

Radical Externalism

Amia Srinivasan

St John's College, Oxford

(Forthcoming in *The Philosophical Review*; please cite published version)

ABSTRACT: This paper presents a novel challenge to epistemic internalism. The challenge rests on a set of cases which feature subjects forming beliefs under conditions of 'bad ideology' – that is, conditions in which pervasively false beliefs have the function of sustaining, and are sustained by, systems of social oppression. In such cases, I suggest, the externalistic view that justification is in part a matter of worldly relations, rather than the internalistic view that justification is solely a matter of how things stand from the agent's individual perspective, becomes the more intuitively attractive theory. But these 'bad ideology' cases do not merely yield intuitive verdicts that favour externalism over internalism. These cases are moreover *analogous* to precisely those canonical cases widely taken to be counterexamples to externalism: cases featuring brains-in-vats, clairvoyants, and dogmatists. That is, my 'bad ideology' cases are, in all relevant respects, just like cases that are thought to count against externalism – except that they intuitively *favour* externalism. This, I argue, is a serious worry for internalism, and bears interestingly on the debate over whether externalism is a genuinely 'normative' epistemology.

It is impossible by a mere individual...effort
to escape from the web of the social lie

TROTSKY¹

1. A new challenge for internalism

Consider the following case:

¹ As quoted and translated by Edmund Wilson (2003), p 438, from Trotsky's *Biography of Lenin, Volume 1* (1936).

RACIST DINNER TABLE: Nour, a young British woman of Arab descent, is invited to dinner at the home of a white friend from university. The host, Nour's friend's father, is polite and welcoming to Nour. He is generous with the food and wine, and asks Nour a series of questions about herself. Everyone laughs and talks amiably. As Nour comes away, however, she is unable to shake the conviction that her friend's father is racist against Arabs. But replaying the evening in her head she finds it impossible to recover just what actions on the host's part could be thought to be racist, or what would justify her belief in the host's racism. If pressed, Nour would say she 'just knows' that her host is racist. In fact the host *is* racist – he thinks of Arabs as inherently fanatic, dangerous and backwards – and as a result did send off subtle cues that Nour subconsciously registered and processed. It is this subconscious sensitivity that led to her belief that her host is racist.

Here is my question: is Nour's belief that her host is racist (epistemically) justified? I think the intuitive answer is yes. Nour's belief, after all, is the product of a sensitivity to racism, a sensitivity that allows her to dependably track whether or not the people she encounters are racist. It would seem odd to say that Nour ought not, epistemically speaking, have formed the belief that her host is racist, or that she did something epistemically impermissible in forming this belief. And it would seem similarly odd to say that, having formed the belief, Nour ought to now, epistemically speaking, give it up. Indeed, it seems right to say that if Nour were to give up her belief in her host's racism, she would be losing an item of *knowledge*. If so, it follows that Nour's belief must be justified.²

Of course, Nour has no awareness, introspective or otherwise, of how her subconscious racism-detection mechanism works – indeed, not even *that* it works. And she is unable to cite anything – any experience or bit of evidence – as *grounds* for her belief. (As she says, she 'just knows'.) We might well want to say that Nour would be better off, epistemically speaking, if she had such a higher-order

² I assume throughout that justification is a condition on knowledge.

awareness.³ Perhaps such an awareness would give Nour's cognitive economy a greater degree of overall coherence,⁴ or a greater robustness against misleading counterevidence.⁵ And yet it seems counterintuitive to infer from the fact that Nour could be epistemically better off in these ways to the conclusion that her belief as it stands is unjustified. For Nour's belief that her host is racist is not only true, but non-luckily so. Nour's subconscious sensitivity to racism means that her belief gets on to the truth not as a mere matter of chance, not as a happy accident, but as a matter of predictable dependability. Surely then, her belief is justified.

Now consider a second case:

CLASSIST COLLEGE: Charles is a young man from a working-class background who has just become the newest fellow of an Oxford college. He is initially heartened by the Master's explicit commitment to equality and diversity. The Master assures him that, though the college is still dominated by wealthy fellows, Charles will be welcomed and made to feel included. Indeed, the Master tells Charles, he too is from a working-class background, and has experienced plenty of discrimination in his time. Charles is confident not only that the college will be a good community for him, but also that the Master is a person of excellent judgment on these matters. However, a few incidents soon disrupt Charles' rosy view of things. At high table, when Charles explains that he went to a state school, a fellow responds with 'but you're so well-spoken!'. At a visit to the pub, a number of young fellows sing the Eton boating song while Charles sits uncomfortably silent. Finally,

³ Just because Nour doesn't know (ex hypothesi) how, or that, her racism-detection mechanism reliably works doesn't necessarily mean that Nour can't know that she knows that her host is racist. On a Stalnakerian view, knowledge iterates automatically, without the operation of inference. Many thanks to [omitted] for discussion of this issue.

⁴ See Sosa 2009 on the virtues of 'reflective' over 'animal' knowledge. See also Goldman 1988 on the distinction between 'weak' and 'strong' justification.

⁵ One might also think that Nour would need such higher-order awareness in order for the fact of her host's racism to constitute what Grice (2001, ch 3) called a 'personal' reason for Nour to act – for example, by declining future supper invitations. (Thanks to [omitted] for this point.) My own view is that Nour's knowing that her host is racist suffices to make this a practical reason for Nour. But nothing I say hangs on this.

Charles hears that the other fellows call him ‘Chavvy Charles’. Charles, who has a dependable sensitivity to classism, goes to the Master to report that he has experienced a number of classist incidents in college. Shocked, the Master asks him to explain what happened. But when Charles describes the incidents, the Master is visibly relieved. He assures Charles that none of these are genuinely classist incidents, but playful, innocuous interactions that are characteristic of the college’s communal culture. He tells Charles that he is sure that Charles himself will come to see things this way once he gets to know the college and its ways better. And finally, he gently suggests that Charles is being overly sensitive – something to which (the Master goes on) Charles is understandably prone to being, given his working-class background. Charles leaves the conversation unmoved, continuing to believe that he has faced classist discrimination in the college, and dismissing the Master’s testimony. Charles meanwhile is unaware that some people from working class backgrounds (e.g. the Master) suffer from false consciousness, distorting their ability to recognise class-based oppression.

Is Charles’s ultimate belief that the college is classist justified?⁶ I think the intuitive answer is, again, yes. Like Nour, Charles forms a true belief on the basis of a dependable sensitivity to instances of classism. Unlike Nour, Charles maintains his true belief in the face of seemingly credible, misleading evidence – namely, the Master’s testimony to the effect that Charles’ belief is actually formed on an unreliable basis, viz. an *oversensitivity* to classist slights. And yet, this does not seem intuitively to affect the justification of Charles’ belief. Charles is not only justified in forming the initial belief that the college is classist; he intuitively remains justified even after the Master gives his misleading testimony to the contrary. Indeed, as with Nour, it seems right to say that Charles *knows* that the college is classist, both before and after the Master’s attempt to explain away Charles’ belief.

⁶ For the sake of simplicity, I am going to use the phrase ‘the college is classist’ as shorthand for the fact that Charles experienced a series of classist incidents in the college. Of course, just what it is for an institution to be classist, racist, sexist, etc., and just how this relates to the classism, racism, sexism, etc. of its constituent members, is a complicated issue.

In dismissing the Master's testimony, we might think that Charles exhibits a mild form of epistemic dogmatism, a dogmatism that could, if indulged in other circumstances, lead him to recklessly dismiss non-misleading evidence and court ignorance. (One way of cashing this out is to note that, if Charles had been in world in which his evaluation of the college was wrong, and the Master's right, he would have stubbornly maintained a false belief.)⁷ We might also think that Charles would be better off, epistemically speaking, if he had available to him the phenomenon of false consciousness as a debunking explanation of the Master's testimony: if he were in a position to *explain away* the Master's testimony, rather than simply dismiss it. Perhaps we even think that Charles is somewhat blameworthy, epistemically speaking, for this act of dogmatism. And yet, none of this intuitively precludes Charles from justifiably believing – indeed, I think, *knowing* – that the college is classist. Charles' belief that the college is classist, like Nour's belief that her host is racist, is true not as a matter of good luck, of happy accident, but as a function of his capacity to dependably get on to the truth. This intuitively seems sufficient to justify it.

Bear with me for a third, and final, case:

DOMESTIC VIOLENCE: Radha is a woman who lives in rural India. Her husband, Krishnan, regularly beats her. After the beatings, Krishnan often expresses regret for having had to beat her, but explains that it was Radha's fault for being insufficiently obedient or caring. Radha finds these beatings humiliating and guilt-inducing; she believes she has only herself to blame, and that she deserves to be beaten for her bad behaviour. After all, her parents, elders and friends agree that if she is being beaten it must be her fault, and no one she knows has ever offered a contrary opinion. Moreover, Radha has thoroughly reflected on the issue and concluded that, given the natural social roles of men and women, women deserve to be beaten by their husbands when they misbehave.⁸

⁷ One might object that there is in fact no metaphysically possible world in which – holding fixed the non-normative facts – the college is not classist. For my purposes, however, what matters is that there *conceivably* is such a world, not that such a world in fact be metaphysically possible. For discussion, see §3.2.

⁸ On contemporary attitudes towards domestic violence in India, see:

Is Radha's belief that she deserves to be beaten justified? I think the answer is: surely not. For Radha's belief is not merely false, but moreover the product of a convincing, and systematic, patriarchal illusion: that it is men's place to subordinate women. This illusion – one that infects not only the testimony of Radha's peers and respected elders, but her moral emotions (shame, remorse) and best attempts at rational reflection – ensures that Radha has no dependable access to the moral facts of her situation. Radha, despite her own best efforts, is tragically cut off from moral reality.

Radha's false belief is hardly her fault; it is not only explained but obviously *excused* by the patriarchal illusion of which she is a victim. Radha is doing, we want to say, the best she can, given her own distorted epistemic connection to the world. Her belief is eminently understandable; we would be naïve to expect anything better of ourselves in Radha's position. And yet none of this is the same as saying that Radha's belief is justified. Indeed, once we draw the distinction between justification on one hand, and excusedness or blamelessness on the other,⁹ it feels intuitive, I think, to say that Radha's belief meets the conditions for the latter, but not the former. Radha's belief is the product of a distorted relationship to reality – a relationship that excuses the falsity of her beliefs, but does not thereby render them justified.

