

Draft of paper to appear in a festschrift for Robert Audi, edited by John Greco, Al Mele
and Mark Timmons

On Being Justified in One's Head¹

Timothy Williamson

0. In 'An Internalist Theory of Normative Grounds' (2001), Robert Audi provides what his title promises.¹ His account is characteristically nuanced and ecumenical; it therefore constitutes an excellent basis for an appraisal that is not merely *ad hominem* of one kind of internalism.

With admirable generality, Audi treats the normative grounds for both belief and action. For simplicity, this paper concentrates on his account of the justification of belief. Its arguments, if sound, extend to the justification of action too.

Audi explains what he means by 'normative' in the case of belief:

cognitive (epistemic) normativity is a matter of what ought to be believed, where the force of the "ought" is in part to attribute liability to criticism and negative (disapproving) attitudes toward the person(s) in question.

(21). He takes the concept of justification as a paradigmatically normative notion for belief (20). Given this normative fix on justification, Audi defines his form of internalism about the justification of belief as:

the view that what justifies a belief, i.e., the ground of its justification, is something internal to the subject

(21). He explains the ‘internal’ as:

the (internally) accessible: that to which one has access by introspection or reflection, where introspection can be simply focusing on what is in consciousness and reflection can be as brief as considering a proposition

(ibid.). He emphasizes that having access to something does not require accessing it; the former is a potential, the latter its actualisation.

This paper is structured as follows. Sections 1 and 2 consider Audi’s clarifications of internalism in response to two familiar objections, and argue that his responses are insufficient to prevent those objections from doing serious damage to its plausibility. Sections 3 and 4 appraise his positive case for internalism, and show how externalism can do justice to all his considerations.

1. One challenge to internalism about justification arises from externalism about content. On the latter view, two internal duplicates may differ in what they believe or have other propositional attitudes towards. In the standard example, Oscar on Earth in 1750 believes truly that there are pools of water. Oscar does not believe that there are pools of twater, where twater is a liquid on counterfactual Twin Earth with the same superficial characteristics as water but an utterly different underlying nature, for in Oscar’s world

there are no pools of twater, nor are there grounds for ascribing a false belief on that matter to him. He has no specific beliefs as to the underlying nature of water, except that it has one. By contrast, on counterfactual Twin Earth, Oscar's duplicate Twin Oscar believes truly that there are pools of twater. Twin Oscar does not believe that there are pools of water, even though he is in exactly the same internal states as Oscar, for in Twin Oscar's world there are no pools of water, nor are there grounds for ascribing a false belief (in his world) on that matter to him. He has no specific beliefs as to the underlying nature of twater, except that it has one.² Of course, some philosophers will dispute this description of the case, but Audi does not. His strategy is to argue that such externalism about content is compatible with his internalism about justification. It is thus legitimate to assume the content-externalist description of the case in discussing Audi's claim that it is compatible with his internalism.

It is not in dispute that we can pick an example in which Oscar's belief that there are pools of water is justified. Perhaps he is swimming in one. Thus Oscar has the justified belief that there are pools of water. But Twin Oscar lacks a justified belief that there are pools of water, because he lacks the belief that there are pools of water. Thus Oscar and Twin Oscar differ in their justified beliefs, even though they are internal duplicates. Likewise, of course, Twin Oscar has the justified belief that there are pools of twater, while Oscar lacks a justified belief that there are pools of twater, because he lacks the belief that there are pools of twater: that is just another difference in justified belief between Oscar and Twin Oscar.

Given Audi's initial characterization of the internal in terms of a special sort of accessibility, it is not immediately obvious why such examples should constitute a threat

to his internalism. What rules out differences between internal duplicates in what is accessible in the special way? Externalism about content is arguably compatible with one's privileged, non-observational access to one's mental states with such content.³ Why can't Oscar introspect that he has reason to believe that there are pools of water, while Twin Oscar introspects that he has reason to believe that there are pools of twater? However, Audi construes his internalism as ruling out differences between internal duplicates in justification in some sense, and we shall see that doing so is important to his motivation for his view. Let us suppose for the sake of argument that Audi can stipulate a sense for 'internal access' in which differences between internal duplicates are not internally accessible.

In response to concerns about content externalism, Audi first considers the postulation of 'narrow contents' of belief in common between Oscar and Twin Oscar (32). Such contents remain the pious hope of some internalists, but there is little evidence that we believe any such things. Moreover, it is not clear how they would meet the original challenge. For even if Oscar and Twin Oscar do have some justified beliefs in common, the externalist argument was that since they do not have *all* their justified beliefs in common, justified belief does not supervene on internal states. The postulation of additional contents would not undermine the point that Oscar has, while Twin Oscar lacks, the justified belief that there are pools of water. Merely restricting the internalist claim to the justification of belief in narrow contents threatens to be *ad hoc*, for no such restriction is built into Audi's positive case for internalism about justification.

Audi prefers a different response, one that does not offer a hostage to fortune by postulating narrow contents:

Whether the internalist can find some proposition that both duplicates actually believe is not crucial; the point is that if the two are internally identical, then what they are justified *in believing* is the same.

(32). One may not believe everything that one is (or would be) justified in believing. In terminology that Audi also uses, his internalism is a doctrine about what one has justification *for* believing, not about what justified beliefs one has (43, n. 2). That seems to fit his characterization of his theory as one about normative *grounds*.

