15

Reply to Williamson on Intuition, Understanding, and the A Priori[1]

Paul Boghossian

In his 'Reply to Boghossian on Intuition, Understanding, and the A Priori' (Chapter 14) Williamson takes issue with the two main claims I make in Chapter 13.

The first of these claims is that there are two types of understanding-based justification, the Constitutive type and the Basis type. Williamson doesn't dispute the aptness of this distinction; rather, he rejects that there are any instances of either type.

The second claim he disputes is that intuitions, understood as *sui generis* states of intellectual seeming, must play a role in any adequate epistemology of the a priori.

Constitutive Accounts

Let's begin with Williamson's discussion of the Constitutive type of understanding-based justification, according to which your assent to a proposition, p, is justified in virtue of the fact that your understanding of p is partly constituted by your disposition to assent to p.

> (Understanding–Assent Link, UAL)   To understand p, X must have an (underived) disposition to assent to p[2]

On this view, as it is often elaborated, your disposition to assent to p is part of the *conceptual role* of one of p's ingredient concepts.

In Chapter 13, I revisited a criticism of Williamson's well-known, expert-based recipe for generating counterexamples to any proposed UAL (see Williamson 2007a). I had previously hinted at this criticism in my 2012, but wanted here to develop it further.

According to a conceptual-role semantics, the dispositions to assent to p that are constitutive of understanding p are the dispositions that a thinker would find *primitively compelling*: that is, those that he would find compelling when considering p on its own and without the use of any collateral considerations.[3] In (UAL) above, I put this in terms of X possessing an *underived* disposition to assent to p; but it might be more revealing to use Peacocke's (1992) terminology: X's disposition to assent to p must be based on X's finding p primitively compelling.

In the case of *conjunction*, for example, a conceptual-role theorist might identify your understanding of *and* in part with your finding it primitively compelling to infer *A* from *A and B* (Conjunction Elimination, CE). She will not identify your understanding of *and* with dispositions that you might find compelling as a result of other background beliefs; otherwise, she won't end up with only those dispositions that are genuinely constitutive of mastery of *and*. (Similarly, she might identify your understanding of *all* in part with your finding it primitively compelling to infer from *All Fs are Gs* to *This F is a G*. And so on for the other constants.)

This appeal to inferences or truths that are primitively compelling poses a prima facie problem for Williamson's recipe for generating a counterexample to any proposed UAL. His recipe depends on imagining an expert on, for example, *conjunction* who develops theoretical misgivings about CE and, as a result, comes to reject it. However, an expert who comes to reject CE as a result of theoretical considerations is an expert who comes to find CE *derivatively* uncompelling; and that is, of course, perfectly consistent with her continuing to find it primitively compelling.

In his reply, Williamson says that he anticipated this objection and dealt with it:

In my previous critique, I consider actual or possible people who came to understand *p* in the normal way but then developed sophisticated theoretical doubts about *p*, on the basis of which they now dissent from *p*, while still understanding *p*. These people are experts; their understanding of *p* is non-deferential. Boghossian concedes that such cases can arise. However, he rightly notes, dispositions can be retained while inhibited. Handcuffs may inhibit a prisoner's disposition to hit out. Similarly, an expert may retain the disposition to assent to *p*, while inhibiting it on theoretical grounds. Examples are common. Thus the mere combination of understanding with refusal to assent does not refute Boghossian's constitutive account in its current form. So far so good.

Naturally, in writing my original critique, I was well aware of the possibility of inhibited dispositions to assent, and addressed its relevance explicitly (Williamson 2003: 254; 2007a: 99–105). In particular, I argued that the expert may gradually lose even the *disposition* to assent to *p*, while still retaining understanding of *p*. After all, many acquired dispositions are *habits*, which can be lost as well as gained over time. Since having the underived disposition to assent to *p* entails having the disposition to assent to *p*, the expert lacks the underived disposition to assent to *p* but still understands *p*, and thus is a direct counterexample to Boghossian's present constitutive account.   (Chapter 14, p. 209)

Williamson's reasoning here is not appropriately responsive to the point I was making. Let me see if I can explain myself better.

