1. Introduction
The present paper will be in three parts: the first part concerns the reliability of intuitions, the second part concerns the idea of philosophy as conceptual engineering, and the third deals with certain possible limits to the idea of conceptual engineering, what I will call conceptual fixed points. Let me begin with a roadmap.

In recent years, philosophers have taken a greater interest in what intuitions are supposed to be, and whether they are reliable. Some of these philosophers are self-styled rationalists, seeking to defend reliance on intuitions; others seek to debunk the practice of relying on intuitions. Prominent in the latter camp are experimental philosophers, who have conducted empirical studies that, among other things, appear to show that intuitions vary between groups. One thing I am concerned to stress in this paper is that both friends and foes of intuitions overlook important distinctions between types of intuitions. Roughly, the idea is this. The reliability of some intuitions, what I will call competence intuitions, can be defended by appeal to how, if we didn’t have these intuitions, we wouldn’t have the concepts we have. But one cannot defend the reliability of intuitions in this way across the board. Distinguishing between different kinds of things that can full under the heading of intuition is also relevant to evaluating skepticism as to whether there even is such a thing as intuition.¹

The discussion up until this point (through section 3) will have largely amounted to a defense of competence intuitions, and the neat way to continue the discussion would have been to defend a conception of philosophy as conceptual analysis, an enterprise we successfully engage in by consulting competence intuitions. But for one thing, even if there are competence intuitions and they are reliable, such intuitions are only potentially relevant to a small subset of philosophical questions. Moreover, and as I will discuss at greater length (section 4), even in cases where the methodology of relying on competence intuitions seems the most promising, competence intuitions only give word of features of our actual concepts – our actual concept of knowledge, or our actual concept of meaning, or… – or features of the properties that these concepts stand for. And surely it would be rather

¹ Many thanks to Christopher Daly, David Liebesman, Folke Tersman, an audience at the University of Vermont, and the participants in a workshop on conceptual revisionism at the University of Oslo for helpful feedback.

¹ See especially Cappelen (2012) for such skepticism.
miraculous if, say, our actual concept of knowledge was the most useful and appropriate for the epistemic theoretical purposes to which this concept is put. Compare perhaps: the folk concept of weight is not the concept best suited to theorize about the relevant physical phenomena. Philosophy should rather be thought of as conceptual engineering. We should consider what the best concepts to employ are like. But this means that, unfortunately, we cannot conceive of philosophy as just relying on competence intuitions. The question of what we draw on when engaged in conceptual engineering is a separate one.

The easy route would have been to stop here: philosophy is, or ought to be, conceptual engineering. But again (section 5) there are complications. In some cases, I argue, there are principled difficulties applying the idea of conceptual engineering; some concepts are, as we may put it, conceptual fixed points. (I say “difficulties”, and I say so advisedly. What I have to say, even if correct, does not immediately show that conceptual engineering is impossible in these cases; it shows only something rather more subtle.) That is the summary. Let me now turn to execution, starting with questions about the reliability of intuitions.

2. Kinds of intuitions
There has recently been increased interest in the question of the reliability of intuitions. In part this is related to a general interest in meta-issues and the methodology of philosophy. In part this relates to the surge of so-called experimental philosophy. Eventually I will seek to relate to the empirical studies done under the heading of experimental philosophy. But first I will devote some time to describing the lay of the land from a theoretical perspective and to discussing in general terms how empirical studies might in principle be relevant.

While there is a tendency in the literature to speak in general terms about ‘intuition’, and to speak generally about kinds of intuitions, there are many different phenomena that tend to be grouped together under that heading. Consider the following list of purported intuitions:

(1) That “three philosophers carried two pianos” is ambiguous
(2) That “three philosophers two carried pianos” is not well-formed
(3) That such-and-such a mathematical proposition is provable.
(4) That it is possible to trap the opponent’s queen.
(5) That if a man is a bachelor then he is unmarried.
(6) That “Feynman”, as used by a person who associates with “Feynman” only the property of naming a famous physicist, refers to Feynman.
(7) That assuming the watery stuff around us is H2O, watery stuff with different chemical composition found on a different planet is not water.
(8) That modus ponens is valid.

(9) That the subject in a Gettier case does not have knowledge.

(10) That it is wrong to push the fat man in the relevant trolley case.²

(11) That there is no object composed of Rand Paul’s nose and the Eiffel Tower.

Some of these cases, like cases (3) and (4), don’t have immediate bearing on philosophy at all. There are, to be sure, things going on in philosophy which are in some ways parallel, like a philosopher’s sense that there are counterexamples to be found to a given theory. Some philosophers undeniably have more of a nose for that sort of thing than others, and are more apt at making judgments about which theories there are counterexamples to - given accepted methodology regarding how to evaluate whether something is a counterexample. But a skeptic about the use of intuition in philosophy will of course question the basis of this accepted methodology.

