Centrality and Marginalisation*

Brian Weatherson

2014

A commentary on Herman Cappelen's "Philosophy without Intuitions".

1 Welcome to the History of Late Analytic Philosophy

It's a good time to be doing history of late analytic philosophy. There is flurry of new and exciting work on how philosophy got from the death pangs of positivism and ordinary language philosophy to where it is today. Some may see this as a much needed gap in the literature. Indeed, there are a couple of reasons for scepticism about there being such a field as history of late analytic philosophy, both of which are plausible but wrong.

One reason is that it is too recent. But it can't be too recent for general historical study; there are courses in history departments on September 11, so it's not like looking at philosophy from thirty to forty years ago is rushing in where historians fear to tread. And indeed, if logical positivism could be treated historically in the 1960s, and ordinary language philosophy could be treated historically at the turn of the century, it seems a reasonable time to look back at the important works of the 1970s that established the contemporary era in philosophy.

Another reason is that we all know it so well. We are still so engaged with the key works by Kripke, Lewis, Burge, Perry, Thomson and so on that we don't need to also look at them the way we look at Descartes, Locke and Hume. But this, it turns out, is not true. Books by Daniel Nolan (2005) and Wolfgang Schwarz (2009) changed the way that some philosophers, even those who knew the Lewisian corpus fairly well, changed the way they read Lewis. There has also been a minor flurry of work on how important the Gödel/Schmidt case is to the argument of *Naming and Necessity* (Devitt 2011; Ichikawa, Maitra, and Weatherson 2012; Machery et al. 2012).

But that's nothing compared to the bombshell that is *Philosophy Without Intuitions*. (Cappelen (2012); all page citations, unless otherwise noted, to this book.) Herman Cappelen shows, extremely convincingly to my eyes at least, that intuitions play a much smaller role in late analytic philosophy than many philosophers thought. Indeed, there is a lot of textual evidence both for the claim that intuitions don't do much philosophical work, and for the claim that many people have said that they do. The first of these

*Thanks to Herman Cappelen and Ishani Maitra for many discussions about the material in this paper.

claims is all to the good, says Cappelen, since there isn't a particularly good epistemological defence of the use of intuitions.

The evidence for Cappelen's claims comes in two parts. The first part, which I won't discuss much here, is an extended argument that words like 'intuitively', or 'counterintuitive', as they appear in philosophical discourse, don't in general function to pick out, or even draw attention to, any distinctive kind of mental state we could call an 'intuition'. The second part argues that when we look at the actual introduction of thought experiments into late analytic philosophy, we don't see the appeal to intuitions that many philosophers seem to think go along with thought experiments. Rather, we see a whole host of interesting philosophical moves. Sometimes a thought experiment functions to highlight an explanandum. Sometimes it gives us a *prima facie* plausible thesis that we then argue for (or against) at great length. Sometimes it just raises a puzzle.

One upshot of this historical work, one that Cappelen I think does a good job highlighting, is that contemporary philosophy is much more *interesting* than its practitioners sometimes take it to be. Philosophy is a way of investigating hard questions about the world, often at great expense in terms of human capital, but with thankfully little in the way of other expenses. It isn't a matter of tidying up conceptual space. Thinking of philosophy this way should, I think, help us see why so many different kinds of projects are philosophically important.

2 Centrality and Its Discontents

The big goal of Cappelen's book is to refute the view, which he dubs Centrality, that intuitions (of a certain kind) are central to analytic philosophy, and in particular that they are a primary source of evidence for analytic philosophers. The intuitions that he has in mind have these three characteristics. (The quotes are from pages 112-3, where these features are articulated.)

F1: Phenomenology "An Intuitive Judgment has a distinctive phenomenology" .

F2: Rock "An intuitive judgment has a special epistemic status ...Intuitive judgments justify, but they need no justification".

F3: Conceptual A judgment is an intuition "only if it is justified solely by the subjects' conceptual competence".

There's some more detail on F2, but we'll get to that in Section 6. And there's a fourth characteristic of intuitions that I want to add.

F4: Speed Intuitions are rapid reactions..[1]

I'm going to spend much of this paper defending a view that intuitions characterised by F2 and F4 do play a role, though perhaps not a *central* role, in philosophy. But I do

[1] My own views about the importance of this, as well as much else in this paper, owe a lot to Jennifer Nagel (2007, 2013).

think that intuitions characterised by F1 and F3 are just not important to philosophy. Indeed, I think it's a very important fact that they are not that important.

The claim that intuitions have a distinctive phenomenology is mostly harmless but, it seems to me, false. I certainly don't find anything in common when I introspect my judgments that, say, no set is a member of itself, or that losing a limb would seriously reduce my happiness, or that the only language I think in is English. It will fall out of the view I'm defending that the best intuitions have no phenomenology, but I don't think that's a particularly important fact about them.

But the claim that intuitions derive solely from conceptual competencies, plus the claim that these are the central source of evidence in philosophy, is both wrong and dangerous. If that conjunction were true, we'd expect most philosophical conclusions to be conceptual truths (whatever those are). I'm not going to take a stand on whether there are conceptual truths, but I think it is pretty obvious that conceptual truths won't help much resolve the following debates. (Compare the list E1-E6 on pages 200-201, which I'm basically just extending.)

- Do bans on pornography involve trading off speech rights versus welfare considerations, or do they just involve evaluating the free speech interests of different groups?
- Is it permissible to eat whales?
- Under what circumstances is it permissible to end a terminally ill patient's life, or to withhold life-saving treatment?
- Is all context dependency in language traceable to the presence of bindable variables?
- Does belief have a phenomenology?
- Which animals (and which non-animals) have beliefs?