On Williamson's description of the case, the expert, let's call her 'Vanna', comes to have serious theoretical doubts about p and so stops assenting to p, while retaining the disposition to assent to p. Later, she gradually loses even the disposition to assent to p.[4] All the while she retains her mastery of p. Williamson takes the possibility of this scenario to be sufficient to secure the counterexample he is after; however, it is *not* sufficient.

The reason is that, in general, you can lose a disposition to assent to p in one of two ways. On the one hand, you can lose the disposition to assent to p because, *all things considered*, p no longer seems plausible to you; on the other hand, you can lose the disposition to assent to p because, *independently of any other considerations*, p no

longer seems plausible to you. (As we will see, Williamson's failure to draw this distinction will crop up again when we turn to the discussion of intuition later on in this chapter.)

To get a counterexample to the claim that CE is a UAL for *and*, you need a case where Vanna loses the disposition to assent to CE in the *second* of these two ways, without reliance on theoretical considerations, while retaining full mastery of *and*. Hence, merely emphasizing, as Williamson does, that Vanna loses not merely lose her practice of assenting to CE, but also her disposition to assent to CE, is not responsive to this point: it doesn't distinguish between the two ways in which Vanna might have come to lose her disposition to assent to CE, only one of which would constitute a counterexample to the view under consideration.

You might think, Well, what's the big problem? Why can't Williamson simply stipulate that, in addition to Vanna's coming to find CE objectionable on theoretical grounds, she also comes to find it utterly implausible when it is considered on its own and totally independently of the theoretical considerations she has developed?

Williamson is free to stipulate that, of course. What he isn't free to stipulate is that, when Vanna comes to find CE utterly implausible, when it is considered on its own and independently of any other considerations, she continues to retain full mastery of *and*, as opposed to mastery of some closely related concept.[5]

Williamson says,

The core of Boghossian's reply is his claim that once the expert loses her underived disposition to assent to *p*, it is no longer plausible that she still understands *p*. He treats her as if she had simply forgotten what proposition *p* is. He does not consider what the expert can still do, having lost the disposition to assent to *p*. She can still *remember* assenting to *p*; she knows just what it was like to exercise the disposition to do so. She can also *imagine* just what it is like to exercise the disposition, for others who still have it. She still expects *them* to assent to *p*. She feels to herself like someone who has painfully struggled free of the disposition to commit a common fallacy, or to reason according to the prejudices of the bigoted community in which she was reared. In these ways, her situation is very unlike that of 'an entirely unsophisticated person, a child, for example', to which Boghossian compares it. Unlike the child or unsophisticate, she can fluently converse with others when they assert *p*, perhaps explaining to them why, in her opinion, they are mistaken.   (Chapter 14, pp. 209–10)

Some aspects of Williamson's response here are a bit mystifying. For example, simply to assert, as he does, that the expert who has lost her underived disposition to assent to p can still enjoy a number of p-involving propositional attitudes is, in the present context, simply to beg the question. Whether a person who has come to find p primitively uncompelling is still a master of p is the very question at issue.

If we set these question-begging claims to one side, how clear is it that Williamson has offered us a counterexample of the required type? To see that he hasn't, let us unpack his example in terms of the relevant parameters.

Initially, Vanna will understand perfectly well why folks are tempted to believe that *Mary ate the apple* follows from *Mary ate the apple and the pear*. She herself will still understand that it is hard to see how Mary could have eaten *both* the apple *and* the pear, yet failed to have eaten the apple. However, she will try to explain to the folks who care that, while this inference may seem entirely non-optional when considered

on its own, it runs into trouble when it is applied in some remote theoretical domain, say, quantum mechanics.[6]

In the next stage, we are invited to imagine that Vanna comes to lose any sense even of initial plausibility for the claim that Mary's having eaten both the apple and the pear necessarily entails Mary's having eaten the apple. When she looks at this inference now, she feels no temptation to endorse it. She realizes, of course, that others are trained to make this inference; but she can no longer explain, to herself or to others, why this inference seems compelling.