Some items on the list, (5) and (6), might sound odd to classify as “intuitions”. The Feynman case is one of the cases famously used in Kripke (1980) to argue against the description theory. Kripke’s arguments against the description theory are routinely described as intuition-based, but what are focused on are the more fanciful scenarios concerning Gödel, not the Feynman point. (In the Gödel case we are supposed to imagine that the man we knew of as “Gödel” was an imposter and the incompleteness theorem was proved by a man who went by the name of “Schmidt”. The question is which of these men our name “Gödel” would semantically refer to. Kripke says it would be the former, while the description theory’s prediction is that it would be the latter.)

The linguistics-oriented examples (1) and (2) can be relevant for philosophers’ theories about language, insofar as philosophers as much as linguists might be in the business of giving semantic theories for natural languages. But one thing that sets them apart from typical philosophy-related examples is that the objects of the intuitions – what is judged – is overtly metalinguistic.

Turn now to the most paradigmatically philosophical of the examples, (7)-(11). I believe there is an important difference between (7) and (8) on the one hand and (10)-(11) on the other, with (9) occupying a middle ground. (I would classify (9) along with (7) and (8) but there is a complication making this classification less than obvious.) Suppose (for now) that our intuitions about these cases are fairly uniform. I believe that this supposition would provide good grounds for taking intuitions (7) and (8) to be veridical in a way in which it does not provide good grounds for taking (10)-(11) to be veridical.

² This version of the trolley case is due to Judith Thomson. A trolley is hurtling down a track towards five people. You are on a bridge under which the trolley will pass, and you can stop it by dropping a heavy weight in front of it. As it happens, there is a very fat man next to you. Your only way to stop the trolley is to push him over the bridge and onto the track. You then kill him but in doing so you save the five people. Should you proceed?”
Focus on (7). And compare us with a different possible community, one whose members uniformly have different intuitions from us in this case. The members of this community use “water” in much the way that we do in other respects, but they have an intuition that they give voice to by saying things like “the XYZ on this other planet would be water”, and this intuition is stable in the sense that they retain the intuition under conditions of improved information. I will mention some complications shortly, but I think a natural verdict would be that the people in this community simply mean something (slightly) different by “water” from what we do. Generally, there are different possible meanings for “water” to have, and the fact that we have the intuitions we have is linked to the fact that given how we are disposed to use “water” we use it with such a meaning as to make our sentence “the XYZ on this other planet wouldn't be water” true, and their having the intuitions that they have is naturally seen as being linked to the fact that given how they are disposed to use “water” they use it with such a meaning as to make their sentence “the XYZ on this other planet would be water” true.

There are here the materials for an explanation of why the intuitions in the relevant class are reliable. They are reliable because a community’s having the intuitions they have is linked to their thinking the contents they do and using sentences with the meaning they do. It is, partly, by virtue of our dispositions to use “water” they way we do that “water” means something such that our sentence “the XYZ on this other planet wouldn’t be water” expresses something true, and this explains why our intuitions, insofar as they accurately reflect these dispositions to use, are reliable. Call intuitions whose reliability can be explained this way competence intuitions: (This is just a label. I don’t mean to pack much into talking about “competence” here. See below.)

If we and some other community stand in the corresponding relation with respect to cases (10)-(11) the verdict would be different. Focus on (10). Suppose we uniformly have the intuition that it would be wrong to push the fat man, but there is a different linguistic community, supposed to be such that they uniformly have different intuitions from us in this case. The members of this community use “wrong” in much the way that we do in other respects, but they uniformly give voice to what they intuit by saying things like “pushing the fat man would not be wrong”. I don’t think the natural verdict would be that the members of this other community mean something other by “wrong” than what we do. Nor do I think this would be the natural verdict even if their differing intuition was about general principles rather than about a specific case like that of the fat man on the bridge. Similar remarks apply to the case of a community supposed to be such that they have different intuitions than us about (11). If (10)-(11) differ from (7)-(8) in this way, then the offered explanation of the reliability of (7) and (8) does not generalize to (10) and (11). Call intuitions pertaining to philosophical matters whose reliability cannot be explained in the way the reliability of competence intuitions can rational intuitions. (“Rational” to suggest a certain kind of rationalism. But
again, this is a mere label.) The reason (9) is hard to classify is that “knows” arguably is normative: it conveys some particular kind of positive epistemic assessment. This means that there might be some pressure to classify it alongside paradigmatic substantive normative intuitions such as (10).

Suppose that there indeed is this difference between cases like (7) and (8) and cases like (10)-(11) that I have argued there is. What would explain this difference? I will first present what I believe to be a much simplified account: intuitions are sometimes concept-constitutive in the sense that possessing the concept in question requires one to have the intuition. If one didn’t have the intuition, one wouldn’t have the concept. Focus on (8). The suggestion might be that if one didn’t have the intuition that from \( P \) and \( P \to Q \), \( Q \) follows, then one’s concept \( C \) wouldn’t be a concept of the conditional. Further, if this rule for the conditional (or the intuition that it is valid) is in this way constitutive of the concept of the conditional, then the concept’s semantic value is such as to make this rule truth-preserving. If the account just given is correct, then there is an easy explanation available of the reliability of the intuition that this rule is valid: that the thinker has this intuition helps ensure that the intuition has its content something true. By contrast, in the case of (10)-(11), the intuitions in question aren’t concept-constitutive.

