BLIND REASONING

by Paul Boghossian and Timothy Williamson

I—Paul Boghossian

ABSTRACT The paper asks under what conditions deductive reasoning transmits justification from its premises to its conclusion. It argues that both standard externalist and standard internalist accounts of this phenomenon fail. The nature of this failure is taken to indicate the way forward: basic forms of deductive reasoning must justify by being instances of 'blind but blameless' reasoning. Finally, the paper explores the suggestion that an inferentialist account of the logical constants can help explain how such reasoning is possible.

I

The Question. I'm in the mood for some music; what, I wonder, is on offer today in Carnegie Hall? A quick check of the schedule reveals that Martha Argerich is scheduled to play on the 20th. As a result, I come to believe that:

(1) If today is the 20th, then Martha Argerich is playing today in Carnegie Hall.

A glance at the calendar reveals that in fact:

(2) Today is the 20th.

With these two beliefs in place, I move immediately to the conclusion that:

(3) Martha Argerich is playing today in Carnegie Hall.

And I pick up the phone.

If, prior to making this modus ponens inference, I already believed (presumably with only a low level of confidence) that Martha Argerich was playing today in Carnegie Hall, then the inference looks to have strengthened whatever justification I had for that belief. If, prior to making the inference, I had no views about who was playing today in Carnegie Hall, the inference looks to have augmented my beliefs with a further justified belief. Whichever scenario obtained, how did my two premises contribute to justifying the conclusion that I drew on their basis? Under what conditions does an inference transfer justification in this way?

Clearly, at the very least, the following two conditions must be satisfied. First, the thinker must be justified in believing the premises. Second, his justification for believing the premises must not depend on his being *antecedently* justified in believing the conclusion.

Equally clearly, though, these conditions do not suffice for the inference to transfer justification. In addition, the premises must bear an appropriate relation to the conclusion they ground. And my question is: What is that relation?

In this paper, I am going to restrict myself to asking this question about *deductive* inference, leaving it an open question to what extent what is said here generalizes to other cases of justification- or warrant-transfer.[1] In a deductive inference, the thinker takes his premises to justify his conclusion in part because he takes them to necessitate it.[2]

II

Inferential Externalism. The simplest possible answer to our question is this:

> **(Simple Inferential Externalism):** A deductive inference performed by S is warrant-transferring just in case (a) S is justified in believing its premises (b) S's justification for believing its premises is suitably independent of his justification for believing the conclusion, and (c) the implicated pattern of inference is valid—necessarily such as to move S from truth to truths.[3]

1. Even with this restriction in place, the paper covers a lot of ground rather quickly. I am aware that many of its claims need more detailed support than is possible within present limits. My aim is to offer a broad view of the terrain with the hopes of distinguishing the dead ends from the promising pathways.

2. It's a tricky question how this 'taking' is to be understood, but I can't pause to consider the matter here. In this paper, I shall use the terms 'justification,' 'warrant' and 'entitlement' interchangeably.

3. I shall soon be contrasting this externalist conception of inference with a diametrically opposed internalist conception. The idea of converging onto the (hopefully) correct view of inference by picking a course between these two traditionally opposed extremes was first suggested in my 'How are Objective Epistemic Reasons Possible?' reprinted in J. Bermudez and A Millar (eds.): *Reason and Nature* (Oxford: Oxford University Press, 2002), pp. 1–47; see pp. 35–39. In his commentary on that paper, 'Basic Logical Knowledge: Reflections on Paul Boghossian's 'How are Objective Epistemic Reasons Possible?', also reprinted in Bermudez and Millar, Crispin Wright suggested the labels 'simple internalism' and 'simple externalism' for the traditionally opposed extremes, labels which I am happy now to modify and adopt.

I should emphasize that I am asking *by virtue of what facts* a deductive inference transfers warrant, and not just under what conditions it does so.

This can't be a good answer: large numbers of inferences that we are in no intuitive way justified in performing satisfy the stipulated conditions. For example: it is easy for me to be justified in believing any particular claim of the form:

x, y, z, and n are whole numbers and n is greater than 2.

If I inferred from this proposition that:

$x^n + y^n$ is not equal to z^n

I would have performed an inference that is, as we now know, reliably truth-preserving. But it would be absurd to suppose that anyone making such an inference would be drawing a justified conclusion, whether or not they knew anything about Andrew Wiles's proof of Fermat's last theorem or had checked each individual inequality. Being valid (in conjunction with the other two conditions) clearly does not *suffice* for being warrant-transferring, even if it may be necessary.[4]

A number of philosophers seem to believe that this objection can be easily met. All we need to do, they say, is to restrict the Reliabilist claim to those inferences that are sufficiently 'simple'.[5]

But what could 'simple' possibly mean here if not something like: an inference whose validity it is easy to 'take in'? If that isn't what 'simple' means, why should an inference's simplicity be relevant to the question of justification? If that is what 'simple' means, then it's this presumed 'taking in' that's doing the relevant explanatory work, and not the assumed reliability of the simple inference. In a moment, we will consider views according to which an inference is entitling just in case the thinker can 'take in' its validity, but this is not where such views belong.

