**Internalism and externalism about justification**

Theories of epistemic justification can be divided into two groups: internalist and externalist.

Internalist theories of justification say that whatever can justify A’s belief that \( p \) must be something that is internal to A. “Internal” means “inside A’s mind,” or “directly accessible by A’s mind.” For example, A’s experiences or memories are internal to A (they are in A’s mind), and internalists will say that these can potentially justify A’s beliefs. Internalism does not say that A can just make up justification for their beliefs, or that any beliefs that A thinks are justified automatically are justified. Rather, internalism says that if anything justifies A’s belief that \( p \), it is some mental state of A’s; but not every belief of A’s is justified, and not every mental state is a potential justifier for beliefs. So, for example, Joan might form some inferential belief that \( x \) is true, where the inference involves some complicated reasoning using conditionals. Every step in the inference process is in Joan’s mind, and she thinks every step is proper. But, let’s say that Joan does not fully understand how conditionals work, and so one of the steps in her inference is not logically appropriate. Joan’s belief about \( x \) is not justified. Further, internalists believe that justification can be defeated (as we’ve been discussing in class). So, even if some mental state can usually justify beliefs, that justification can be defeated in particular cases.

There are disagreements among internalists about many issues. For example, there is disagreement about what is means for something to be internal to A, or about which things internal to A can justify A’s beliefs (e.g. internalists might disagree about whether hearing testimony can by itself justify beliefs (i.e. in a basic way) or if inference is required). Another thing internalists discuss is whether something internal to A can justify A’s belief if A is incapable of recognizing it as a reason for that belief. For example, I know some mathematical axioms, and I know I know them. So these axioms are cognitive accessible to me (I have beliefs about them
inside my mind). Imagine that, as a matter of fact, these axioms entail some further, very complex, mathematical truth – e.g. Goldbach’s conjecture. However, I don’t realize that they do, because I’m not capable of doing the relevant proofs. Some internalists might say that I have reasons to believe in Goldbach’s conjecture, although I don’t realize that I have these reasons. After all, I have access to the relevant axioms, and the axioms are evidence for Goldbach’s conjecture, even though I don’t realize it; so the reasons to believe Goldbach’s conjecture are inside my grasp. However, other internalists might say that my beliefs in these axioms don’t give me reasons to believe Goldbach’s conjecture, because, while I can access the axioms, I can’t tell that they are reasons to believe Goldbach’s conjecture.

Externalist theories of epistemic justification say that, in some cases, what justifies A’s belief that \( p \) can be something that A is not internal to A. Note that externalists don’t say that all things that justify beliefs are external to us, just some of them. To illustrate externalism, consider a form of externalism called process reliabilism. Process reliabilism is the view that, if A’s belief that \( p \) is produced by a reliable process, then the belief is justified (unless that justification is defeated…). A reliable process is one that typically produces true beliefs. Note that this view does not require that A knows, or is able to know, that their belief is produced by a reliable process. That is what makes process reliabilism an externalist view: facts about A’s reliability are not internal to A’s mind. Instead, A’s reliability is about the relationship between A’s mind and the outside world. Process reliabilism would say that, if I have good vision, my beliefs that are based on vision are justified. But, if my vision is unreliable in some way that I am not aware of, then vision does not justify my beliefs.

Externalists can, and do, think that some aspects of justification have do with one’s internal mental states. I’ll give you an example, again using process reliabilism. Imagine that Marge is at the zoo. She sees an elephant. Her vision is reliable, so she is justified in believing
that there is an elephant. Bart tells her that she is hallucinating. Forming beliefs by trusting testimony is also a reliable belief forming process. So she is also justified in believing that she is hallucinating. These two beliefs – that there is an elephant, and that she is hallucinating – contradict each other. Most reliabilists will say that Marge’s justification for believing that there is an elephant is defeated. Notice that what defeats her justification is that she believes she is hallucinating; that is, her justification is defeated by some mental state of hers.

Externalists disagree about many things. For example, are external facts always relevant to justification, or only sometimes? For example, some externalists say that inferential beliefs are justified by mental states (an inferential belief gets its justification from the beliefs it is based on), and that is only basic beliefs that are justified by external facts. Another thing externalists disagree about is which external facts are relevant to justification. For example, reliabilists think it is facts about reliability. Something can be reliable without being 100% reliable. So, reliabilists think that some of our justified beliefs can be false. Other externalists think that justified beliefs have to be true (so, it is facts about the truth of \( p \) that justify beliefs about \( p \)).