Temporarily accepting that some intuitions have the status of being concept-constitutive, one might next go on to ask what explains why some intuitions have this status and some don’t. This is a natural question to ask, but I am not sure that the friend of the account in question needs to admit that it has a very interesting answer. Perhaps no deeper explanation can be given than one that appeals to how we simply treat some intuitions as concept-constitutive and others not. Further to illustrate: The friend of the concept-constitutivity idea could allow that there is a community making judgments not on the face of it different from our judgments in (7)-(11), but where – simply because of differences in what they take to be concept-constitutive – use different concepts from ours. Maybe they take the preparedness to make certain particular positive existence claims to be constitutive of having their concept of existence. The friend of the line we are now considering can say that they then simply use a different concept of existence from ours, and that the relevant judgments of theirs have different contents from ours.

There are a number of ways in which what has been sketched regarding competence intuitions is too simple. First, even in the most favorable cases it is natural to think there are scenarios where someone fails to have the intuition but still can have the concept in question.\(^3\)

\(^3\) As has been prominently stressed by Williamson, e.g. in his (2007).
with that simple model. The to my mind undeniable kernel of truth underlying the simple account just canvassed is that meaning is determined by use. The meanings of linguistic expressions are determined by how they are used; and the contents of someone’s concepts is determined by facts about what goes on in her mind and how she acts. How exactly this happens is of course a vexed issue. But for whatever reason, and in whatever way, certain aspects of how an expression is used are more closely linked to the meaning the expression has than other aspects are; the expression’s meaning is stable under some changes in its use but not under others.

There are complications having to do with how intuitions relate to actual dispositions to use. One’s intuitions need not match ones actual dispositions to use, and intuitions can for that reason be at a remove from what constitutes meaning. Suppose some community of thinkers is such that in some cases what the members of this community judge concerning a given hypothetical situation differs from what they actually would say when in that situation. For example, they say of subjects in hypothetical Gettier cases (cases where the subjects have justified true belief but, according to philosophical orthodoxy, do not have knowledge) that they do have knowledge; but in fact, in such a situation they would deny that the subject has knowledge. Then it seems the latter judgments are a better guide to what they actually mean than their judgments about hypothetical situations. Given how intuitions are discussed in the literature, it is more common and natural to call judgments about hypothetical situations ‘intuitions’ than it is to call a speaker’s judgment, in a real Gettier situation, that the thinker who has justified true belief does not have knowledge an ‘intuition’. If we talk this way about intuitions one can say: what has just been suggested is that intuitions are not as closely tied to meaning-determination as are judgments made in actual speech situations. However, while this is a genuine complication, I doubt that it matters much in practice; for I doubt that there are widespread mismatches of this kind. Another complication regarding how intuitions relate to what determines meaning is the following. Intuitions are supposed to be characterized by a certain kind of immediacy. But what about when there is a mismatch between our immediate judgments and what we arrive at after reflection? This is a hard question. There may be some reason to think reflective judgments trump unreflective intuitions. One’s view upon reflection better reflects one’s true outlook on things. On the other hand, consider a community whose members are happy to say “it is possible for there to be married bachelors”. If we find out about this community that its members have the same intuitive reaction to the question of whether ‘bachelors’ can be married as we do but say what they say because they have been persuaded by theoretical arguments to the effect that ‘bachelors’ can indeed be married, contrary to what intuitions say, we might be more inclined to take this community to mean bachelor by ‘bachelor’. I won’t attempt to resolve the question of what should take priority. The present point is that there is a question of whether to give priority to reflective judgments or
unreflective intuitions, and if one gives priority to the former then intuitions are not strictly reference-determining.

A different sort of complication is that one must distinguish between on the one hand an argument to the effect that some intuitions are guaranteed to be intuitions of something true or other and on the other hand an argument, of some specific contents that we intuit, that those contents are true. The suggested defense of the reliability of competence intuitions at least in the first instance only provides the former.

The critical discussion of intuitions found in the work of experimental philosophers has tended to focus very much on intuitions in epistemology; for example intuitions about Gettier cases and the Truetemp case. Such cases are arguably not the best ones to focus on in a discussion of methodology, for they present special problems of their own. On the one hand, our intuitions about these cases can seem like competence intuitions; on the other hand, epistemological concepts are normative concepts, and as I mentioned, paradigmatic normative intuitions are paradigmatic examples of rational intuitions, given the distinction between competence intuitions and rational intuitions. It could be that good arguments could be given for a given classification of the relevant intuitions. But however that may be, the point remains that there are important distinctions to be drawn within the class of intuitions, and it only obscures things to focus on cases that are hard to classify.

