

Why externalism?

I want to know what I, and others, should believe. The “should” here refers to a should of a certain kind. We sometimes worry about what is morally right for us to believe, or what beliefs are in our best interests, but typically when one asks what one should believe, one wants to know what one *epistemically* should believe. Just what this means is the subject of some debate, but there’s a decent amount of agreement that the epistemic should has something to do with the end (or goal or value) of having true beliefs and avoiding false ones.¹ To figure out what it is that we should believe (epistemically speaking), we need to figure out what it is that makes a belief a belief that one should have. Another way of putting this is that we need to figure out what it is that justifies beliefs, or what constitutes an epistemic reason to believe something. Theories of justification can be divided into two groups: internalist and externalist. This paper will argue that some externalist theory of justification is better than any internalist theory. I won’t advocate for any particular flavor of externalism, although what I will say on the behalf of externalism generally will suggest constraints on what a good externalist theory of justification will look like.

Internalist accounts of justification say that whatever can justify A’s belief that *p* must be something that is cognitively accessible to A, something that A can be directly aware of. This doesn’t require that A realize that his or her reasons for believing *p* are reasons for believing *p*.

1 A competing account of the epistemic should is that it has to do with achieving *empirically adequate beliefs*, which are beliefs that won’t be shown false by experience. In my opinion, while it may be plausible that science pursues empirically adequate beliefs, it isn’t plausible that belief about ordinary things, or belief in philosophical claims, aims at empirical adequacy. Ordinarily, one wants to believe things about the external world that are true, not just ones that you won’t discover are false. Similarly, one wants to know if God really exists, or the actual nature of moral rightness, not just arrive at a theory that won’t get falsified. I’m interested in what justifies philosophical beliefs, and I think that we can learn about this by studying the justification of ordinary beliefs. So I will assume that the epistemic should has to do with the end of achieving true, and avoiding false, beliefs.

For example, I know some mathematical axioms, and I know I know them, and it might turn out that these axioms entail the truth of Goldbach's conjecture; however, I might not realize that these axioms are evidence for the conjecture because I'm not smart enough to do the proofs. Even so, an internalist could count my beliefs about these axioms as reasons to believe the conjecture, because they are things I am aware of.

Externalist accounts, on the other hand, say that, for some agents *A* and propositions *p*, what justifies *A*'s belief that *p* is something that *A* is not directly aware of. Note that externalists don't say that *all* reasons are out of our awareness. Plausible externalist accounts will want to say that some of our beliefs are inferentially justified. Even so, externalists leave open the possibility that some of the reasons that justify our beliefs are things that are outside of our cognitive possession. While I won't advocate any particular externalist account in this paper, I'll use *simple process reliabilism* as an illustration. Simple process reliabilism is the view that (for any agent *A* and proposition *p*) *A*'s belief that *p* is justified if the belief is produced by a reliable process.² This is an externalist account because it is possible that *A* might not know or even have any evidence that their belief is produced by a reliable process; the claim is that it is the reliability of the process that justifies the belief, and not any thing that is in *A*'s cognitive grasp.

The initial attractions of externalism

What might attract us about externalism? The question we started with is, what makes a belief justified, given the end of believing what is true and not what is false. It would be nice if whatever it is that justifies belief in this sense is something that makes beliefs more likely to be true. This would make it quite straightforward why we ought to have beliefs that have properties which make them justified, and why we ought to avoid beliefs which lack these properties. We

² Goldman is the originator of the view. Later versions of reliabilism have become more sophisticated, saying, e.g. that a belief must be produced by a reliable process and undefeated in order to be justified.

can capture this idea with the slogan, “Justification should be truth conducive.” With that slogan in mind, we can see a problem with internalist theories. On internalist theories, what justifies belief is something that is cognitively accessible to the believer, something in the believer’s mind. But there’s nothing about our mental states themselves that guarantees that what is in our minds matches the world. In fact, it’s not clear that there’s anything about mental states themselves that makes it *more likely than not* that they match the world. That’s not to say that our mental states are or even could be systematically mistaken.³ We may live in a world, and inhabit bodies, such that our experiences tend to reflect reality (I’m not crazy, so I think we do live in such a world and in such bodies). But what makes it the case that our experiences tend to be correct is not the experiences themselves, but rather something about things outside of our heads, something about the causal relationship between us and the world. That is to say that it’s not anything directly accessible to us that makes our mental states truth conducive. If that’s right, then whatever it is that justifies belief on any internalist theory must be something that is not, itself, truth conducive. It’s just, in some sense, a lucky coincidence that our justified beliefs tend to be true. Internalist justification can’t live up to the slogan that justification should be truth conducive.

