive template—specifically the claim that when strong

externalism is assumed, freedom by me and my speech community from all historical contact with water would suffice to induce content-defectiveness into our purported water thoughts. That is unquestionably so when the alternative scenarios considered are of the Dry Earth type. But what justifies us—theorists—in restricting attention to those? If Twin-Earth scenarios are considered as well, then our purported water thoughts would suffer not divestment of content but change of content. Had our actual encounters been with twater, not water, the relevant B would have been false. But the relevant C would not have been true: my term, 'water', would not have been divested of content but would have expressed the concept, twater, instead. So the template, it seems, does not cleanly fit the case.” (2001, p13).

As we saw in Chapter 2 Wright’s template has recently been criticised. In Chapter 2 I took the most plausible amendment to be Brown’s changing of condition (ii) but had some misgivings about whether this change was exactly correct. But the main point here is that even on revised versions of the template condition (iv) remains unchanged. So it looks like on Wright’s template non-empirical warrant for believing (W1) will transmit across the a priori known entailment (W2). Hence, S will have a non-empirical warrant that he is in a water-world and hence know non-empirically that he is in a water world. Is this conclusion problematic? Judging from the quote above from Wright it seems not – all non-empirical knowledge that I am in a water world amounts to is that I know non-empirically my thought expresses some kind of content – rather than no content at all.

This point is put well by Noordhof:

“In another sense, though, my a priori knowledge is not based on a capacity to identify something highly specific about the world at all... [Given externalism] subjects cannot only distinguish between their [water thought which has as its extension H₂O] and [their water thought that has as its extension XYZ], they also cannot distinguish their thought that water [which has as its extension H₂O] is wet from the thought that water is wet where ‘water’ picks out electrical stimulations brought about in a brain in a vat, or the machinations of a malicious demon, or a whole range of other phenomenally similar but microstructurally different substances in different possible worlds. By the same token subject’s a priori knowledge that they are in a [water-world] is compatible with their ability to distinguish between all the scenarios just described” (2004, p55).

Hence, when if have the intuition that we cannot know non-empirically that we are in a water-world. Then we may have in mind that the empirical proposition we know is very specifically identifying an environmental condition. But when we realise such knowledge is indistinguishable from changes of content and illusions of content, then us having non-empirical knowledge of a very general empirical proposition is not intuitively intolerable.

It would be interesting to see how things work out on Davies' limitation principle, if we allow for illusions of content (Note that this is not Davies' position, since as we shall see below Davies does not allow for illusions of content.)

Recall in Chapter 2 I modified Davies template as follows:

(L-modified) 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 premiss counts as a warrant only against the background of certain assumptions and acceptance of those assumptions cannot be rationally combined with *any kind of* doubt about the truth of the conclusion.

The question is: if I doubt I am in a water-world, then can I rationally combine my doubt with the background assumptions required for me to have a non-empirical warrant for me to think that water is wet? This will depend on your view about what the background assumptions for self-knowledge are. Suppose one of those background assumptions is I'm not having an illusion of content about water. Is such a background assumption rationally combinable with some kind of doubt I'm not in a water world? I think the answer is 'yes' suppose I don't doubt I'm having a content illusion but I have a mundane doubt about whether my thoughts have some kind of content. Then it seems here, I can rationally combine this doubt with my non-empirical warrant that I am thinking that water is wet. This result is enough to secure the result that there is no failure of transmission of warrant in the SK paradox. But it does not mean to say the proponent of the SK paradox wins. Our investigation as to why the transmission failure compatibilist solution fails ends up revealing why the bullet biting compatibilist strategy seems correct. Again however, this result is conditional, on the SK paradox allowing for illusions of content.

§3 Davies: An epistemic principle which denies the possibility of content illusion

In this section I want to review Davies (2000)'s actual strategy for dissolving the SK paradox.

Davies' first move to resolving the SK paradox is to claim that the following condition, (P), is necessary for any subject S to have any warrant for a proposition:

(P) "[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." (p410).

Davies' reasons for this claim are:

"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 or to confirm as the proposition A, then there could be no question of anything justifying my believing A." (p410).