Is it clear that later-stage Vanna, who can see no plausibility whatsoever in the claim that Mary's eating the apple and the pear entails Mary's eating the apple, still has the same concept of *and* that we do, as opposed to some closely related concept? It's certainly not clear to me.

If she were to say 'I completely understand why you are tempted by this inference, it's just that it runs into trouble when applied in quantum mechanics', that would be one thing. But when she says 'I have no idea why, apart from societal indoctrination, anyone would find this inference compelling', we have every right to question whether she is truly deploying the same concept as we are.

All these points are obscured by Williamson's stipulation that Vanna is an *expert* who understands *and* if anyone does. The stipulation is supposed to lull us into thinking that there is no way such a person could lose the concept *and*; anything she comes to think about *and* must be construed as an insight into *and* rather than a rejection of it. As Williamson is fond of saying, 'The last thing our expert needs is a language lesson.'

However, as I pointed out in my 2012, in the relevant sense, it is not insulting to Einstein's intelligence to point out to him that, in his Special Theory of Relativity, the word 'simultaneous' does not end up expressing the same concept (that of a two-place relation, rather than a three-place relation) that it did in the past.

The ineffectiveness of his example is obscured by Williamson's conflation of Vanna's finding CE primitively uncompelling with her losing the disposition to assent to CE. Since it is true that having an underived disposition to assent to CE entails having the disposition to assent to CE, and since it's clear that Vanna can come to lose the disposition to assent to CE, while still retaining mastery of *and*, it can look as though the example works.

However, as I've emphasized, Vanna can lose the disposition to assent to CE because she has come to find CE all-things-considered false, while still finding it primitively compelling; so merely securing that she has lost the disposition to assent to CE is not enough. We have to ask, Has Vanna lost the disposition to assent to CE because she finds it all things considered false, or because she finds it primitively false? If the former, it's irrelevant to its intended target; if the latter, it is unconvincing as a counterexample.

As I said in Chapter 13, to which Williamson is responding, if the expert's sophisticated considerations were really playing no role in persuading us that she continues to be a master of *and*, we ought to be able to describe a counterexample to CE in which such considerations are completely absent.

And the clearest way of doing that would be to describe someone with no sophisticated views about logic—a child, for example—*non-deferentially* acquiring

and for the first time, while refusing to endorse CE. Such a child would all along refuse to see any plausibility at all in its following from *Mary ate the apple and the pear* that *Mary ate the apple* while still clearly being a master of *and*.

This, however, Williamson has not come close to doing. And there is every reason to believe it can't be done.

Casalegno Revisited

In this respect, Paolo Casalegno's (2004) recipe for generating counterexamples to proposed UALs is more promising than Williamson's, since Casalegno tries to make it plausible that an unsophisticate, indeed someone suffering from an inferential disability, could clearly understand *and* without having so much as the ability to perform CE inferences. Williamson mentions Casalegno's recipe with approval and has defended it in his 2012 against the objections to it that I laid out in my 2012. Let me take this opportunity to continue the debate.

Here, once again, is Casalegno's recipe:

Apart from this difficulty, the idea that, given a logical constant *C*, there is a well-defined set R of rules of inference such that a subject cannot be regarded as knowing the meaning of *C* unless she accepts the rules in R is intrinsically problematic. No matter how a rule of inference is chosen, it seems to me that we can imagine situations in which we would be disposed to say that a subject knows the meaning of *C* although the subject does not accept the rule in question. Suppose Mary suffers from a cognitive disability which makes her completely incapable of performing anything which could be counted as an inference. Nevertheless, she is able to use logically complex sentences to describe visually presented scenes. (I do not know whether this kind of cognitive disability has ever been observed; but it is no doubt conceivable.) For example, we can imagine that Mary, although unable to perform conjunction-introductions and conjunction-eliminations, would be able to assert, in appropriate circumstances, 'The box is red *and* the book is blue', or things like that. It seems to me that it would then be possible to say that, in spite of her disability, Mary knows the meaning of the word 'and'. At least: my intuition is that we would spontaneously adopt a homophonic translation for sentences such as 'The box is red and the book is blue' uttered by Mary; and wouldn't that be a way of acknowledging that, in Mary's mouth, the word 'and' has the same meaning it [has] for us?