Externalist justification can, however. The externalist allows what justifies our beliefs to be stuff that is not directly accessible to us. This allows makes possible theories that say that there is something necessarily truth-conducive about justified beliefs. Process reliabilism is a nice example of this. Since what justifies belief, on that theory, is the reliability of the process that produced it, we are guaranteed that justified beliefs will tend to be true. Note that the mere fact that a theory is externalist doesn’t guarantee that its notion of justification is truth conducive.

³ When I say “could be” here I am speaking of nomological possibility – what is possible given the laws of nature. It’s obviously logically possible that our experiences are systematically mistaken.

There is stuff that is not accessible to us that is also not truth-conducive. It's rather that any theory according to which justification is truth conducive *must* be externalist, and not that all externalist theories *must* say that justification is truth conducive.

This may seem like a fine, and not very interesting, difference. Almost everyone thinks that our sensory experiences are generally a good guide to the truth. Sensory experience (and the beliefs based on it) is available to us, and is the sort of thing that internalists tend to see as justifying belief. So, whether or not what internalists see as justifying beliefs is *necessarily* truth conducive, we all think that it's *actually* truth conducive. Why isn't that good enough?

It's not good enough because, in my opinion, the interesting goal of epistemology is not really to work out whether or how our empirical beliefs are justified. There is a surfeit of theories of justification according to which (assuming there actually is no evil demon) our empirical beliefs are justified, and (again on that assumption) everyone agrees that our empirical beliefs are typically true. If all we wanted out of a theory of justification was that it didn't entail skepticism about the external world, there'd be nothing to make us choose between these theories, other than a need to publish to keep our jobs. But that isn't all there is to epistemology (in my opinion). What we really are interested in is working out how some more difficult class of beliefs is justified (I'm interested, for example, in philosophical beliefs). So, perhaps we can dismiss worries about truth-conduciveness of what justifies empirical beliefs as merely skeptical worries. But similar worries about, let's say, religious or moral beliefs are not merely skeptical. This is because we have much better reasons to suspect that what is in our minds does not reflect religious or moral reality than we do to suspect that what is in our minds doesn't reflect empirical reality. It's simple to see this: almost no one systematically disagrees about chairs or colors, but there is a ton of systematic disagreement about religion, politics, and so forth. It's

clearly harder to get things right in these realms, and so we can't be so blasé about the truth-conduciveness of reasons; we want an account of justification that makes it fairly certain that justified beliefs about *anything* will tend to be true. Some externalist theories give us that, but no internalist theory can. This is the attraction of externalism.

The problem with externalism

What's, then, the problem with externalist theories? Why does anyone lean towards internalism? The advantage of externalism – that it allows us to say that a connection to reality is what justifies our belief, making justification necessarily truth-conducive – is also the disadvantage. Since, on externalist theories our beliefs can be justified by things the believer has no cognitive access to, a believer cannot always tell when a belief is or is not justified. To illustrate, consider again process reliabilism. Let's say that Fred is systematically deceived by an evil demon. However, his set of experiences is indistinguishable from yours, so as far as he can tell he's doing a great job forming beliefs. Since none of his beliefs are reliably formed, none of them would be justified according to the process reliabilist. But what could he have done differently? Nothing, because he has no way of telling that he's done wrong. Conversely, imagine that Sam has completely reliable ESP, even though he has no evidence whatsoever that ESP is possible. If he believes everything his ESP tells him, without confirming any of it, then according to process reliabilism Sam is forming justified beliefs. But he's got no idea – he can't tell that he *should* be trusting his ESP.