The counterexamples to Simple Inferential Externalism echo a broader class of objections to Reliabilist accounts of justification more generally. Since many of these examples are so well known, I will not discuss them here in detail except to say this. I am not so impressed with those counterexamples that rely on the subject's having a justified belief to the effect that some reliable belief-forming method of his is not reliable; I think a Reliabilist

---

4. I shall come back to the question whether it is necessary.
5. I have often encountered this suggestion in conversation; I didn't manage to track down a published reference.

can handle those cases by imposing a 'no-undermining belief' condition. But I am inclined to regard as decisive those counter-examples in which the reliability of the relevant method is *not* subjectively undermined—either because the subject has no *justi-fied* belief about its reliability, or because he has no belief about it at all. Laurence Bonjour has described just such a case.

> Norman, under certain conditions which usually obtain, is a com-pletely reliable clairvoyant with respect to certain kinds of subject matter. He possesses no evidence or reasons of any kind for or against the general possibility of such a cognitive power or for or against the thesis that he possesses it. One day Norman comes to believe that the President is in New York City, though he has no evidence either for or against this belief. In fact the belief is true and results from his clairvoyant power under circumstances in which it is completely reliable.[6]

Our robust response to this case is that Norman is not justified. And a plausible and widely-accepted diagnosis of our response is that we are reluctant to regard someone as justified in holding a given belief if they are being epistemically *irresponsible* in hold-ing that belief. Being justified is, at least in part, a matter of being epistemically blameless.

What lesson should we draw from these counterexamples? The moral seems almost forced. We need to ensure that being justified excludes being epistemically blameworthy. Mere reliability doesn't do that. How is that to be ensured except by insisting that, if a subject is to be justified in believing some proposition p, he must have to hand a *reflectively accessible* warrant for the proposition that p?

It looks, in other words, as though the counterexamples to Reliabilism motivate an *Access Internalism* about justification: S is justified in having the belief that p only if S is in a position to know,[7] by reflection alone, that he has a warrant for the belief that p. If S is to have genuine justification, it must be a reflec-tively *transparent* justification.[8]

6. See L. Bonjour: *The Structure of Empirical Knowledge* (Cambridge: Harvard Uni-versity Press, 1985), p. 41.

7. In insisting on knowledge here, as opposed merely to justified belief, I follow standard presentations of Internalism, though my arguments will depend only on the weaker condition.

8. Note that what I am calling 'Access Internalism' is the weaker of the two possible versions of Internalism: it requires only that the epistemic fact be reflectively access-ible, if the person is to be justified, not that the person actually have accessed it.

III

Inferential Internalism. If we go along with this diagnosis of the failure of Reliabilism and apply it to the case we are focusing on—the case of warrant transfer by deductive inference—the thought would then be that, in a deductive inference, the subject must be in a position to know by reflection alone that his premises provide him with a good reason for believing the conclusion, if his inference is to justify his conclusion. Call this

> **(Simple Inferential Internalism):** A deductive inference performed by S is warrant-transferring just in case (a) S is justified in believing its premises (b) S's justification for believing its premises is suitably independent of his justification for believing the conclusion, and (c) S is able to know by reflection alone that his premises provide him with a good reason for believing the conclusion.

There are any number of problems with this idea. Take the case of our original Argerich modus ponens (MPP) inference, from (1) and (2) to (3). What would it be for S to be able to know by reflection alone that his premises provide him with a good reason for believing the conclusion?

Since we are dealing with a deductive inference, the most natural suggestion is that S would know this epistemic fact only if he knew that his premises necessitate his conclusion. So the question becomes: How might S be in a position to know by reflection alone that p and 'p→q' imply q?

There look to be two options: inferential and non-inferential. The inferential route, which would include empirical and broadly pragmatic accounts of knowledge of logical implication, may be ruled out immediately.

The point is a subtle one. It's not that there is necessarily a problem with rule-circular justifications of meta-logical claims. Indeed, I have elsewhere argued that, if we are to have knowledge of basic logical or meta-logical truths at all, it must be via rule-circular reasoning.[9] However, in the present context, where knowledge of the validity of MPP is supposed to be a component in the justification that we have for *reasoning* according to MPP,

---

9. See 'Knowledge of Logic', in P. Boghossian and C. Peacocke (eds.): *New Essays on the A Priori*, (Oxford: Oxford University Press, 2000), and 'How are Objective Epistemic Reasons Possible?', *op. cit.*.

an inferential route to knowledge of the validity would be completely useless. The very sort of reasoning whose justification is at issue would have been presupposed.

What about the non-inferential options? If a given item of knowledge is non-inferential, then it is either justified by observation alone or it is justified by nothing.

For reasons that I don't have the space to rehearse here, it seems to me very implausible that one can be said to know the general proposition that any argument of the form MPP is valid on the basis of nothing, as though all one would have to do to be justified in believing such an ambitious proposition is simply to believe it.[10] If there is to be any hope for the non-inferential option, it must lie along the observational branch.

IV

Rational Insight. Clearly, what is at issue is not ordinary empirical observation. Rather, the idea is that we are equipped with a special capacity for non-empirical observation, a capacity whose exercise is capable of yielding insights into necessary truths. That we have such a capacity has recently been defended by Laurence Bonjour.

> When I carefully and reflectively consider the ... inference ... in question, I am able simply to see or grasp or apprehend ... that the conclusion of the inference must be true if the premises are true. Such a rational insight, as I have chosen to call it, does not seem to depend on any particular sort of criterion or any further discursive or ratiocinative process, but is instead direct and immediate.[11]

In finding the idea of rational insight attractive, Bonjour joins a venerable tradition, one that stretches from Plato through Leibniz to Gödel. Why, then, does the idea have so few supporters these days?