But can that distinction bear the burden that Audi imposes on it? Since Oscar and Twin Oscar are internally identical in Audi's sense, he is committed to saying that what they are justified in believing is the same. Given his disclaimer about narrow contents, the phrase 'what they are justified in believing' should be read as applying to ordinary belief contents, not just to postulated narrow contents. Therefore, since (by hypothesis) Oscar is justified in believing that there are pools of water, Twin Oscar is also justified in believing that there are pools of water (although he does not in fact believe that there are pools of water). Since there are no pools of water in Twin Oscar's world, there is a false proposition that Twin Oscar is justified in believing. Similarly, since Twin Oscar is justified in believing that there are pools of twater, Oscar is also justified in believing that there are pools of twater (although he does not in fact believe that there are pools of twater). Since there are no pools of water in Oscar's world, there is a false proposition that Oscar is justified in believing.

Twin Earth scenarios can be multiplied indefinitely. Thus Audi's proposal seems to imply that for very many of our justified beliefs there are masses of closely related false propositions that we are justified in believing, although we do not in fact believe

them. The problem here for Audi is not the general idea that one can be justified in believing a false proposition, for like many epistemologists he will take that as obvious. Rather, the problem is that justification for believing something false does not come anything like as cheap as Audi's internalism makes it come. If one is justified in believing a falsehood, one's epistemic situation is in some way misleading. Yet nothing that we assumed about Oscar's epistemic situation seemed to imply that it was in any way misleading at all. He just went swimming in some rock pools. Although someone from a twatery environment who was suddenly switched with Oscar without realising it might thereby be placed in a misleading epistemic situation, that does not imply that *Oscar's* epistemic situation is misleading; whether one is in a misleading epistemic situation depends in part on one's past history. Of course, Oscar's total evidence is in many respects incomplete. For example, he has no evidence that water is H₂O (he does not even have the concept *H₂O*). But incomplete evidence need not be misleading, for it can justify agnosticism. He has no evidence that water is not H₂O. In any case, Oscar has quite enough ordinary evidence to justify him in believing that there are pools of water. It is deeply implausible to suggest that the very same evidence also justifies him in believing any number of false propositions of a similar form.

Could Audi retreat to the weaker internalist thesis that although internal duplicates may differ in what *gets justified*, they do not differ in what *does the justifying*? On such a view, Oscar's internal state would justify him in believing that there are pools of water but not in believing that there are pools of twater; the very same internal state would justify Twin Oscar in believing that there are pools of twater but not in believing that there are pools of water. One trouble with the envisaged fallback position for Audi's

purposes is that it does not fit a major part of his original motivation for his theory, namely, that it

explains why, if our beliefs are suitably based on internal grounds, then even if they are false, say because of bad luck or even because a Cartesian demon has caused us to hallucinate in such a way that our beliefs are “imperceptibly” false, we would remain justified in holding them.

(22-23). Given that you and your twin in the demon world are internal duplicates, the weakened sort of internalism just envisaged implies that you and your twin have the same justifiers, but it does *not* imply that those justifiers justify the same propositions; thus it does not imply that your twin is justified in holding the very beliefs that you hold. So it does not explain the supposed constancy of justification in and out of sceptical scenarios: it lacks the explanatory capacity that Audi claims as a crucial advantage of his internalism. If you and your twin in the demon world are held to be justified in believing the same propositions because you are internal duplicates, then Audi’s explanandum in effect just is his internalist principle that ‘if the two are internally identical, then what they are justified *in believing* is the same’, whose implausible consequences were marked above.

Could Audi restrict his claims to propositions that both subjects can grasp? For it might be argued that Oscar cannot grasp the proposition that there are pools of twater, and Twin Oscar cannot grasp the proposition that there are pools of water. But that feature is not essential to the problem. We can add to the example that a traveller once showed Oscar a tiny phial of twater and told him (truly) that the liquid in it was called ‘twater’, had similar superficial characteristics to water but an underlying nature that was

utterly different (in unspecified ways), and occurred only in minute quantities, not pools (in this world); Oscar believes what the traveller told him and has no beliefs from any other source about twater. Similarly, a traveller once showed Twin Oscar a tiny phial of water and told him (truly) that the liquid in it was called (confusingly for us) 'twater', had similar superficial characteristics to twater but an underlying nature that was utterly different (in unspecified ways), and occurred only in minute quantities, not pools (in that world); Twin Oscar believes what the traveller told him and has no beliefs from any other source about water. Thus Oscar can think of twater and Twin Oscar can think of water. Although Oscar associates twater with different descriptions from those which he associates with water, and likewise for Twin Oscar, externalists about content will typically hold that in such circumstances Oscar can grasp the proposition (false in this world) that there are pools of twater while Twin Oscar can grasp the proposition (false in that world) that there are pools of water. After all, we share the belief that there are pools of water with people whose beliefs about water are quite different from ours. Nevertheless, Oscar is justified in believing that there are pools of water but not in believing that there are pools of twater, while Twin Oscar is justified in believing that there are pools of twater but not in believing that there are pools of water.⁴ Thus the difference in justification between internal duplicates can arise even for propositions that both grasp. Indeed, we can modify the example so that Oscar, in addition to his justified true belief that there are pools of water, has an irrational unjustified false belief that there are pools of twater, while Twin Oscar, in addition to his justified true belief that there are pools of twater, has an irrational unjustified false belief that there are pools of water.