If philosophy uses largely conceptual evidence, these aren't philosophical questions. More generally, if Centrality (in Cappelen's sense) is true of philosophy, then feminist philosophy, legal philosophy, political philosophy, bioethics, philosophy of language and (most of) philosophy of mind are not part of philosophy. (This list is far from exhaustive; making philosophy Centrality-friendly would involve writing out huge swathes of the discipline.)

Modus tollens obviously beckons. But as Cappelen notes (213), one occasional reaction to this is to identify certain parts of philosophy as the 'Core' of the discipline, and say Centrality is true of those. If Centrality is true of the core of philosophy, then feminist philosophy *etc*., are not part of the core of the field. Maybe now some people would be disposed to use modus ponens not modus tollens.

That would be a large mistake. It would have shocked Plato, and Locke, and Hume, and practically every other major figure in the history of philosophy to learn that political philosophy wasn't central to the field. I do think (contra some of what Cappelen says) that some philosophy involves a priori and conceptual investigation. Indeed, I even do some of it. But it's not true that when I'm doing that I'm doing work that's deeper, or more philosophical, or more central to philosophy than the work that, for

example, Rae Langton or Susan Moller Okin or Tamar Szabó Gendler or Sarah-Jane Leslie do.

This reason alone suffices for me to hope that Cappelen's book has a very wide readership. Centrality isn't true, but it is I think widely believed to true of at least some parts of the field. (Cappelen quotes many people endorsing this view.) I suspect that on the basis of this mistake, the parts of philosophy about which Centrality is not obviously false (especially metaphysics and epistemology) have been seen as more central to the discipline than they really ought to be. That's not a bad state of affairs for metaphysicians and epistemologists, but it's not good for philosophy, and I hope that Cappelen's book helps put a stop to it.

3 Intuitions in Detective Work

Despite my very broad sympathy with Cappelen's project, I do think there's a role for intuitions of some kind in philosophy. Just what this kind is, and what this role is, will take some spelling out to avoid Cappelen's arguments. So that's what I'll do for the next few pages.

The intuitions I have in mind are characterised by F2 and F4; they are default justified, and they are fast. Here's how I think these kinds of intuitions could matter philosophically.

When humans are growing up, they develop a lot of cognitive skills. Some of these skills are grounded in specific bits of propositional knowledge. We learn to count in part by learning that 2 comes after 1, and 3 comes after 2, and so on. But not all of them are. We learn how to tell causation from correlation, at least in simple cases, by developing various heuristics, none of which come close to a full theory of causation. Indeed, none of these heuristics would even be true, if stated as universal generalisations. But this ability to pick out which of the many predecessors of an event is its cause is one we develop very early (Gopnik 2009, 33–44), and it is vital to navigating the world.

I think we develop a lot of skills like that; skills which either go beyond our propositional knowledge, or at the very least are hard to articulate in terms of propositions. That we have these kinds of skills should hardly be news to philosophers; under the label 'heuristics' they have become quite familiar thanks to the work of, among others, Daniel Kahneman. They occasionally get a bad press, because one central way in which psychologists detect them is by seeing where they lead to errors that careful thought would correct. (For instance, our heuristics sometimes say that a conjunction is more probable than one of the conjuncts, and careful thinking would correct this.) But this should not blind us to the fact that these incredibly fast heuristics are often very reliable; reliable enough to be an independent check on our theorising.

The use of the term 'intuition' to pick out these heuristics isn't particularly idiosyncratic; Kahneman (2011) himself moves back and forth freely between the two terms. He approvingly cites Herbert Simon's remark that "intuition is nothing more and nothing less than recognition", which I think is basically right. We intuit that *a* is *F* by recognising that it has the tell-tale signs of *F*hood. Of course we're a million miles from con-

ceptual or *a priori* reasoning here; as I said, I agree entirely with Cappelen that F3 is not a feature of any philosophically significant source of evidence. Here are a couple of cases, one real life and one fictional, that draw out far removed intuitive thinking can be from a priori or conceptual thinking. The first is from Kahneman's description of a case reported by Gary Klein (1999); the second is from (Norwegian) crime novelist Jo Nesbø (2009). First Kahneman,

> A team of firefighters entered a house in which the kitchen was on fire. Soon after they started hosing down the kitchen, the commander heard himself shout "Let's get out of here!" without realizing why. The floor collapsed almost immediately after the firefighters escaped. Only after the fact did the commander realize that the fire had been unusually quiet and that his ears had been unusually hot ... He had no idea what was wrong, but he knew something was wrong. (Kahneman 2011, 11)

Now Nesbø. In the story, Harry is the hero, Harry Hole, and Beate is a talented forensic detective.

> 'Forget what you have or haven't got,' Harry said. 'What was your first impression? Don't think, speak.'
>
> Beate smiled. She knew Harry now. First, intuition, then the facts. Because intuition provides facts too; it's all the information the crime scene gives you, but which the brain cannot articulate straight off. (Nesbø 2009, 126)

There's at least a family resemblance between Harry Hole's instruction here and Lewis's instruction to his readers at the start of "Elusive Knowledge" (Lewis 1996).

> If you are a contented fallibilist, I implore you to be honest, be naive, hear it afresh. 'He knows, yet he has not eliminated all possibilities of error.' Even if you've numbed your ears, doesn't this overt, explicit fallibilism still sound wrong? (Lewis 1996, 550)

Reviewers of Nesbø's books often describe his hero as 'intuitive'. That's a little misleading; Harry Hole thinks intuition has a key role to play in detective work, but the adjective suggests that he relies heavily on his own intuition. That's not right; he's just as often badgering his colleagues to give him their impressions of a crime scene, or an interview subject. In these scenes he reminds me of no one so much as a colleague constantly wanting to know what one thinks about some thought experiment or variation on a familiar case. (These are often the best kind of colleague - full of inspiring ideas!)