(2004: 407)

What Casalegno is claiming, in other words, is that, for any rule, R, and constant, C, we can cook up a counterexample to the claim that inferring according to R is required for possession of C, by imagining someone who lacks the ability to infer with R but who nonetheless understands C because she uses it competently in sentences used to describe visually presented scenes.

There are several difficulties with this strategy for refuting a conceptual-role semantics for the logical constants.

First, recall that we are talking about 'full mastery' of a logical constant, the sort of mastery characteristic of experts, and not mere 'possession', which we might accord someone even if they fell short of full mastery. It's hard to reconcile someone's being an expert on logic with their suffering from severe inferential disabilities.

Second, the proposed strategy seems limited in its applicability. Even if we conceded that it produces a counterexample to the necessity of CE for mastery of

and, how would this strategy work for the other logical constants? How would competence with the description of visually presented scenes be manifest for *if*, for example?

Third, it is important not to confuse the general idea of a conceptual-role semantics, which commits itself to there being UALs of some type or other, with the more specific idea of an inferential-role semantics, which commits itself to there being UALs that involve *inferences*. Casalegno's case raises an interesting question *within* conceptual-role semantics: should competent use of *and* to describe visually presented scenes also be regarded as necessary for mastery of *and*, in addition to whatever inferential abilities we may insist upon? The answer to this may well be 'yes'. But this would be a *contribution* to a conceptual-role semantics, not a refutation of it.

Fourth, notice how Casalegno's strategy proceeds. It questions whether an inferential competence is necessary for mastery of C by suggesting that a different *competence*, one involving using C to describe visually presented scenes, is sufficient. So, we haven't gotten away from the idea that some competence or other with C is necessary for mastery of C. Thus, even if we granted Casalegno's claim, we would still have it that *either* a UAL involving visually presented scenes, or one involving inference, is necessary. And it is hard to see what other kinds of competence might be relevant.

Finally, it is not out of the question that there might be more than one route to the mastery of a concept, as I pointed out in my 2011a. In his 2011 (Chapter 5, footnote 3) reply to me, Williamson responded to this suggestion dismissively:

In a footnote, Boghossian suggests that even if all understanding–assent links in the strict sense fail, 'A friend of epistemological analyticities might well be satisfied with the existence' of clusters of links of the form 'necessarily, anyone who understands w accepts $S(w)$ or $S'(w) \ldots$ or $S^*(w)$'. To be a genuine alternative, the latter must mean a disjunction of acceptances, not acceptance of a disjunction. Such a disjunctive link will not confer *a priori* status on knowledge of any one of the disjuncts (or on knowledge of their disjunction). Since Boghossian does not explain how it would serve his overall argumentative purposes, I discuss it no further.

(2011: 502)

Williamson's observations here take us away from what the focus of the discussion has been so far—namely, the existence of UALs—to the very different question of how such UALs might ground a priori justification, if they existed. As I have had occasion to note, Williamson has focused his fire exclusively on the existence of UALs and not on the nature of the bridge principles that might connect a particular UAL with the epistemic status of the assent it concerns.

It would take us too far afield to explore this issue in detail. Suffice it to say that, if it is at most a disjunction of acceptances that is necessary for mastery of a given concept, there are well-known ways to build a theory of a priori justification on that basis.

The idea I have in mind here is hardly without precedent. A similar possibility arose in connection with the description theory of proper names, when Searle (1958) proposed, contra Frege, that although there was no particular description that a subject need associate with a given name in order to count as understanding it, some

cluster of descriptions in a disjunction of such clusters is necessary. Kripke (1980: 71–2) explains how that fact could be parlayed into saying that, if a particular cluster D* was the one that was associated with a given name by a given thinker, then D* would be a priori for that thinker.

Similarly, suppose that Mary's particular route to understanding **w** is through S*(**w**), as opposed to any of the other routes available to her. Then, by relativizing the notion of a priori status to individuals, we could say that S*(**w**) is a priori for Mary. This is hardly the place to delve into all the attendant complexities; I mean only to indicate in a general way that there is space for this possibility.