Thus the difference in justification between internal duplicates can arise even for propositions that both believe.⁵

We can imagine various further manoeuvres on Audi's behalf in response to such cases. But they will typically be manoeuvres of very much the sort to which internalists about content are forced to resort. They will involve the postulation of something like a taxonomically basic level at which content is narrow, so that the narrow content underlying Oscar's justified belief that there are pools of water is the same as the narrow content underlying Twin Oscar's justified belief that there are pools of twater and different from the narrow content underlying Twin Oscar's unjustified belief that there are pools of water. The obstacles to that content-internalist programme are forbidding.

Audi could still claim that Oscar and Twin Oscar are equally justified in their beliefs, although with respect to different beliefs: our overall normative assessment of them as believers is the same. That would be a long step back from his assertion that '*what* they are justified in believing is the same' (32, italics moved). Moreover, we shall see in section 4 that it is very doubtful that our overall normative assessment of internal duplicates as believers always is the same.

Contrary to Audi's claims, his internalism about justification is in serious conflict with externalism about content. Given the strength of the arguments for externalism about content, that is bad news for his internalism about justification.

2. A different challenge to internalism about justification, forcefully pressed by Alvin Goldman (1999), arises from bare factual memory. Many of our factual memories come without any particular supporting phenomenology of memory images or feelings of

familiarity. We cannot remember how we acquired the information, and it may be relatively isolated, but we still use it when the need arises. Although few if any memories stand in total isolation from the rest of our conscious lives, very many memories are too isolated to receive impressive justification from other internal elements. One may believe that one's memory has been fairly reliable in the past; of course, its actual past reliability is irrelevant to justification on an internalist theory such as Audi's. In such cases, there is a shortage of internal justifiers for the memory-based belief.

Audi's response is to suggest that in such cases we have knowledge without justification (30-31), for he combines internalism about justification with externalism about knowledge (26).⁶ For example, I know that Trincomalee is in Sri Lanka, but I am not justified in believing that Trincomalee is in Sri Lanka. Is that treatment adequate?

Audi's talk of knowledge without justification may sound acceptable if one thinks of it as knowledge that the subject cannot back up with a discursive defence; it is plausible that we have such knowledge. But Audi's primary characterization of justification is normative rather than dialectical. Thus in postulating knowledge without justification, he is saying something negative about the normative status of the knowledge. In trying to elucidate the normativity of justification, Audi suggests:

justified beliefs are those one is in some sense in the right in holding:
holding them is normatively appropriate given one's sensory impressions,
rational intuitions and other internal materials. In the language of desert,
one does not deserve criticism (from the point of view of the effort to
reach truth and avoid falsehood) for holding them.

(29). Therefore, since my belief that Trincomalee is in Sri Lanka is not justified, it is not one of those that I am ‘in the right in holding’; that is, I am not in the right in holding it. I do ‘deserve criticism (from the point of view of the effort to reach truth and avoid falsehood)’ for believing that Trincomalee is in Sri Lanka, even though I know that Trincomalee is in Sri Lanka. Such claims seem quite implausible. One ought to give up a belief once one realizes that one is not in the right in holding it. But it would be silly for me, having read Audi, to *give up* my belief that Trincomalee is in Sri Lanka. That would be to give up some of my knowledge. If believing that Trincomalee is in Sri Lanka is not ‘normatively appropriate given my sensory impressions, rational intuitions and other internal materials’, that merely goes to show that the internal (as Audi understands it) is the wrong place to look for the justification of my belief.

With an eye to the case of young children’s beliefs, Audi insists that ‘not justified’ does not imply ‘unjustified’ (31). But the word ‘unjustified’ has not been used here. Rather, he has sketched a necessary and sufficient condition for justification in normative terms: where he denies justification, he should therefore be taken to hold that his normative condition does not obtain. The trouble is just the implausibility of the claim that the normative condition does not obtain in cases of bare factual memory knowledge.

Audi’s proposal also faces a structural difficulty. Suppose that S knows p without being justified in believing p . It may easily happen that at the same time S has a justified belief in some conditional $p \rightarrow q$ without knowing $p \rightarrow q$, for Audi denies that justified belief entails knowledge.⁷ Now S applies modus ponens to $p \rightarrow q$ and p , competently deduces q , and comes to believe q on that basis alone. What is the status of her belief in q ? It seems to be good in some way, since it is held on the basis of competent deduction

from premises each of which was either known or believed with justification. However, S does not know q , for S does not know one of the premises on which her belief in q is essentially based. Indeed, q may even be false, for S may have a justified false belief in the conditional. Equally, her belief in q is not justified, for S lacks justification for one of the premises on which her belief in q is essentially based. After all, given Audi's internalism about justification, it is irrelevant to whether S is justified in believing q that she has knowledge where she lacks justification. Thus her belief in q constitutes neither knowledge nor justified belief. At the very least, Audi seems to be missing an epistemological category with which to classify such cases. Since he structures his epistemological discussion around the two contrasted categories of internalist justification and externalist knowledge, something is amiss with the structure of his epistemology. By separating the internal and external perspectives in this way, he makes it hard to see what is good about S's belief in q . If one takes the internal perspective, one cannot see what is good about S's belief in q because S was not justified in believing p . If one takes the external perspective, one cannot see what is good about S's belief in q because S did not know $p \rightarrow q$.