Note that (P) denies us the possibility of an illusion of content. This is because an illusion of content would allow that one can have a warrant for believing one's 'seeming thoughts that water is wet'. (P) denies this possibility: in order to have a warrant for 'I am thinking water is wet', then 'I am thinking that water is wet' must be a proposition. By the externalist thesis (W2) if I am thinking that water is wet, then I am in a water-world. Hence (P) has the consequence that in order to have a warrant for 'I am thinking water is wet', then it must be true that I am in a water-world.

Davies' next step in resolving the paradox is to transpose his limitation principle (L) in Chapter 2 into a slightly different limitation principle which I shall label (L').

(L') "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 premiss and (ii) of the warrants for the other premises cannot be rationally combined with doubt about the truth of the conclusion." (p412).

What is the difference between (L') and Davies' limitation principle in Chapter 2 (L).

Recall that (L) is:

(L): "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 premiss 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." (p402).

One difference is that (L') does not speak about background assumptions in general but individuates the one particular background assumption namely the background assumption (P) there must be such a proposition as for the knower to think of as that premise, if the knower is to have a warrant for that premise. The second difference is that (L') is meant to apply to arguments with more than two premises.

(L') is quite clearly met in the case of the (W) argument if I doubt that I'm in a water-world but my having a non-empirical warrant for (W1) and externalism has the consequence that I am in a water-world, then the two would not be rationally combinable. This is because if doubt a consequence of (W1) and grant externalism, then if I am rational that would lead me to doubt (W1). Now if I doubted (W1) the rational thing for me to do would be for me not to take my non-empirical warrant for (W1) as a warrant. This is because there would be no such thing for me to think as water is wet and I couldn't have a warrant for some sort of content-illusion token of 'I'm thinking water is wet' because the possibility of content illusion has been denied.

In so far under the assumption that possibility of content illusion is denied I think this is the correct response. Moreover, once this view is combined with the result in Chapter 2 that many of the criticisms of Davies limitation principles have failed it lands the proponent of the SK paradox in the following position. It establishes that it is not the case that S knows non-empirically that (W3) by way of his non-empirical knowledge that (W1) and a priori knowledge that (W2). The natural move for the proponent of the SK-paradox would then be to criticise Davies' arguments for his limitation principles but as saw in Chapter 2 at least some of these criticisms fail. Once again this is a conditional conclusion, it depends on whether the principle (P) is true:

(P) “[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.” (p410).

I won't try to settle the truth of (P) here. I will only note that if the proponent of the SK paradox denies (P) then he allows for the possibility of an illusion of content. But if the proponent of the SK paradox allows for an illusion of content, then as we saw in §2 this gave the bullet biting compatibilist good reason to explain why non-empirical knowledge that I'm in a water-world is not absurd.

§4 Summary

In this chapter I have shown that a number of flaws in the transmission failure compatibilist response to the SK paradox have revealed why S's non-empirical knowledge that (W3) (I am in a water world) is not absurd. On one view, provided the principles about warrant (WRD) and (ROA) are denied, such knowledge is not absurd because it is harmless knowledge in that it cannot be used to convince a rational doubter about (W3) or used to rule out alternatives to (W3). If we allow for illusions of contents in our thoughts, then non-empirical knowledge that I'm in a water-world is not absurd because it only allows S to know that his thought has some kind of content but this knowledge does not allow S to distinguish between being in a water-world and having a content illusion. Hence, on these views, subject to certain assumptions, even though transmission failure compatibilism is wrong, the SK paradox is also wrong to claim (4) S does not have non-empirical knowledge that (W3) (I am in a water world).

I have also shown that if we allow for the principle about warrant (P), then Davies' compatibilist solution seems to be sound. Moreover, if the proponent of the SK paradox were to then try and refute some part of Davies' argument for these limitation principles, then this may be a difficult thing for him to do, given Davies has been adequately defended against some criticisms in Chapter 2. And also if the proponent of the SK paradox denies (P) as a response to Davies, he just invites the bullet-biting compatibilist's response that this shows why non-empirical knowledge that I'm in a water-world is unspecific and harmless.