This is a problem. It's widely accepted that we should be able to use normative truths (truths about what we should or shouldn't do) to guide us in the right direction. If we can't tell that we've done or believed wrongly, then we can't guide our actions in the correct direction. We can't intentionally do a good job, we can't intentionally avoid doing a bad job. How can we

make sense of shoulds that we can't intentionally live up to? Let's use the slogan, "Justification should be action guiding," to capture the notion that we must be able to intentionally act so as to conform to the advice given by epistemic shoulds. The worry about externalist theories is that, if they were correct, some epistemic shoulds would not be action guiding. This would be because some of our reasons would be outside of our cognitive access, and so we couldn't always tell when there were or were not justifying reasons for our beliefs. Shoulds that cannot be acted on aren't really shoulds. That's a problem for externalism.

In the rest of the paper I will argue that, for some form of externalism, justification *can* be action guiding, given a wide enough notion of whose actions shoulds are supposed to guide. I'll argue that this wider notion is the correct notion. In order to argue this about externalism and epistemic justification, I'll illustrate it by talking about another type of should.

Legal shoulds and action guidingness

Sometimes, legally speaking, we should do x , and sometimes, legally speaking, we should not. Put another way, some actions are legally permissible and some actions are legally wrong. What makes an action legally permissible or legally wrong? The law does (whether it is legislator made, judge made, or made by regulators). In order to understand a contentious question about one type of should, it can be helpful to think about a simpler kind of should. We know quite a lot about legal shoulds, and we can use this to shed light on epistemic shoulds.

Are legal shoulds always action guiding? Yes and no. There are possible situations in which it is legally wrong for A to do x at time T, yet A cannot avoid x -ing at T. Because of this, there will be situations in which A cannot act on a legal should that applies to him or her; in such cases, A's actions are governed by a should, but A's actions cannot be guided by that should. For example, it used to be illegal to not pay one's debts when they were due. It was illegal

whether or not one had the money to pay the debts in question. So, if Sam owed Shepard \$10,000, due on Tuesday, but Sam only had \$0 on Tuesday, Sam would break the law even though it was impossible for Sam to comply with the law. The law gave Sam no advice on Tuesday that Sam could follow. Yet this doesn't mean that the law no longer applies to Sam: when he fails to pay his debt, he's done legal wrong.

So, in some sense, legal shoulds are not action guiding. Yet I imagine that many of you are thinking right now that Sam *could* have avoided this problem: he could have avoided borrowing the money from Shepard. So, in some sense, the law *can* guide his actions. I agree, and I'll say a bit more about that. Imagine a society that had only this one law – that said only that people must pay their debts when they come due. Now, Sam is considering taking a loan from Shepard, a loan that Sam knows he will not be able to pay back. Does Sam have legal reasons to not take this loan? Yes he does, even though there is no law about *taking out* loans. Where do those legal reasons come from? They come from the fact that, if he takes this loan, he will end up doing legal wrong. The wrong that Sam wouldn't be able to avoid in the future gives him advice now, and that advice now is, "Don't take this loan." Let's imagine that Sam takes this loan and fails to pay. Now he's on trial for violating the one law. Does the judge have legal reasons to make any particular judgments? Sure: the judge has legal reasons to find Sam guilty, reasons, even though there is no law that says the judge must. Where do these reasons come from? They come from the legal wrong that Sam has done in the past.⁴

So, we have a law that, when it applies to Sam, will give him advice that he can't follow (the law only applies to Sam once the loan comes due, and at that point he won't be able to

4 You might think that the judge isn't legally obligated to convict Sam, because there is no law saying that the judge must do so. Perhaps that's right. Even so, I think it would be bizarre to say that the judge has no legal *reasons* to convict Sam, even if he lacks a legal obligation to do so. If I'm right, then these reasons must come from somewhere, and the only place they can come from is the one law that Sam broke.

follow it). Yet the law can still guide one's actions; it still gives one legal reasons that one can act on. It advises Sam before it applies to him, saying, "Don't get in a position where you will break me." It advises other people after the fact, saying, "Treat the person who has broken me in certain ways." It might give other sorts of advice, as well. Perhaps, once Sam has broken the law, he has reasons to try to rectify the situation, reasons that stem from the fact that he has done wrong. Or perhaps other people have reasons to not assist Sam in breaking the law, reasons that come just from the fact that Sam's behavior will be illegal if he does it. If so, then the law gives people action guiding reasons, *even though they are not in a position to break it*.