Basis Accounts

Williamson also seeks to refute Basis accounts of understanding-based a priori justification. In cases of a priori justified assent that is best explained by a Basis model, the understanding is not constituted by the assent in question, but rather serves as a good *epistemic basis* for it. But how can the understanding serve as such a basis, if an expert can understand p without assenting (or being disposed to assent) to it?

As Williamson rightly says, the answer I propose is that, even in such cases, the understanding does provide the requisite justification: it's just that the expert's theoretical misgivings preclude her from appreciating that justification. The expert has propositional justification for p, but can't convert it into a doxastic justification for p, a situation we are quite familiar with from other epistemic settings. Williamson objects:

> However, given what has just been argued, we may assume that the expert understands *p* while having no disposition, derived or underived, to assent to *p*. In the key cases, she has no way of getting to *p* just by steps which strike her as cogent. Does *her* understanding of *p* provide her with a good justification for believing *p*? For her, such a justification is out of reach, not just out of sight.   (Chapter 14, p. 210)

Williamson's only argument against the existence of Basis-style explanations of a priori justification rests on his argument against the existence of UALs. Since we have found reason to reject that argument, we may continue to suppose that Vanna continues to find p primitively compelling, even as she has no disposition to assent to p, because she finds p derivatively uncompelling. This fits hand in glove with my description of Vanna as having propositional justification, but lacking the means to turn it into a doxastic justification. Hence, Williamson's case against Basis-style explanations also fails.

Intuitions and the A Priori

Let us turn to Williamson's discussion of my appeal to intuitions in the theory of the a priori. Williamson has two central complaints. First, that intuitions, on my account, even if they were conceded to exist, and were accorded justificatory power, cannot play a distinctive role in explaining a priori justification. Second, that it is doubtful

that there are any mental states that fill the job description that I specify for intuitions. I will take these complaints in turn.

Williamson says that there is nothing in my account of intuitions that explains why intuitions are the source of distinctively a priori justification. He imagines you receiving an invitation from Donald Trump to his birthday party and having the direct and unmediated intuition that you should not accept it. And yet, he says, by ordinary standards, how to respond to a birthday invitation from Donald Trump is not an a priori matter.

Williamson's thought experiment is unrealistic. People don't tend to have unmediated *de re* moral intuitions about persons per se. They tend to have unmediated moral intuitions about their attributes or characters. If it were to seem true to me that I shouldn't accept an invitation from Trump, that would typically be grounded in my intuition that, other things being equal, I ought not to celebrate the birth of someone so evil (fill in your favorite Trump description). And *that* proposition *is* a priori.[7]

Well, perhaps most ordinary folks are as I say. But couldn't there be someone who did have a direct and unmediated intuition that he ought not to accept an invitation from Donald Trump, one that did not depend on any view about his character or attributes? And wouldn't it then follow that this person's empirical belief is justified by his intuition?

No, because the friend of the a priori need only have committed himself to its being a priori propositions that are (directly) justified by intuitions. An a priori proposition is a proposition—for example, *If it's sunny then it's sunny*—that *can* be justified non-perceptually. By contrast, an a posteriori proposition is one—for example, *It's sunny*—that *cannot* be justified non-perceptually. (This is not discussed in the chapter to which Williamson is responding; it is discussed in Boghossian (forthcoming).)

Is this a natural restriction or a merely ad hoc one? Why is intuition well suited for justifying only a priori propositions? Why is it a good way of justifying the moral *principle* I cited above, but not the empirical proposition that Trump is a pathological liar, or that I could levitate if I tried?

The answer is just what you would expect. Perceptions, being the product of a cognitive mechanism that is causally sensitive to facts about its environment, are well suited for justifying contingent propositions about the actual world. Intuitions, not being the product of such a causally sensitive mechanism, if they are well suited for justifying anything at all (not in dispute just now), are well suited for justifying propositions that are either necessary, or anyhow insensitive to the features of the actual world.[8] For promising examples of accounts along these lines, see McGinn (1975) and Peacocke (1993a), although there is much that remains to be worked out.