Conclusion

In Chapter 1 I defended the view that the most plausible SK paradox is as follows:

The paradox is that it is possible for any subject S to think through the (W) argument resulting in:

- (1) S knows non-empirically that (W1) I am thinking that water is wet.
- (2) (W2) If I am thinking that water is wet, then I'm in a water-world: (i.e. if (W1), then (W3)).
- (2a) S knows a priori that (W2)

Therefore,

- (3) S knows non-empirically that (W3) I'm in a water-world.

(3) then conjoins with (4) below:

- (4) It is not the case S knows non-empirically that (W3) I'm in a water-world.

In order to derive the absurd conclusion (5):

- (5) S knows non-empirically that (W3) and it is not the case that S knows non-empirically that (W3).

I identified the principle which derives (3) from (1) and (2a) as being the *Warrant Transmission Principle*.

For any subject S and any proposition P and any proposition Q: if S has a warrant, *w*, for believing P and S knows P entails Q, then S has a warrant for believing Q by way of his warrant *w* for believing P and knowledge that P entails Q.

And showed that the compatibilist needs to deny the relevant instance of the principle for the (W) argument. I identified the compatibilists strategy for doing this as making two broad steps:

Step 1': A certain epistemic structure is present in the cases where we intuitively think a certain kind of perceptual warrant fails to transmit across a known entailment.

Step 2': The epistemic structure identified in Step 1' is present in the case where we have non-empirical knowledge that (W1) (I am thinking water is wet) and a priori knowledge that (W2) (if I am thinking that water is wet, then I am in a water-world).

Step 2' allows the transmission failure compatibilist to conclude S's non-empirical warrant that (W1) fails to transmit across S's a priori knowledge that (W2). Hence, according to the compatibilist, S cannot gain a non-empirical knowledge sufficient warrant for the consequent of (W2) (which is (W3)). Therefore, according to the compatibilist, S does not know (W3) non-empirically resolving the SK paradox.

In examining the compatibilist's Step 1'. I identified two arguments (Argument 1 and Argument 2) that the warrant transmission principle fails for certain cases of perceptual warrant. I argued that Wright can be approximated as holding Argument 1 and Davies could be approximated as holding Argument 2. I noted that Davies' argument could be seen as the much stronger argument in the sense that it remained neutral over what account of warrant is correct and replied on Davies' behalf to a number of criticisms. In discussing Arguments 1 and 2 that one could respond to either of the arguments and allow the transmission principle true for every kind of perceptual knowledge by denying the following principles about warrant:

Warrant and rational doubt principle:

(WRD)

For any prima facie warrant w for a proposition P: if w is a genuine warrant for believing (or to believe) P, then w is able to resolve the doubt of a rational subject about P.

The ruling out alternatives claim:

(ROA)

For any warrant prima facie warrant w for a proposition P: if w is a genuine warrant for believing (or to believe) P, then w is able to rule out some alternatives to P.

I then applied the these results to the SK paradox above and concluded that:

First, if the principles (ROA) and (WRD) is denied and the (W) argument has the correct epistemic structure for transmission failure, then the compatibilist cannot reach his desired conclusion since we do gain very limited non-empirical knowledge by way of the (W) argument. But such knowledge is not absurd because it is harmless knowledge in that it cannot be used to convince a rational doubter about (W3) or used to rule out alternatives to (W3).

Secondly, provided we allow for illusions of content, then the compatibilist strategy fails on both Wright and Davies' own templates for transmission failure. But again knowledge that I'm in a water-world is not absurd in this case because it only allows S to know that his thought has some kind of content rather than no content at all. Hence, on these views, even though transmission failure compatibilism is wrong, the bullet biting compatibilist is correct.

Thirdly when the epistemic principle (P) is affirmed:

(P) 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

Then the possibility of an illusion of content is denied, and Davies' limitation principle does in fact stop (3) being derived from (1) and (2) in the SK paradox.

It is worth thinking about these results from the perspective of the proponent of the SK paradox. In order to stop the bullet-biting compatibilist response he needs to affirm the epistemic principles (WRD) and (ROA) but he also needs to deny the (P) principle to block the transmission failure compatibilist response from Davies. But also the second result has it that content-illusion needs to be denied to stop the bullet-biting compatibilist response. Therefore the proponent of the SK paradox has to find a way of denying the possibility of a content illusion and deny (P) given he is affirming (WRD) and (ROA).

