PHIL 640: Week Three- Williams

Brian Weatherson

September 20, 2021

Big Picture

Most writers we look at will fall neatly enough into one of two categories.

1. They take thick concepts seriously, i.e., they think that they are philosophically consequential, and one of those consequences is that they undermine philosophical theories which rely on (anything like) a fact/value distinction.
2. They think that thick concepts are at most an interesting bit of language to be explained away, and do not threaten the fact/value distinction that they rely on in their theories.

Most of the people in category 2 rely on a fact/value distinction either in their metaphysical picture, or their picture of mind and language. Occasionally there are people who have a fact/value distinction in their epistemology but not in the metaphysics or theory of content. (Possibly that's literally just me.)

Williams, despite being the person everyone cites as being the historically central figure in the discussion of thick concepts, doesn't fall neatly into either camp. He does have a fact/value distinction - or at least something very much like it. So you might think he'd be in category 2. But in fact he thinks thick concepts are very significant to philosophy - though in his case the significance is a distinct kind again. He thinks that thick concepts are particularly significant ethically.

But he also thinks that a kind of ethical relativism is true, and he does not think that any kind of scientific relativism is true. So he needs some kind of distinction. And when thick concepts come in, they actually seem like a kind of problem for his view - since they suggest that the kind of objectivism about ethics he disparages might be defensible.

The Distinction

I've been calling it roughly the fact/value distinction, but Williams spends a lot of ink on why that's not, in his opinion, the right way to conceive of it.

It's not the fact/value distinction because not all questions about value are on the right hand side of the distinction. In particular, a lot of questions in aesthetics are about value, but they are not (I think) on the right hand side of the distinction for Williams. He also thinks 'fact' isn't right for the left-hand side, but I couldn't really see why. It's possible that he's assuming some kind of minimalism about facts, so there can be ethical facts, it's just that they aren't scientific facts.

It's not the theoretical/practical distinction either. Not all questions about what to do are ethical questions. Children fighting over a cake are engaged in a practical dispute, but not an ethical one. (Though either may try to turn it into an ethical one by invoking principles of justice.) And I think Williams wants to say that not all factual questions are theoretical ones either. There is something funny about describing my wondering where my car keys as a theoretical inquiry, though I also don't think we go seriously wrong by this way of speaking.

What he does ultimately think is that it is a distinction between objective and non-objective fields. And the explanation is interestingly epistemological. A field is objective if we can expect that there will be convergence in people's opinions, and this convergence is explained by the opinions being correct. Or at least that's the initial theory - I'm not sure how consistently Williams sticks to it. But even this theory is worth spelling out a bit, and it's worth first taking a step back to something that is very familiar.

Williams's Background Theory of Knowledge

So as many of you know, contemporary epistemology starts with the publication of a three page paper in 1963: Edmund Gettier's "Is Justified True Belief Knowledge?" Gettier wanted to reject the follow class of theories of knowledge.

S knows that p just in case these three conditions are met:

1. S believes that p.
2. This belief is true.
3. This belief is 'justified.'

You can treat the third as a placeholder for any notion that assigns positive normative status to the belief. So we could just as easily say that the belief is sensible, or reasonable, or held for good reasons, or etc. And the worry is that cases where 2 and 3 are both true, but they are true for completely unrelated reasons, don't seem like knowledge. It wasn't immediately realised, but there are cases of this form, actually used for refuting roughly this theory, from a few places in the history of philosophy. The following two examples are both taken from Ichikawa and Steup (2013).

> Imagine that we are seeking water on a hot day. We suddenly see water,

> or so we think. In fact, we are not seeing water but a mirage, but when we reach the spot, we are lucky and find water right there under a rock. Can we say that we had genuine knowledge of water? The answer seems to be negative, for we were just lucky. (From the 8th century North Indian philosopher Dharmottara)
>
> Let it be assumed that Plato is next to you and you know him to be running, but you mistakenly believe that he is Socrates, so that you firmly believe that Socrates is running. However, let it be so that Socrates is in fact running in Rome; however, you do not know this. (From the 14th century Italian philosopher Peter of Mantua)

It is important in these cases to assume that the false belief is justified, reasonable, held for good reasons, whatever you think should go in 3. But it should be possible to have cases that are like that.

So after 1963 there were a flood of ideas about how to either add a fourth condition to this account of knowledge, or change the 3rd condition in a way that made it immune to these problems. The general idea was that we had to somehow link the 2nd and 3rd conditions, and I think something like this is basically right.