The single most influential consideration against rational insight theories can be quite simply stated: no-one has been able to explain, clearly enough, in what an act of rational insight

10. See my 'How Are Objective Epistemic Reasons Possible?'.
11. L. Bonjour: *In Defence of Pure* Reason, (Cambridge: Cambridge University Press, 1998), pp. 106–107

could intelligibly consist. No-one denies, of course, that we can *think about* properties and relations, including logical properties and relations, and that, as a result, we can reason our way to general conclusions about them, just as we can reason our way to conclusions about other topics. The question is whether we can be said to have some sort of *non-discursive*, *non-ratiocinative*, *insight* into their natures, an insight that would disclose immediately, and without the help of any reasoning whatsoever, that all instances of MPP are truth-preserving.

Bonjour does not succeed in convincing us that there is. On reading his book, one can't help but be struck by how little progress a resolute and resourceful defender of rational insight was able to make.

There are many issues that could be discussed with profit; here I have space only to raise the fundamental issue of cognitive access. What sort of relation obtains between a thinker and a property, when the thinker has 'rational insight' into its nature? Or, to pick a more immediately relevant example, what relation obtains between a thinker and the conditional, when the thinker has a rational insight into its nature, when he is simply able to *see* that when his MPP premises are true so must be its conclusion?

The analogy with sense perception encourages us to think that the relation between the thinker and the conditional is causal. But that is impossible since the conditional is just an abstract object. Bonjour wrestles with this issue and comes up with a three-part response.

The first is to downplay the analogy with sense perception. The second is to propose the revival of an Aristotelian conception of mental content, according to which when we think about a given property—triangularity, for example—that very property itself is *instantiated* in our thought. Finally, Bonjour suggests assimilating rational insight to a form of introspective examination of that thought content.

His proposal raises more questions than it answers. First, we are not really told how the property of triangularity could be instantiated by my thoughts without their being—absurdly—triangular as a result. Second, we are not told how this neo-Aristotelian conception of mental content would help, even if it were not mystifying. Suppose my thoughts about a particular

triangle do instantiate the property of triangularity. How does this help explain how I am able to directly divine the nature of triangularity without the help of any reasoning whatsoever? The belief that the sum of the interior angles of any triangle is 180 degrees and the belief that it is 320 degrees both instantiate triangles. Why is the first justified and the second not?

We are left staring at the problem with which we began, rather than feeling that we have been placed on the path to real understanding.

V

Rational Insight and Carrollian Circularity. Even if we waived worries about the very cogency of the capacity for rational insight, however, it's not clear how rational insight into the validity of MPP could help vindicate Simple Inferential Internalism, and this for reasons that are reminiscent of Lewis Carroll's famous note 'What the Tortoise Said to Achilles.'[12]

According to Simple Inferential Internalism, a thinker's move to (3) on the basis of (1) and (2) counts as justified only if he is in a position justifiably to believe that his inference is valid. Let us suppose, for the sake of argument, that a capacity for rational insight explains how he is in a position justifiably to believe that, if his MPP premises are true, then so must be his conclusion. How might this explanation go?

For obvious reasons, it's not plausible to think of this capacity for rational insight as operating on individual inferences one by one, generating for each of them the insight that if its premises are true, then so is its conclusion. Rather, we suppose that rational insight equips the thinker to arrive at the wholly general insight that MPP is valid, that is that:

For all p, q: Necessarily: If both p and 'p→q', then q.

Now, however, we need to ask how such a justified belief in the validity of all arguments of the form MPP could help a thinker be justified in performing any *particular* MPP inference, for example, the Argerich inference with which we began.

12. *Mind*, 4, 1895, pp. 278–280.

To bring this knowledge to bear on the justifiability of that inference will, it would seem, require the thinker first to establish its relevance to that inference, by reasoning as follows:

(i)   Any inference of the form MPP is valid.
(ii) This particular inference, from (1) and (2) to (3) is of MPP form.

Therefore,

(iii) This particular inference from (1) and (2) to (3) is valid.[13]

Rational insight, we are conceding, gets us as far as the general propositional knowledge that all arguments of MPP form are valid. However, to bring this knowledge to bear on the justifiability of any particular inference will require the thinker to be able justifiably to infer the validity of that particular inference from the validity of all arguments of MPP form. And this will require him to be able to reason according to MPP justifiably.

Now, however a fatal circularity looms. To infer from (1) and (2) to (3) justifiably, I must be able justifiably to believe that the inference from (1) and (2) to (3) is valid. To be able justifiably to believe that this inference is valid, I must be able justifiably to infer that it is valid from the general proposition that all inferences of its form are valid. To be able justifiably to infer that it is valid from the general proposition that all inferences of its form are valid, I must be able justifiably to infer according to MPP. So, on the picture on offer, my inference from (1) and (2) to (3) will count as justifying only if I am already able to infer according to MPP justifiably. The very ability we are trying to explicate is presupposed by the internalist account on offer.

At this point, an internalist might be tempted by the following thought. So long as we are being so concessive about rational insight, why can't we grant thinkers rational insight into the validity of *specific* inferences, and not require that this be derived from some general knowledge of the validity of all inferences of the form MPP. Perhaps this more general knowledge could be

_____

13. This argument was originally presented in my lecture at the Pacific APA in Albuquerque in April 2001.

arrived at later, by using the knowledge gained through these acts of individual insight?