The idea of knowledge without justification should strike us as anomalous, at least when 'justification' is understood in the normative way that Audi sketches. Knowing p is the central, unproblematic case of normative appropriateness in believing p . Although one may know p without being able to make explicit an elaborate discourse of justification for p , it would be naïve to regard such an ability as necessary for normatively appropriately believing p , and Audi does not so regard it (27). However tempting it may be for some epistemologists to think that those who lack the capacity for epistemological

reflection are not entitled to believe anything, the thought is a mere professional deformation. Justification becomes problematic only for beliefs that do not constitute knowledge.

Audi's misalignment of the concepts of knowledge and justification results from his combination of externalism about knowledge with internalism about justification. By concentrating the normative aspect of the appraisal of beliefs into his conception of justification and the truth-tracking aspect into his conception of knowledge, he arrives at an overall view on which justification and knowledge look more independent of each other than they really are. But the two aspects are not orthogonal dimensions; 'from the point of view of the effort to reach truth and avoid falsehood', knowledge is not normatively neutral.

3. We have seen that Audi's attempts to reconcile his internalism about justification with content externalism and the phenomenon of bare memories have committed him to some exceedingly implausible claims. Only very strong arguments could justify persisting with such a view. It is time to examine Audi's positive case for internalism.

According to Audi,

Internalism is motivated by at least two ideas. One is that what justifies a belief is somehow available to the subject — through consciousness or reflection — to use in justifying it

(22). The other motivating idea is that our justified beliefs remain justified in sceptical scenarios (23). Let us examine the relation between justification and availability to the

subject first. In the subsequent section, we will assess how much changes when we add the second motivating idea.

Properly understood, the claim that only what is available to a subject justifies her belief is an attractive one, for we may interpret the question of justification precisely as the question whether the belief is a good one for the subject to hold, given what is available to her. ‘Availability’ here means some sort of epistemic accessibility. In a general sense, whatever the subject is in a position to be aware of is epistemically accessible to her. On that reading, the claim that only what is available to a subject justifies her belief follows from the claim that only what the subject is in a position to be aware of justifies her belief; I will not challenge the latter claim.⁸

Audi does not intend his claim in that general sense. Rather, he has in mind a special sort of epistemic accessibility, which he qualifies parenthetically as ‘internal’. He explains it thus:

To have (internal) access to something is either to have it in consciousness or to be able, through self-consciousness or at least by reflection, whether introspective or directed “outward” toward an abstract subject matter, to become aware of it in the (phenomenal) sense that it is in one’s consciousness.

(21-22). I have epistemic access in the general sense to the presence of a computer screen before me, because I know by sight that this computer screen is before me. Have I internal access in Audi’s sense to the presence of the computer screen before me? One might argue that I have, because I am conscious of the presence of this computer screen before me, and therefore have it ‘in consciousness’ in some sense. However, Audi’s

whole discussion makes it clear that he does not intend his words to be applied so widely: if they are, nothing distinctive remains of his internalism. Instead, he takes for granted that the contents of consciousness are confined to more traditional items, such as mental images. On his view, an experience as of a computer screen before me may be ‘in consciousness’, but not the computer screen itself. To vary the example, I have epistemic access in the general sense to Trincomalee’s being in Sri Lanka, because I know by memory that Trincomalee is in Sri Lanka. On Audi’s view, I lack internal access to Trincomalee’s being in Sri Lanka, for otherwise such cases would not exemplify knowledge without justification. Although I consciously remember that Trincomalee is in Sri Lanka, what is in my consciousness in his sense is not that Trincomalee is in Sri Lanka, but only that it seems to me that Trincomalee is in Sri Lanka (or that some related narrow content seems to me to be the case).

It is far from obvious that even introspection is restricted in the way that Audi requires. Pretheoretically, it is natural enough to say that I can introspect that I am occurrently aware that this computer screen is before me, or that I occurrently remember that Trincomalee is in Sri Lanka. Since ‘aware’ and ‘remember’ are factive, the deliverances of introspection would not then be neutral in respect of the external environment. Audi does not consider this possibility.

Let us suppose for the sake of argument that Audi has somehow defined the phrase ‘internal access’ so that it applies to just the sorts of case that he envisages, even if his actual definition can be read otherwise. The phrase ‘epistemic access’ will henceforth be used for epistemic access in the more general sense. The question immediately arises: why accept Audi’s restriction of what justifies the subject’s belief to what she has internal

access to, when that excludes all sorts of other truths to which she has epistemic access? What is so special about internal access with respect to justification?⁹ Unfortunately, Audi does little to address this question. Yet without a proper answer, his first motivating idea for internalism looks arbitrary and question-begging. On the face of it, I *can* use truths to which I have epistemic but not internal access to justify my beliefs. For example, I can justify my belief that my computer has not been stolen by inference from premises one of which is that this computer screen is before me; I can justify my belief that Trincomalee is in Asia by inference from premises one of which is that it is in Sri Lanka.

Audi might reply that what *ultimately* justifies my belief that my computer has not been stolen does not include the premise that this computer screen is before me, and that what *ultimately* justifies my belief that Trincomalee is in Asia does not include the premise that it is in Sri Lanka. To restrict his internalism to ultimate justification would already be a significant concession, since his paper advertises a theory of justification *tout court*, not merely a theory of ultimate justification. Moreover, we have as yet been given no reason to think that even ultimate justifiers must be internally accessible.