Williams endorses a specific version of this kind of strategy, the one defended by Robert Nozick in his *Philosophical Explanations* (1981). Simplifying a lot, Nozick wanted to replace the third condition with the following counterfactual.

- If p were not true, S would not have believed it.

This feels like it gets a lot of everyday cases right. We say when this condition is met that S's belief **tracks** the truth. And this is a good property for beliefs to have. Indeed, you can see a lot of safeguards we put on scientific experiments as aims to make sure the results of experiments, and hence of experimenters' beliefs, track the truth about the experimental situation.

But it can't be right as a theory of knowledge. A lot of problems were raised for it almost immediately by Saul Kripke, who naturally didn't publish these problems for three decades. (See Ichikawa and Steup (2013) for more about Kripke's lectures.) Here's one kind of case that seems to be devastating for this view. Imagine that I'm currently completely sober, but when I drink my judgment goes awry in all sorts of ways. Then the following claims could both be true.

1. I know that I'm sober enough to drive.
2. If I were not sober enough to drive, I would still believe that I'm sober enough to drive.

And there are many more problems where that comes from. While Nozick's theory was popular for a while, it is now as universally accepted as anything ever is accepted in

philosophy that it simply fails.[1]

Is it a problem that Williams relies on a false theory here? I don't think it is. Because while Nozick's theory fails, he doesn't need that theory to be true, and there are near enough theories that are true.

As he notes at a few times, Williams ultimately isn't relying on theories of knowledge here. He wants an account of what objectivity is, and he allows that people can have knowledge of the non-objective. So he could just drop the references to Nozick, or treat Nozick's view as at best an inspiration.

Or he could adopt the kind of theory that virtue epistemologists such as Ernest Sosa and Linda Zagzebski have endorsed. They think that the core extra condition is something like that what justifies the belief is in part the truth of the belief. We'll say more on this when we do more virtue epistemology.

Scientific Realism

I found the discussion of 'science' in chapter 8 very tough going, in part because I don't know what he even means by science. It's this very classically English philosophy of science that doesn't get its hands dirty by engaging with any actual scientific work. When you look at quantum physics, do you see one step of convergence after another? Or macro-economics?

Williams's theory is that we can accept a kind of realism in science because in science we have theorists whose beliefs converge, and the convergence is explained by their truth. It seems easy pickings for a dedicated scientific anti-realist, like say Fraassen (1980), to reply to this. Maybe we get convergence because theories do better and better at predicting the macroscopic phenomena. But since there are so many possible theories that predict the macroscopic phenomena, and since so many theories that approximately predicted macroscopic phenomena in the past have been rejected, that's no reason to believe in the truth of contemporary theories.

What is Objectivity for Williams?

Throughout chapter 8 there are a number of overlapping ideas that come up as being criteria for a field being objective. And I wanted to lay some of them out, because they didn't all seem to be the same to me.

Convergence on Reality It is reasonable to expect that people's beliefs will converge, and that what will explain the convergence is that the beliefs are true.

[1] See, there is philosophical progress!

This is the original, and I think official, definition of objectivity. As I said in the previous section, I think it's actually a fairly bold claim to think that this is true of science. But it isn't the only thing Williams puts forward.

Convergence on Justification It is reasonable to expect that people's beliefs will converge, and and that what will explain the convergence is the same thing as what justifies the people in having their belief.

This is closer to what we see when we get onto the stuff about thick concepts later in the chapter. It's how I interpret the slogan he repeats a few times "What explains is what justifies." And while it's similar to the first one, it isn't the same. Various anti-realist approaches to science, especially those driven by anti-metaphysical views, might think that scientists' views will converge, and that this convergence will be explained by the same phenomena (especially observational data) that justifies the scientists in their beliefs. But you can say all this while still thinking it not clear whether scientific theories are even the right kinds of things to be true.

Co-Tenability Either one of the convergence criteria are met, or the end state will not be convergence, but a partition of space into people with different views. But if the latter obtains, then there will be an external viewpoint from which any person in any part of the partition can view the other cells as correct from their standpoint.

I think, think, that this is what is going on in the color case. The discussion of color is very confusing. It reads to me like there is a small set of theories that are acceptable-in-Oxford-in-1985, and there is a tacit assumption somewhere that the disjunction of these is correct. There are various moves of the form *If not this, then that* which I couldn't make sense of really on their own (because 'this' and 'that' are defined in terms of those theories) but also couldn't see why they were the only alternatives.