It's important to appreciate that this manoeuvre will not help overcome Simple Inferential Internalism's difficulties with the problem of circularity. For, once again, we can ask how my knowledge of the validity of the inference from (1) and (2) to (3) is supposed to bear on my warrant to infer (3)?

According to Simple Inferential Internalism, this inference will be justified only if I am able justifiably to believe that that my premises provide me with a good reason for drawing the conclusion. But it is very hard to see, once again, how my putatively justified judgment that my premises entail my conclusion could bear on my entitlement to draw the conclusion in anything other than inferential form, thus:

(iv)   This particular inference from (1) and (2) to (3) is valid.

(v)   If an inference is valid, then anyone who is justified in believing its premises and knows of its validity is justified in inferring its conclusion.

Therefore,

(vi)   Anyone who is justified in believing the premises of this inference is justified in inferring its conclusion.

(vii) I am justified in believing the premises (1) and (2).

Therefore,

(viii) I am justified in inferring (3).[14]

Even if we conceded, then, that we have rational insight into the validity of specific inferences, we do not escape the threat of

14. The idea that Internalism suffers from the difficulties outlined in the preceding two arguments was first presented in 'How are Objective Epistemic Reasons Possible?'. In that paper, though, I followed Carroll in supposing that the inference to which the internalist is committed is an inference from knowledge of the validity of the MPP inference, along with its premises, P and 'If P, then Q,' to the original target conclusion Q. In his commentary on my paper, Crispin Wright suggested that this second anti-internalist argument would be more effective if it claimed not that the internalist is committed to inferences to the target conclusion itself, but rather to the conclusion: the inference to Q is justified. I am not entirely persuaded that Wright's is a better way of running this second argument; but I am persuaded that it raises fewer distracting objections than my original presentation, and so I adopt it for the purposes of this paper.

circularity that afflicts the internalist account. Once again, an ability to infer justifiably according to MPP is presupposed.

Commenting on an earlier presentation of this argument, Crispin Wright observes:

> It is clear how the simple internalist must reply to Boghossian. To staunch her view against all threat of Carrollian regress, she must insist that recognition of the validity of a specific inference whose premises are known provides a warrant to accept a conclusion not by providing additional information from which the truth or warrantedness of the conclusion may be inferred, but in a direct manner ... In effect, and paradoxically, her view must be that warrants acquired by inference are, in a way, a subspecies of non-inferential warrant in general: that an appreciation that a conclusion follows from warranted premises confers, when it does, a warrant for an acceptance of that conclusion in no less direct a fashion than that in which a visual appreciation of the colour of the sky confers warrant for the belief that it is blue.[15]

I agree with Wright that this represents the only escape route available to the simple internalist. Wright himself does not endorse it, or even present it as an especially attractive option. He claims for it only that it remains undefeated by the sorts of considerations adduced so far.

No doubt there remains scope for discussion. But it is very difficult to see, it seems to me, how the inferential case is to be plausibly assimilated to the admittedly non-inferential warrant provided for the belief that the sky is blue by the observation of a blue sky under favourable circumstances.

Admittedly, we don't have very refined ways of deciding when a warrant for a particular belief is direct and when it is fundamentally inferential in nature. What we seem to operate with is a rough-and-ready criterion which says in effect: when the gap between the content of an apparent observation and the content of the belief that it is supposed to justify is too large, the justification must be inferential in nature, even if that may not be apparent from its presentation in everyday conversation.

So, for example, in response to the question 'How do you know it is going to rain?' I may simply point to the dark and

15. 'Basic Logical Knowledge: Reflections on Paul Boghossian's "How Are Objective Epistemic Reasons Possible?"', p. 81.

threatening clouds. But as everyone would agree, the observation of the dark and threatening clouds doesn't justify the belief in rain all by itself, but only by way of an inference in which the content of that observation serves as a premise. The gap between the content of the observation and the content of the belief it is supposed to ground is simply too large.

Similarly, I say, in the case before us. The gap between the content of the apparent observation

If (1) and (2) are true, then (3) must be true

and the belief:

I am justified in believing (3)

is simply to large for the warrant to be direct, even if in most conversational contexts the inference could be left unsaid.[16]

To sum up. In order to ensure that a thinker's inference from particular MPP premises to a particular MPP conclusion not be blameworthy, the simple inferential internalist insisted that the inference's justifiedness be transparent to the thinker—the thinker has to be in a position reflectively to appreciate that his inferring this conclusion from these premises is justified. But this runs into two major problems. First, it requires us to take seriously a notion of rational insight, a notion that no-one has been able to render respectable. Second, and even if we waived this first worry, the aimed-for transparency will still be unattainable, since the only way to attain it will require that the thinker use such knowledge as rational insight is able to afford him as the basis for an inference to the justifiedness of his conclusion. So no matter how concessive we are about rational insight and about the knowledge of logical implication that it is supposed to engender, there seems to be no way to satisfy the transparency insisted upon by Simple Inferential Internalism.