So I think a lot of philosophical progress is made by drawing on, and drawing out, these skills. But isn't this just to say something uncontroversial and uninteresting, namely that philosophy relies on implicit knowledge? As Cappelen puts it,

> It is not controversial that conversations have propositions in the common ground. Nor is it controversial that all arguments start with premises that are not argued for. (155)

Well, there's something a bit interesting here, namely that the 'common ground' and the 'not argued for' premises have much greater overlap in philosophy than in other fields. A book starting with observations about the Galápagos Islands starts with premises that are not argued for, but are asserted on the basis of observations. These premises surely weren't in the common ground before the 'conversation' starts. I'll say more about this in the next section.

Because first I want to fuss a little about just what 'common ground' is. We'll start with an observation Cappelen makes about the Ginet/Goldman case of Henry and the fake barns (Goldman 1976). Many philosophers take it to be an interesting fact that in one scenario, Henry knows there's a barn, while in another he does not. Cappelen says that these facts are "presented as being pre-theoretically in the common ground" (172). That seems false at first blush. Before reading Goldman's paper, it's not clear philosophers are in a position to form singular thoughts about Henry. That's an uncharitable reading though. A more plausible claim is to say that we are pre-theoretically disposed to accept some long sentence that roughly says that an agent in such-and-such scenario knows there is a barn, while an agent in a slightly different scenario does not.

We might gloss that last claim as saying that we implicitly knew something about these scenarios. I'm not sure that's right though. We do surely have lots of implicit knowledge. I know, and so do you, that the Sydney Opera House is south of the Royal Albert Hall, even if you'd never articulated that thought to yourself or another. But do our dispositions to respond to quite finely drawn, and often reasonably long, vignettes count as implicit beliefs, or should they count as things we were in a position to know, but only learned once a philosopher had done the work of drawing the vignette? I can see merit in both positions, and don't see firm grounds for preferring one.

Let's introduce some terminology to avoid taking a stance on this question. Say that a subject has *Socratic knowledge* that p when the following two conditions are met:

1. Once the agent is asked to consider p in the right way, they will come to know p.
2. The evidential basis for this knowledge that p is not the asking itself.

The first clause says that anyone who reacts to a Gettier case with "Oh, of course that's justified true belief without knowledge" has Socratic knowledge that such a case is a counterexample to the JTB theory of knowledge. And they have this Socratic knowledge before the case is even raised. The second clause says that if the person reacts instead with "Oh, some philosophers use thought experiments that don't make sense unless you know which cars come from which countries", that *won't* count as Socratic knowledge. They would be expressing some knowledge, to be sure, but the telling of the example would play an evidential role.

If you are very liberal about which dispositions count as implicit beliefs, and implicit knowledge, then Socratic knowledge will just be a special kind of implicit knowledge. But if you think considering examples can lead to learning new facts, not just drawing out dispositions, then you will think 'Socratic' is like 'alleged', a non-factive modifier. As I've defined it, once you hear Gettier cases once, that they are counterexamples to the JTB theory ceases to be Socratic knowledge, and becomes regular knowledge. Note also

that we can make sense of some implicit states being more or less Socratic than others; some dispositions to assent require very careful work to trigger.

Why is the class of propositions that we Socratically know so rich and fertile? It's because of the central role of heuristics in our cognitive lives. Our interactions with the world don't just furnish us with a set of truths about the world. They also furnish us with skills that we can apply to generate more truths. I suspect that something like this observation is at the heart of the endorsement of F3, that intuitions reveal conceptual truths. When we intuit that *p*, we don't always merely recall a prior belief that *p*, or infer *p* from what we antecedently explicitly knew. But nor do we observe that *p*. So what is it? It must be something internal, but not memory or inference. Conceptual competence isn't a bad first guess, but Cappelen shows that isn't the right answer. I think the right answer has to do with cognitive skills, i.e., heuristics.

4 Philosophy: A Negative Characterisation

So intuitions matter because they reveal Socratic knowledge, and Socratic knowledge, when made explicit, is a very good guide to the world. That implies that intuitions should not be confined to philosophy. And, indeed, they are not. If an economic theorist claimed the standard of living among English men was higher in 1915 than in 1935, it would be perfectly reasonable to reply that intuitively that cannot be right, because in 1915 a rather large number of English men were living on the Western Front in catastrophically poor conditions. What is distinctive of philosophy then?

We need to clarify this question before we can answer it. Philosophy is both a discipline with a history over many millennia, and an organisational unit inside modern universities. These two things overlap well, but not perfectly. Once we note that they are distinct, we can separate out the following three questions.

1. What questions are philosophical questions?
2. What questions are, within the academy, primarily addressed by researchers in philosophy departments?
3. What questions should be, at least within the academy, primarily addressed by researchers in philosophy departments?

The three questions don't overlap. When Milton Friedman (1953) writes about economic methodology, I think he's addressing a philosophical question, but work like this is, and probably should be, carried out in economics departments. Questions about professional ethics are philosophical questions that I think should be researched in philosophy departments, but in the United States at least typically receive more attention in professional schools. Let's focus on the third question; what should a philosophy department do?

My colleagues at Michigan and St Andrews work on an incredibly wide range of questions, from the interpretation of quantum physics through history of logic through moral psychology and so on. And I think philosophy departments should have this range of interests. But what do all these questions have in common?

It's not anything to do with necessity or a priority. Those categories seriously cross cut philosophy, as Cappelen points out. Historical investigations into disputes about the parentage of various might-have-been-royals, or mathematical investigation into the nature of the primes are not philosophical, but have to do with necessity and a priority. Whether there's a language of thought is contingent, a posteriori, and almost paradigmatically philosophical.

It's not really anything to do with *depth*, at least on a natural understanding of that. Why pandas have thumbs, and humans have appendices, turn out to be reasonably deep questions, but they are for biologists, not philosophers. Under what circumstances is democracy compatible with a strong executive is, at least to me, an incredibly deep and important question, but it's a question to be answered, primarily, in history and political science departments.[2] On the other hand, whether we can tell a plausible supervaluationist story about belief reports is not particularly deep, but a perfectly good subject for a philosophical inquiry as in Weatherson (2003a).