Before continuing I should note that even examples that might appear to be clear cases of rational intuitions can be classed as cases of competence intuitions. Eli Hirsch has prominently argued that friends of what on the face of it appear to be different ontological views are just using different concepts of existence; their different judgments about ‘what there is’ reflects their using different concepts. Frank Jackson’s neo-descriptivist view on ethical discourse arguably yields that ethical intuitions are best thought of as what I have here called competence intuitions. A very ambitious claim would be that all manner of intuitions relied upon in philosophical debates can be classed as competence intuitions: their reliability can always be defended in the way sketched above. When offering (7) and (8) as examples of one kind of intuition and (10) and (11) as examples of another, I meant only to operate with some natural classifications. Nothing in what I have said hinges upon which classifications are in the end correct.

3. The importance of the distinctions

---

4 See e.g. Weinberg et al (2001) and Swain et al (2007). The Truetemp case is a famous counterexample to reliabilism, due to Keith Lehrer (1990).
5 See the essays collected in Hirsch (2011).
I believe the distinctions drawn above are of relevance to the kind of criticism of reliance on intuitions that experimental philosophers have presented. This criticism is based on studies that seem to show that among ordinary speakers, intuitions vary considerably and seem to a considerable extent to be determined by things like cultural background.7

Exactly how is this kind of criticism relevant to reliance on intuitions? What I have here called rational intuitions are often taken to be problematic quite independently of the empirical findings. Even if our intuitive judgments were in agreement with each other when it comes to substantive matters in ontology and ethics, reasonable skeptical questions about the reliability of rational intuition could be raised. The criticism is the most relevant when it comes to what I have called competence intuitions. If there are the kinds of variations in intuitions that the empirical studies appear to point to, that does present genuine problems concerning reliance on competence intuitions. But given the link between competence intuitions and the meaning-determining facts, general, principled skepticism about competence intuitions is out of the question. I will illustrate this point by discussion of one of the more theoretically sophisticated experimentalist discussions of intuition, Weinberg (2007). Weinberg’s target is intuition generally, and his work is meant to bolster the skepticism of intuition associated with experimental philosophy, but if one makes use of the distinction between rational and competence intuitions, central points Weinberg makes lose their force.

Weinberg says that “intuitions are odd critters”, and approvingly refers to Robert Cummins (1998) making the point that “[o]ur best possible accounts of where intuitions can come from do not square well with our hopes that they are any sort of reliable guide to a truth beyond themselves.”8 Weinberg’s article is focused on a comparison between intuition and perception. Some arguments for skepticism about intuition appear to generalize so as to, if sound in the first place, yield also skepticism about perception. Weinberg seeks an argument for skepticism about intuition that does not so generalize. The argument he proposes is centered on the notion of hopefulness, where “A source of evidence that is not practically infallible is hopeful to the extent that we have the capacity to detect and correct its errors”.9 He argues that intuition is hopeless but perception is not.

What Weinberg is concerned with here is – in part – something that parallels what other authors have said in discussions specifically focused on slightly different matters. Here is Crispin Wright, from an article on modality, on a case (supposedly) analogous to that of modality:

Why does the suggestion that the sense of humour is a cognitive faculty seem so outlandish? At bottom, it is because our conception of ourselves and of our knowledge-acquiring powers is

---

7 For some paradigmatic and prominent criticisms of this kind, see Mallon et al. (2009) and Weinberg et al. (2001).
broadly naturalistic. We are content to regard something as a cognitive faculty only so long as there is the promise of a proper explanation of its physiological basis, an account which enables us to see its output as the product of physical interaction between our bodies and the environment. There simply is no such promise in the case of amusingness; we do not have the slightest idea what the amusingness of a situation could physically consist in (contrast redness and the emission of light waves of such-and-such frequency).10

The concerns that Wright raises are not directed against a target like ‘intuition’; rather, they are about knowledge of a supposed special region of reality, the region postulated by the metaphysical realist about modality. That said, the concerns they raise can be raised concerning rational intuition, given what rational intuition is supposed to be used for. We lack a satisfying explanation of how rational intuition could be “adequate to its task” – that is, detecting things like fundamental normative and ontological facts.

Insofar as Weinberg is concerned to argue that rational intuition is hopeless, I am somewhat inclined to be sympathetic. But importantly, a successful argument to the effect that rational intuition is hopeless raises many questions. We do seem to have knowledge of some things such that philosophers have postulated rational intuition to explain our knowledge of those things. The skeptic about rational intuition seems to owe us either an alternative explanation of our seeming knowledge or to defend the view that we do not really have knowledge of those things.