But as best as I can make out, what's important to Williams is that the people involved can see each other as being part of a broader system, and the system justifies (and makes true?) their own beliefs, but also justifies (and makes true?) the beliefs of the others.

I'd like to say exactly how this is meant to work in the case of color, but I honestly can't tell precisely what theory of color Williams is working with here. I can't even really tell if he takes the canonical sentences to be things like *That's green*, or *Those two things are the same color*. These raise quite different issues. You probably can't use the word 'green' unless you have some acquaintance with our color practices. But you can say that things are the same color. And you can say that, or deny it, in cases where you have very different metamers to typical humans. (There is also the tricky fact that there is huge variation in color vision within humans, some of it surprisingly correlated with biological sex, so what 'our' practice is here is a bit elusive.)

There are two other things that get mentioned at various points, neither of which seem exactly like any of the three things said so far.

Explanation of Error An explanation of why most of us converge should also explain why some people make errors, and what those errors consist in.

Repeating Term in Explanation The explanation of convergence shouldn't change the subject, and could use the term whose convergence we're explaining.

The pattern analytic philosophers have of suspecting that if someone isn't saying precisely something, they aren't saying anything at all, is sometimes rather annoying. But the alternative, of saying allusive things but not quite sticking to your own terms, isn't fun either.

Aside on Mathematics

Is mathematics objective in Williams's sense? As I read the text, he's assuming that it must be, and that it's a constraint on a decent theory of objectivity that it turns out to be objective. And he's right to be scared of this, because it really isn't obvious on his theory that mathematics is objective.

There is a big question here - do we believe mathematical claims because they are true? This is hard to adjudicate because we normally understand those 'because' claims in terms of counterfactuals, and counterfactuals about math being false are literally incoherent. So there is a worry that mathematics won't be objective on Williams's view.

Williams suggests another way to be a realist (or objectivist) about math - all the parts of it could in principle be believed together. But I think this is true only on a particular view about what we're doing in math, what we might call 'if-then-ism.' On the face of it, it isn't true that you can believe all of mathematics at once. Different set theorists, for example, start with different axioms, and naturally prove different things. You can't believe all of them. You can only have Williams's attitude towards math if you think that these folks aren't really proving theorems are true, they are just proving that certain theorems follow from particular axioms. That's not an absurd theory of mathematics - some people endorse it - but it should be controversial. Like with the discussion of science, this reads like someone who never actually engaged with real cutting edge technical work, and is just going off what they learned in school.

Knowledge Involving Thick Concepts

Williams thinks that at least some people in cultures unlike ours can have knowledge involving thick concepts. Let's take a particular instance of this, involving the character Laertes from *Hamlet*. Hamlet himself is exactly the kind of person who in questioning those norms that hold society together undermines the confidence that Williams is so

keen on. But Laertes seems a pretty good 'un by Williams's lights I reckon. Or at least not a bad 'un, and that'll be enough. So think about (1).

1. Laertes knows that it is honorable to kill Hamlet to avenge his father's death.

Williams is interested in three challenges to this claim: one from tracking, one from truth, and one from reflection.

The tracking concern is relatively easy to deal with. Laertes' beliefs about honor do track something reasonably well. They aren't random. They do accord with the cultural norms of honor. They aren't good beliefs. They encourage revenge murder. As Ophelia points out, they lead to hypocritical treatment of men and women - Laertes can have all the affairs he wants off in Paris, but if she has an affair with Hamlet, that's dishonorable. But still, Laertes isn't being random or capricious here. We'll come back to whether conformity to culture is enough, but Laertes does seem to be tracking.

The truth concern is a little trickier. If you understand honor as meaning 'good in virtue of X' for some descriptive property X, then you will not think that 1 is true. It isn't good in virtue of anything for Laertes to kill Hamlet. And here Williams simply wants to deny that's how thick concepts work. They are action-guiding all right - Laertes kills Hamlet because that's the honorable thing to do. But they are more directly action-guiding. They don't tell you - oh this thing is honorable, so it's good, so do it! Rather, they just say - oh this thing is honorable, so do it! So Williams thinks at least this reason to think that Laertes' belief is false fails. But what does he think the truth conditions of Laertes' belief are? I'm not sure, and we'll come back to this.