Another reply on behalf of Audi in the same spirit starts from the other end: the subject has epistemic access *primarily* to internal elements of actual or potential consciousness and only secondarily to premises such as that this computer screen is before me or that Trincomalee is in Sri Lanka. He would then have to explain why we need primary access to justifiers; why is secondary access insufficient? Moreover, we have as yet been given no reason to think that we do have primary access only to internal elements of actual or potential consciousness. To invoke such a conception of ultimate justification or primary

access as rooted in elements internal to actual or potential consciousness would involve Audi in explaining and defending deeply contentious epistemological premises of a kind that would obviously beg the question in an argument for internalism about justification. Since he makes no attempt to address the difficulties, it may be uncharitable to attribute such replies to him.

Internalists might try to support the talk of ultimate justification by arguing that if serious doubts arise as to whether I really am aware that a computer screen is before me or really do remember that Trincomalee is in Sri Lanka, I am likely to fall back on defensive remarks such as ‘At least, it seems to me that I am aware that a computer screen is before me’ or ‘At least, it seems to me that I remember that Trincomalee is in Sri Lanka’. But it is a fallacy to assume that retreats in the face of doubt always reveal a pre-existing structure of justification. Someone may be simultaneously disposed to retreat to premises about appearances if put under pressure by idealists about the external world and to retreat to premises about brain scans if put under pressure by eliminativists about the mind. In responding to a doubt, we look for ground that it does not undermine, but where that ground is depends on the doubt. That we can be made to retreat to a place does not show that it is where we started from.

We can sharpen our sense of the dialectical position by briefly considering an externalist rival to Audi’s internalism about justification. On the view that I defend, what justifies beliefs is the subject’s total evidence, which is the total content of the subject’s knowledge (Williamson 2000). A belief that constitutes knowledge is *ipso facto* justified. A belief that does not constitute knowledge is justified, to whatever degree it is justified, by its relation to beliefs that do constitute knowledge. Thus chains of justification can

terminate in knowledge of any kind; they do not all lead back to *recherché* internal elements of actual or potential consciousness. Of course, knowing something depends on various preconditions: I know that this computer screen is before me, or that Trincomalee is in Sri Lanka, only if all sorts of other things are true of me too. But those preconditions for knowledge need not themselves be justifiers in the sense in which subjects must have access to their justifiers. In the most straightforward way, this externalist account satisfies the requirement that subjects have epistemic access to their justifiers, for their justifiers are truths that they know. Obviously, the account does not satisfy Audi's requirement that subjects have *internal* access to their justifiers. But that simply makes more pressing the question: why is epistemic access insufficient?

Someone might doubt that such an externalist theory of justification really does satisfy even the epistemic access requirement. For in a sceptical scenario we lack knowledge without being in a position to know that we lack it; less obviously, we sometimes have knowledge without being in a position to know that we have it. Consequently, given that one's total evidence is one's total knowledge, something can be excluded from one's evidence even though one is in no position to know that it is excluded from one's evidence, and something can be included in one's evidence even though one is in no position to know that it is included in one's evidence. But what these points show is that something can be a non-justifier although one is in no position to know that it is a non-justifier, and something can be a justifier although one is in no position to know that it is a justifier. They do not show that something can be a justifier although one is in no position to know *it*. If one is in no position to know it then one does not know it and so, on this externalist account, it is not one of one's justifiers. One can

have epistemic access to a justifier without having epistemic access to the further truth that it is a justifier.

Of course, someone might want to impose the additional requirement that one must be in a position to know of any given item whether it is one of one's justifiers. But Audi does not impose that further requirement, and wisely so. For it implies that the condition that p is a justifier and the condition that p is not a justifier are *luminous*, in the sense that whenever the condition obtains one is in a position to know that it obtains. Elsewhere, I have argued that only trivial conditions are luminous (2000: 93-113). The condition that p is a justifier and the condition that p is not a justifier carry substantive information about one's current epistemic state on any serious view of justification; for example, on Audi's view, they may concern whether one is currently experiencing a mental image of a certain kind. Therefore those conditions are not trivial in the relevant sense. Therefore they are not luminous. Neither internalism nor externalism can meet the further requirement. Audi provides no hint of an objection to the anti-luminosity argument. Thus the further requirement should not be imposed on his behalf.

The question remains: in motivating his internalism, what entitles Audi to invoke his requirement of internal access on justification, rather than the requirement of epistemic access that the externalist theory satisfies? At one point, he suggests an explanation for the supposed internal grounding of normativity across diverse domains: our practices of justification

must have criteria usable in the everyday contexts in which we engage in these practices. One must, given how we are built, apply them from the inside.

(40). But the requirement of epistemic access already secures usability in everyday contexts. If we have knowledge then we can use it in justification. Of course, we will sometimes be mistaken as to whether something is part of our evidence, but no everyday practice requires infallibility about anything, and even Audi refuses to endorse the claim that our self-ascriptions of mental states are infallible (44, n. 11). In fact, it is internalism, not externalism, that risks losing touch with usability in the everyday contexts in which we engage in practices of justification. For justification is typically a *social* practice: we try to justify our beliefs to others in response to their challenges. But if what I use in justifying my belief to you is internal to me, how much use will it be to you? When George W. Bush and Tony Blair were expected to justify their belief that there were weapons of mass destruction in Iraq, what was needed was publicly verifiable evidence, not mental images or feelings of conviction. Of course, internalists will concoct a story about public justification as consisting in the causation by the justifier of suitable internal elements in the consciousness of the hearer, but no such story is motivated by the ordinary criteria for public justification. An externalist account of justification as rooted in the content of knowledge in general is far better adapted to our ordinary practice of justification.¹⁰

My present concern is not to develop the arguments for a particular externalist theory of justification in detail, since I have done so elsewhere. The immediate point is simply that at least one externalist theory of justification meets at least as strong an access constraint as may legitimately be assumed without serious argument. In imposing his requirement of internal access, Audi starts from an assumption that lacks the

pretheoretic plausibility needed if his case for internalism is to do more than preach to the converted. We must therefore turn to Audi's second motivating idea for internalism.