Better, I think, is to say that philosophical questions are those where implicit or Socratic knowledge, including crucially intuitions, can plausibly play a large role in getting to an answer. Philosophy is a little recursive, so it includes investigations into its own investigations, including historical work and metaphilosophical work. (Two fields which, prior to Cappelen's book, had surprisingly little interaction.) That's not to say we're always right that Socratic knowledge can answer the questions philosophy sets. Maybe some questions in mind and language are best answered with the aid of neurological or phonological work that requires powerful measuring devices. But the questions are ones where starting with the knowledge and skills we already have seems like a plausible starting point, or at least not entirely crazy. This makes philosophy distinct from, say, history. We use intuitions in history too, especially intuitions about what explains what. But we need more; intuition won't help if you want to know how many troops Henry had at Agincourt.

This hypothesis explains, I think, one of the historically important facts about philosophy. Philosophy gives birth to disciplines. Physics, economics, psychology and cognitive science were all, at one time, part of philosophy. In some cases, the split was very recent. The economics tripos at Cambridge only split from philosophy in 1903 (Tribe 2002). The *Australasian Journal of Philosophy* was the *Australasian Journal of Psychology and Philosophy* until 1946. Why does philosophy give rise to disciplines like these?

I think having a negative characterisation of philosophy helps explain it. Philosophy has a lot in common, methodologically, with physics, economics, psychology and so on. All those fields use intuitions and other forms of Socratic knowledge. But the other fields use other things too, especially observation. It's when it becomes clear that armchair methods play too small a role in the research that the field leaves philosophy.

Of course, philosophers care more about their questions than their methods, so when the need for non-armchair methods becomes pressing, some of the individual

[2] This is not to say that political philosophers couldn't help with this question. There are lots of questions that should have as their research centre some other department, but to which philosophers can usefully help. Indeed, the examples from economic methodology and evolutionary explanation I just mentioned are two more such questions.

philosophers will go along, picking up more and more observational knowledge and experimental skills. Note how much more empirical research informs the recent work by (for example) Gilbert Harman, Kim Sterelny and Peter Carruthers, compared to their earlier work (Harman 1973; Kilkarni and Harman 2011; Devitt and Sterelny 1987; Sterelny 2012; Carruthers 1990, 2011). From the other direction, our armchairs come with more knowledge now than they used to, which is partially why engaging with Laura Ruetsche's work in philosophy of science requires more empirical knowledge engaging with William Whewell's (Ruetsche 2011; Whewell 1840). But still I think the general picture holds; a question is fit for philosophy iff it is plausible that the intuitive, armchair methods which are part of every academic's toolkit can, on their own, generate serious progress on the question.

5 Letting Go

I've said that Lewis's instruction at the start of "Elusive Knowledge" is to look to intuitions, not to theoretical beliefs. But that might involve reading more into Lewis than is really there. What he literally asks the reader is to not appeal to their preferred theory of knowledge. Is that the same as an appeal to intuitions?

It need not always be. Sometimes, asking people to let go of their prior theory involves asking them to engage in a complex cognitive task. In Meditation One, Descartes has us go through quite a lot of thoughts before we can be pre-theoretical in the way he wants us to be.

But I don't think that's what's going on with Lewis. For one thing, he doesn't guide us back to a pre-theoretic naïveté the way Descartes does. But more generally, I think getting snap judgments is a way of letting go of some prior theories.

The picture I have here, and it is nothing more than a picture, is that intuitions are judgments delivered by heuristics, heuristics are deployed by Fodorian modules, and Fodorian modules are informationally encapsulated (J. A. Fodor 1983; J. Fodor 2000). That is, when we rely on a heuristic, we don't use all of the information at our disposal. The classic example of this is eyesight; we may know that there are no elephants on Market Street in St Andrews, but given the right visual stimuli, our eyes will still insist that there is an elephant *right there*. The background theory about the spatial distribution of elephants isn't encoded into the visual module. More generally, to rely on a heuristic just is to make a judgment using a part of our mind that doesn't believe some of the things that we do. And that's good, because it is a kind of independent check on the beliefs we have.[3]

But isn't the idea that snap judgments are essential to philosophy inconsistent with the fact that we work very hard on getting our examples just right, and (as Cappelen shows), argue at great length over what to say about various examples? I think it isn't,

[3] Philosophers sometimes understate the importance of independent checks. We can know a scale is working, but if we want to check its reliability we don't use it, we use something else. I suspect that a certain amount of theory-independence is part of the explanation of the value of intuitions.

because there are two respects in which our practice reveals a sensitivity to snap judgments, and a respect for their use as a check on theorising.

Let me tell you a small secret. I haven't heard anything that even sounds like a counterexample to the broadly Stalnakerian theory of indicative conditionals that I like for about a decade. That's not because there aren't any intuitive counterexamples. It's just because my intuitions have been trained to accord with this kind of theory.[4] So what do I do? Do I give up on the use of intuitions as a test of theory? No, I ask colleagues for their intuitions. Sometimes I ask them a lot of different questions, and sometimes I work rather hard on refining the question, or (when they sadly disagree with my theory) finding ways to undermine their intuitions. Given the number of similar questions I get from other colleagues, I don't think my methodology here is distinctive. In short, we can work very hard before and after getting the snap judgments, while giving those judgments a role.