VI

Blind Yet Blameless Inference: *Deflationary Options*. The question is where we go from here. Simple Inferential Externalism is

---

16. What if we imagined that we are able to gain rational insight not only into the validity of specific inferences but directly into their justifiedness? Not only is this unimaginable but I believe that a Carroll-style problem will arise for it as well. I hope to elaborate on this elsewhere.

false. Simple Inferential Internalism, construed as requiring some form of reflectively accessible warrant, is unsatisfiable. We know, furthermore, that we cannot say that deductive inferences do not transfer warrant; that would be not merely implausible but self-undermining. Hence, unless we are to admit that our epistemic system is subject to deep and crippling paradox, there had better be a stable and coherent account of what the conditions for warrant transfer are.

In searching for a solution, we must respect the following facts. On the one hand, the failure of Simple Inferential Internalism teaches us that it must be possible for certain modes of reasoning to be entitling without our knowing, or being able to know, anything about them. I'll put this by saying that it must be possible for certain inferences to be *blind* but justifying.[17]

On the other hand, the counterexamples to Reliabilism teach us that the way *not* to accommodate this phenomenon is through Simple Inferential Externalism. So our question is: How should we construe warrant transfer consistent both with Simple Internalism's and Simple Externalism's falsity?

If there is to be a way forward, the following had better be true: the anti-Reliabilist examples, properly understood, don't really motivate Access Internalism, even though they seem to do so. Rather, they motivate something weaker which *can* be reconciled both with the falsity of Reliabilism and with the falsity of Access Internalism. What could that intermediate position be?

The minimal lesson of the anti-Reliabilist examples, as we saw, is that being justified cannot coexist with being epistemically blameworthy. To get from here to Access Internalism you need to assume furthermore that what *makes* a belief epistemically blameworthy is the *absence* of a reflectively appreciable warrant

---

17. The allusion here is to Wittgenstein's remark at *Philosophical Investigations* 219:

> When I obey a rule, I do not choose.
> I obey the rule *blindly*. (219)

One of Wittgenstein's fundamental insights, it seems to me, was to realize that we must be capable of a form of blameless reasoning that did not depend on any other cognitive state of the thinker's and, in particular, not on anything analogous to sight. As he put the matter in *On Certainty*:

> Giving grounds, however, justifying the evidence, comes to an end—but the end is not certain propositions' striking us immediately as true, i.e., it is not a kind of *seeing* on our part; it is our *acting*, which lies at the bottom of the language game.

for it. And although all the known examples uniformly support this construal, it's not actually forced. For all that the examples show, in other words, it is possible that there should be some *other* way in which a belief might be held blamelessly other than by being supported by some reflectively appreciable warrant.

Some philosophers are inclined to think that there isn't much of a problem here because they think that it isn't all that hard to be epistemically blameless. Gilbert Harman, for example, thinks that just about any belief, or method for forming beliefs, that one cares to have is blameless, at least initially. He writes:

> What I take to be the right theory of justification goes something like this (Goodman, 1995; Quine, 1960a; Quine and Ullian, 1978; Rawls, 1971). In deciding what to believe or what to do, you have to start where you are with your current beliefs and methods of reasoning. These beliefs and methods have a privileged status. You are justified in continuing to accept them in the absence of a serious specific challenge to them, where the challenge will typically involve some sort of conflict in your overall view. Conflict is to be resolved by making conservative modifications in your overall view that makes your view more coherent in certain ways. Your goal in resolving conflict is to reach what Rawls calls a 'reflective equilibrium', in which your various views are not in tension with each other ... The crucial point is that, to a first approximation, continuing to accept what you accept does not require justification. What requires justification is making changes in your view.[18]

On this view, which Harman dubs 'General Conservatism', we have a quick and painless answer to our question. We are now justified in using MPP because MPP is one of the methods with which we 'start' and we have, so far, encountered no incoherence in our overall view to which the best response would have been to reject or modify it.

The principal thought behind general conservatism is an 'innocent until proven guilty' model of epistemic justification. It doesn't matter what beliefs or methods one starts with—they are all *prima facie* justified. What matters is how one changes one's view in response to a developing incoherence.

But this is all very misleading, for the notion of 'coherence', is empty unless it embeds a specific conception of logical consequence and logical consistency. (Actually, it would probably

18. G. Harman: 'The Future of the *A priori*', ms, p. 5.

have to include not just that but conceptions of probabilistic consistency, and a great deal more, but I'll let that pass.)[19] That in effect implies, however, that talk about 'coherence' *presupposes* an answer to our question, rather than providing one. You need to have figured out which deductive rules are justifying in order to have a substantive coherence theory rather than the other way around.

This brings us to the second 'deflationary' answer to our question: a *list*. You want to know which inference patterns are permitted to be blind? These ones: Modus Ponens, Non-Contradiction, and a few others. Don't ask *why* it is precisely these inference patterns that are sanctioned. There is no deep answer to that question; there is just the list.

What makes this brand of deflationary answer unsatisfactory is that it is hard to believe that the property of being warrant-transferring is simply a primitive property that an inference pattern either has or fails to have. Surely, if an inference pattern is warrant-transferring there must be some property by virtue of which it is warrant-transferring. And our question is: What, in the most basic cases, in which reflectively available support is not possible, could that property be?

VII

Blind Yet Blameless Inference: *Concept Constitution*. An important question—which I don't wish to prejudge for present purposes—is whether the validity of the inference is a *necessary condition* on warrant transfer. What I will be exclusively concerned with in the remainder of this paper is the question of blamelessness.

The 'inflationary' answer to that question that I want to explore may be roughly formulated as follows (we'll see how to refine it later):

> A deductive pattern of inference P may be blamelessly employed, without any reflective appreciation of its epistemic status, just in case inferring according to P is a precondition for having one of the concepts ingredient in it.