Bibliography

- Beebe, H. (2001) 'Transfer of warrant and Begging the Question' *The Philosophical Quarterly* 51 pp356-374.
- Boghossian, P. A. (1997) 'What the externalist can know a priori' in Wright, C.; Smith, B.C. and Macdonald, C.M. (eds) *Knowing Our Own Minds*, pp 271-284 Oxford: Oxford University Press.
- Brewer, B. (2000) 'Externalism and a priori knowledge of empirical facts' in Boghossian, P.A. and Peacocke, C. (eds) *New essays on the a priori*. New York: Oxford University Press.
- Boghossian, P.A. (2000) 'Knowledge of Logic' in Boghossian, P.A. and Peacocke, C. (eds) *New essays on the a priori*. New York: Oxford University Press.
- Brown, J. (forthcoming) *Anti-individualism and Knowledge*
- Brown, J. (2004) 'Wright on transmission failure' *Analysis* 2004 pp57-67
- Brown, J. (1995) 'The incompatibility of anti-individualism and privileged access'
- Brueckner, A. (1992) 'What an anti-individualist knows a priori' *Analysis* 52 pp111-118.
- Burge, T. (1988) 'Individualism and self knowledge' in Cassam (ed.) (1994) *Self-knowledge* Oxford: OUP.
- Davidson, D. (1988) 'Reply to Burge' *Journal of Philosophy* Volume 85 pp664-665.
- Davidson, D. (1987) 'Knowing one's own mind' in Cassam (ed.) (1994) *Self-knowledge* Oxford: Oxford University Press.
- Davies, M. (2000) 'Externalism and Armchair Knowledge' in *New essays on the a priori* (ed.s Boghossian and Peacocke). New York: Oxford University Press.

Davies, M. (1998) 'Externalism, Architecturalism, and Epistemic Warrant', in Wright, C.; Smith, B.C. and Macdonald, C.M. (eds) *Knowing Our Own Minds*, pp321-361.
Oxford: Oxford University Press.

Dretske, F. (1970) 'Epistemic Operators' *Journal of Philosophy* 67 pp1007-1023.

McKinsey (2003) 'Transmission of Warrant and Closure of Apriority' in Nuccetelli, S. (ed) *New Essays on Semantic Externalism and Self-Knowledge* pp101-125
Cambridge: MIT Press.

McKinsey (2002) 'Forms of externalism and privileged access' *Philosophical Perspectives* 16 pp199-224.

McKinsey (1991) 'Anti-individualism and privileged access' *Analysis* 51 pp9-16.

McLaughlin (2003) 'McKinsey's challenge, warrant transmission and scepticism' in Nuccetelli, S. (ed) *New Essays on Semantic Externalism and Self-Knowledge* pp101-125
Cambridge: MIT Press.

Moore, G.E. (1959) 'Proof of an External World' in *Philosophical Papers* London: Allen & Unwin pp127-50.

Noordhof, P. (2004) 'Outsmarting the Brown-McKinsey reductio' *Analysis* 2004 pp50-56

Pollock, J. L. (1974) *Knowledge and Justification* (Princeton: NJ)

Pryor, J. (ms) 'Is Moore's Argument an Example of Transmission-Failure?' available at <<http://www.princeton.edu/~jimpryor/papers/Transmission.pdf>>

Pryor (2001) 'The sceptic and the dogmatist' *Noûs* 54 pp517-459.

Silins, N. (ms) 'Transmission Failure Failure'

Warfield, T.A. 'When epistemic closure does and does not fail.' *Analysis* 2004 pp36-40

Wright (2002) 'Anti-Sceptics: Simple and Subtle: Moore and McDowell' *Philosophy & Phenomenological Research* 65 pp330-348.

Wright, C. (2001) Remarks for APA Central Divisional Meetings, Symposium on *Externalism, Self-knowledge, and Skepticism*

Wright, C. (2000) 'Cogency and Question-begging: Some reflections on McKinsey's Paradox and Putnam's Proof' *Philosophical Issues* 10 pp140-163.

Zalabardo, J.L. (ms) 'Externalism, Scepticism and the Problem of Easy Knowledge.'