However, Weinberg’s target is intuition generally, and also what I have called competence intuitions clearly fall under the purview of his argument. The sorts of empirical studies he mentions present prima facie problems also for competence intuitions, and specific arguments for taking intuition to be hopeless that Weinberg relies on are relevant also to the case of competence intuitions. He offers three specific reasons for taking intuition to be hopeless. (a) We do not know when intuition can be trusted, for example when intuitions about what to say concerning highly unusual situations can be trusted. (b) We lack an understanding of how intuitions work, where what Weinberg means by this is we lack an understanding of the psychological processes underlying intuition. (c) We lack means of resolving cases where our intuitions don’t agree with each other, and assumptions to the effect that our intuitions converge upon reflection lack justification.11 In addition to making points (a)-(c), Weinberg also stresses that intuitions do not enjoy much by way of external

10 Wright (1986), pp. 206-7. (Compare too Yablo (1993), p. 3f.) I think Wright is correct in saying that the claim that a sense of humor is a cognitive faculty would be outlandish. But he does not well pinpoint what makes it so. He says, “we do not have the slightest idea what the amusingness of a situation could physically consist in”, but the same goes for all sorts of high-level properties instantiated in the natural world. What does, say, a recession, or a cold war, physically consist in?
corroboration (although he also mentions some exceptions – cases where intuitions do enjoy such corroboration, like logic and mathematics). Each of these points applies equally well to competence intuitions.

How can we evaluate the case for the hopelessness of competence intuitions? It is worth comparing a different case, and using that case to draw lessons regarding Weinberg’s discussion. It is a relatively popular view that some concepts are response-dependent: that some concepts C are such that something is C iff we would (under suitable conditions) judge that it is C, and that somehow the right hand sides of these biconditionals are prior in the order of explanation. Wright’s example, the comic, is something sometimes brought up here: the concept of being funny is sometimes treated as one plausible example of such a concept. Like any other example, this can be debated; but let us work with it. Given a response-dependence account of what is funny, there is no principled mystery concerning how our judgments about what is funny can track the facts about what is funny. Indeed, and relating to the quoted passage from Wright, one attractive feature of response-dependence accounts is that they promise to avoid epistemological problems that otherwise arise. But Weinberg’s worries regarding intuition arise equally for our judgments concerning what is funny, even given a response-dependence account. Weinberg’s (b) and (c) have rather obvious analogues. We arguably don’t know much about what psychological processes underlie our judgments about what is funny. And we certainly make different judgments about what is funny, and there is no ready process for deciding who is right and little justified optimism that our judgments about what is funny will converge after reflection. Weinberg’s (a) does not in the same way have an obvious analogue. We do not often consider what would be some highly unusual cases of funniness. But it is natural to speculate that our judgments would differ significantly concerning highly unusual cases, and that there would be no ready means to decide whose judgments were correct.

But in response-dependence cases the analogues of (a)-(c) do not appear very worrisome. Even if we are ignorant of the psychology underlying judgments about what is funny, that doesn’t undercut reliance on these judgments when we figure out what is funny. Likewise the unavailability of a ready means or settling disputes is not a concern. And in the case of response-dependent concepts it is not clearly relevant to ask for such corroboration. If our judgments about what is funny play the role in determining what is funny that they do on a response-dependence account, it is unclear what would be asked for if an external corroboration of those judgments were demanded.

These points are of relevance to what Weinberg says about intuition. Weinberg’s case against intuitions, including competence intuitions, generalizes to where there is an obvious justification for reliance on our judgments: cases where our judgments are constitutive of the facts we are dealing with. We seem to be able to know the relevant facts, and there is no independent method that would work better. Moreover, while there are differences between the case of judgments about what is F,
where F is a response-dependent concept, and the case of competence intuitions – see the discussion of competence intuitions early on – the cases are sufficiently similar that assimilating the two cases is not entirely frivolous. Although, for reasons noted earlier, it is not entirely accurate to say that having the right intuitions is a necessary condition on competence, it is sufficiently close to accurate that the lessons from the discussion of response-dependence apply.\textsuperscript{12}

All this said, however, there is an obvious point to be made. Even given that the distinction between competence intuitions and rational intuitions can be satisfactorily spelled out and defended, the results of the experimental philosophers do present some problems for philosophers’ reliance on competence intuitions. Even if, modulo the important complications noted, competence intuitions are meaning-determining, the relation between meanings and individual intuitions is a complex and thorny one, and there can be all sorts of reasons why individual intuitions do not accurately reflect meaning. This should make us cautious about drawing immediate conclusions from individual intuitions. That these complications regarding competence intuitions remain does not mean that the above discussion responding to Weinberg is otiose. The important lesson is this. There is world of difference between the charge that intuition is not to be trusted since it is a mystery how an intuition can be sensitive to the facts it is about, and the charge that because of variability and shiftiness in intuitions one must be careful about drawing general conclusions from individual intuitions. (Reasons have been presented in the literature for why the relevant results of the experimental philosophers should be taken with a grain of salt.\textsuperscript{13} I do not want to enter into that debate here. All I am concerned with is the question of what kind of relevance the work of the experimental philosophers in principle can be taken to have.)

These points having been made, still of course the question remains: to what extent do the problems raised concerning competence intuitions present serious problems regarding philosophers’ reliance on such intuitions? There some different suggestions I wish to make here. (And when it comes to the points to follow, they are equally relevant to reliance on rational intuitions as they are to reliance on competence intuitions.)