These three cases – RACIST DINNER TABLE, CLASSIST COLLEGE and DOMESTIC VIOLENCE – together present, I want to suggest, a serious challenge to a widely-held view in epistemology. According to epistemic *internalism*, justification is a matter of a subject's (non-factive) mental states: 'internal' duplicates, the internalist says, do not differ in justification. A typical internalist says that epistemic justification is a matter of fit with one's evidence, or with one's epistemic reasons, or more generally with how things look from one's own perspective on the world – where it is presumed that such facts are facts about one's (non-factive)

<https://dhsprogram.com/pubs/pdf/frind3/frind3-vol1andvol2.pdf>, pp 475ff.

⁹ The commonsensical distinction between justification and excuse is one to which Austin (1956) famously exhorted philosophers to attend. If I run over your dog while carefully backing out of my drive, I might be excused for killing him, but I certainly wouldn't be justified for so doing.

mental states.¹⁰ Epistemic *externalism*, meanwhile, denies that epistemic justification supervenes solely on such ‘internal’ facts: ‘internal’ duplicates might well differ in justification. The externalist says that epistemic justification is at least partly a matter of facts that lie beyond one’s mental states – for example, whether one’s belief exhibits an appropriate causal connection to its content,¹¹ or is a product of a reliable or safe method.^{12,13}

It is generally thought, by internalists and externalists alike, that intuitive reflection on a range of well-known cases – brains-in-vats, unwitting clairvoyants, dogmatists, and so on – supports internalism over externalism. But the three cases I described above – RACIST DINNER TABLE, CLASSIST COLLEGE and DOMESTIC VIOLENCE – disrupt this tidy view of things. For externalism has a much easier time of vindicating what I take to be the intuitive verdicts on these cases than does internalism. In RACIST DINNER TABLE and CLASSIST COLLEGE, the subjects have a belief that is, *ex hypothesi*, reliably and safely connected to the truth. It is thus no mystery, from the externalist perspective, how such beliefs could be justified, since

¹⁰ One can have a view on which one’s epistemic reasons or one’s evidence is not a matter of one’s non-factive mental states, e.g. one can think, with Williamson, that one’s evidence is what one knows (Williamson 2000, ch 9).

¹¹ Goldman 1967.

¹² On reliability theories see Armstrong 1973 and Goldman 1975. On safety theories see Sosa 1996 and 2000, Williamson 2000 and Pritchard 2005.

¹³ One might worry that the debate between internalism and externalism is merely verbal, since many externalists wish to maintain an internalistic notion within their epistemic taxonomies, e.g. Goldman has a notion of ‘weak’ justification (1988), Lasonen-Aarnio has the notion of ‘reasonableness’ (2010), and Wedgwood has ‘rationality’ (2002). However, such externalists insist that the notion of justification relevant to *knowledge* is externalistic, not internalistic, rendering their opposition to internalism substantive. A different way of dissolving the internalism/externalism debate is by adopting a *pluralist* metaepistemology, on which there are different sorts of epistemic justification, and (thus) different sorts of knowledge. (A pluralism about justification alone will not dissolve the internalist/externalist debate.) Elsewhere I argue that there are no non-trivial generalisations about the supervenience base of the relevant internalistic notion ([omitted]). Many thanks to an anonymous referee and [omitted] for urging me to address these issues.

they straightforwardly satisfy the typical externalist conditions on justification.¹⁴ In DOMESTIC VIOLENCE, meanwhile, the subject fails to exhibit such a connection between her belief and the truth, even while doing as best as she can by her own lights, believing in accordance with her evidence, and so forth. Thus it is again no mystery, from the externalist perspective, how her belief could fail to be justified.

The internalist, meanwhile, has a more difficult time of it. Internalists divide into two kinds, depending on how they understand what it is to be an ‘internal’ duplicate. According to *access* internalism, A and B are internal duplicates just in case they are identical with regard to their *introspectively accessible* mental states. According to *mental state* internalism, A and B are internal duplicates just in case they are identical with regard to their (non-factive) mental states, whether those states are accessible or not.¹⁵ For the sake of simplicity, I will focus on the challenge that my cases present to access internalism. (In §5, however, I will briefly show how my challenge extends to mental state internalism as well.)

How could Nour or Charles be *justified*, according to the (access) internalist¹⁶, given that neither has any awareness of the reliable grounds of his or her belief – and indeed, in Charles’ case, has strong internalistic reason to think he is *not* so reliably grounded? And how could Radha be *unjustified*, according to the internalist, given that she believes in accordance with her (rather misleading) evidence, with what she has reason to believe, and with how things seem to her? The internalist appears to be faced with a choice between biting the bullet on these cases – conceding that they provide intuitive support for externalism – or trying to find a way to vindicate the ‘externalistic’ intuitions in a way consistent with internalism.

But the real challenge for internalism lies elsewhere. My three cases are not merely recalcitrant to internalistic treatment. They are moreover analogous

¹⁴ Externalistic treatment of cases like CLASSIST COLLEGE is complicated by the question of whether justification can be *defeated* by misleading evidence, a question on which externalists differ. I will discuss these complications shortly. Suffice it to say for now that at least some externalists would say that Charles’ belief retains its justification in the face of the Master’s testimony.

¹⁵ On the distinction between access and mental state varieties of internalism, see Conee and Feldman (2001).

¹⁶ From now on, when I refer to ‘internalists’ and ‘internalism’ I mean access internalists and access internalism, unless stated otherwise.

to those very cases that internalists have canonically presented as counterexamples to *externalism*. Consider, for example, one of Laurence Bonjour's famous such counterexamples:

CLAIRVOYANT: Norman, under certain conditions that usually obtain, is a completely 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 reliable¹⁷

Because Norman's belief is based on a reliable method, the typical externalist will say that it is justified.¹⁸ But many find this externalist verdict absurd. Bonjour writes that Norman's belief is 'epistemically irrational and irresponsible, and thereby unjustified'.¹⁹ For it is part of one's epistemic duty, he goes on, to 'reflect critically upon one's beliefs, and such critical reflection precludes believing things to which one has, to one's knowledge, no reliable means of epistemic access'.²⁰ Norman's belief, Bonjour says, is from Norman's own perspective nothing more than an 'unfounded hunch' (ibid). Thus Norman's belief, *pace* the externalist, is unjustified.

It is worth noting that not all internalists will agree with Bonjour's account of why Norman is unjustified. For Bonjour is invoking a very strong internalist condition on justification, according to which S's belief that *p* is justified iff S has (independent and undefeated) reason to believe that her belief was formed on a reliable basis. For Bonjour, agents need to be not only aware of their grounds, but also aware *that* their grounds are their grounds. Most internalists will reject this strong demand, insisting, more minimally, that agents must be aware of their

¹⁷ Bonjour 1980, 62.

¹⁸ Bonjour intends CLAIRVOYANT as a counterexample to reliabilist versions of externalism, but it can be thought of as a potential counterexample, *mutatis mutandis*, to other varieties of externalism as well.

¹⁹ Ibid 63.

²⁰ Ibid.

grounds in order to have a justified belief. The problem with Norman for this more common type of internalist is that there appears to be nothing of which Norman is aware that *could* serve as the grounds for his belief. While BonJour doesn't explicitly say so, we can stipulate that Norman's clairvoyance is accompanied by no experiences or phenomenology; rather than clairvoyance being a sort of quasi-perceptual seeming, it is simply a subconscious power that produces reliably true beliefs that 'pop' into Norman's head. Norman's belief can't be justified, on the typical internalist view, because Norman is bereft of any mental state to serve as its grounds.

And yet, CLAIRVOYANT is analogous to RACIST DINNER TABLE, in which, recall, Nour's belief is intuitively *justified*. Both Norman and Nour exhibit a sensitivity to the truth, a sensitivity of which they are unaware but that nonetheless produces reliably true beliefs. From her internal perspective, Nour's belief is no better, BonJour would presumably say, than a 'hunch', and must be therefore unjustified. And, like Norman, Nour has nothing that is introspectively available to her – no experiences or phenomenology – that could potentially serve as the grounds for her belief, since it is only her subconscious that detects the subtle cues of her host's behaviour. And yet Nour's belief that her host is racist seems eminently *justified*.

But how could it be that Norman's belief is any less justified than Nour's? Indeed it cannot be so. Insofar as these cases are analogous in the relevant respects, Nour and Norman's justification must stand or fall together.²¹ It is not enough, then, for the internalist simply to bite the bullet on RACIST DINNER TABLE, concluding that it intuitively supports externalism while cleaving nonetheless to internalism. If the internalist wants to continue to use CLAIRVOYANT as evidence against externalism and in favour of internalism, she needs to say something about RACIST DINNER TABLE. Either the internalist needs to tell us why RACIST DINNER TABLE is in fact relevantly disanalogous to CLAIRVOYANT, or she needs

²¹ While I myself am inclined to endorse the claim that Nour is justified iff Norman is justified, all that matters is there are no relevant disanalogies between the two cases to which the *internalist* can point. An externalist might well think that Nour is justified while Norman is not, for reasons unavailable to the internalist: for example, because Nour's belief-forming method is naturalistically explicable, while Norman's is not. For the sake of ease I will assume an externalism on which the two cases are relevantly analogous, but nothing in my argument hangs on this. Many thanks to an anonymous editor for this point.

to offer us an error theory as to why the intuitions elicited by RACIST DINNER TABLE are not to be trusted. Of course, the externalist who wants to use RACIST DINNER TABLE as part of her case against internalism faces a symmetric challenge: that is, she is under pressure to explain why RACIST DINNER TABLE, but not CLAIRVOYANT, elicits unreliable, internalistic intuitions about justification. And indeed, I will offer such an error theory on behalf of the externalist in the course of this paper.

For now let me return to CLASSIST COLLEGE and DOMESTIC VIOLENCE, which, I want to argue, are also analogous to cases traditionally thought to favour internalism over externalism. Recall that in CLASSIST COLLEGE, I claimed, Charles has an intuitively justified (and indeed intuitively knowledgeable) belief that his Oxford college is classist, despite the misleading testimony of the Master. Charles' belief is intuitively justified, I suggested, because it is based on Charles' dependable sensitivity to instances of classism. But CLASSIST COLLEGE is analogous to a kind of case that is canonically thought to elicit the opposite intuition, and thereby impugn externalism. Compare:²²

DOGMATIST: At a time t_1 Mary walks into an art gallery and sees a red sculpture. There is nothing abnormal about Mary's perceptual faculties or the lighting conditions in the gallery. Thus she forms a true belief that the sculpture is red. At a slightly later time t_2 a gallery assistant tells Mary that the sculpture is not red, but illuminated by a hidden red light, such that any object it shines on would look red even if it weren't. Mary ignores the misleading testimony and continues to believe, on the basis of her reliable perceptual faculties, that the sculpture is red. What Mary does not know is that the exhibition – including the gallery assistant's misleading testimony – is being put on by a famous artists' collective dedicated to epistemic hoaxes.