Intuitions or Inclinations to Judge?

Williamson's more important objection to my appeal to intuitions stems from his doubts that anything can fill the job description that I specify for them.

Williamson is perfectly happy to grant that, in responding with the judgment that p to a thought-experimental question, its seeming true to you that p will precede and explain your judging p. What he denies is that we have any reason to suppose that

this seeming true that p is anything other than a *conscious inclination* to judge that p. By contrast, the intuitional picture holds that seemings true are quasi-perceptual states that precede and explain both the judgment *and* the inclination to judge. Hence, our disagreement may be pictured thus:

Williamson's picture: Thought-experimental scenario/question → Inclination to judge p → (which, when not defeated) → Judging p

Intuitional picture: Thought-experimental scenario/question → Intuition that p → (which, when not defeated) → Inclination to judge p → (which, when not defeated) → Judging p

How should we decide which picture is correct?

About the Trolley judgment that it would be wrong to throw the large man off the bridge, Williamson says,

Of course, for all I have proved, in Boghossian's case you may also be in some further mental state which both presents the proposition that it would be wrong to throw the man off the bridge and precedes even the disposition to make that judgment. But how is one supposed to know whether one is in such a mysterious extra state? In my own case, for all introspection tells me, there is only the consciously inhibited disposition to judge.

Could Boghossian just postulate that we are in these mysterious extra states of intuiting? The trouble is that, given his internalism, any difference in justification presumably corresponds to a consciously available difference in mental state. Thus, if the difference between the presence and absence of the extra state of intuiting is not consciously available, it makes no difference to justification. Positing a consciously unavailable difference will not serve his purposes.

(Chapter 14, p. 212)

Searching his mind, Williamson finds only the conscious inclination to judge and no consciously available impression of truth that precedes that inclination and explains it. If we don't find such impressions via introspection, might we not have reason to *postulate* them? Williamson says that, even if a case could be made for postulating them, that path is not available to me, since, on my view, intuitions are needed to serve an internalist epistemological purpose, and so they need to be consciously available. His overall argument, then, goes like this:

1. To have reason to believe in a certain type of mental state, it must either be consciously available or reasonably postulated.
2. Intuitions are not consciously available (mere introspection does not reveal them).
3. Hence, to have reason to believe in intuitions, they must be reasonably postulated.
4. If a type of mental state is postulated, then it is not consciously available.
5. If a type of mental state is to serve an internalist purpose, it must be consciously available.
6. Hence, even if we had reason to believe in intuitions, they would not be able to serve an internalist purpose.

There is an enormous amount of interesting philosophy here that is worth discussing; in this reply, I will only be able to address a few central points.

In my view, much of Williamson's argument is vitiated by the fact that he treats the categories of *consciously available* and *postulated* as dichotomous. However, these categories are not mutually exclusive. Rather, what we need is a threefold distinction between those mental state types that are

(INT)   consciously available to introspection *without* theoretical guidance;

those that are

(INT+)   consciously available to introspection (typically) *only with* theoretical guidance;

and those that are

(MP)   merely postulated.[9]

Some non-representational phenomenal states—like pains or tickles—will impress themselves upon your attention without any help, just by the sheer strength and vivacity of their phenomenal properties.

However, in the case of experiences with representational content—the experience of blue, for example—their 'diaphanousness' can get in the way of introspecting them, as Moore (1903b) pointed out:

... the moment we try to fix our attention upon consciousness and to see *what*, distinctly, it is, it seems to vanish: it seems as if we had before us a mere emptiness. When we try to introspect the sensation of blue, all we can see is the blue; the other element is as if it were diaphanous ...   (450)

As Moore himself went on to observe, though:

Yet it [the sensation of blue] *can* be distinguished if we look attentively enough, and if we know that there is something to look for.   (450)

Even with such phenomenal states as seeing blue, then, you might need guidance if you are to find them through introspection.