What is the lesson we learn from this discussion of legal shoulds? The lesson is this: there are two ways that a should can be action guiding. The fact that doing x is wrong can guide your actions by telling you to not do x when you are in a position to do so. And the fact that doing x is wrong can guide the actions of people who are not in a position to do x . Note that the wrongness of x can be action guiding even if those in a position to x can't tell that x is wrong, as long as someone is in a position to act on the reasons given by the wrongness of x . If Sam doesn't know the law, and doesn't realize that it is illegal to not paying Shepard, the illegality of Sam's actions can still give reasons to act to others, such as judges.

Back to epistemic shoulds

How does this apply to epistemology? Well, let's say that it is epistemically wrong for Jill to believe that p , but Jill can't tell that this particular belief is wrong for her to have. So, she can't be guided in forming beliefs about p by this wrongness. However, this wrongness can still be action guiding. Jill has reasons to avoid situations where she's likely to form wrong beliefs, even if she won't be able to tell in those situations that she's doing wrong. Imagine, for example, that Jill knows that when she gets drunk, she forms false beliefs, although when she's

drunk she can't tell that they are false or that she's drunk. Were these beliefs epistemically wrong to form, she would have epistemic reasons to avoid getting drunk. The wrongness of beliefs can also give third parties reasons to act. If Esther knows that Jill is forming impermissible beliefs, but doesn't know whether these beliefs are true or false, then Esther knows that she should trust Jill less.

Externalist accounts of justification say that A's justification for p might arise from stuff that is outside of A's cognitive grasp. Thus, at the moment where A believes that p , A won't always be able to tell that A's belief is justified. Thus, we won't always be able to act on our justification: when we believe wrongly (according to the externalists), we won't always be able to realize this and change. But this wouldn't mean that justification is not action guiding. On at least some versions of externalism, the wrongness of a belief will be able to guide to belief formation of third parties, or the inquiry of agents who might form these unjustified beliefs. This can be true even if agents will not be able to tell, at the moment they form unjustified beliefs, that they are unjustified. We will be able to tell sometimes that certain situations lead to wrong beliefs, or that others have believed wrongly. Knowing this will give us epistemic reasons to believe or act in certain ways. So, even if we won't be able to tell, at the moment we form beliefs, that they are wrong to form, we can be guided in other ways. This is enough to make externalist shoulds action guiding enough.

Externalist accounts of justification can make justification truth-conducive. They can also make justification action guiding. Of course, the mere fact that a theory of justification is externalist won't guarantee that it gets us these two desiderata. But *some* externalist theory will, while no internalist theory can. That's why we should be externalists.⁵⁶

5 You can read this final claim in two ways. One is that we should use externalist epistemic shoulds, rather than internalist ones, when forming and evaluating beliefs. Another is that the true account of what we should believe

References

As I'm a bit rushed for time, I'll put a bibliography in later; note that *you* are not allowed to do this when you turn in a paper to *me*.

is externalist, not internalist. I slightly lean towards the first, because I'm open to the possibility that all (coherent) sets of norms exist and are true, so the real question in thinking about norms is which are the ones to use. But I'm not strongly committed to this. If it turned out false, then only the best account of epistemic norms would be the true account, in which case it would be some externalist account.

⁶ Here are two objections worth considering. First, if externalism is correct, then why doesn't the justification of a belief reduce to its truth (that a belief is true would be the best possible reason to believe it)? Second, when A is deciding at a moment whether or not to believe *p*, all A has to go on is their available evidence. Does this mean that our own evaluation of our own beliefs must always be according to internalist shoulds?