19. See, for example, L. Bonjour, *The Structure of Empirical Knowledge*, Ch. 5, for a discussion of coherence conceptions of justification.

Now, in some sense this is a very old answer to our question. It falls into what may broadly be called 'analytic' explanations of the *a priori*. Previous versions of such views, however, have suffered, I believe, in two respects.[20] First, they did not adequately distinguish between questions concerning our entitlement to certain logical beliefs, and questions concerning our entitlement to certain belief-forming methods of inference (like the one I have dubbed MPP). Second, they did not adequately explain what concept-constitution has to do with the epistemology of blind inference. In both of these respects, I hope here to do a little better.

Prima facie, there is a difficulty seeing how appeal to concept constitution can help with our question. What is the connection supposed to be?

The thought is this. Suppose it's true that my taking A to be a warrant for believing B is constitutive of my being able to have B-thoughts (or A-thoughts, or both, it doesn't matter) in the first place. Then doesn't it follow that I could not have been epistemically blameworthy in taking A to be a reason for believing B, even in the absence of any reason for taking A to be a reason for believing B? For how could I have had *antecedent* information to the effect that A is a good reason for believing B, if I could not so much have had a B-thought without taking A to be a reason for believing B in the first place? If inferring from A to B is required, if I am to have the ingredient propositions, then it looks as though so inferring cannot be held against me, even if the inference is blind.

Applied to the case of deductive inference before us, then, the thought would be that we would have an explanation for the blameless blindness of MPP if it's constitutive of having the concept *conditional* that one take p and 'p→q' as a reason for believing q.

Now, of course, the idea that, in general, we come to grasp the logical constants by being disposed to engage in some inferences involving them and not in others, is an independently compelling idea. And the thought that, in particular, we grasp the conditional just in case we are disposed to infer according to MPP

20. See, for example, Christopher Peacocke, 'How Are *A Priori* Truths Possible?'.

is an independently compelling thought. So, if the meaning-entitlement connection that I've gestured at is correct, it looks as though we are in a position to mount an explanation of the blameless blindness of MPP that we were after.

VIII

Problems for the Meaning-Entitlement Connection. Unfortunately, matters are not quite so straightforward. If we spell out the principle underlying the meaning-based explanation of blameless blindness gestured at, it would be this:

> (**Meaning-Entitlement Connection, or MEC**): Any inferential transitions built into the possession conditions for a concept are *eo ipso* entitling.

And the trouble is that, at least as stated, there seem to be clear-cut counterexamples to the MEC: it doesn't in general seem true that if my taking A as a reason for believing B is constitutive of my believing B, that this *automatically* absolves me of any charge of epistemic blameworthiness. For there seem to be clear cases where the acceptance of some inference is written into the possession of a given concept but where it is also clear that the inference isn't one to which the thinker is entitled.

One famous illustrative case is Arthur Prior's connective 'tonk'.[21] To possess this concept, Prior stipulated, a thinker must be willing to infer according to the following introduction and elimination rules:

(Tonk) $\dfrac{A}{A \text{ tonk } B}$ $\dfrac{A \text{ tonk } B}{B}$

Obviously, no-one could be entitled to infer any B from any A; but this entitlement appears to flow from the possession conditions for 'tonk' along with the MEC.

A similar conclusion can be drawn from the case of racist or abusive concepts, for example the concept *boche* discussed by Dummett.[22] Plausibly, a thinker possesses the concept *boche* just

---

21. A. Prior, 'The Runabout Inference Ticket', *Analysis* 21 (1960), pp. 38–9.
22. M. Dummett, *Frege's Philosophy of Language* (New York: Harper and Row, 1973), p. 454.

in case he is willing to infer according to the following rules:

(Boche) $\underline{\text{x is German}}$ $\underline{\text{x is boche}}$
             x is boche   x is cruel

Yet no-one is entitled—let alone simply as the result of the intro-duction of a concept into the language—to the view that all Ger-mans are cruel. How should we think about such cases?

Robert Brandom has this to say about *boche*-like concepts:

> The use of any concept or expression involves commitment to an inference from its grounds to its consequences of application. Critical thinkers, or merely fastidious ones, must examine their idioms to be sure that they are prepared to endorse and so defend the appropriateness of the material inferential commitments implicit in the concepts they employ ... The proper question to ask in evaluating the introduction and evolution of a concept is not whether the inference embodied is one that is already endorsed, so that no new content is really involved, but rather whether the infer-ence is one that *ought* to be endorsed. The problem with 'boche' is not that once we explicitly confront the material inferential com-mitment that gives the term its content it turns out to be novel, but that it can then be seen to be indefensible and inappropriate— a commitment we cannot become entitled to.[23]

From the standpoint of a proponent of the MEC, there is nothing in this passage that helps protect it from the threatening examples. It's no answer to the challenge they pose to observe that whatever entitlement concept possession gives rise to, it can be defeated by further considerations. No-one should expect more than a defeasible entitlement, even from concept pos-session; and what's implausible in the case of 'tonk' and 'boche' is that there is any entitlement there at all, defeasible or no.

If we are to save the MEC, it seems to me that we must do one or both of two things: either restrict it to certain concepts from which entitlement really does flow, or restrict what we count as a genuine concept. I will advocate doing the former.