First, while intuitions form a starting point for philosophical inquiry, overall fit with intuitions is arguably not all that matters. Rather, what happens can – no doubt very simplistically – be represented as follows. Theories are built that try to respect some given intuitions. These theories are also evaluated by their theoretical virtues. If the theories that seek to respect some given intuitions

\textsuperscript{12} I am a bit uncertain about Weinberg’s intentions – whether he wants to argue for principled skepticism about intuition, or only that there are methodological problems with our current practice. Different aspects of his text point in different directions. The important point is this: regardless of what Weinberg’s own intentions are, it is important to distinguish clearly between the kind of case that can be mounted against reliance on rational intuitions and the kind of case that can be mounted against reliance on competence intuitions.

\textsuperscript{13} See, for example, Cullen (2010), Kauppinen (2007), and Martí (2009).
are not virtuous enough, the task rather becomes one of explaining away these intuitions. Theory choice is a more holistic affair than any simple view on the relevance of intuitions would have it. This does not mean that intuitions have no role to play. At the same time, shiftiness is not immediately a great problem, for intuitions provide only some of all the relevant considerations. (My main concern is to defend what I called competence intuitions. But the friend of rational intuitions could adduce the same considerations, as far as the present issue is concerned.) Second, the problems would be more worrisome for what is regarded as traditional methodology if this methodology were that of taking all that matters to the acceptability of a given theory to be overall fit with intuitions. But even then, their significance ought not to be overstated. When one thinks of Kripke’s case against descriptivism and for a causal view on reference as relying on intuitions, one thinks of his claims regarding the Gödel/Schmidt case and not of his claims to the effect that “Feynman”, as used by an ordinary speaker without identifying knowledge of Feynman, refers to Feynman. There is a general point to be made here. As Williamson (2007) says, “Philosophical controversy will naturally make the unclear cases salient. That should not blind us to the wide range of clear cases (talking donkeys are possible)”.

A related response to the problems posed by experimental philosophers’ data concerning the variation in competence intuitions is that the actual philosophical methodology is fundamentally one of comparison and not merely of assembling intuitions about particular cases. Let me explain. If, considering a description of a given case, philosopher X has a given intuition, that is just a starting point in the process. The next step is to ask a number of questions about this reaction, such as: if one has that should say what X’s intuition says about this case, what does that seem to commit one to with respect to other cases? might X’s having the intuition she has be a matter of reading more into the description of the case than is strictly warranted? or might X’s having the intuition she has be a matter of her not being sufficiently attentive to a given detail, whether it’s a matter of simply overlooking the detail or a matter of not seeing its importance?

Here is an illustration. When people are presented with a teletransportation case, of the kind often discussed since Parfit (1984), people’s initial judgments concerning whether one survives teletransportation naturally differ. Suppose that your initial judgment is that one does not really survive teletransportation, but others make a different judgment. When engaging in a philosophical dispute, you will not simply let the variation in intuitions be the end of discussion. One thing you may do is to seek to influence the views of your opponents by comparing other cases. When it comes to this matter, there is, famously, a particular type of different case that can be used to this effect: teletransportation with branching, where at the end of the process we have two people

---

psychologically continuous with the original person. One’s judgment about this case can reasonably affect one’s judgment about the initial case. Consideration of branching makes vivid facts about the relation of psychological continuity – that a person at t1 can be fully psychologically continuous with more than one person at a later time t2 – such that when these facts are vivid one can be less inclined to say that one survives ordinary teletransportation. Return now to the experimental philosophers’ surveys indicating that there is great variety in intuitions about cases. What these surveys show is that there is great variety in initial reactions to these cases. To focus only on the initial reactions is to overlook the process of philosophizing. (Cappelen (2012) seeks to cast doubt on the idea that philosophy relies on intuitions. One point he seeks to bring home is that what one finds in canonical texts of analytic philosophy is not intuition-mongering but careful argumentation. What I see as right is encapsulated in what I have just stressed about the process of philosophizing. But one cannot disregard questions about the input to the process and doubts concerning the reliability of these sources.)

This last lesson takes us back to Weinberg’s discussion. I mentioned Weinberg’s points (a)-(c) above, but did not say much about how Weinberg attempts to support these points. Let me pause on (c), that we lack means to resolve cases where our intuitions don’t agree and are not justified in assuming that our intuitions converge upon reflection.

Weinberg supports the latter contention by citing the studies showing that there is variability in intuition, saying “we have little reason to expect to mitigate the fallibility of intuition by appeal to intersubjective agreement”. As against this my immediate protest is that the way to resolve disagreements isn’t by seeing whose intuition has the most support by others – by a kind of vote in effect – but by going through the process of reflection described above. In relation to this it is noteworthy how Weinberg argues his case:

…we may suspect that the uniformity of philosophers’ judgments about personally important cases is an artifact of their theoretical commitments, and not the product of their “pretheoretic” intuitive capacities. This need not be so; perhaps thinking hard about a domain puts one in better touch with one’s “real” intuitions. But this is a mere empirical possibility, and there is some evidence that tells against it.