This might be more idiosyncratic, but I also do a bunch of things in papers to draw out snap judgments. The main idea is to distract the reader from the fact that they are about to be prompted for an intuition, one that may not accord with their preferred theory. So I'll use deliberately absurd props (like Vinny the talking vulture), or start an example without flagging that it is an example. My favourite move along these lines is to set up an example in such a way that the example doesn't make sense unless some theoretical claim I want to argue for is true. Then, after much discussion of the correct verdict on the case, I can announce that the very sensibility of the prior discussion is proof that, at least intuitively, the theory I'm pushing must be true.

We're going to come back to this theme a bit later, because I think it's rather important. The cases you can remember from papers are probably not the ones where intuition mattered. The big role for intuition in philosophy (and in many other disciplines) is in checking the small steps along the way. That's why I join Cappelen in opposing the methodological rationalists; I don't think intuitions are distinctive to philosophy, and these small steps don't have much of a phenomenology. But that doesn't mean they are unimportant.

6 Strength and Fragility

One of the big trends in late 20th Century epistemology has been the separation of two senses of *strength of evidence*. This might mean

1. How strong a doxastic state is supported by the evidence.
2. How resilient the force of the evidence is in the face of counterevidence.

One thing that conservative epistemologies (e.g., Harman (1986)) and dogmatic epistemologies (e.g., Pryor (2000)) have in common is that sources which might be very strong in the first sense might be very weak in the second sense. In particular, there can be sources of evidence that ground knowledge, and hence be rather strong in tthe

[4] Relatedly, I haven't seen Liverpool get awarded an undeserved free kick for about that long.

first sense, but easily overturned by conflicting evidence. I prefer to reserve the terms 'strong' and 'weak' for the first sense, and use the terms 'resilient' or 'fragile' for the presence or absence of the second property. In that language, the important insight of the conservatives and dogmatists is that evidence can be strong but fragile.

That's roughly how I think of intuitions – they are strong but rather fragile. So they can be unjustified justifiers, which is how I read Cappelen's feature F2 (i.e., Rock).[5]

Cappelen notes it is hard to tell whether something is being used as a starting point, or an unjustified justifier, so he gives three diagnostics for this. I mostly agree with one, and disagree with the other two. I agree that intuitions are non-inferential, and they aren't based on any particular experience, which is his criteria F2.1. (Though they usually are based on experiences taken collectively.) But I would alter the following suggestion, which he gives as a second diagnostic.

> **F2.2 Evidence Recalcitrance** Intuitions are evidence recalcitrant; i.e., holders of them are not disposed to give them up even when their best arguments for those intuitions are shown to fail. (Compare pg 112)

I would rather offer something normative here. What's true of intuitions is that they might provide a stronger ground for belief than the best evidence we can offer for them. Compare the case of Gettier. As Cappelen carefully notes (194n3), Gettier doesn't appeal to a raw intuition. He gives an argument that his subjects don't know. Unfortunately, it isn't a compelling argument, since it takes as a premise that we can't get knowledge from a false belief, and that isn't quite right (Warfield 2005). But Gettier was, to some extent, justified in believing these subjects didn't know to a greater degree than he was justified in believing this argument was sound. And that, I think, is not uncommon.

This is why I don't think Cappelen's 'Rough Guide to Rock Detection' (121), the third of the diagnostics, is perfectly reliable. He says that if evidence is given for p in a context, that's evidence that p isn't an unjustified justifier in that context. But sometimes we give arguments for judgments that we think could rest without them. Compare this little dialogue.

[5] There's an ambiguity in Cappelen's text that I'm not sure I'm interpreting the right way. Let's that someone intuits that in a particular case, c doesn't cause e. Call the content of that intuition, i.e., what is *intuited*, p_d. And call the proposition that the person has this intuition, i.e., the event of the intuiting, p_g. Plausibly both p_d and p_g could be evidence in the right cases, though most of the time the salient evidence will be p_d. I think p_d can be an unjustified justifier in the sense that other beliefs, e.g., that a particular theory of causation is false, can be justified on the basis of p_d, but no other beliefs the agent has justify p_d. But you might want a stronger sense of 'unjustified', where it means not just not justified by anything else, but not justified *at all*. I think in these kinds of cases, p_d is justified, just not justified by anything else. And the justification is, as I'll get to below, strong but fragile. If when Cappelen says that intuitions, according to Centrality, are unjustified justifiers he means that the belief that p_d is unjustified, then I'm not defending Centrality. I just mean that the agent need not have any other mental states which justify the belief p_d, or indeed any access to anything that justifies p_d. But for all that it might be that the belief that p_d is justified, and the grounds for the justification include what the agent learned about causation as a child, plus perhaps her competence in distinguishing causes from non-causes.

A: Is 'John happiness' a well-formed sentence?
B: No; it doesn't have a verb.

Here B gives a judgment, then offers a little argument for it. The argument has a strong premise, namely that all sentences have verbs. That's debatable; 'Lo, gavagai!' may be a counterexample. But B's judgment isn't undermined by examples that undermine her argument. As in the Gettier case, we may give an argument that doesn't capture the full normative force of the judgment.

To say that intuitions are unjustified justifiers is not to say they are particularly special. If some conservative or dogmatic epistemology is true, there will be other unjustified justifiers. And if not, then this story about intuitions will be pretty implausible.

This picture of intuitions as strong but fragile meshes well, I think, with the picture from section Section 5. There I said the important intuitions are the ones you barely notice or remember. That's because the intuitions are fragile; if you remembered them enough to argue about them (or experimentally test them), the fragility conditions had probably been triggered, and the intuition probably wasn't doing much argumentative work.[6]

But why not think that intuitions are so fragile that they have no use in any philosophical debate? This question deserves more space than I can give it, but here are three sketches of answers.