Among reliabilists, there is disagreement about what it takes to be reliable. For example, reliability might vary from context to context. Let’s say a process is, on average, 70% reliable, because it is 90% reliable in one context, but only 50% reliable in another context. When we ask if it is reliable, do we use the average reliability, or the reliability in the particular context? If the latter, how specific does the context get? (The worry here is that, if we get too specific, then we have to ask, “Is this process reliable at forming this particular belief right now?” and we end up saying “Only true beliefs are justified,” and reliabilists don’t want to say that)

Internalism and externalism are views about justification, and not views about what is true or false, or about what makes things true or false. For example, both internalism and externalism are compatible with the idea that most truths are culturally relative, and also with the view that
most truths are objective. Combining externalism and culturally relativism, for example, we’d get the view that what is true varies from culture to culture, and that beliefs about what is true are only justified if they reliably identify what one’s culture says is true. For another example, combining internalism plus the idea that moral truths are objective, we’d get the view that moral claims are true whether or not anyone believes they are true, but that a person’s justification for believing moral truths depends just on what information that person has available to them. Further, internalism and externalism make no claims about how reliable or biased we are. Internalism can be true even if we are very unreliable, but so can externalism (if we are very unreliable, and externalism is true, then most of our beliefs would not be justified).

**Justification and the pursuit of truth**

Epistemic justification has something to do with the pursuit of truth. It would be nice if whatever it is that justifies belief is also something that makes beliefs more likely to be true. If that were so, then most of our justified beliefs would be true. With that 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. Hopefully our mental states do typically match the world. But they might not. For example, if we were dreaming right now, all of our experiences would be misleading. So, if internalism is true, there is no guarantee that justified beliefs will tend to be true beliefs.

Externalist theories, however, can guarantee that justified beliefs tend to be true. This is because, according to externalism, what justifies our belief can be partly outside of our minds, so externalist theories can make justification necessarily connected to external reality. Process reliabilism is a nice example of this. Since what justifies belief, according to process reliabilism, is the reliability of the process that produced it, we are guaranteed that justified beliefs will tend to
be true.

**Justification and guidance**

The advantage of externalism – that it allows us to say that a connection to reality is what justifies our belief – is partly a disadvantage as well. On externalist theories, 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 aliens, who have scooped his brain out of his head and hooked it up to a computer that produces convincing, but false, experiences. 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. One of the reasons we care about justification is that we want to use the notion of justification to guide our reasoning; we want to try to be reasonable, so we can be better thinkers. If we can’t tell that we’ve formed unjustified beliefs, then we can’t guide our reasoning in the correct direction. If externalism is true, then in *some* cases, people won’t be able to tell that they are forming justified or unjustified beliefs, and so they can’t use the notion of justification to properly guide their reasoning. This won’t be in every case – we often can tell that we are reliable or unreliable (e.g. if we constantly contradict ourselves, we are being unreliable). But we can’t always tell, and so, if externalism is true, we can’t always tell when we are justified in our beliefs.

One last thing to consider. Pretend externalism is true. Even if I can’t tell that my belief
about $p$ is justified or unjustified, you might be able to tell that it is. You might be able to detect that I am reliable even if I can’t. So, even if I can’t (in some case) use the notion of justification to tell when I am reasoning well or not, you might be able to tell when I am reasoning well or not. We often deploy the notion of justification not just to check our own beliefs, but to check others’ beliefs as well. That might make externalism more attractive in some ways. But it’s also worth noting that, when we check others’ mental-processes for reliability, we still do so from our own internal point of view, so we can’t completely escape this problem.

**Theory choice in epistemology**

The above considerations are a starting point for your thinking about justification. There are several questions to answer. Are some internalist theories better than others in producing true beliefs? (That is, are beliefs that are justified according to some theories more likely to be true than beliefs justified according to others?) Are some externalist theories more guiding than others? (That is, is it easier to tell when we are justified according to some externalist theories?) To what extent can we make a theory that is the best of both worlds – such that justified beliefs tend to be true, and also that we can guide ourselves towards justified beliefs?

If we can’t get an “ideal” theory, what should we give up – should we prefer a theory that is more aimed at truth, or that is more guiding? Finally, is the answer different in different contexts, or for different people? That is, might there be some topic $X$ such that the correct theory of what justifies beliefs about $X$ is internalist, and another topic $Y$ such that the correct theory of what justifies beliefs about $Y$ is externalist?