The reflection concern is in two parts. First, Laertes could not hold on to this belief on critical reflection. Second, a belief is only knowledge if it could survive this kind of reflective process. I want to say more soon about reflection. Williams connects it to the Socrates, and we could think of it as being the attempt to survive the process of questions and answers that you see in the Socratic dialogues. But you could also think of it in terms of what Hamlet does - spending hours pacing around wondering about the nobility or whatever of various acts you're contemplating. Anyway, Williams thinks that you can know things, even if this knowledge wouldn't survive reflection. As he says, reflection can destroy knowledge. So the fact that Laertes is not reflective, and that he would have different beliefs if he were reflective, doesn't threaten his knowledge.

But even though 1 is true, Williams wants to reject 2.

2. It is an objective fact that it is honorable for Laertes to kill Hamlet to avenge his father's death.

Why? Well it fails most of the conditions for objectivity mentioned above. We can explain why so many people in the castle, and perhaps in the broader culture, agree about honor without invoking honor. The best explanation will be the social utility of

concepts like honor. (Or at least their utility to the powerful people.) So this won't be an objective fact.

But there is this further question about what those of us who reject the honor culture should say about what it is honorable for Laertes to do. Williams's own answer is hard to parse. He gives this super weird example involving schoolkids, and then hand waves about it being something vaguely like that. But the example doesn't make a whole lot of sense. I'm used to the idea that there are people who can mention but not use a particular charged term. (That's a view about slurs, though one that isn't as popular now as it was 10 years ago.) But what is going on with a term that some of us are only allowed to mention or use in indirect speech reports? Could we also use it in describing the beliefs that the schoolkids have? I don't get it.

And this is the super-hard question. There are two obvious options around here.

1. Say that honor-talk is evaluative, and since we reject the evaluations, we don't go in for that kind of talk at all - even to the extent of attributing knowledge to Laertes. I think that's Foot's view.
2. Say that honor-talk is purely descriptive, though it maybe implicates or something that we endorse an evaluative picture. So we won't use it unless there is a way to cancel that implicature. That's Pekka's view more or less.

Now Williams I think wants to slide between these two views. But I don't see the space there. This is surely more a reflection of my lack of insight than anything else. But I don't know what the positive picture is supposed to be.

Two Notions of Reflection

Williams talks a lot in these chapters about reflection. But as I read it, there are two different kinds of reflection, and they play different theoretical roles.

- One kind is reflection by theorists who are looking to explain the uses of a thick term, and especially to explain convergence in its use.
- Another is reflection by users of the term who are worried about whether they ought to be using the term.

The first kind of reflection plays a major role in chapter 8. It is because our reflections on honor-talk go a certain way that honor-talk is not objective in Williams's sense. The second kind of reflection plays a major role in chapter 9. It's what Hamlet does, and it's what leads to the terrible crisis of confidence that Williams is upset about at the end of the chapter.

Now it is possible that these don't go the same way, especially on some ways of solving the problems at the end of the last section. In particular, some ways of solving those problems will end up with us saying sure they converge on beliefs about honor because

those beliefs are true. Now what makes those beliefs true will not be remotely what they think makes those beliefs true. It will be something about social practices not something about true morality or whatever. But that's ok - as David Chalmers points out, our beliefs about tables, chairs and beer mugs aren't made true by what we think they are made true by.

The Hamlet-style reflection by users is a different story. (I'm maybe cheating a little bit here. Hamlet worries about, e.g., whether suicide is noble, not about whether nobility itself is good. But I think the way he does it really does undermine the very notion of nobility.) That does seem to require stepping back and using very different notions, in a way that maybe (maybe) reflection on our color talk does not.

But I was surprised at how quickly Williams moves to thinking that we'll have to use thin concepts to judge thick concepts. We could just use other thick concepts. And in fact we often do. We often ask whether the way we use certain terms or concepts is **just**. When Laertes sees the error of his ways, and perhaps of the honor code he was working under, he says that he is *justly* slain, not *rightly* slain. More recently, Miranda Fricker (2007) talks about judging epistemic concepts by the standards of epistemic **justice**. Maybe that's wrong, we should not focus on epistemic justice and injustice but, as Kristie Dotson suggests, epistemic **oppression**. There are lots of thick concepts we have around for use in judging other thick concepts.

Justice in general is a bit of a problem for Williams, and maybe we can talk about that more in class. It doesn't really fit into his discussion of either thin concepts or thick concepts, and this I suspect reveals a weakness in the whole structure.