1. Intuitions might be valuable checks on theory, and might be resilient enough to perform a valuable checking role.
2. Just like heuristics have characteristic errors, it might be that careful reasoning has characteristic errors, and there are cases where our first impressions are more reliable. See Gladwell (2005) for a summary of some relevant evidence.
3. Somewhat surprisingly, there may be cases when it is best to trust the less reliable source. The case for this is a bit detailed, and not original to me, so I'll just include a brief footnote for those interested.[7]

[6] I'm simplifying a little here. My preferred position is that intuit*ed*s provide strong but fragile evidence, while intuit*ing*s provide weak but resilient evidence. The reason this is relevant is related to footnote 7.

[7] At one point in Ben Levinstein's doctoral dissertation, he considers whether there's a general rule for deciding which of two conflicting sources we can trust. There turns out to be very little in general one can say. In particular, *trust the more reliable source* turns out not in general to be good advice. If sources have characteristic errors, it might be that given what the two sources have said, it is better on this occasion to trust the less reliable source, because the verdicts the sources deliver provide evidence that we are seeing one of the characteristic errors of the more reliable source. It takes more space than I have here to fill in the details of this argument, and most of the details I'd include would be Levinstein's not mine. But here's the big conclusion. Assume that intuitions are often wrong, but rarely dramatically wrong. The reason for that is that heuristics are bad at getting things exactly right, and good at getting in the ballpark. And that careful reasoning is often right, but sometimes dramatically wrong. This is trickier to motivate, but I think true. Then when intuition dramatically diverges from theory, and we don't have independent reason to think that intuition is mistaken about the kind of case that's in question, we should trust the intuition more than the theory.

7 Some Lewisian Case Studies

I've described one kind of mental state that deserves the name 'intuition', and which could play a role in philosophical activity. But, as Cappelen presses, we have to work to convert that 'could' to a 'does'. Do we really rely in intuitive, or heuristic-driven, judgments about cases in analytic philosophy?

As Cappelen shows, the answer is "A lot less than you may have guessed." We argue a lot more than we intuit, especially about the famous cases.[8] The bit of analytic philosophy I'm most familiar with is David Lewis's corpus, and since that doesn't play much of a role in Cappelen's story, I'll illustrate his point with some examples from it.

Going from memory, I would have guessed the clearest example of a case refuting a theory was the use of finkish dispositions to refute the conditional analysis of dispositions. But go to the opening pages of "Finkish Dispositions" (Lewis 1997a), and you find not an intuition about a case, but an argument that finks are possible. And even though that argument is followed up with more cases, Lewis rather explicitly *argues* for his conclusions about each one. See, for example, the glass loving sorcerer on page 147. Lewis doesn't avert to an intuition that the loved glass is fragile, rather he "wield[s] an assumption that dispositions are an intrinsic matter." (Lewis 1997a, 147)

The discussion of causation turns out to be a little more fertile. From (the longer version of) "Causation as Influence", I count the following appeals to intuitions about cases.

- The chancy bomb example which shows simple probabilistic analyses of indeterministic causation won't work (Lewis 2004a, 79).
- The Merlin and Morgana example which shows that trumping is possible, and matters for what is the cause (Lewis 2004a, 81).
- The variant on Billy and Suzy that raises problems for quasi-dependence (Lewis 2004a, 83).
- The crazed President example which shows that causation by double prevention is possible, and that causation is not an intrinsic relation (Lewis 2004a, 84).
- The Frankfurt example which shows we can have causation without dependence (Lewis 2004a, 95).

There's a strong sense, I think, in which none of the judgments in these cases are argued for. Indeed, they arise as *problems* for theories that are otherwise doing rather well. If there was an argument around, it would be for the negation of the intuited judgment. So I think there's a role for intuition here.

But we should not imagine that this is normal for philosophers, or even for Lewis. Cases, it is true, play a large role in Lewis's writing. But they are very rarely simple refutations of existing theories. We could perhaps distinguish four roles that cases play, or perhaps four types of philosophical cases.

[8] There is interesting work to be done on the relative role of intuitions and arguments about *principles*, but I'm going to leave that for another day, and focus here on cases. The principles/cases distinction can be a bit slippery, but paradigm cases are easy to identify, and we'll be working with fairly paradigmatic cases here.

1. Refutation of theories, as in these causation cases.
2. Illustrations that help explain what's going on in an argument, as in the examples from "Finkish Dispositions". For a more extensive version of this, see Lewis's version of Puzzling Pierre (Lewis 1981).
3. Tools for showing that we must distinguish various concepts, such as the discussion of Ned Kelly's proof that there's no honest cop (Lewis 1988).
4. Simplified versions of the real world, on which we can test various explanatory hypotheses, such as the footy and rugby people in "Naming the Colors" (Lewis 1997b).

And that list is probably incomplete. The last is fairly fascinating as a case study actually.[9] Some of you may have had the following experience when programming, or indeed doing anything that looks like working with code (such as writing in LaTeX). A bug arises. It helps to find a minimal example in which the bug arises, i.e., a smallest program that produces the same bug. This helps you spot what's going on, and if you still need help, it helps your interlocutors focus on the central problem. It's important that you haven't changed the problem; the example must be of the same kind as what you started with. But the example could be much simpler than the case you're most interested in. Some philosophy examples are, I suspect, like that. Their value lies in revealing that some striking feature of reality would persist even if the world were simpler. So, probably, the explanation of the feature lies in some respect the real messy world shares with the simple example world. (Compare Cappelen's discussion of Perry's messy shopper in section 8.1.)

It is perhaps no coincidence that the easiest place to find examples of type 1 in Lewis's work is in the papers on causation. Lewis thinks there is no such thing as causation (Lewis 2004a, 2004b). Whatever our theory of 'causes' should be, it shouldn't match that verb with a binary property. Rather, the aim of philosophical work on causation is to give a reductive analysis of causal thought and talk. In such a project, judgments about how we use 'causes' are more likely to be central.

It's also not coincidental that when an example is central to a paper, such as the 'dishonest' cop and Puzzled Pierre, they really don't look like type 1. That's one big and important lesson from Cappelen's work. Philosophers do use examples to refute theories, but they are rarely the big famous examples. If an example is central to a philosophy paper, it typically plays one of the other three roles.