---

15 My own view happens to be that one does actually survive ordinary teletransportation, and that it’s important here to keep in mind that also physical criteria of identity allow for branching cases. But however it may be with that, the points made in the main text remain. The crucial point is that the dialectical move described is a reasonable one, and characteristic of discussions where intuitions about cases vary. That further moves can be made is a different matter.
Strikingly, Weinberg speaks of the possibility of resolution of cases where intuitions don’t agree as if it is a matter of changing the intuitions (as a matter of getting in touch with one’s real intuitions), and as if favoring the idea that such resolution can be achieved by reflection is bound up with the idea that the intuitions of reflective experts are more reliable. But emphasizing the process of reflection, as illustrated above, is a different matter. It is that of emphasizing what is done with an unreflective judgment in the process of philosophizing. Resolution needn’t be a matter of getting intuitions to agree but that of arriving at a reasoned judgment in face of conflicting intuitions. What reflection offers is not in the first instance better intuitions but instead better considered views in light of the intuitions.

4. What to focus on?
One theme thus far has been that of defending one kind of intuition, competence intuitions, from the attacks of critics, like experimental philosophers. It would be nice to go on to present a picture of philosophy according to which philosophy tends to rely on competence intuitions, and should rely on competence intuitions. But the take home message cannot be this nice simple one. First, even if there are competence intuitions and they are reliable, it is a further question in how many kinds of cases discussed in philosophy competence intuitions provide significant evidence. This is an important point. I will not pause much on it. There is a second reason for resisting the picture of philosophy just presented, and to that I now turn.

On one common – and often implicit – way of looking at what much of philosophy is doing, we are studying our ordinary concepts: our concept of truth, our concept of knowledge, our concept of justice, and so on and so forth. Sometimes the ‘our’ is not explicit: the target is described as the concepts of truth, of knowledge, of justice, etc. But these are our concepts. The topic of the study is these concepts and not some variant concepts we do not actually employ when describing and explaining the world. If one doesn’t like the talk of concepts in this connection – and there are many reasons to be unhappy with such talk – one can say that what we study rather is truth, knowledge, justice, etc., and mean not the concepts concerned but the properties and relations they stand for: the property truth, the relation of knowledge that sometimes holds between thinkers and propositions, the property justice, etc. But when it comes to what I currently want to focus on, the reformulation is irrelevant. Even someone who insists that she rather is focused on properties and relations is focused on the properties and relations that the ordinary concepts stand for, and not the different properties and relations that some alternative concepts would stand for.

But I think that already a little reflection shows how – what’s the best way to put it? – navel-gazing this kind of study is. The concepts we have are the ones we have ended up with due to various biological and cultural factors. By some measure they have proven themselves. But still, why should
the concepts we actually have be the best conceptual tools for describing and theorizing about the
relevant aspects of reality? Maybe philosophy should rather be seen as conceptual engineering as a study
of what concept best plays the theoretical role of our concept of truth and what features this concept
has, what concept best plays the theoretical role of our concept of knowledge and what features this
concept has, etc.\(^\text{18}\)

When stating this alternative conception of what we should focus on, I have again slid into
talking of philosophy as concerned with concepts. But again the point is easily restated in terms of
properties and relations: why study the properties and relations that our ordinary concepts stand for
rather than the most theoretically important properties and relations in the vicinity?

It may be easiest to motivate the conception of philosophy as conceptual engineering by
comparing how an opposing conception is motivated. Here is Antti Kauppinen (2007):

Moral responsibility, for example, is not a technical notion, though some terms that philosophers
use in explicating it may be. Indeed, why should anybody care about what philosophers do if
they just argued about their own inventions? People want to know if they have moral
responsibility or knowledge of other minds in the very sense in which they ordinarily talk about
responsibility or knowledge, and to get at that sense one must work with the folk’s own
concepts. By and large, philosophers oblige; revisionism is a last resort, to be used only when one
is convinced that the folk concept is hopelessly confused or too imprecise for one’s purposes.\(^\text{19}\)

Here is one point Kauppinen makes: the philosophical questions we start out with are questions that
arise out of ordinary reflection, and are stated using ordinary concepts. We wonder about moral
responsibility, not about (some feature picked out by) some related concept. But, first, it could be
that although we state our concerns using a notion of moral responsibility, what we are at bottom
concerned with would be better stated using some improved notion. Compare perhaps the
development of physics, from rudimentary folk physics through today. Some of our original
questions may have had to do with weight; but what we are at bottom concerned with is such that
our concerns are better stated using a more theoretically sophisticated notion. Second, even if
Kauppinen is right in that we have a special kind of interest in questions best stated using ordinary
concepts and not some related counterparts, it doesn’t follow that the only, or the most, interesting

\(^{18}\) In an introduction to philosophy (1999), Simon Blackburn explicitly characterizes philosophy as conceptual
engineering. Robert Brandom (2001) likewise uses this label. More recently (2012), David Chalmers has
defended this conception of philosophy at some length but without using the “engineering” label. Alexis
Burgess and David Plunkett (2013, 2013a) characterize “conceptual ethics” as a research area, meaning by
conceptual ethics normative questions about which concepts to employ.