4. According to Audi, we retain our justification in cases of perfect hallucination. He supports the point by appeal to normative features of our situation:

I am in no way at fault for believing what I do, nor do I deserve any criticism (at least on the non-skeptical assumption that we may generally trust our senses in this way). Far from it. I am like a surgeon who skillfully does all that can be expected but loses the patient. There I should feel regret, but not guilt; I should explain, but need not apologize; and when we know what my evidence was, we approve of what I did. We consider it reasonable.

(23). This passage mixes various normative claims, some more plausible than others. It does scant justice to the complexity of our normative thinking.

We can readily agree with Audi that the victim of a paradigmatic sceptical scenario is not to be *blamed* for forming false beliefs under the misapprehension that they constitute knowledge. The subject has a cast-iron excuse for having formed those beliefs. But the subject in the corresponding non-sceptical scenario whose beliefs constitute knowledge needs no excuse for having formed those beliefs; there is nothing to excuse. That the two subjects are equally blameless does not imply that there is no normative difference between them. The subject who blamelessly forms false beliefs has not done as *well* as the subject who gains knowledge.

Suppose that the patient in Audi's simile would not have died if the surgeon had not operated, just as error would have been avoided in the sceptical scenario by suspension of belief. Of course, the surgeon followed the best available guidelines in deciding to intervene, and operated with state-of-the-art skill and technology. Nevertheless, it turns out that it would have been better for the patient if the operation had not been performed. To say, in Audi's words, that we *approve* of what the surgeon did hardly captures the complexity of our normative response to the case. We absolve the surgeon of all blame, while also thinking that in some sense the decision to operate turned out to be wrong. We praise the surgeon less than we would have done had the operation saved the patient's life rather than causing her death.

Audi concedes that the sceptical and non-sceptical scenarios can differ normatively. For the content of an obligation can be external:

Suppose I promise to care for a child. Is it enough that I do all I can, so that if a Cartesian demon makes me hallucinate success, while the child is in fact left alone, I have fulfilled my obligation? To say so would be to parlay excusability for not fulfilling an obligation into its actual fulfilment.

(37). Excusable failure is not normatively equivalent to success. Nevertheless, Audi maintains:

the *content* of my obligation can be external while its grounds are internal, just as the content of my belief about the water in my glass can be external though the grounds of that belief are internal.

(37-38). Given the results of section 1, this comparison with externalism about content will not help Audi. As for the claim that the grounds of the obligation are internal, it is at

best misleadingly expressed: how can the grounds of my obligation to care for the child be independent of whether the child exists and of whether I made a promise *to someone* to protect the child? Moreover, if obligations can differ between the sceptical and non-sceptical scenarios, why shouldn't epistemic obligations do so in particular? For example, if I have an obligation to believe only what I know, then I fulfil that obligation in the non-sceptical scenario but not in the sceptical scenario, and my beliefs differ normatively between the two scenarios as a result.

Presumably, Audi will restrict 'justification' to a special sort of normativity with respect to which the sceptical and non-sceptical scenarios are equivalent, although that is already to qualify the apparent claims of his paper. Furthermore, if he is not merely to have made an uninteresting stipulation about the use of the word 'justification', some argument is needed that there really is an important normative respect in which any two such scenarios are equivalent. The normative distinctions that we make between them show that the existence of such a respect cannot simply be taken for granted.

One might regard the normative differences between the two scenarios as something like moral luck (Nagel 1976; Williams 1976). Then the challenge to the internalist would be to demarcate a luck-free zone as the realm of justification. Suggestive though this way of putting the matter may be, the notion of epistemic luck is too contested to figure in a clear statement of the issue. Consider knowledge reached by a thoroughly reliable process outside the subject's ken. Externalists will tend to say that the truth of the subject's belief is not a matter of luck, because the process is so reliable. Internalists will tend to say that the truth of the belief is a matter of luck, because in some sceptical scenario the belief is false and merely appears to have been reached by that

process. Let us stick to the question whether Audi can isolate an important normative respect in which a sceptical scenario is always equivalent to the corresponding non-sceptical scenario.

It may seem that the internalist can easily meet the challenge by appealing to Audi's first motivating idea for internalism: access. The proposal would be that the two scenarios are alike *in all normative features accessible to the subject*: more precisely, the subject in the non-sceptical scenario and the subject in the sceptical scenario have access to exactly the same normative truths. For agents have to act on the basis of the limited information available. To assess the proposal, let us ask a prior question: do the subjects in the two scenarios have access to exactly the same truths quite generally?