8 Summary

Let's take stock. I've argued for the following theses:

1. Socratic knowledge is important to philosophy.
2. The distinctive feature of philosophy is that it addresses questions that can, at least *prima facie*, be productively worked on while relying primarily on Socratic knowledge.

[9] See Sugden (2000, 2009) for much more on this use of thought experiments.

3. Intuitions are manifestations of cognitive skills, and much Socratic knowledge is constituted by the possession of such cognitive skills.
4. Like other forms of Socratic knowledge, intuitions are mostly *a posteriori*, and have roles outside philosophy as well as inside it.
5. Intuitions are default justified; that is, they can be unjustified justifiers.
6. This default is very weak; intuitions can easily be overridden by other considerations.
7. Relatedly, it is rare for any one intuition to be central to a philosophical work; philosophical intuitions mostly concern the little cases we see along the way to larger projects.

I also hinted at, without developing, an argument for

8. The right intuition can stop even a plausible theory dead in its tracks; and we have (thanks to Ben Levinstein) a mathematical model for why this can be so even if intuitions are much less reliable than theories.

I opened with a discussion of why it matters to philosophy's self-conception that point 4 is correct. Since Cappelen also endorses 4, I probably don't need to say more about that here. But I think there is more to say about 7.

The first thing I want to note is that 7 is of course consistent with Cappelen's textual research on important work in late analytic philosophy. In just about any thought experiment that you can remember, the intuitions about it don't carry much philosophical weight in the work in which it is introduced. The intuitions that matter are the little ones, the ones that go by so quickly that no one questions them and are largely forgotten by all but the *cognoscenti* in that field. Even these intuitions aren't *that* common. There are less of them in Lewis than I would have guessed.

Still, I disagree with Cappelen that philosophy is without these intuitions. And so I disagree that there's no role for double checking, experimentally if need be, whether these intuitions are really intuitive. If a well run survey showed that most people disagree with Lewis's judgment about, say, the chancy bombs example, I'd reconsider my views about probabilistic causation. But I'd be really surprised to see this.

The second thing to note is that while 7 is true, it's not the case that intuitions about one case are never central to a philosophical project. There is one big counterexample: the Gettier literature. Like Cappelen (194n3), I think this literature is incredibly unrepresentative of philosophy. And I think that's in part because it was methodologically flawed. I tried to make this point in an earlier paper (Weatherson 2003b), but I didn't get it quite right. (What I should have said was more like what Elijah Chudnoff (2011) does say.) When we saw the Gettier example, this should have been an invitation to try and find out what feature of knowledge was driving the fact that the belief in the main examples didn't amount to knowledge. Gettier suggested it was inference from a false premise, but that doesn't quite work (Warfield 2005). You might think it is insensitivity, but that doesn't quite work. At this point there should have been one of two paths taken - attempts to find some other explanation of the data, or a reconsideration

of whether our initial judgment about the case was wrong. That's what the picture of philosophy sketched here would have predicted, and (this is the point I was trying but failing to make in the earlier paper) that's what reflection on our successes in other areas of philosophy would have recommended. But the first kind of project ended up intertwined with attempts to analyse knowledge, and stalled for decades. And the second project wasn't seriously undertaken, with some honorable exceptions such as Sartwell (1992) and Hetherington (2001). Now eventually this didn't matter, because we discovered that safety based explanations of the Gettier case would work, even if there is no safety based analysis of knowledge, and even if there is some work to be done in getting the safety condition just right (Williamson 1994, 2000; Sainsbury 1995; Lewis 1996; Weatherson 2004). So if we strengthened 7 into a universal claim it would be false – thirty years of epistemological struggle attest to this. But it was really when epistemology fell into line with practice in other fields of philosophy that it made progress on the Gettier case.

References

Cappelen, Herman. 2012. *Philosophy Without Intuitions*. Oxford: Oxford University Press.

Carruthers, Peter. 1990. *The Metaphysics of the Tractatus*. Cambridge: Cambridge University Press.

———. 2011. *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford: Oxford University Press.

Chudnoff, Elijah. 2011. "What Should a Theory of Knowledge Do?" *Dialectica* 65 (4): 561–79. doi: 10.1111/j.1746-8361.2011.01285.x.

Devitt, Michael. 2011. "Experimental Semantics." *Philosophy and Phenomenological Research* 82 (2): 418–35. doi: ppr201182222.

Devitt, Michael, and Kim Sterelny. 1987. *Language and Reality: An Introduction to the Philosophy of Language*. Cambridge, MA.: MIT Press.

Fodor, Jerry. 2000. *The Mind Doesn't Work That Way*. Cambridge, MA: MIT Press.

Fodor, Jerry A. 1983. *The Modularity of Mind*. Cambridge, MA: MIT Press.

Friedman, Milton. 1953. "The Methodology of Positive Economics." In *Essays in Positive Economics*, 3–43. Chicago: University of Chicago Press.

Gladwell, Malcolm. 2005. *Blink: The Power of Thinking Without Thinking*. New York: Little, Brown.

Goldman, Alvin I. 1976. "Discrimination and Perceptual Knowledge." *The Journal of Philosophy* 73 (20): 771–91. doi: 10.2307/2025679.

Gopnik, Alison. 2009. *The Philosophical Baby: What Children's Minds Tell Us about Truth, Love, and the Meaning of Life*. New York: Farrar, Straus; Giroux.

Harman, Gilbert. 1973. *Thought*. Princeton: Princeton University Press.

———. 1986. *Change in View*. Cambridge, MA: Bradford.

Hetherington, Stephen. 2001. *Good Knowledge, Bad Knowledge: On Two Dogmas of Epistemology*. Oxford: Oxford University Press.