The latter strategy is suggested by the work of Christopher Peacocke who has long urged that we should require that the

23. R. Brandom: *The Articulation of Reasons* (Cambridge, MA: Harvard University Press, 2000), pp. 70–72.

meaning-constituting rules of a genuine concept be truth-preserving.[24]

If we adopt this requirement, we can say that what's wrong with both 'tonk' and 'boche' is precisely that there is no concept that those terms express, for there is no reference for 'tonk' and 'boche' that's capable of making all of their constitutive rules truth-preserving.

While this might seem to yield the right result for 'tonk' it doesn't yield the right result for 'boche': it's hard to believe that racists who employ boche-like concepts fail to express complete thoughts.

And even if we were to put this complaint to one side, it seems clear that truth-preservation alone will not suffice for dealing with our problem about the MEC.

Imagine someone theorizing about water and coming to believe, for whatever reason, that the way in which it is correct to say that water is composed of H_2O is that there is some *other* stuff that composes water and that *it* is composed of H_2O.[25] So he introduces a term—'aqua'—to name this stuff and he stipulates that it is to be governed by the following introduction and elimination rules:

(Aqua) $\dfrac{\text{x is water}}{\text{x is aqua}}$ $\dfrac{\text{x is aqua}}{\text{x is } H_2O}$

Unlike the case of 'boche', we have no independent reason for thinking that these rules are not truth-preserving. But there is clearly something fishy about this concept. And however one feels about that, there is certainly a problem for the MEC, given only the resources that we've been accorded so far: for no-one could think that the mere act of introducing the concept *aqua* into one's repertoire could give one *a priori* entitlement to the inference from x's being water to x's being H_2O.

Or consider the concept *flurg* individuated by the following introduction and elimination rules:

$\dfrac{\text{x is an elliptical equation}}{\text{x is flurg}}$ $\dfrac{\text{x is flurg}}{\text{x can be correlated with a}}$
modular form

24.   See Peacocke 'Proof and Truth', in *Reality*, *Representations and Projection*, eds. J. Haldane and C. Wright (New York: OUP, 1993). I myself took this line in 'How Are Objective Epistemic Reasons Possible?', so the present paper represents a change of heart.

25.   For the purposes of this example, I am assuming that composition is not identity.

It turns out to be a result that Wiles had to prove on the way to proving Fermat's Last Theorem that every elliptical equation can be correlated with a modular form (the Taniyama-Shimura conjecture). Once again, therefore, we have no independent reason to think that these introduction and elimination rules are not necessarily truth-preserving. But it's hard to see that one is *a priori* entitled, merely on the basis of introducing the term 'flurg', to the Taniyama-Shimura conjecture. So there is still a problem for the claim that entitlement flows from meaning-constitution, given only the requirement that a concept's introduction and elimination rules be truth-preserving.

IX

Defective Concepts and Blameless Inference. I would like to propose a different diagnosis of what has gone wrong with concepts such as *aqua* and *flurg*, one that doesn't depend on denying that they constitute genuine thinkable contents.

That denial can be sustained, I believe, in the case of 'tonk', but that is a rather extreme case. No-one can actually possess the concept allegedly expressed by 'tonk', because it isn't possible for someone to follow the rules that are constitutive of that connective. For to follow those rules one would have to be prepared to infer anything from everything, and that is no longer recognizable as belief or inference. But no such extreme claim can be made with respect to all the other examples that have been causing problems for the hypothesized connection between meaning and entitlement.

Start with the example of 'aqua'. The theorist who has conceived the need to introduce a term for the concept *aqua* has come to hold the following *theory*: 'There is some stuff, distinct from water, that composes water and that is itself composed of H_2O. Let me call it "aqua".' Such a theorist already believes in water and H_2O, we may suppose. He has come to hold an additional belief about the world, namely, that it contains another substance, one that is related to water and H_2O in the specified way.

Now, the way we have written down the inferential rules for 'aqua' essentially amounts to insisting that, in order to have the concept *aqua* you must be prepared to *believe* this little *aqua*

theory. Given that you already believe in water and H_2O, the only way for you to acquire the concept *aqua*, on this account of its inferential rules, requires you to *believe* that there is such a thing as aqua. One cannot so much as have the concept of *aqua* without being prepared to believe that there is such a thing.

And although it seems to me that one *can* define and then think in terms of such a concept, it does seem like an epistemically questionable thing to do. Even if the aqua theorist were *certain* that there is such a thing as aqua, he should want the concept he expresses by that term to leave it open whether there is. He should allow for the conceptual possibility that he is mistaken; and he should certainly allow others intelligibly to disagree with him about aqua's existence. The concept itself should not be designed in such a way that, only those who believe a certain creed are allowed to possess it.

Ordinary scientific terms in good standing—'neutrino' for example—are held to have just this feature, of intelligibly allowing for disagreement about their extension. Thus, we don't think of the rules which correspond to our possession of the concept *neutrino* as consisting in the propositions that would actually be believed by a proponent of neutrino theory, but rather as corresponding only to what someone would be willing to believe who was *conditionalizing* on the truth of neutrino theory.