²² There are many versions of this case, but mine follows most closely Lasonen-Aarnio's 'Trick on Suzy' (2010, 1). See also Chisholm 1966: 48, Bonjour 1980: 59–60, Pollock 1995: 41 and Pollock and Cruz 1999: 44. Lasonen-Aarnio, unlike the other authors, embraces the (pure) externalist verdict that her protagonist continues to *know* in the face of the misleading testimony – a case of what she calls 'unreasonable knowledge'.

The standard intuitive verdict on DOGMATIST is that Mary's belief, while initially justified at t_1 , loses justification at t_2 , when the misleading testimony is delivered. After all, how could Mary's belief that the sculpture is red be justified after she is told by a seemingly reliable expert that her perceptual capacities are unreliable? In continuing to believe that the sculpture is red, doesn't Mary ignore evidence that, from her perspective, bears squarely and damningly on her belief? Isn't her belief, even if true, both irresponsible and blameworthy? And if so, how could it be justified?

Indeed, the intuition that Mary (and other similar dogmatists) are unjustified is generally thought so compelling that most externalists feel pressure to modify their externalism in order to vindicate it. What we might call a *pure* externalism says that the satisfaction of the externalist condition (e.g. reliability, safety) is both necessary and *sufficient* for justification. According to pure externalism, Mary's belief at t_2 – because it is still based on her reliable (safe, etc.) perceptual faculties – retains its justification.²³ Cases like DOGMATIST prompt most externalists to reject pure externalism, instead favouring a modified, internalistically-inspired theory according to which S's belief is justified just in case S's belief satisfies the externalist justification-condition *and* S is not in possession of strong misleading evidence to the contrary.²⁴ This modification allows the moderate externalist to vindicate the intuition that Mary's belief is justified at t_1 but loses its justification at t_2 . Meanwhile, only a small minority of externalists are willing to bite the bullet on DOGMATIST-type cases in order to

²³ A pure externalist could argue that Mary loses justification at t_2 (and *mutatis mutandis* for other DOGMATIST-type cases) because the method on which Mary's belief at t_2 is based (which involves dismissing misleading evidence) is in fact unreliable or unsafe. Thus DOGMATIST turns out to be a case of Mary switching from a justification-conferring method to a justification-depriving method. I set aside this possibility in what follows, assuming that Mary's belief-forming method is stable from t_1 to t_2 .

²⁴ For some examples of moderate externalists, see Alston 1988, Bergmann 2006, Goldman 1986: 62-3 and 111-2, and Nozick 1981, 196. Some externalists (e.g. Bergmann and Goldman) endorse defeat conditions that are even more liberal than those of the moderate externalism I describe, allowing that even *unjustified* higher-order beliefs (e.g. beliefs generated by baseless paranoia) are sufficient to destroy justification.

maintain pure externalism.²⁵ DOGMATIST-type cases are widely thought to show that pure externalism must be false.

And yet, DOGMATIST is analogous to CLASSIST COLLEGE, in which Charles' belief, recall, intuitively *retains* its justification, despite the misleading evidence, just as the pure externalist predicts. How could Charles and Mary differ in justification? Both Charles and Mary use their properly functioning capacities – to detect classism, to detect colour – to arrive at their respectively true beliefs. They are then both met with testimony, from sources they have strong antecedent reason to believe to be trustworthy, to the effect that their belief-forming capacities are in fact unreliable. They both dismiss this misleading evidence, despite the fact that neither has available to her the proper explanation of why the evidence is misleading – that the Master is suffering from false consciousness, that the 'gallery assistant' is part of the art piece – maintaining their original beliefs. So it seems that Mary's belief cannot be any less justified than Charles'. Either both beliefs are justified, or neither is. The internalist who wishes to use DOGMATIST as a counterexample to (pure) externalism will have to explain why it is that an analogous case appears to be a counterexample not to externalism, but internalism.

Third and finally, recall that in DOMESTIC VIOLENCE, Radha has an intuitively unjustified belief that she deserves to be beaten, a belief that is a symptom of a systematically distorted relationship to reality. Now consider the following case:

BRAIN-IN-A-VAT: Jane_{BIV} is a handless brain-in-a-vat, subjected to a compelling, electrochemical illusion to the effect that she is a normally embodied person

The external world sceptic asks how it is that Jane, a normally embodied person, can know that she has hands given that, for all Jane knows, she could be Jane_{BIV}. The externalist answers that because Jane is, *ex hypothesi*, a normally embodied person – i.e. because Jane is in fact not Jane_{BIV} – Jane's belief that she has hands enjoys a (reliable, safe, etc.) connection with the external world which in turn secures justification. While Jane and Jane_{BIV} are internal duplicates, the externalist insists, their beliefs enjoy different justificatory standing. In turn this explains, the

²⁵ e.g. Plantinga 1986, Lasonen-Aarnio 2010, Williamson 2014.

externalist goes on, how it is that Jane can know she has hands despite the fact that Jane_{BIV}'s situation is subjectively indiscriminable from Jane's.²⁶

But this capacity to brush off sceptical threats comes at a price, one that strikes most internalists as extortionate. For typical externalist theories imply that Jane_{BIV}'s belief that she has hands is *unjustified*: for Jane_{BIV}'s belief, unlike Jane's belief, lacks the appropriate externalist connection to the world. But how could it be, the internalist asks, that Jane and Jane_{BIV} differ in justification, when (it is granted by all parties) things appear just the same for both Jane and Jane_{BIV}; when Jane_{BIV} is entirely blameless for her belief; when Jane_{BIV} seems to be acting no less responsibly, epistemically speaking, than Jane; and when it is just a matter of bad luck that Jane_{BIV} is envatted rather than embodied? Stewart Cohen famously called this the 'new evil demon problem' for externalism.²⁷ 'It strikes me as clearly false' Cohen writes, 'to deny that [the brain-in-a-vat's] beliefs could be justified. If we have every reason to believe e.g., perception is a reliable process, the mere fact that unbeknown to us it is not reliable should not affect [our] justification' (ibid, 281-2). Cohen explains that this judgment 'hinges' on his 'viewing justification as a normative notion' (ibid 282). So long as one's belief is 'appropriate to the available evidence', he says, one cannot be 'held responsible for circumstances beyond [one's] ken' (ibid). Justification, in other words, cannot be a matter of facts unavailable to one: the externalist's verdict on BRAIN-IN-A-VAT must be false.

And yet, BRAIN-IN-A-VAT is analogous to DOMESTIC VIOLENCE, where the intuitive verdict was, I suggested, that Radha's belief is *not* justified. Radha and Jane_{BIV} are both victims of systematic illusions; Jane_{BIV} is literally envatted, while Radha is, as it were, envatted in patriarchal ideology. Both of their beliefs are internalistically impeccable: both Jane_{BIV} and Radha believe in accordance with how things seem to them, do not neglect any evidence, and do as well (epistemically speaking) as we would in their shoes (or vat). Why then should Radha's belief that she deserves to be beaten be any less justified than Jane_{BIV}'s belief that she has hands? The internalist who wants to use BRAIN-IN-A-VAT as

²⁶ Is Jane's situation also subjectively indiscriminable from Jane_{BIV}'s situation? The typical externalist will think not. Jane is able to know that she has hands, and from this (the externalist thinks) she can knowledgeably infer that she is not Jane_{BIV}. Thus the externalist will object to the sceptic's initial claim that Jane, 'for all she knows', could be Jane_{BIV}.

²⁷ Cohen 1984.

part of her argument against externalism – who wants, that is, to insist that the ‘new evil demon problem’ really is a problem – needs to explain just why it is that our intuitions about DOMESTIC VIOLENCE appear to favour externalism over internalism.

Together, DOMESTIC VIOLENCE, RACIST DINNER TABLE and CLASSIST COLLEGE present a serious challenge for internalism. As I have said, these cases are more straightforwardly and intuitively handled by externalism than by internalism. More pressingly, these new cases are analogous to precisely those familiar cases that are generally thought to be counterexamples to externalism. The internalist who wants to continue to treat the traditional cases as evidence against externalism is thus under pressure to say something about my new cases. Here she has two options: to show that my new cases are in fact relevantly disanalogous to the traditional cases, or to show how the externalistic intuitions the new cases elicit can be explained away by a compelling error theory.

Having set out this new challenge to internalism, the remainder of this paper proceeds as follows. In §2 I offer a diagnosis as to why our intuitive verdicts diverge between these two sets of cases, old and new. In the new cases, subjects are operating under what we might call conditions of *bad ideology*: that is, conditions in which pervasively false beliefs have the function of sustaining, and are in turn sustained by, systems of social oppression. When we consider subjects operating under such conditions, I want to suggest, the externalist verdict that justification is a matter of an agent’s relationship to the world becomes much more intuitively appealing than the internalist verdict that justification is a matter of how things stand from the agent’s individual perspective. Externalism, but not internalism, allows us to treat justification as a structural matter: that is, a matter of the larger systems in which agents are embedded.

I then go on to canvass strategies for an internalist response to my challenge. In §3 I raise, and respond to, the objection that my new cases are importantly disanalogous to the old cases. In §4 I discuss the internalist’s prospects for an error theory: that my ‘bad ideology’ cases are too morally/politically charged to generate reliable intuitions. In §5 I briefly show how my challenge applies to mental state internalism, and thus to internalism generally. In §6 I conclude with a discussion of how my challenge bears on the debate over externalism’s status as a genuinely ‘normative’ epistemology.

2. A diagnosis: justification and bad ideology

What explains the divergence in our intuitions across the two sets of cases I discussed in §1? The new cases are what we might call ‘bad ideology’ cases – that is, cases that feature subjects who exist in conditions in which pervasively false beliefs have the function of sustaining (and are in turn sustained by) systems of social oppression: patriarchy, racism, classism. In such cases, I want to suggest, the salient epistemological question becomes not whether subjects are blameworthy or praiseworthy for their beliefs – whether their beliefs are reasonable by the subjects’ own lights – but how these beliefs relate to a system whose function it is to distort subjects’ access to the truth for the purposes of oppressing them. What intuitively matters most in such cases is whether the subject’s truth-tracking capacities are distorted by ideological forces, or whether the subject is endowed with capacities that allow her to pierce through ideological distortion.