If 'access' means epistemic access, then one should resist any temptation to answer in the affirmative, for that way leads to scepticism. If we are not sceptics, we will insist that the subject in a suitable non-sceptical scenario knows that the child is playing.¹¹ The subject in the corresponding sceptical scenario does not know that the child is playing, because the child is not playing. Thus the subject in the non-sceptical scenario has access to something to which the subject in the sceptical scenario lacks access. But once it is granted that the two scenarios differ in some features accessible to the subject, what can stop them from differing in some *normative* features accessible to the subject? For example, the subject in the non-sceptical scenario may know that he has fulfilled his obligations. The subject in the sceptical scenario does not know that he has fulfilled his obligations, because he has not fulfilled them.

If one's total evidence is one's knowledge, then the two scenarios differ in the subject's evidence. That the child is playing is part of the subject's evidence in the non-

sceptical scenario but not in the sceptical scenario. Of course, the subjects in the two scenarios share *some* evidence. For example, that it perceptually appears that the child is playing may be part of the subject's evidence in both scenarios. On this view, the difference is that in the non-sceptical scenario the subject's total evidence *entails* that the child is playing, whereas in the sceptical scenario the subject's total evidence does not entail that the child is playing, although it does (misleadingly) make it probable that the child is playing. Consequently, the belief that the child is playing is *more justified* in the non-sceptical scenario than it is in the sceptical scenario, even though in both scenarios it has considerable justification.¹²

Of course, Audi will not accept that theory of evidence and justification. But he cannot simply assume its falsity if he is to give a non-circular motivation for internalism about justification. It is not simply obvious that there is no subtle normative difference of the kind that it postulates in the status of beliefs between the two scenarios.

Audi will say that the two scenarios are exactly alike with respect to both the normative and the non-normative features *internally* accessible to the subject, and perhaps he can explain a sense of 'internal' on which that is so. But that reply merely raises again the question to which section 3 found no answer: what is so important about internal access in Audi's special sense as opposed to any other kind of epistemic access? If agents have non-internal access to some information to which they lack internal access, what is the point of assessing their beliefs and actions as if they were based only on the information to which the agents have internal access? Why not assess their beliefs and actions in relation to all the information to which they have access?

A different strategy for Audi is to try to specify which dimension of normativity he intends to be internalist about in directly normative terms. The most promising candidate is the dimension of blame, since it was granted that the subject in a paradigmatic sceptical scenario is blameless, like the subject in the ordinary case. Audi's own analogy of the surgeon suggests that there are differences between the two scenarios in praiseworthiness. Thus 'justification' in Audi's sense would reduce to blamelessness. However, it is doubtful that this idea generalizes in the way that Audi requires, for several reasons.

First, if Audi concedes externalism about content, as he seems willing to do, then the blameworthiness of belief with a given content can vary across internal duplicates. For instance, section 1 sketched a story on which Oscar blamelessly believes that there are pools of water but blameworthily believes that there are pools of twater, while Twin Oscar blameworthily believes that there are pools of water but blamelessly believes that there are pools of twater.

Second, if justification is defined in terms of blame, then presumably to be justified is simply to be blameless. Thus Audi's classification of bare memory knowledge as unjustified belief implies a category of blameworthy beliefs that constitute knowledge, which exacerbates the puzzles of section 2.

Third, paradigmatic sceptical scenarios correspond to non-sceptical scenarios in which the subject is blameless for her beliefs. The internalist must also hold that mildly blameworthy true beliefs in a non-sceptical scenario would have been no more blameworthy if false in the corresponding sceptical scenario. In Audi's simile, that is to say that the slightly clumsy surgeon who nevertheless performs the operation

successfully would have been no more blameworthy if, with the same slight clumsiness, he had killed the patient, which is hardly a plausible claim. The surgeon who causes his patient's death through negligence has more to be blamed *for* than the equally negligent surgeon who saves his patient's life.

Similar problems arise if we replace the concept of blame with other familiar normative concepts. Thus Audi is driven back to characterizing the relevant normative dimension in epistemic terms. We have already seen how hard it is to motivate his approach from an epistemological starting-point. To conclude, we will consider a quite general problem for that approach, no matter what normative dimension Audi associates with justification.

5. On Audi's conception, the sceptical and non-sceptical scenarios are exactly the same in the relevant normative respect (whatever that is) because the difference between them is not internally available to the subject for normative use. Let us call situations 'indiscriminable' if and only if any difference between them is not accessible in the relevant way to the subject, and 'normatively identical' if and only if they are the same in the relevant normative respect. Then Audi's approach commits him to this principle:

Normative Tolerance Indiscriminable situations are normatively identical.

Is Normative Tolerance true? Imagine a sorites series of possible situations, each indiscriminable from its immediate neighbours, where the first and last members are very different, in particular in the relevant normative respect; the details will depend on the

selected sorts of access and normativity. For example, we can define such a series by taking successive situations with a very short constant time interval over a long process of very gradual change. Successive members of the series are indiscriminable because they are too similar for the subject in them to be aware by introspection or reflection that there has been any change at all, just as the subject in the sceptical scenario is not in a position to be aware by introspection or reflection that the situation is different from the non-sceptical scenario. Therefore, according to Normative Tolerance, each situation is normatively identical with its neighbours. But normative identity, like exact sameness in any other given respect and unlike indiscriminability, is transitive. Therefore the first and last members of the series are normatively identical, contrary to hypothesis.¹³

Since Normative Tolerance is false, there can be normative differences between indiscriminable situations. Once that is conceded, it is quite illegitimate to argue for internalism from the premise that there is no normative difference between the sceptical and non-sceptical scenarios. Thus Audi's argument for internalism fails. Given the implausible consequences of his internalism noted in sections 1 and 2, it should therefore be rejected.