Relativism

This has gone on a bit, but I want to end with a discussion of one brief argument in chapter 9 against a certain kind of relativism.

> Consider once more the hypertraditional society, and suppose that it does have some rules expressed in terms of something like "right" and "wrong." When it is first exposed to another culture and invited to reflect, it cannot suddenly discover that there is an implicit relativisation hidden in its language. It is will always be, so it speak, too early or too late for that. It is too early, when they have never reflected or thought of an alternative to "us." (A question from Chapter 7 applies here: how could this have come into their language?)

So this I think is badly mistaken, but for interesting reasons. I think there is an implicit two step argument here.

1. The locals are in control of at least the syntax of their language, of how many

argument places their predicates have.
2. If relativism is true, then terms like "right" and "wrong" have one more argument place than people thought they did.
3. So, relativism is wrong.

Premise 1 is fairly debatable - I think Williams is assuming some kind of internalism about the language, at least at some level. And that's controversial. You certainly don't want an individual level version of that principle, as we can see by thinking about time zones. But set worries about premise 1 aside, and focus on premise 2.

Think about the (very) ordinary sentence *It's raining*. Call it R. We know what it is for an utterance of R to be true. It has to be raining

- In the place of the utterance;
- At the time of the utterance;
- In the world of the utterance.

When I utter R on Christmas Day in Alice Springs, it doesn't matter whether it is or isn't raining somewhere else, or on some other day, or in some other possible world. As long as it is raining at that place, at that time, in that world, my utterance is true.

But this massively underdetermines what the content of R is. In general we want to say that a sentence is true if its content corresponds to the way things are. But when you have something like R, it is tricky to know what to put into the content, and what to put into the way things are. And in fact there is a choice to be made on each of the three bullet points.

Most theorists, in fact everyone I know, thinks that the place goes into the content. That is, an utterance of R expresses a content that has the place of the utterance as a constituent, and the truth of that content is measured against a 'place-neutral' world. The alternative would be to say that the content is just that rain is happening, and it is measured against a world that looks a bit like a map with a "You are here" icon, and the utterance is true if it is raining at the icon. This is coherent, but I don't think anyone thinks it. (In part because it yields weird predictions about embeddings.)

Most theorists, but not quite everyone I know, thinks that the world does not go into the content. An utterance of R does not express that it is raining (in Alice Springs or wherever) *in this world*. Rather, which possible world we're in is part of the facts that the content is measured against. We could go the other way around, and say that an utterance of R has a world-specific content. And there is some evidence for that approach. But mostly people think it's wrong.

And there is no consensus about how to handle time. *Eternalism*, in this context, is the view that the content of R includes the time of utterance, and that content is measured against a 'timeless' world. *Temporalism* is that the content is just that it is raining (at a

place), and the world it is measured against has an icon or something saying "This is now." And what matters is that it is raining (at a particular place) when that icon occurs. Among philosophers of language, Eternalism is the majority view, but Temporalism has a number of adherents, including some very prominent ones.

Why does this all matter? Because we can do the same trick about cultural relativity. Even if sentences like "It is honorable to kill people who kill your father" are culturally relative, we can ask whether the relativity goes in the content, or in the world they are measured against. And Williams's argument assumes, without any argument at all, that the relativity goes in the content. And this is probably the wrong form of relativism.

Well this isn't a philosophy of language seminar (really it isn't), so I won't keep going. If you're interested (or even remotely curious) you should read John MacFarlane's excellent book *Assessment Sensitivity*, which as well as going into great detail on this, will also suggest ways in which the setup I've given here is somewhat mistaken. But that's a debate for another, very different, class. For now I'll just note that there are a lot more options for relativism that are taken seriously now than there were when Williams was writing, and some of the arguments here fall short because of this restricted vantage.

Fraassen, Bas van. 1980. *The Scientific Image*. Oxford: Oxford University Press.

Fricker, Miranda. 2007. *Epistemic Injustice*. Oxford: Oxford University Press.

Ichikawa, Jonathan Jenkins, and Matthias Steup. 2013. "The Analysis of Knowledge." In *The Stanford Encyclopaedia of Philosophy*, edited by Edward N. Zalta, Fall 2013. Metaphysics Research Lab, Stanford University. http://plato.stanford.edu/archives/fall2013/entries/knowledge-analysis/.

Nozick, Robert. 1981. *Philosophical Explorations*. Cambridge, MA: Harvard University Press.