Thus in DOMESTIC VIOLENCE, what seems to intuitively matter most, epistemically speaking, is not that Radha’s belief is perfectly reasonable by her own lights – a thought the externalist is happy to capture by saying her belief is *excused*²⁸ – but that her belief is the product of an ideologically-distorted mechanism. This,

²⁸ On the externalist notion of excuse, see Williamson 2007.

we instinctively feel, is what matters for justification.^{29,30} Meanwhile, in the case of Nour and Charles, we have subjects who are able to reliably get on to the truth,

²⁹ What should we say about *true* ideological beliefs, of the kind that arise because of self-fulfilling processes (see Haslanger 2007)? For example, what shall we say about a case in which a man's belief that his wife is submissive is reliably true, but as a result of his treating her with the expectation that she will be submissive? (This sort of belief is usually accompanied by a further false belief that women are submissive by *nature*. But here I am concerned with the straightforwardly true belief.) A standard externalism might count such a belief as justified, since (with regard to self-fulfilling ideological beliefs) believing in accordance with patriarchal ideology is a reliable, truth-tracking, safe, etc. method. I am inclined to accept this verdict, and indeed to say that part of what is troubling about self-fulfilling ideology cases is precisely that oppression can give rise to not only true but moreover *justified* beliefs. If this is right, then it seems that externalism can only explain the deficiencies of *false* ideological beliefs, not true ones. Of course, an externalist need not say this: she might argue instead that believing in accordance with patriarchal ideology is in fact an unreliable, unsafe, etc. method – a method that yields nearby false verdicts in cases of non-self-fulfilling beliefs. (That it is an open question whether ideological mechanisms are reliable, safe, etc. in such cases is an instance of the well-known generality problem for externalism.) Thanks to [omitted] for calling my attention to this issue, and to an anonymous referee for further drawing me out.

³⁰ An anonymous referee presses the following challenge. Imagine a case, DOMESTIC VIOLENCE*, in which Radha* exists in a non-patriarchal society, but because of a quirk of her neurophysiology, believes she deserves to be beaten by her husband. Is Radha* any more intuitively justified than Radha? If not, then it would seem that the right explanation for why we are inclined to judge Radha unjustified in the original case has nothing to do with the presence of bad ideology. I am not convinced by this objection. First, cases of neurophysiological dysfunction are not obviously cases that meet the standards for internalistic justification: some internalists will simply want to say that only agents with minimally competent cognition are candidates for justified belief. If so, then these two cases are not analogous tests of internalistic intuition. Second, it's not clear that we *can* entertain a case featuring a woman believing she deserves to be beaten by her husband without tacitly invoking bad ideology. Consider a case that features a false belief about dessert that is (in our world) clearly non-ideological. Suppose Radha** falsely believes that she deserves the last slice of cake because all her evidence suggests that this is right. After all, all her friends and family says she does, as do her gut instincts and her careful deliberations. Is her false belief justified? I suspect that internalists will be much more inclined to say yes than in the original DOMESTIC VIOLENCE case.

despite their bad ideological circumstances, thanks to a socially-endowed capacity to pierce through ideological illusion. For Nour and Charles, because of their respective social positions – as an Arab, and as a member of the working class – share an ability to reliably apply concepts (*racist*, *classist*) in ways that contest the dominant ideology.³¹

Externalistic verdicts in bad ideology cases are attractive, I want to suggest, because what intuitively matters in such cases is not how things seem from the agent’s own (often limited) perspective, but how the agent relates to the epistemically distorting systems in which they are embedded.³² Justification, for the internalist, is a sort of meritocratic good: it is available to all minimally competent agents, regardless of their circumstances, distorted or veridical. All that is required to be internalistically justified is individual conscientiousness. (Thus Radha must be justified.) Moreover, internalistic justification is not something that one can possess through the good luck of veridical circumstances: agents who are ‘internally’ equal will also be equal in justification. (Thus Nour and Charles must be unjustified.) For the externalist, meanwhile, justification can come apart from questions of personal responsibility and blamelessness. To be externalistically justified requires, in part, the cooperation of the external world: one must have an undistorted relationship to the relevant bit of reality, which is not something entirely within one’s control. Thus Radha is unjustified, despite her individual conscientiousness and blamelessness, because of her ideologically-distorted connection to reality. What is more, the externalist maintains that one can, as it were, ‘stumble into’ justification, by being felicitously connected to the relevant bit of reality. Thus Nour and Charles are justified, despite the fact that they each have an internal duplicate who *falsely* believes that, respectively, their host is racist and

³¹ Radha is a member of an oppressed social group – women – who is not so similarly endowed. Mere membership in an oppressed class does not suffice for the ability to see through bad ideology. A discussion of this point, and its connection with standpoint epistemology, follows shortly.

³² To be clear, I am not arguing that epistemic justification is in any way *context-dependent*, shifting with whether a subject exists under conditions of bad ideology or not. I am instead merely offering an explanation as to why it is that, for many, the externalistic verdict becomes more intuitively attractive when we shift from the old (non-ideological) to the new (ideological) cases.

their college is classist.³³ This is because (so says the externalist) Nour and Charles have the good luck – thanks to their social positions – of being properly connected to the world in a way that is conducive to knowledge. In all three cases, whether the subjects are justified or not turns, in part, on factors that are not within their individual control.

In short, externalism – in its insistence that justification can supervene on facts external to the agent’s own ken – is poised to vindicate what we might think of as a *structural* rather than merely *individualistic* notion of justification. A structural explanation gives an account of its explanandum by averting to the larger system of which the explanandum is a part, rather than (solely) adverting to features of the explanandum itself. To explain that the dutiful housewife does the lion’s share of the domestic labour because she prefers it that way is to give an individualistic explanation of her behaviour; to explain that the dutiful housewife does the lion’s share of the domestic labour because that is what is socially expected of women is to give a structural explanation. Both explanations might well be true – social expectations can produce adaptive preferences – but in certain contexts it is clearly the latter sort of explanation that is important. Similarly, to employ a structuralist account of justification is to explain a subject’s justificatory status in terms of the broader epistemic system in which she is embedded. Radha’s belief is unjustified *because she is the victim of bad ideology*. Nour and Charles’ beliefs are justified *because their group membership allows them to pierce through bad ideology*. Meanwhile, the internalist – in her insistence that justification supervenes on a subject’s mental states – is not poised to underwrite a structural notion of justification. Instead, she can only explain an agent’s justificatory status in terms that are intrinsic to the agent herself. Externalism, but not internalism, is poised to vindicate Trotsky’s claim that ‘escape from the web of the social lie’ is more than a matter of ‘mere individual effort’.

As the reference to Trotsky suggests, my way of thinking about what might be deeply at stake between internalism and externalism – that is, the ability to

³³ One might object that Charles does *not* in fact have an internal duplicate who falsely believes that the college is classist, since any possible college in which such events took place would be classist (cf. fn 7). Similarly, one could object that Radha does not have an internal duplicate who truly believes that she deserves to be beaten – and that this in turn shows that Radha’s belief is not even internalistically justified. I take up this line of objection in §3.3.

vindicate a structural epistemology – has resonances with Marxist standpoint epistemology. For Marx, the proletariat’s relationship to the means of production confers on it, as a class, an epistemic privilege vis-à-vis society’s economic relations. While the ‘Free-trader Vulgaris’ sees the marketplace as ‘a very Eden of the innate rights of man’ where ‘alone rule Freedom, Equality, Property and Bentham’, those who are forced to sell their labour are positioned to see the material reality under the ideological appearance: to see that the marketplace is a site of exploitation, where they have ‘nothing to expect but...a hiding’ (Marx 1867/1887, 123).³⁴ Likewise, for feminist standpoint epistemologists, it is women’s relationship to the means of reproduction – women’s role in childrearing, nurturing, and caregiving – that confers on them, as a class, an epistemic advantage in understanding the real material relations of society under the ideological (patriarchal) appearance.³⁵ As both Lukács and Hartsock stress, the proletarian and feminist standpoints, respectively, are to be *achieved*, and are not something automatically given in virtue of one’s status as an oppressed subject under capitalism or patriarchy. Piercing the ideological appearance requires an overcoming of false consciousness and the achievement of revolutionary consciousness, in turn a matter of both political analysis and political action. But neither revolutionary analysis nor revolutionary action is easily achieved by the individual alone: thus the emphasis on *collective* consciousness raising in both Marxist and (especially) feminist politics. This is the second sense, for standpoint epistemologists, in which the overcoming of the ‘social lie’, as Trotsky says, is not a merely individual achievement. First, where one finds oneself in the social ordering (one’s place, say, in relation to the means of production or reproduction) confers on one epistemic advantages or disadvantages, putting one in a better or worse place to recognise the truth under the lie. Second, even if one is, epistemically speaking, advantageously positioned, one’s likelihood of achieving the privileged standpoint turns on one’s relationship to others – specifically, whether one enjoys the sort of political community that can together create a rival to the dominant ideology.

³⁴ The classic elaboration of Marxist standpoint theory is Georg Lukács’ (1923/1971).

³⁵ The *locus classicus* of feminist standpoint epistemology is Hartsock 1983. See also Harding 1983 and 2004, Jameson 1988a, Collins 2000, and Wylie 2003. Note that the particular understanding of the feminist standpoint as grounded in women’s relationship to the means of reproduction, due to Hartsock, is not shared by all proponents of feminist standpoint theory.

There is a natural kinship, I want to suggest, between Marxist standpoint epistemology and externalism. Both stress the way in which the distribution of epistemic goods turns, in part, on factors beyond subjective control, producing a deep epistemic asymmetry between agents in internally analogous positions. Thus the proletarian, like Charles, is able to know something of the truth about his society, despite the fact that he is surrounded by the misleading counter-testimony of those who see the world as the reigning ideology dictates – and despite, moreover, the fact that such dogmatism, in someone not so reliably situated vis-à-vis the reigning ideology, would lead to error. And, like Nour, the proletarian is able to know even without knowing the grounds of her belief – even though such credulity would, again, lead to error in someone not so reliably situated. Meanwhile Radha, like some members of the proletariat, is a classic victim of bad ideology, believing that the world is just as it presents itself as being. Hartsock writes that the capitalist’s vision of the world ‘cannot be dismissed either as simply false or as an epistemology relevant to only a few’. For, she goes on, ‘the worker as well as the capitalist engages in the purchase and sale of commodities...and [as] material life structures consciousness, this cannot fail to have an effect’ (Hartsock 1983, 288). This is not to say that Radha – or anyone else suffering from false consciousness – could *never* come to know the truth of her situation. But it is to say that, for many victims of false consciousness, coming to know the truth would require something more than more assiduous reflection. Most obviously, it would require that she have her consciousness raised through political engagement.³⁶