In rejecting Audi's internalism about justification, we need not reject his insight that what does the justifying must be accessible to the subject, provided that we do not restrict the accessible to what is internally accessible in his special sense. We can put the natural externalist alternative by saying that p is available to you as a reason if and only if you know p .¹⁴ That alternative applies to reasons for action, not just to reasons for belief. In a given case, the right reason for helping someone may be that they need help, not that you are in a certain internal state (the latter anyway sounds like an unattractively self-

regarding reason). You can act for that reason only if you know that they need help. If it only seems to you that you know that they need help, then it only seems to you that you can act for that reason.

On a robustly realist conception of the normative, we will expect the distinction between appearance and reality to apply to every dimension of the normative, just as it applies to everything else (including appearances themselves). That does not make the normative any more mysterious or unknowable than it already was. On every dimension of the normative, there is a difference between the facts and what the subject is in a position to know of the facts, even though the subject is often in a position to know the most important facts. Once we have acknowledged that difference, we should no longer be tempted by the futile project of trying to carve out a dimension of the normative that is perfectly accessible to the subject, either internally or externally. We acknowledge that the normative is imperfectly accessible to the subject. That motivates an externalist theory of normative grounds.

Notes

1 I replied to an earlier version of that paper at a meeting of the Philosophical Society in Oxford, and am grateful for having had the opportunity to debate the ideas in it with Robert Audi. Numbers in parentheses are page references to the published version. Thanks to audiences at Oxford and King's College London for discussion of earlier versions of this paper.

2 The example from Putnam 1973 is reconceptualized here as usual on the lines of Burge 1979, in terms of thought concepts rather than just linguistic meanings. For those worried about snow and other H₂O that isn't liquid water, or dirty water that isn't H₂O, or H₂O in Oscar's body but not Twin Oscar's, the example can easily be changed.

3 Ludlow and Martin 1998 provides an introduction to the debate on whether content externalism is compatible with privileged access.

4 For present purposes it matters neither whether Oscar is justified in believing that there are no pools of twater nor whether Twin Oscar is justified in believing that there are no pools of water.

5 The arguments of this paragraph, unlike the earlier ones about what Oscar or Twin Oscar is justified in believing but does not in fact believe, assume that belief contents are Russellian rather than Fregean: their individuation is coarse-grained

(roughly: determined by the referents of the terms with which we express the contents) rather than fine-grained (roughly: determined by senses of those terms). However, the neo-Fregean form of externalism arguably also implies differences in justification between internal duplicates, since sameness of sense and therefore the validity of arguments is sensitive to external factors; see Campbell 1987-88.

6 Audi 1995 has more extended discussion of such cases.

7 Read → as material rather than strict implication.

8 To be in a position to be aware of something is to have some sort of potential for awareness of it; how remote the potential can be is an important question which Audi discusses but is not of immediate concern here.

9 Goldman 1999 presses a similar question.

10 The remarks in the text should not be taken as endorsing the over-intellectualized view that the justification of a belief, in the sense of its normative status, is determined by the believer's capacity to respond in dialectically effective ways to challenges. Beliefs that constitute knowledge are in good standing even if the slow-witted knower is easily out-manoeuvred in dialectic; see section 2. In the quoted passage, Audi is proposing an aetiology, not an analysis, of our concept of justification.

11 If you are a contextualist, pick a suitable context of utterance for that sentence.

12 The difference in justification is explained more rigorously within a framework of evidential probability at Williamson 2004: 313-14. In addition to the degree to which the subject's evidence supports the given belief content, there is also the question of how far the subject's belief is in fact causally based on the evidence. This is an issue for externalists and internalists alike (Audi notes it at 43 n. 2). On the externalist view in the text, the subject in the sceptical scenario does worse in this respect too, since the belief is largely based on what the subject falsely regards as evidence.

13 Closely related arguments are developed in much greater detail in Williamson 2000.

14 On a more liberal conception of availability, p is available to you as a reason if and only if you are in a position to know p .

References

- Audi, R. 1995. 'Memorial justification'. *Philosophical Topics* **23**: 31-45.
- Audi, R. 2001. 'An internalist theory of normative grounds'. *Philosophical Topics* **29**: 19-46.
- Burge, T. 1979. 'Individualism and the mental'. *Midwest Studies in Philosophy* **4**: 73-121.
- Campbell, J. 1987-88. 'Is sense transparent?'. *Proceedings of the Aristotelian Society* **88**: 273-292.
- Goldman, A. 1999. 'Internalism exposed'. *Journal of Philosophy* **96**: 271-293.
- Ludlow, P., and Martin, N., (eds.) 1998. *Externalism and Self-Knowledge*. Stanford: CSLI.
- Nagel, T. 1976. 'Moral luck', *Aristotelian Society* sup. **50**: 137-155.
- Putnam, H. 1973. 'Meaning and reference'. *Journal of Philosophy* **70**: 699-711.
- Williams, B. 1976. 'Moral luck', *Aristotelian Society* sup. **50**: 115-135.
- Williamson, T. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.
- Williamson, T. 2004. 'Replies to commentators', *Philosophical Books* **45**: 313-323.