If we follow Russell, Ramsey, Carnap and Lewis, and represent neutrino theory

   T(neutrino)

as the conjunction of the two propositions

   (S) $(\exists x)$ Tx

and

   (M) $(\exists x)$ Tx $\rightarrow$ T(neutrino)

then the point is that we think of possession of the concept *neutrino* as requiring someone to affirm only M and not S as well.[26]

---

26. This paragraph follows Paul Horwich's discussion of the conditional nature of semantic stipulation in 'Stipulation, Meaning and Apriority', in Boghossian and Peacocke: *New Essays on the A Priori*. I ignore various complexities that a thorough discussion of the representation of scientific theories would require.

Now, someone could certainly introduce a concept that did not have the conditionalized structure that I've claimed is actually true of *neutrino*, but which consists rather in the inferences that are characteristic of neutrino theory unconditionalized. Call this *neutrino+*. Such a person would insist that it is a condition on having *his* concept of neutrino that one be willing to endorse the characteristic claims and inferences of neutrino theory, and not merely the conditionalized claim captured in (M). But, for the reasons previously articulated, there would be something epistemically *defective* about this concept, even if its constitutive rules turned out to be truth-preserving.

Flurg, *aqua* and *neutrino+*, then, all suffer from the same problem: they are all unconditional versions of a concept, when only its conditionalized version would be epistemically acceptable. I don't think we should put this by saying that they are not real concepts. Concepts are relatively cheap. But they are *defective* concepts. They are structured in such a way that perfectly reasonable questions about their extensions are foreclosed.

Under what conditions is only a conditionalized version of a concept acceptable? Here I want to make two claims: one bold, one sober.

(Bold)   Whenever both a conditional and an unconditional version of a given concept are available, it is the conditional version that ought to be used. Given the availability of both versions, the unconditional version counts as epistemically defective.

(Sober)   In the case of some concepts, only the unconditionalized version will be available.

Start with Bold. Whenever a conditionalized version of a concept is available, that is the version that ought to be used. In those contexts, an unconditionalized version would be defective.

The argument for this is quite straightforward and recapitulates the considerations we have just been looking at. You don't ever want the *possession conditions* for a concept to foreclose on the possible falsity of some particular set of claims about the world, if you can possibly avoid it. You want the possessor of the concept to be able coherently to ask whether there is anything that falls under it, and you want people to be able to disagree about whether there is.

If in a certain range of cases, however, it is logically impossible to hold the governing theory at arm's length then, in those cases, obviously, it can hardly be a requirement that one do so. But in all those cases where that is possible, it ought to be done.

What about Sober? It should be clear, given the kind of conditionalization that is in view here, that not every meaningful term in a language can be thought of as expressing a concept that conditionalizes on the existence of an appropriate semantic value for it. Take the case of 'flurg'. The stipulation that would correspond to a conditional version for 'flurg' would amount roughly to this:

> If there is a property which is such that, any elliptical equation has it, and if something has it, then it can be correlated with a modular form, then if x has that property, x is flurg.

The corresponding introduction and elimination rules that would specify the possession condition for it would therefore be:

$$\frac{(\exists F)\,[T(F)\ \&\ Fx]}{x \text{ is flurg}}$$

and

$$\frac{x \text{ is flurg}}{(\exists)\,[T(F)\ \&\ Fx]}$$

As this makes clear, the only thinkers who could follow such rules—and, hence, the only thinkers who may be seen as implicitly conditionalizing on the existence of an appropriate semantic value for 'flurg'—are those who (a) possess a basic set of logical constants and (b) are able to refer to and quantify over properties in particular, and semantic values, more generally.

It follows, therefore, that conditional counterparts for one's primitive logical constants will not be available and hence that one could hardly be blamed for employing their unconditionalized versions. In particular, if the conditional is one of your primitive logical constants, you couldn't conditionalize on the existence of an appropriate truth function for it, for you would need it in order to conditionalize on anything. In such a case, there is no alternative but to accept 'conditional theory'—Modus Ponens in effect—if you are so much as to have the conditional concept. It thus couldn't be epistemically irresponsible of you just to go ahead and infer according to Modus Ponens

without conditionalizing on the existence of an appropriate truth function for it—that is simply not a coherent option in this case.

What is the full range of those concepts for which conditional counterparts would not be available? To answer this question, one would need to have a clear view of what the minimal logical resources are that are needed to conditionalize in the envisaged answer on the truth of an arbitrary theory, and I don't have a systematic theory of that to present at the moment. What does seem clear, though, is that some set of basic logical constants would have to be presupposed and that is enough to get me the result that inference in accord with their constitutive rules can be entitling even though blind.

If we go back to the MEC, then, it seems clear what we should say: Any rules that are written into the possession conditions for a *non-defective* concept are *a fortiori* entitling.

With that in hand, we have the answer to our question: how could MPP premises warrant MPP conclusions while being blind? Answer: they do, because they are written into the possession conditions for the conditional, and the conditional is a non-defective concept.

X

Conclusion. If we are to make sense of the justified employment of our basic logical methods of inference, we must make sense of what I have called *blind but blameless* reasoning—a way of moving between thoughts that is justified even in the absence of any reflectively appreciable support for it.

In this paper, I have attempted to sketch the outlines of an account of this phenomenon, one that avoids the pitfalls both of an overly austere Reliabilism and an overly intellectualized Internalism. The account seeks to revive and exploit two traditionally influential thoughts: first, that following certain inferential rules is constitutive of our grasp of the primitive logical constants; and, second, that if certain inferential rules are constitutive of our grasp of certain concepts, then we are *eo ipso* entitled to them, even in the absence of any reflectively appreciable support.[27]