³⁶ What is the externalist to say about the false ideological beliefs of those who *benefit* from the dominant ideology? What should we say, for example about the Wall Street trader who believes he deserves his wealth because he works so hard? Naturally, the externalist will say that this belief is not justified (again, I think this is the intuitively correct verdict), since it is based on a faulty mechanism – viz. the ideology of meritocracy. But is the trader *excused* for believing he deserves his wealth? What about the false ideological beliefs of those who are also (but differentially) oppressed, e.g. the out-of-work coal miner who blames his joblessness on immigrants? My own intuition is to say that the out-of-work coal miner’s false belief is more excused than that of the Wall Street trader, but that perhaps even the latter has some degree of excuse. The details matter: what evidence do the trader and coal miner have? Are they really like Radha, who is fully immersed in a self-confirming ideological scenario? (I do not think that all victims of oppression are like Radha; many might well be *blameworthy* for their false ideological beliefs, because they fail

What about Charles and Nour? Are they not able to achieve the relevant standpoints all by themselves?³⁷ To think so is to make two mistakes. First, insofar as Nour and Charles are able to reliably apply the concepts *racist* and *classist*, respectively, it is because they have learned those concepts (directly or indirectly) from counter-hegemonic political communities. What is more, their reliable sensitivity to racism and classism, respectively, is also due to their particular social positions – as an Arab woman and a working class man, respectively – which in turn are given, not chosen. Second, neither Charles nor Nour *has* achieved something like a full proletarian standpoint. To carry on with Trotsky’s image, they have begun to disentangle themselves from the web of the social lie, but they are still largely its captives. Charles does not yet have full revolutionary class consciousness, and nor does Nour have full race consciousness. Indeed, one can easily imagine Charles and Nour losing the epistemic gains they have made, their initial confidence shaken by further misleading testimony or ideological re-entrenchment. What they need is more help from the outside: others with whom, together, they can articulate a coherent and stable worldview that can withstand the onslaught of the dominant ideology.

One feature of the externalist treatment of Charles and Nour is that it can achieve a balance between two rival, attractive thoughts. On one hand, we have Marx’s recognition that the position of the oppressed can afford a dispensation from some of the epistemic ills of the oppressors. And on the other, we have the thought – on which much emphasis has been put in recent years by analytic philosophers³⁸ – that oppression can deprive subjects of epistemic goods like justification and knowledge. A balance between these two thoughts might be achieved by recognising the ways in which the knowledge afforded by oppressed social positions might yet fail to be ideal forms of knowledge: by recognising, for example, that Nour would be better off if she were aware of her own reliable sensitivity to racism, or that Charles would be better off if he were able to explain away the

to believe in accordance with their evidence.) In any case, I am not trying to offer a general theory of when one’s false belief is excused by ideological circumstances; rather, my point is simply that externalists are able to make sense of beliefs that fall short of justification but are nonetheless blameless. Thanks to an anonymous referee for pushing me to address this point.

³⁷ With thanks to an anonymous editor for discussion of this point.

³⁸ e.g. Fricker 2007, Stanley 2015. Cf. Medina 2013.

Master's misleading testimony as a product of false consciousness, and that both would be better off still if they had worked out worldviews to rival the dominant ideology. The externalist verdict that Nour and Charles nonetheless know something of their social reality allows us to ward off scepticism about the social world while still duly noting the epistemic costs of oppression.³⁹ This, I want to suggest, is part of why the externalist verdict is attractive in these cases: just as externalism allows us to know empirical truths about the external world despite the spectre of our envatted brain doppelgängers, externalism allows us to know about the social world despite the spectre of bad ideology.

I have here suggested a diagnosis as to why the externalist verdicts are more intuitive in my new cases than they are in the analogous, more familiar cases. My diagnosis – that in cases featuring bad ideology we are more prone to think in structural rather than individualistic terms, and thus more pulled towards externalistic rather than internalistic notions of justification – is, to repeat, merely psychological. It does not, by itself, constitute an argument in favour of externalism or against internalism. Or rather, it does not so constitute an argument on the presumption of metaepistemological realism, according to which there is simply a fact of the matter as to whether internalism or externalism is the correct theory of justification. In §4, I will leverage my diagnosis to offer an argument for externalism and against internalism, one that will be congenial to the realist. But it is worth saying that, for a certain sort of epistemologist, my diagnosis already takes us a long way towards such an argument.

On a *pragmatist* approach to metaepistemology (Greco forthcoming and Shafer 2014), the internalism/externalism debate is substantive to the extent that the two views have different practical cash values. On an *ameliorative* view of

³⁹ A similar approach is taken by Fredric Jameson in his treatment of conspiracy theory, which he calls 'the poor person's cognitive mapping in the postmodern age...a degraded figure of the total logic of late capital, a desperate attempt to represent the latter's system' (1988b, 356.) Jameson's point is that conspiracy theorising is at best a second-best attempt on the part of the poor to come to terms, epistemically speaking, with an oppressive capitalist system: a 'degraded' and 'desperate' figure of their reality, but a representational one nonetheless. Conspiracy is neither an ideal form of representation – Jameson says it is 'marked by its slippage into sheer theme and content' (ibid), variations on worn clichés that preclude genuine analysis – but nor is it a mere symptom of late capitalism.

epistemology (Haslanger 2000), the choice between internalist and externalist notions of justification should be guided by the question: which view would be most morally and/or politically useful? Both metaepistemological approaches would take the fact that externalism allows us to explain agents' justificatory statuses in structural terms as potentially speaking in favour of externalism over internalism. For it is eminently plausible that the ability to count members of oppressed groups such as Charles and Nour as knowers, and oppressed people like Radha as being robbed of justification by bad ideology, speaks *practically* in favour of externalism. The dispute between internalism and externalism will not only turn out to be substantive, but will moreover turn out to be a dispute between epistemologies of different political value.⁴⁰

Before moving on, let me say, to those who are left unsatisfied by my diagnosis, that my overall challenge to internalism does not crucially hang on it. Even if I am wrong about why it is that our intuitions shift across the old and new cases, my cases still present a *prima facie* problem for internalism. For, insofar as the argument for internalism rests on the intuitive support it receives from cases such as CLAIRVOYANT, DOGMATIST and BRAIN-IN-A-VAT, the internalist must find something to say in response to my new, apparently analogous cases. The most obvious thing for an internalist to say is that my cases are not, despite appearances, analogous. How the internalist might argue for this claim is the topic of the next section.

⁴⁰ My thanks to [omitted] for discussion of these issues.

3. In search of a disanalogy

Perhaps my cases are relevantly disanalogous to the traditional ones, differing in a way that makes a difference for justification. If so, then the internalist could *accept* the verdicts that Radha is unjustified, and that Nour and Charles are justified, without having to abandon her internalism. My challenge would be thereby disarmed.

3.1 RACIST DINNER TABLE and CLAIRVOYANT⁴¹

Clairvoyant Norman isn't justified, according to internalists, because there is nothing introspectively available to him that could plausibly serve as his grounds for believing the president is in New York. He does not have anything like a quasi-perceptual vision of the president's being in New York; at best he has, BonJour says, a 'hunch'.⁴² But in Nour's case, internalists might protest, there is something that serves as her introspectively available grounds: for Nour *experiences* her host being racist. Why is it not this experience that, for the internalist, justifies her belief?

It is true that Nour experiences her host: she sees him, hears him, speaks with him. But does she experience her host being racist? As I have described the case, Nour does not have a conscious experience as of her host being racist: she enjoys no 'my-host-is-racist' phenomenology. Nevertheless, the internalist might press on, does Nour not experience the subtle behaviours – the verbal and physical cues – on which her subconscious racism-detecting mechanism picks up? Does she not thus have *evidence* that her host is racist, albeit evidence that she subconsciously rather than consciously processes?

Certainly, there is a version of RACIST DINNER TABLE in which this is precisely what happens. But in the version of the case that I have described, Nour does not have conscious experience of those features of the host's behaviour that trigger her subconscious racism-detecting mechanism. It is not that she hears a

⁴¹ I am indebted to an anonymous referee for pressing the following objection.

⁴² Presumably the internalist does not think that mere hunches are sufficient to ground justification. If they were, then everyone who believed they were going to win the lottery would be justified in so believing.

certain inflection in the host's voice, and sees certain fleeting micro-expressions, which in turn are processed by her subconscious, ultimately delivering the (consciously available) verdict that her host is racist. Rather, the detection of her host's subtle behaviours – the 'seeing' and 'hearing' – is itself subconscious, or what psychologists call 'preattentive'. Preattentive or subconscious processing of perceptual information, especially information related to environment threats⁴³ and emotions⁴⁴, is a broadly recognised phenomenon, though not an uncontroversial one.⁴⁵ Whether one could have the sort of subconscious processing of perceptual information that I ascribe to Nour's is of course a further issue, one that I cannot possibly resolve. But even if it turns out that it is not possible for humans to subconsciously process subtle cues of racism, this just puts Nour in the same boat as Norman, who enjoys a power that no human actually has.

Let me offer one further brief reply to this internalist line of defence. Suppose it really is a conscious experience – of vocal hesitations, flickerings of the eyes, etc. – that is then subconsciously processed by Nour, which in turn results in her belief that her host is racist. Now imagine Nour*, who has an internally identical experience to Nour, but whose host is just a bit socially awkward, and not at all racist. Is Nour*'s belief that her host is racist justified? My intuition here, for what it is worth, is that she is not. If this is right, then it means that even in this version of the case – where Nour does have some conscious experiences that could potentially serve as the grounds for belief – it is not these conscious experiences, but something else, that renders Nour's belief justified. The externalist offers us an easy answer as to what that something else might be: namely, Nour's reliable connection to the truth.

3.2 CLASSIST COLLEGE and DOGMATIST

Dogmatist Mary truly believes – based on her reliable colour perception, which she has reason to believe is unreliable – that the sculpture is red. Similarly, Charles truly believes – based on his reliable judgment, which he has reason to believe is

⁴³ See e.g. Gray 1982 and 1995.

⁴⁴ See e.g. Balconi and Mazza 2009.

⁴⁵ See e.g. Block and Phillips 2016.

unreliable – that the college is classist. The internalist might protest that the cases are nonetheless disanalogous. For while Mary believes a proposition that is only contingently true (the sculpture is red), Charles believes a proposition that is true by necessity. For the college – holding fixed the non-normative facts about it – is classist in all metaphysically possible worlds.