Many other mental state or event types fall into the INT+ category. Take, for example, occurrent judgments. We certainly don't believe in them simply because we stumble across them when we introspect. They can't be said to impress themselves upon our attention in the way that pains do. We believe in them in part because we know that there are states that play a specific type of role in the explanation of action. However, once we know what to look for, our occurrent judgments are readily introspectable.

In contrast with both of the preceding cases, there are such states as unconscious desires, states whose existence was established by Freud among others, that are merely postulated.

Intellectual impressions of truth belong to the INT+ category: they are consciously introspectable, though only if you know what to look for. What considerations might help us see that, in addition to conscious inclinations to judge, there are impressions of truth that precede such inclinations and help explain them?

It's obvious, I trust, that in standard empirical cases, we have the three-part impression–inclination–judgment structure outlined above. In a Müller–Lyer case,

for example, there is the visual impression of the lines as being of unequal length; there is the consequent inclination to judge that they are unequal in length; and, finally, there is the judgment that they are unequal in length. No one element can be reduced to the others, and, in particular, the impression can't be identified with the inclination to judge. This is obvious from the fact that, when I discover that the lines are not unequal in length, I lose both the judgment, and the inclination to judge, that they are; the impression, however, remains.

Williamson's view, then, is that in classical thought experiments, and in other a priori cases, with an exception being made for geometry, the impression element goes missing: in such cases, there is just the inclination to judge, and, if it is not inhibited, the judgment itself. As a result, there is no consciously available explanation for the inclination to judge. It must be experienced by the subject as coming out of nowhere.

However, this doesn't seem true to the phenomenology of intuitive responses to classical thought experiments. The conscious inclination to judge the Trolley proposition doesn't present itself as coming out of nowhere; rather, it presents itself as a response to its *seeming true* that you ought not to throw the large man off the bridge.

To see this, contrast your judgment in a Trolley case with that of Norman the clairvoyant, made famous by Laurence Bonjour (1985). In Bonjour's imagined scenario, Norman has a clairvoyant ability that he does not know about. One day, out of the blue, he finds himself with the inclination to believe that the President is in New York City. He has no consciously available evidence that bears on the President's whereabouts. His conscious inclination to belief is the product of his reliable clairvoyance, though Norman knows nothing about that. He just finds himself with that inclination; he can identify nothing on which it might be based.

Bonjour uses this example effectively against reliabilist theories of justification. But it is also useful for our purposes now, in illustrating what it would be like to come to be inclined to believe something out of the clear blue. If you ask Norman *why* he believes that the President is in NYC, he would say that he has no idea why. If you ask him if it *seems* true to him that the President is in NYC, he would say (exactly what Williamson says about thought experiments in general): 'Apart from the fact that I find myself inclined to believe it, there is no seeming.'

Williamson-type inclinations to believe, then, ones that are not based on anything, and in particular not on seemings, are hardly inconceivable. Indeed, they fit Norman's case perfectly. However, our typical responses to classical thought experiments are not like that. When we find ourselves inclined to judge that we ought not to throw the large man off the bridge, or that Mr. Smith has a justified true belief but does not know, these inclinations don't feel as though they are coming out of the blue in the way in which Norman's clairvoyant beliefs do. They present themselves as based on their seeming true. Thus, if we are to adequately capture how classical thought experiments work, we need to recognize not only conscious inclinations to judge but intellectual seemings as well, which would explain and justify those inclinations.

A second consideration in favor of the intuitional picture returns us to the observation that there are many cases where a proposition's seeming true survives the discovery that the proposition is false. For example, even after we discover that the Müller–Lyer lines are not unequal in length it still seems true that they are. And

as Williamson acknowledges, this happens in a priori cases as well: even after we discover that Naïve Comprehension is false, it continues to seem true.

In the Müller–Lyer case, of course, the seeming true that survives the discovery that the lines are not unequal in length is neither the judgment that they are unequal in length, nor the inclination to judge that they are. Once I know that they are not unequal in length, I lose all inclination to judge that they are. What survives is just *the visual impression* of the lines, which continues to present them as being unequal in length, despite the fact that we now know they are not.