Ichikawa, Jonathan, Ishani Maitra, and Brian Weatherson. 2012. "In Defence of a Kripkean Dogma." *Philosophy and Phenomenological Research* 85 (1): 56–68. doi: 10.1111/j.1933-1592.2010.00478.x.

Kahneman, Daniel. 2011. *Thinking Fast and Slow*. New York: Farrar, Straus; Giroux.

Kilkarni, Sanjeev, and Gilbert Harman. 2011. *An Elementary Introduction to Statistical Learning Theory*. Hoboken, NJ: Wiley.

Klein, Gary A. 1999. *Sources of Power*. Cambridge, MA.: MIT Press.

Lewis, David. 1981. "What Puzzling Pierre Does Not Believe." *Australasian Journal of Philosophy* 59 (3): 283–89. doi: 10.1080/00048408112340241. Reprinted in his *Papers in Metaphysics and Epistemology*, Cambridge: Cambridge University Press, 1999, 408-417. References to reprint.

———. 1988. "The Trap's Dilemma." *Australasian Journal of Philosophy* 66 (2): 220–23. doi: 10.1080/00048408812343301. Reprinted in his *Papers in Ethics and Social Philosophy*, Cambridge: Cambridge University Press, 2000, 95-100. References to reprint.

———. 1996. "Elusive Knowledge." *Australasian Journal of Philosophy* 74 (4): 549–67. doi: 10.1080/00048409612347521. Reprinted in his *Papers in Metaphysics and Epistemology*, Cambridge: Cambridge University Press, 1999, 418-446. References to reprint.

———. 1997a. "Finkish Dispositions." *The Philosophical Quarterly* 47 (187): 143–58. doi: 10.1111/1467-9213.00052. Reprinted in his *Papers in Metaphysics and Epistemology*, Cambridge: Cambridge University Press, 1999, 133-151. References to reprint.

———. 1997b. "Naming the Colours." *Australasian Journal of Philosophy* 75 (3): 325–42. doi: 10.1080/00048409712347931. Reprinted in his *Papers in Metaphysics and Epistemology*, Cambridge: Cambridge University Press, 1999, 332-358. References to reprint.

———. 2004a. "Causation as Influence." In *Causation and Counterfactuals*, edited by John Collins, Ned Hall, and L. A. Paul, 75–106. Cambridge: MIT Press.

———. 2004b. "Void and Object." In *Causation and Counterfactuals*, edited by John Collins, Ned Hall, and L. A. Paul, 277–90. Cambridge: MIT Press.

Machery, Eduoard, Ron Mallon, Shaun Nichols, and Stephen Stich. 2012. "If Folk Intuitions Vary, Then What?" *Philosophy and Phenomenological Research* 86 (3): 618–35. doi: 10.1111/j.1933-1592.2011.00555.x.

Nagel, Jennifer. 2007. "Epistemic Intuitions." *Philosophy Compass* 2 (6): 792–819. doi: 10.1111/j.1747-9991.2007.00104.x.

———. 2013. "Defending the Evidential Value of Epistemic Intuitions: A Reply to Stich." *Philosophy and Phenomenological Research* 86 (1): 179–99. doi: 10.1111/phpr.12008.

Nesbø, Jo. 2009. *The Redeemer*. London: Vintage Books.

Nolan, Daniel. 2005. *David Lewis*. Chesham: Acumen Publishing.

Pryor, James. 2000. "The Sceptic and the Dogmatist." *Noûs* 34 (4): 517–49. doi: 10.1111/0029-4624.00277.

Ruetsche, Laura. 2011. *Interpreting Quantum Theories*. Oxford: Oxford University Press.

Sainsbury, Mark. 1995. "Vagueness, Ignorance and Margin for Error." *British Journal for the Philosophy of Science* 46: 589–601. doi: 10.1093/bjps/46.4.589.

Sartwell, Crispin. 1992. "Why Knowledge Is Merely True Belief." *Journal of Philosophy* 89 (4): 167–80. doi: 10.2307/2026639.

Schwarz, Wolfgang. 2009. *David Lewis: Metaphysik Und Analyse*. Paderborn: Mentis-Verlag.

Sterelny, Kim. 2012. *The Evolved Apprentice: How Evolution Made Humans Unique*. Cambridge, MA.: Bradford.

Sugden, Robert. 2000. "Credible Worlds: The Status of Theoretical Models in Economics." *Journal of Economic Methodology* 7 (1): 1–31. doi: 10.1080/135017800362220.

———. 2009. "Credible Worlds, Capacities and Mechanisms." *Erkenntnis* 70 (1): 3–27. doi: 10.1007/s10670-008-9134-x.

Tribe, Kevin. 2002. "The Cambridge Economics Tripos 1903–55 and the Training of Economists." *The Manchester School* 68 (2): 222–48. doi: 10.1111/1467-9957.00191.

Warfield, Ted A. 2005. "Knowledge from Falsehood." *Philosophical Perspectives* 19: 405–16. doi: 10.1111/j.1520-8583.2005.00067.x.

Weatherson, Brian. 2003a. "Many Many Problems." *The Philosophical Quarterly* 53 (213): 481–501. doi: 10.1111/1467-9213.00327.

———. 2003b. "What Good Are Counterexamples?" *Philosophical Studies* 115 (1): 1–31. doi: 10.1023/A:1024961917413.

———. 2004. "Luminous Margins." *Australasian Journal of Philosophy* 82 (3): 373–83. doi: 10.1080/713659874.

Whewell, William. 1840. *The Philosophy of the Inductive Sciences, Founded Upon Their History*. London: John W. Parker.

Williamson, Timothy. 1994. *Vagueness*. Routledge.

———. 2000. *Knowledge and its Limits*. Oxford University Press.

Published in *Philosophical Studies*, 2014, pp. 517-533.