To answer these questions, you have to think about what is important about beliefs and belief formation. And you should try to start with easy cases – examples of beliefs or believers where you are very confident they are doing well or doing poorly – and use these to try to learn about the hard questions we’ve been discussing.
A quick note about objectivity and justification

One might wonder at this point whether there is any objective truth about what beliefs are and are not justified. That is a good thing to be wondering about. I am not going to say either way. But I will bring up some points that you should consider when you are thinking about this.

When you think about whether there are any objective truths about justification, you should first think a bit about how, for any \( x \), you can tell if there are, or are not, objective truths about \( x \). In philosophy, we always start with easy cases and see what we can learn about them, and then apply that to harder cases. So, let’s start with an easy case where we know that there are objective truths, and ask how we know there are objective truths.

We know that there are objective truths about tables and chairs and other relatively ordinary objects. How do we know? Through experience, testimony, memory, and certain kinds of inference. Now we ask: do experience, testimony, and inference indicate that there are objective truths about justification (memory is probably irrelevant to this issue)? It does seem to me that there are clear cases of beliefs that are obviously and objectively unreasonable. So my experience seems to tell me that there are at least some objective truths about justification. Further, other people seem to agree with me about some of these cases, so testimony is in agreement with memory. Finally, I seem to be able to draw inferences about what is and is not justified, although I haven’t completely worked out everything there is to know on the subject. So, there is at least some reason to think that there are objective truths about justification (or so it seems to me).

But that doesn’t settle the issue. You might be starting to get the sense that there is a lot of disagreement about what is justified or unjustified. After all, the whole point of this reading was that there is disagreement among philosophers about internalism and externalism. Views on
what is justified to some extent vary from time to time and society to society (e.g. with some societies putting more emphasis on trust in political leaders than others, or less trust in scientific reasoning). We know that people have biases in their thinking about what they should and should not believe (that’s the whole point of the Fricker reading). Individual people make inconsistent judgments about what is justified (that is, they say that x is justified sometimes, and then contradict themselves at other times). And, since we don’t completely know what justification is, we don’t know how it is that our experience could reliably detect it.

These are all good reasons to wonder about whether there is an objectively correct theory of justification. But do they show that there is not? Again, let’s consider a relatively easy case, a case that we know a lot about, and see what this can tell us about the hard thing we are wondering about.

For most of human history, people did not know what blueness is (it was not until relatively recently that we learned that color has to do with wavelengths of light). Nor did they have much of an idea of how we detected blueness, other than that we used our eyes somehow. For most of human history, we had no way of checking to see that our eyes were getting things right, other than by talking to other people. People disagree about what is blue. This is very common for “edge” cases of blue – cases that are somewhere between blue and some other color. But there is also a very significant rate of colorblindness, where people are bad at detecting “easy” cases of blue. Non-colorblind people also have systematic biases in their ability to detect blueness reliably; people make mistakes of all sorts about colors in different sorts of situations (e.g. in poor lighting). Color detection varies from culture to culture: most cultures through human history had no word for blue, and members of some cultures are systematically better at telling
I believe that there are objective truths about blueness. I suspect you agree with me. And I think people knew that there were objective truths about blueness even before modern science figured out what blueness is. This means that we should not be too quick to say there are no objective truths about justification. All of the things that make us worry about justification are also present for blueness. And we can give many different examples like blue, where the same worries arise but it is clear that we can still have knowledge of objective truths. (As an aside, we can have the exact same discussion about objective moral truths – there’s a lot of disagreement, and cultural variation, and bias in our beliefs about moral truths; but it is not clear that there is so much of these that we should reject objective moral truth)

What does this mean? You can still argue that there are no objective truths about justification (I think it is a view worth taking very seriously). But, to give the argument, you need a lot of compelling evidence. The sort of evidence I have just sketched is a start, but it is not enough by itself to settle the issue. Further, to argue that there are no objective truths about justification, you need to provide defeaters for the evidence that there are such truths. My point here is not to tell you that there is, or is not, an objectively correct theory of justification, but rather to teach you to think well about such questions, and to show you what you would have to do to answer them.

---

1 http://www.sciencealert.com/humans-couldn-t-even-see-the-colour-blue-until-modern-times-research-suggests