Why might this difference matter? Because one might think that beliefs in necessary truths are *immune from defeat*: that they cannot lose their justification through the acquisition of misleading first-order evidence (to the effect they are false) or through the acquisition of misleading higher-order evidence (to the effect that they are based on unreliable, unsafe or otherwise faulty mechanisms). Nagel, for examples, argues that beliefs in very simple logical and mathematical truths cannot be defeated by either first-order or higher-order evidence, because it is inconceivable that they are false (1996, 62ff). Because one cannot meaningfully entertain, Nagel says, the sceptical possibility that one's judgment is false or based on an unreliable mechanism, defeat of such beliefs is impossible.⁴⁶ But Nagel concedes that this isn't the case with all necessary truths; in particular, he thinks that most ethical truths are such that we can conceive the possibility that our judgment about them is unreliable.⁴⁷ Presumably, *that the college is classist* is one

⁴⁶ Nagel's argument glosses over an important distinction in how a belief might be unreliable or unsafe. On a standard safety account, S's belief that *p* is justified just in case it is based on a method that does not produce a false belief that *p* in a sufficiently nearby world. A belief in a necessary truth trivially satisfies this test because there is no world – and *a fortiori* no sufficiently nearby world – in which S's belief that *p* is false. But a refined safety account – one meant to apply to necessary as well as contingent truths – will say that S's belief that *p* is justified just in case it is based on a method that does not produce a sufficiently nearby *false belief*. The refined account implies that S's belief that *p* is unjustified if S believes *not-p* in a sufficiently nearby world. So even if Nagel is right that one cannot meaningfully entertain the proposition that *2 plus 2 doesn't equal 5*, so long as one can entertain the proposition that one *could have easily believed that 2 plus 2 doesn't equal 4*, one's belief could be still defeated by higher-order evidence. But I leave aside this complication in the following. I take it that, insofar as the Master's testimony is (higher-order) defeating, it is because it constitutes evidence that Charles could easily falsely believe that the college is classist, not that he could easily falsely believe that the college is not classist.

⁴⁷ In such cases Nagel says we must weigh 'the plausibility of the debunking explanation against the plausibility of the ethical reasoning at which it is aimed' (ibid 115). Nagel's formulation here is somewhat misleading, since one could have high credences in both an

such proposition: it does not enjoy the self-evidentiary quality of simple mathematical and logical propositions. So Nagel's account of why certain beliefs in necessary truths enjoy immunity from defeat will not serve the internalist here. Nor will Plantinga's insistence that what he calls 'properly basic' beliefs are immune from higher-order defeat, since Plantinga includes both perceptual beliefs and ethical beliefs in this category.⁴⁸ Of course, one can simply *insist* that beliefs in all necessary truths – or just all necessary normative truths – are immune from defeat.⁴⁹ This would be to draw a sharp disanalogy between CLASSIST COLLEGE and DOGMATIST, one that an internalist could leverage to explain why it is that, consistent with her internalism, Mary's belief loses its justification but Charles' belief remains justified. But absent a convincing explanation as to why Charles' belief is so immune, this response is suspiciously ad hoc.

3.3 DOMESTIC VIOLENCE and BRAIN-IN-A-VAT

Jane_{BIV} has an internal duplicate who knows that she has hands: normally embodied Jane. But Radha, it seems, does *not* have an internal duplicate who knows that she deserves to be beaten. Why might this disanalogy matter? It matters, the internalist might argue, because it suggests that – contrary to what I have claimed – Radha's evidence does not in fact give internalistic support to her belief that she deserves to be beaten. If Radha's evidence is not metaphysically compatible with a possible world in which she deserves to be beaten, how can it be that it supports her being in such a world? If this is right, the internalist has a ready explanation for why Radha is not justified: her belief is not (unlike with Jane_{BIV}) supported by her evidence.⁵⁰

ethical proposition and the proposition that one's ethical reasoning is unreliable. This would put one in a position of *epistemic akrasia* – a position that is not necessarily irrational. On epistemic akrasia see Coates (2012), Greco (2014), Horowitz (2014) and Sliwa and Horowitz (2015).

⁴⁸ Plantinga 1986, 311.

⁴⁹ One might, for example, insist that we should assign credence 1 to all necessary truths. But this feature of standard Bayesian accounting is a notoriously unrealistic, and thus unattractive, way of thinking about rational belief in necessary truths/falsehoods.

⁵⁰ I am indebted to [omitted] and an anonymous referee for raising this objection.

Let us grant for the sake of argument that there is no metaphysically possible world in which Radha, fixing her mental states, deserves to be beaten.⁵¹ Why should this imply that Radha's evidence cannot give internalistic support for the belief that she deserves to be beaten? Imagine an early modern scientist before the discovery of the chemical composition of water. Despite his assiduous laboratory procedures, a sample of his water gets accidentally contaminated with xyz; as a result, tests on this sample issue in results incompatible with the sample being (pure) H₂O. Does the scientist thereby receive misleading evidence that water is not H₂O? Surely the answer is yes. But there is no metaphysically possible world in which water is not H₂O. So it would follow, from the principle that evidence cannot internalistically justify a metaphysically impossible proposition, that the scientist cannot get evidence that water is not H₂O.

Something has gone awry. Our mistake was to think that relations of evidential support are constrained by metaphysical possibility. Bodies of evidence can be *evidentially compatible* with (and *evidentially supportive* of) propositions with which they are not metaphysically compatible. What matters for evidential compatibility is not metaphysical possibility but *conceivability*. It is conceivable that water is not H₂O, even though water is necessarily H₂O; thus it is possible to get misleading evidence that water is not H₂O. Likewise, it is conceivable that Radha deserves to be beaten – indeed, Radha is just one of many who believes that disobedient wives deserve to be beaten – so it is possible for Radha to get misleading evidence to the effect that she deserves to be beaten. This is in spite of the fact that Radha has no metaphysically possible internal duplicate who knows she deserves to be beaten.

An alternative way for the internalist to press a disanalogy between Radha and the brain-in-a-vat is to insist that Radha in fact *is* in a position to know that she doesn't deserve to be beaten. For – a certain sort of internalist might say – all (minimally competent) humans are capable of knowing certain moral truths (e.g. that wife-beating is wrong) through *a priori* moral reflection. If Radha believes she deserves to be beaten, this just goes to show that she has not adequately reflected on the question. In this way Radha is importantly unlike the brain-in-a-vat, for whom no amount of assiduous reflection will reveal the truth that she doesn't have hands. Such an internalist could *agree* with the externalist that Radha's belief that

⁵¹ As an anonymous editor points out, epistemic internalism goes naturally hand-in-hand with an ethical view on which internal duplicates can't differ in their moral properties.

she deserved to be beaten is unjustified, but for reasons quite different from those given by the externalist.⁵² DOMESTIC VIOLENCE would then be no threat to internalism.

I will call this the ‘strong rationalist’ response. A strong rationalist does not merely insist that some moral truths are *a priori* knowable. A strong rationalist insists that certain moral truths are *a priori knowable* to every minimally competent agent. (Just because some very complex mathematical truth is *a priori* knowable doesn’t mean that it’s knowable for every competent agent.)⁵³ Is such a view plausible?⁵⁴ To be a strong rationalist is to effectively deny the possibility of DOMESTIC VIOLENCE.⁵⁵ For I stipulated that Radha thoroughly reflects on her situation and the moral issues at stake, and nonetheless – precisely because she has so thoroughly internalised the patriarchal ideology that surrounds her – is unable to see that she does not deserve to be beaten. The strong rationalist simply denies this possibility: he thinks that the truth that one does not deserve to be beaten by one’s husband is always yielded, to anyone, by proper *a priori* reflection, regardless of her epistemic situation.⁵⁶

⁵² Thanks to both [omitted] and [omitted] for raising this objection.

⁵³ I’m grateful to [omitted] for this point.

⁵⁴ For a defence of strong moral rationalism, see Harman (2011, pp 460ff and m.s. p 22). Julia Markovits defends a strong form of moral rationalism in her (2014), though elsewhere she appears to hold back from endorsing the thesis that all minimally competent agents, regardless of their enculturation, can come to know the necessary ethical truths. In ‘Acting for the Right Reasons’ (2010), Markovits writes: ‘Agents can act from the motive of duty only if their moral reasoning is good...How good our moral reasoning is will depend on many factors that are beyond our control, including the quality of our moral education and...the culture in which we live’ (212). For defences of strong rationalism about the truths of rationality – what we might call ‘strong rationality rationalism’ – see Titelbaum (2015 and m.s., section 4) and Littlejohn (2018).

⁵⁵ Or to deny that Radha’s belief that she deserves to be beaten is false. See Titelbaum m.s., section 4, for a parallel dialectic.

⁵⁶ It’s uncontroversial that such a woman could come to know that she doesn’t deserve to be beaten by joining a feminist consciousness-raising group or reading some Catharine MacKinnon. But the rationalist must say something much stronger: that such a woman, without changing her epistemic situation – without gaining new experiences, new testimony, or new evidence – could come to know that she doesn’t deserve to be beaten. The rationalist might reply by drawing on a distinction between *justifiers* and *enablers*.

I for one find strong rationalism implausible. It seems to me plainly possible for there to be someone like Radha, who grasps the relevant concepts, reflects as thoroughly as she can, and still comes to the false belief that she deserves to be beaten. Indeed a case like DOMESTIC VIOLENCE seems to me quite pedestrian, and no less conceivable than, say, BRAIN-IN-A-VAT. (I don't think there are any real-world brains-in-vats, but I do think there are quite a few real-world Radhas.) Insisting on the impossibility of DOMESTIC VIOLENCE strikes me, then, as a large bullet for the internalist to bite. That said, for the internalist willing to do so, the threat represented by DOMESTIC VIOLENCE can be neutralised. Just how many internalists will find this an attractive option remains an open question.

4. In search of an error theory

The diagnosis I offered in §2 as to why our intuitions differ across the traditional and new sets of cases suggests a promising avenue for an internalist error theory of the intuitions elicited by my cases: political confounds. Specifically, the internalist might argue that our intuitions in the 'bad ideology' cases are unreliable *precisely* because they are cases about bad ideology. Perhaps it just feels too

While Radha has, in her current situation, all that is epistemically required for justifying the belief that she doesn't deserve to be beaten (*viz.* her capacity for moral reason), she lacks the external conditions that would *enable* her to exercise that capacity. On this view, an activity like consciousness-raising would not directly change Radha's epistemic situation, but would instead serve a role analogous to that served by pen and paper in completing a math problem. Does such a response help the internalist who wants to explain why Radha is unjustified? I'm not sure it does. Suppose Robert is trying to solve a difficult math problem. He carefully and assiduously reasons his way through it, multiple times, and consistently gets an answer that appears to him to certainly follow, but which is in fact wrong. Robert is told by his math teacher and his peers that the answer is in fact right. Further suppose that Robert *would* (unbeknownst to him) have got the right answer had he been able to use pen and paper, which weren't available. Is Robert's belief in the wrong answer internalistically justified? I strongly suspect that the internalist will want to say it is. If so, the absence of enablers – pen and paper, consciousness-raising – doesn't suffice to render false beliefs internalistically unjustified. So this sort of strong rationalism won't help the internalist explain DOMESTIC VIOLENCE. Thanks to [omitted] for prompting me to address this point.

politically unsavoury to say that a woman can justifiably believe that she deserves to be beaten, or that victims of racism or classism can be unjustified in believing themselves to be such victims. For political reasons we want to resist the internalist verdicts, and side with the externalist. But this is to allow – the objection goes – our political judgments to contaminate our epistemological ones.