Why does the visual impression persist in depicting the lines as being of unequal length, despite our discovery that they are not? The answer, presumably, is that perception, being to a considerable degree modular, is not penetrable by the background knowledge about the length of the lines. Hence, the impression of inequality persists despite our discovery.

What about Naïve Comprehension? Here, too, there is a seeming that survives the discovery that it is false. Williamson thinks that he can account for this by claiming that the conscious inclination to believe Naïve Comprehension survives the discovery that it is false, the discovery serving only to inhibit its exercise.

As I've just been arguing, however, inclinations to believe tend not to work like that. They are cognitively penetrable and tend to disappear when their propositional objects are discovered to be false. In general, they are not like habits that are hard to shake. (The qualification allows that there might be some, perhaps very basic inclinations to judge, that are sticky—examples might include the inclination to believe that there is an external world, or that animals have propositional attitudes. But such special cases would not serve Williamson's purpose.)

The only good way to account for its continuing to seem that Naïve Comprehension is true, even after it is discovered to be false, is to recognize that there is an intuition of the truth of Naïve Comprehension, which, by virtue of being cognitively impenetrable, retains its misleading content.

The third argument in favor of recognizing intuitions is *epistemic* in nature. Since I've discussed it in Chapter 13, I won't rehearse it here at length. By internalist standards, it is not possible to explain the (very strong) justification that we have for typical a priori claims simply by invoking the fact that we are inclined to make them. In general, your being inclined to judge p is no justification for your judging p. On the other hand, as I've argued, p's seeming true to you is a very good candidate for providing at least prima facie justification for believing p.

In sum, it's quite clear that, in the types of contexts we have been looking at, in addition to the conscious inclination to judge, there is something pre-judgmental that behaves in just the way you would expect an impression of truth to behave.

Williamson also raises a question about the justificatory power of intuitions. He asks, 'Why should they be impervious to all the usual distortions from ignorance and error, bigotry and bias?' (Chapter 14, p. 213).

I do not provide a positive theory of the justificatory power of intuitions in my already lengthy essay for this volume. I do so in Boghossian (forthcoming), attempting to add to the excellent contributions made by such philosophers as Bealer (2000), Bengson (2015), and Chudnoff (2013). This is a rich area for epistemology, with much left to be studied and understood.

Apropos of the particular point that Williamson brings up, no one thinks (or anyhow should think) that just because visual perception is open to distortion by bias and bigotry (for examples, see Siegel (2017)) that it cannot be thought to provide prima facie justification for belief.[10]

Notes

1. This paper is published here for the first time.
2. This, of course, is overly simplified. Some UALs might involve inferences involving p rather than assents to it. What is meant will be clear in each case. The use of the expression 'underived' in this formulation will be further discussed below.
3. It should be reemphasized that 'understanding p' in the present context means 'full mastery of p' and not mere 'possession'.
4. Later on, I will be raising a question about Williamson's view that dispositions to assent to p can persist well after a thinker comes confidently to believe that p is false. For now, I will let it pass.
5. Williamson sometimes complains that this appeal to a 'closely related concept' is empty without a specification of what the alternative concept in question is. It's unclear what he is asking for. Obviously, we lack a word in English whose meaning is that concept. But we can, of course, specify the concept in question in terms of its conceptual role, one which will overlap with that for conjunction in certain respects, but deviate from it in others. (For an analogy, think here of the transition from a two-place concept of simultaneity to a three-place concept, necessitated by Einstein's Special Theory of Relativity.)
6. This, of course, is what some philosophers actually argued in the case of the distributive principles. See, for example, Putnam (1968).
7. Note, by the way, that this is a remark about the architecture of justification, not a remark about the psychological processes that lead to someone's making the moral judgment at issue, for a variety of reasons. For one thing, most of us will have automated our responses to various types of familiar situation so that we know what to think without having to run through an explicit inference in our heads. For discussion see Boghossian (2016).
8. Fixedly actually true, in the terminology of Davies and Humberstone (1980).
9. Some of the material that follows overlaps with Boghossian (forthcoming).
10. Thanks to Ned Block, Christopher Peacocke, and Crispin Wright for helpful comments.