But it's not at all clear that we can do our epistemological theorising free of ethical or political confounds. After all, it is precisely the thought that it would be *unfair* to withhold justification from the brain-in-a-vat, or *unfair* to confer justification on Norman the clairvoyant, that motivates many to endorse internalism. Indeed Jennifer Nagel et al found that people's willingness to attribute justification to protagonists in sceptical cases correlated positively with measures of empathy.⁵⁷ Of course, two wrongs don't make a right. Just because the internalist might depend on ethical confounds for her case doesn't mean the externalist should follow suit.⁵⁸ So let us consider a case in which the protagonist *truly* believes something that those who share my leftist political sensibilities would think false:⁵⁹

ABORTION: Thomas is a young man who has grown up in a community gripped by bad leftist ideology. All his life he has been surrounded by pro-choice peers and adults, and exposed only to pro-choice arguments. But Thomas cannot shake the conviction that abortion is, despite what everyone says, wrong. His elders and peers tell him that this must be a manifestation

⁵⁷ Nagel et al 2013s

⁵⁸ Thanks to [omitted] for pushing me on this point.

⁵⁹ Since moral truths are presumably necessary truths, if (as I think) abortion is morally permissible, then it is necessarily the case that abortion is morally permissible. Thus to entertain a case in which abortion is morally impermissible, as in ABORTION below, is to imagine a counter-possible case. I take it that such a world is at least conceivable. Alternatively, one could refashion the case to involve a contingent but nonetheless politically-loaded belief, e.g. about anthropogenic climate change. Thus one could imagine a world (non-actual but certainly possible) in which anthropogenic climate change was nothing but a liberal conspiracy, and in which a protagonist – through careful first-hand evaluation of the evidence and in the face of overwhelming testimony and gaslighting – formed a reliable true belief that anthropogenic climate change was a myth. For my part I have the intuition that such a belief would be justified, though I am certain of the reality of anthropogenic climate change.

of his deep-seated hatred of women. But in fact it's the result of his genuine sensitivity to the moral truth, namely the truth that abortion really is categorically wrong. Despite all the misleading evidence from his peers and elders, Thomas cleaves to his belief that abortion is wrong.

Is Thomas' belief that abortion is wrong justified? I think the answer is surely yes. If abortion really were wrong, and it really were just a matter of bad ideology that made leftists think otherwise, then an individual who was genuinely sensitive to the badness of abortion, and formed a belief in its badness on such a basis, would – analogously with Charles in *CLASSIST COLLEGE* – be justified. And yet the confound error theory would predict that my intuitions wouldn't go this way, because saying that a young man would be justified in his belief that abortion is wrong grates against my political sensibilities. (It does grate. But I feel the intuition powerfully nonetheless.) *ABORTION* suggests that what is doing the work in the bad ideology cases isn't leftist political sympathies – after all, there is nothing in the notion of bad ideology that is inherently leftist – but the politically-neutral fact that these cases involve subjects forming beliefs under bad ideological conditions.

A different way of pressing the worry about moral/political confounds is to argue that my cases invite a conflation of epistemic justification with moral justification. On this diagnosis, the reason we intuitively want to withhold justification from Radha is because we are wary of endorsing the claim that she would be justified in acting on the belief that she deserves to be beaten. Likewise, we are intuitively inclined to say that Nour and Charles are epistemically justified because we would want to say that their acting on their beliefs would be justified. But this, the internalist might argue, is to conflate epistemic and practical justification: Radha's belief is justified even if she wouldn't be justified in acting on it, and Nour and Charles would be justified in acting on their epistemically unjustified beliefs. Insofar as these internalistic verdicts are counterintuitive, it's because we are conflating two distinct kinds of justification.

This takes us to a deep question that I cannot hope to resolve here, namely the relationship between epistemic and practical normativity. In a recent back-and-forth on the *a priori*, Paul Boghossian and Timothy Williamson come up against just this issue.⁶⁰ Williamson asks us to imagine a perfectly coherent Nazi. Is he

⁶⁰ Boghossian and Williamson (forthcoming).

justified, Williamson wants to know, in believing that homosexuals should be killed? Boghossian, an internalist, assimilates the Nazi to the brain-in-a-vat, and concludes that the Nazi is ‘fully justified’, though of course massively in error, and says that he endorses this seemingly counterintuitive conclusion ‘wholeheartedly’ (ibid). Williamson concedes that one might worry that ‘there is something cheap about using the Nazi’ as a counterexample to internalism, since the ‘toxic nature of Nazism might prompt a moralizing attitude in which we automatically reject any positive description of the Nazi’s beliefs whatsoever’. But this, he goes on, is to miss the crucial connection between justified belief and justified action. If the Nazi is fully justified in believing that he should kill homosexuals, then why – Williamson asks – isn’t he morally justified in killing them?^{61,62}

Williamson, in other words, endorses a tight connection between practical and epistemic justification, which in turn makes reflection on morally and politically charged cases particularly useful for thinking about epistemic justification. On a different view, however, epistemic justification does not enjoy a close connection with moral justification. This is the sort of view that an internalist like Boghossian, who wants to insist that the Nazi is epistemically but not morally justified, should endorse. (Williamson, for his part, says that this sort of view ‘raises a question about the interest of justification in the epistemological sense, once it is severed from justification in the more general normative sense’.) As I said above, I do not hope to resolve this thorny issue here. But I do hope to have shown how the objection to the politically charged nature of my cases presupposes a not uncontroversial view about the relationship of epistemic to practical normativity.

Moreover, there is something to be said for the thought that the political nature of my cases underscores rather than impugns the trustworthiness of our intuitions about them. For my cases are far less *recherché* than their traditional analogues. Subjects operating under conditions of misogyny, racism and classism are a commonplace (or so I think) in the actual world, while brains-in-vats, clairvoyants and epistemology-hoaxing artists are presumably more modally

⁶¹ Like Williamson, I endorse an externalism on which epistemic justification entails truth. Otherwise, there could be cases in which one justifiably believed a false moral truth, and so (assuming a Williamsonian connection between epistemic and practical normativity) would be justified in acting on a false moral claim.

⁶² For a more detailed defence of the use of morally loaded cases in epistemology, see Williamson (forthcoming).

remote. This in turn suggests that if any intuitions should be thrown out as less reliable, it should be those we have in response to the traditional cases, rather than the new, bad ideology cases. The diagnosis I offered in §2, in other words, not only does not impugn the reliability of our intuitions about ‘bad ideology’ cases. It also casts doubt on the trustworthiness of our intuitions about the traditional, humdrum cases we know and love. The challenge to internalism stands.

5. Extending the challenge: mental state internalism⁶³

Mental state internalists claim that justification supervenes on an agent’s mental states, whether those states are accessible to the agent or not. On this view, what we might call ‘total internal duplicates’ – that is, two subjects who are identical in *all* their non-factive mental states – cannot differ in justification. A counterexample to mental state internalism is thus a case in which a subject is justified, but has a total internal duplicate that is not justified.

Are RACIST DINNER TABLE, CLASSIST COLLEGE and DOMESTIC VIOLENCE counterexamples to mental state internalism? The question is straightforward with regard to CLASSIST COLLEGE and DOMESTIC VIOLENCE. If these are counterexamples to access internalism, as I have argued they are, they are also counterexamples to mental state internalism. For neither Charles nor Radha has a non-accessible mental state that would explain why they would differ in justification from a duplicate who shares their accessible mental states. Intuitively, Charles is justified, even while he has a total internal duplicate who is unjustified; and Radha is unjustified, even while she has a total internal duplicate who is justified. So these are potential counterexamples to not just access internalism, but internalism *simpliciter*.

RACIST DINNER TABLE is a slightly more vexed issue. For Nour *does* have a set of non-accessible mental states: namely, the preconscious awareness of her host’s behaviour. Does Nour have a total internal duplicate whose belief that her host is racist is unjustified? Imagine Sarah, who shares all of Nour’s subconscious mental states: she subconsciously ‘sees’ and ‘hears’ subtle facial and verbal behaviours in her host. Except that Sarah’s subconscious is not picking up on

⁶³ With thanks to an anonymous referee and [omitted] for prompting me to consider such an extension.

actual facial and verbal behaviours, but fabricating them: she is subject to a kind of subconscious hallucination, through no fault of her own. (One might object that this surely means that Nour and Sarah have different subconscious mental states; but recall that mental state internalists think that internal duplicates must share only *non-factive* mental states.) Is Sarah's belief that the host is racist justified? I think the answer is, quite plausibly, not. If so, then Nour does have a total internal duplicate who differs in justification – meaning that RACIST DINNER TABLE is a counterexample to mental state as well as access internalism. My challenge thus generalises to internalism *simpliciter*.

6. Conclusion: externalism as a normative epistemology

It's often said that externalism is not a 'normative' epistemology. This is said by both internalists and externalists. Internalists mean it as a criticism: any epistemology worth having must be a normative epistemology, so externalism isn't worth having. Externalists mean it as an explanation: they simply aren't in the game of normative theorising, which is why their theory implies that blameless people (like Jane_{BIV}) can be unjustified and irresponsible people (like Norman and Mary) can be justified. That externalism is not a 'normative' theory is also taken to explain why externalist theories are not 'action-guiding', in the sense of being operationalizable by agents under conditions of uncertainty or ignorance. By contrast, an internalist epistemology, which articulates its justification-norms in terms of the subject's own perspective, can always be used to guide one's own epistemic actions, and neatly tracks whether the subject is blameworthy or blameless.⁶⁴ Internalism treats believers as normative agents, we say, while externalism treats them like mere thermometers.⁶⁵

All this assumes that what it is to be a normative theory is to talk in terms that are familiar to us from ethics: blameworthiness and blamelessness, responsibility, action-guidance. But we might think this an overly restrictive notion of a normative theory (in ethics as well as epistemology). Marxism, for example, is arguably a normative theory, in the sense that it is responsive to the gap between how things are and how things should be. But Marx was uninterested

⁶⁴ Though see [omitted] for an argument that no theory can satisfy this demand.

⁶⁵ A metaphor that originates with Armstrong (1973).

in the questions of what any given individual ought to do or who is to be blamed – concerns that he dismissed as typically bourgeois. Aristotle meanwhile thought the question of whether something is a *good* version of its kind – whether a citizen is a good citizen, or whether a thermometer is a good thermometer – was *the* paradigmatic normative question. And yet Aristotle thought that being a good version of one’s kind is not something that lies solely within the will of that thing. Virtue requires being embedded in a cooperative world. For Aristotle this dependency of the normative goods on ‘external’ facts appears to be a natural consequence of humans’ social existence: the things most worth having are those we can only have through our relationships with others. So too, the externalist thinks, with epistemic goods: the epistemic goods really worth having are those that cannot be had by mere individual effort. It is in this sense that externalism is, or can be, a normative epistemology.

One way of reading this paper is as I’ve presented it – as a straightforward argument against epistemic internalism and, *ipso facto*, for externalism. On this reading, my cases present a challenge to the internalist, one that demands that she say something about why these cases do not threaten her view of things. I would be happy to hear from internalists how this challenge might be met. But I would be just as happy – happier even – to have this paper read as an explanation of why, for at least some of us, externalism is attractive as a genuinely normative theory of justification and knowledge. Insofar as one thinks, as I do, that we live in a world suffused with bad ideology – insofar, we might say, that one’s view of the world is a *radical* one – an epistemology that is capable of operating in terms of structural notions becomes more attractive than one that can trade only in individualistic ones. One might be an externalist, in other words, not in *spite* of externalism’s detachment from the individualistic normative notions we hold dear, but precisely because of it.

Acknowledgments

[omitted]

Works Cited

- Alston, William. (1988). "An Internalist Externalism". *Synthese*, 74: 265–283.
- Austin, J.L. (1956). "A Plea for Excuses". *Proceedings of the Aristotelian Society*, 57: 1-30.
- Armstrong, David. (1973). *Belief, Truth and Knowledge*. Cambridge: Cambridge University Press.
- Balconi, Michela and Mazza, Guido. (2009). 'Consciousness and emotion: ERP modulation and attentive vs. pre-attentive elaboration of emotional facial expressions by backward masking'. *Motivation and Emotion* 33(2): 113-124.
- Bergmann, Michael. (2006). *Justification Without Awareness*. Oxford: Clarendon Press.
- Block Ned and Phillips, Ian. (2016). 'Unconscious seeing—a debate'. In B. Nanay (Ed.), *Current Controversies in Philosophy of Perception*. New York: Routledge.
- Boghossian, Paul and Williamson, Timothy. (forthcoming). *Debating the A Priori*. Oxford: OUP.
- BonJour, Laurence (1980). "Externalist Theories of Empirical Knowledge" *Midwest Studies in Philosophy*, 5: 53–73.
- Cassell, Lisa. (forthcoming). 'Higher-Order Beliefs and the Undermining Problem for Bayesianism'. *Act Analytica*.
- Coates, Allen. 2012. 'Rational Epistemic Akrasia'. *American Philosophical Quarterly* 49(2): 113-24.
- Cohen, Stewart. (1984). 'Justification and Truth,' *Philosophical Studies* 46: 279-296.

Collins, Patricia Hill. 2000. *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Routledge: New York.

Conee, Earl and Feldman, Richard. (2001). "Internalism Defended". *American Philosophical Quarterly*, 38(1): 1-18.

Chisholm, Roderick. (1966). *Theory of Knowledge*. Englewood Cliffs, N.J.: Prentice-Hall.

Christensen, David. (1992). 'Confirmation holism and Bayesian epistemology'. *Philosophy of Science*, 59(4), 540– 557.

Fricker, Miranda. (2007). *Epistemic Injustice*. Oxford: OUP.

Goldman, Alvin. (1967). 'A Causal Theory of Knowing'. *The Journal of Philosophy* 64(12): 357-372.

Goldman, A. (1979). 'What is Justified Belief?' in G.S. Pappas (ed.), *Justification and Knowledge*. Dordrecht: Reidel, 1-25.

Goldman, A. (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.

Goldman, A. (1988). 'Strong and Weak Justification'. *Philosophical Perspectives* 2: 51-69.

Gray, Jeffrey. (1982). 'Précis of the neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system'. *Behavioral and Brain Sciences*, 5, 469–543.

Gray, Jeffrey. (1995). 'The contents of consciousness: A neuropsychological conjecture'. *Behavioral and Brain Sciences*, 18, 659–722.

Greco, Daniel. 2014. 'A puzzle about epistemic akrasia'. *Philosophical Studies* 167(2): 201-219.

Greco, Daniel. (forthcoming). 'Verbal Debates in Epistemology'. *American Philosophical Quarterly*.

Grice, Paul. (2001). *Aspects of Reason*. Oxford: Clarendon Press.

Harding, Sandra. (1983). 'Rethinking Standpoint Epistemology: What Is "Strong

Objectivity”?’ L Alcoff and E Potter, eds. *Feminist Epistemologies*. Routledge: New York, 49-82.

Harding, Sandra. (2004). ‘A Socially Relevant Philosophy of Science? Resources from Standpoint Theory’s Controversiality’. *Hypatia* 19(1): 25-479

Harman, Elizabeth. (2011). ‘Does Moral Ignorance Exculpate?’. *Ratio* XXIV: 443-468

Harman, Elizabeth. (m.s.) ‘Ethics is Hard! What Follows?’

Hartsock, Nancy. (1983). “The Feminist Standpoint: Developing the Ground for a Specifically Feminist Historical Materialism”, in Sandra Harding and Merrill B. Hintikka (eds.), *Discovering Reality*, D. Reidel Publishing Company.

Haslanger, Sally. (2000) “Gender and Race: (What) Are They? (What) Do We Want Them to Be?” *Nous* 34(1): 31-55.

Horowitz, Sophie. 2014. ‘Epistemic Akrasia’. *Nous* 48(4): 718-744.

Huemer, Michael. 2001. *Skepticism and the Veil of Perception*. Lanham: Rowman & Littlefield.

Jameson, Fredric. 1988a. ‘*History and Class Consciousness* as an “Unfinished Project”’. *Rethinking Marxism* 1 (1): 49-72.

Jameson, Fredric. 1988b. “Cognitive Mapping”, in Cary Nelson and Lawrence Grossberg, eds. *Marxism and the Interpretation of Culture*. London: Macmillan.

Lasonen-Aarnio, Maria. (2010). “Unreasonable Knowledge”. *Philosophical Perspectives* 24: 1-21.

Littlejohn, Clayton. (2018). ‘Stop making sense? On a puzzle about rationality’. *Philosophy and Phenomenological Research* 96: 257-72.

Lukács, Georg. (1923/1971). *History and Class Consciousness*. Trans. Rodney Livingstone. MIT Press: Cambridge.

Markovits, Julia. (2010). ‘Acting for the Right Reasons,’ *Philosophical Review*, 119(2): 201-242.

- Markovits, Julia. (2014). *Moral Reason*. Oxford: OUP.
- Marx, Karl. (1867/1887), *Capital: A Critique of Political Economy, Vol. 1*. S. Moore and E. Aveling, trans. F. Engels, ed. Progress Publishers: Moscow.
- Medina, Jose. (2013). *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations*. Oxford: OUP.
- Nagel, Jennifer, San Juan, Valerie and Mar, Raymond. (2013). "Lay Denial of Knowledge for Justified True Beliefs" *Cognition* 129: 652-661.
- Nagel, Thomas. (1996). *The Last Word*. Oxford: Oxford University Press.
- Nozick, Robert. (1981). *Philosophical Explanations*. Cambridge, MA: Harvard University Press.
- Plantinga, Alvin. (1986). "The Foundations of Theism: A Reply," *Faith and Philosophy*, 3(3): 310-312.
- Pollock, John & Cruz, Joseph. (1999). *Contemporary Theories of Knowledge*, 2nd edition, Lanham, MD: Rowman & Littlefield.
- Pollock, John. 1995. *Cognitive Carpentry: A Blueprint for How to Build a Person*, Cambridge, MA: MIT Press.
- Pritchard, Duncan. 2005. *Epistemic Luck*. Oxford: OUP.
- Pryor, Jim. 2000. "The Skeptic and the Dogmatist". *Nous* 34(4): 517-549
- Schafer, Karl. 2014. 'Doxastic Planning and Epistemic Internalism'. *Syntheses* 191(12): 2571-2591.
- Sliwa, Paulina and Horowitz, Sophie. 2015. 'Respecting All the Evidence'. *Philosophical Studies* 172(11):2835-2858.
- Sosa, Ernest. 1996. 'Postscript to 'Proper Functionalism and Virtue and Epistemology'', in J.L. Kvanvig (ed.), *Warrant in Contemporary Epistemology*. Lanham, MD: Rowman and Littlefield, 271-280.

- Sosa, Ernest. (1999). "How to defeat opposite to Moore". *Noûs*, 33: 141-153.
- Sosa, Ernest. 2000. 'Skepticism and Contextualism', *Philosophical Issues* 10: 1-18.
- Sosa, Ernest. (2009). *Apt Belief and Reflective Knowledge*, Volume I and II, Oxford: OUP.
- Stanley, Jason. (2015). *How Propaganda Works*. Princeton: PUP.
- Titelbaum, Michael (2015). 'Rationality's Fixed Point (Or: In Defense of Right Reason)'. *Oxford Studies in Epistemology* 5: 253-294.
- Titelbaum, Michael (m.s.). 'Return to Reason'.
- Wedgwood, Ralph. (2002). "Internalism Explained". *Philosophy and Phenomenological Research*, 65(2): 349-369.
- Weisberg, Jonathan. (2009). 'Commutativity or holism? A dilemma for conditionalizers'. *British Journal for the Philosophy of Science* 60(4): 793-812.
- Weisberg, Jonathan. (2015). 'Updating, undermining and independence'. *British Journal for the Philosophy of Science* 66: 121-159.
- Williamson, Timothy. (2000). *Knowledge and Its Limits*. Oxford: OUP.
- Williamson, Timothy. (2007). 'On Being Justified in One's Head', in Mark Timmons, John Greco and Alfred Mele (eds.), *Rationality and the Good: Critical Essays on the Ethics and Epistemology of Robert Audi*. Oxford: OUP, 106-122.
- Williamson, Timothy. (2014). "Very Improbable Knowing". *Erkenntnis* 79: 971.
- Williamson, Timothy. (forthcoming). 'Morally Loaded Cases in Philosophy'. *Proceedings and Addresses of the American Philosophical Association*.
- Wilson, Edmund. (2003). *To the Finland Station*. New York: The New York Review of Books.
- Wylie, Alison. 2003. 'Why Standpoint Matters', in *Science and Other Cultures: Issues in Philosophies of Science and Technology*, edited by Robert Figueroa and Sandra Harding. Routledge: New York, 26-48.