\(^{19}\) Kauppinen (2007), p. 96.
philosophical questions in the vicinity are formulated using such concepts. Even if we have a genuine interest in questions truly about weight, questions about mass, more attuned to the physical facts, are better to ask when our interest is with physical reality and not just our conception thereof.

Of course, figuring out what our actual concepts are like and what they are true of (or what the properties they stand for are like and are instantiated by) should not be dismissed as devoid of theoretical interest: getting clear on the tools we have come to use to understand the world is obviously a worthwhile project. But distinguish here between what we ought to focus on if our concern is the relevant aspect of reality and what we ought to focus on if our concern is (human) thought about the relevant aspect of reality. If our concern is the latter, of course there is a case for thinking about what our actual concepts are. Preference for philosophy as conceptual engineering relies on thinking of our concern as the former. Compare again the case of folk physics. If our concern is with the physical aspects of reality, as in physics, there is obviously no reason to stick with the folk concept of weight.

Competence intuitions, insofar as they are useful evidence, are useful for the project of studying our ordinary concepts or the properties and relations for which they stand. However we best go about the project of conceptual engineering, it is hardly via relying on competence intuitions, or by studying the judgments of the folk.

5. The limits of conceptual engineering: conceptual fixed points

The conception of philosophy as conceptual engineering is attractive, for reasons given in the previous section. But one may have the vague sense that somehow there must be important limits to conceptual engineering; that certain in some sense basic concepts cannot be replaced. Some motivation for saying this might come either from Kantian philosophy or from a naturalistic philosophy according to which what is innate severely constrains which concepts we can use. I have nothing to say about those kinds of motivation. Instead I will focus on other interesting limitations. I will bring up three cases where we are dealing with intuitively speaking basic concepts and where there do seem to be interesting limitations to conceptual engineering. I have discussed each of these cases elsewhere. The present discussion of these cases will be brief. In each case, I am more concerned to give the gist of the argument than to actually persuasively make the argument in question.20 (I am not going to argue that it is impossible to replace the concepts in question with other ones. The arguments to follow, even if persuasive as far as they go, do not establish anything that strong. The upshots are more subtle.)

20 The argument concerning truth is given in my (2012) and my (forthcoming); the argument concerning existence is given in my (2009) and (2013); and for the argument concerning thin normative concepts see my (2012a).
First, consider *truth*. Here is a reason why the idea of replacing the concept of truth is problematic. Stich (1990) considers certain other properties and asking whether we care whether our beliefs are true rather than have this other property. The aim is to call into doubt whether we should aim for our beliefs to be true rather than have some other property. One particular alternative Stich considers is that of the property *truth***. Truth*** is characterized in terms of reference*** the way that truth can be characterized in terms of reference. And reference*** differs from reference, in that the so-called Lockean view is true of reference***, and ‘water’ refers to H2O and XYZ alike, even while the causal view is true of reference, so that ‘water’ refers only to H2O. Stich seeks to make plausible that we do not, and should not, care about truth over truth***, and seeks to make bigger points about the value of truth, on the basis of that. Here is a fundamental problem with Stich’s argument. Suppose, as should be possible if Stich’s view is true, that there are possible people who care for truth*** the way we care about truth. When we seek to believe and assert what’s true, they seek to believe and assert what’s true***, etc. The question then arises: why should we say that it is the same propositions they believe and assert that we believe and assert? Should we not instead say that they employ a concept *water***, under which both H2O and XYZ fall, rather than the concept *water*? Their use of their word “water” is exactly like how we would use a word expressing such a concept *water***.21

Let me right away stress what this argument does not purport to show. Of course there are other semantic properties besides truth, reference and the like, and we can have predicates standing for these properties. The consideration just presented does not purport to contradict that. The consideration just presented only purports to show that it is difficult to make sense of a property other than truth being the aim of belief and assertion. The consideration presented also does not say that the interpretation of the people concerned as using the concept *water*** rather than the concept water is *preferable* to an interpretation according to which they aim at truth*** rather than truth. At most it yields that Stich’s preferred description of such a case – that the people in question care for truth*** – is not preferable.

The reason the consideration is relevant is this. Suppose we set out to conceptually engineer truth. Insofar as the job description of truth is that of being the property our beliefs and assertions aim at, the engineering project would be that of finding a property more adequate to that job description. But by what has been noted about Stich’s argument, it is hard even properly to conceive of a practice of belief or assertion that is guided by a different property.

One natural but misguided objection at this point is: why make a fetish of belief and assertion? Maybe truth is a constitutive norm of belief, but truth*** is a constitutive norm of belief***. And

---

21 A version of this argument is also given in my (2012).
maybe a different community can have a practice of beliefs and assertions instead of belief and assertions. However, this objection does not get around the underlying issue: why should we say that theirs is a practice of beliefs and assertions rather than beliefs and assertions, and not instead say that they use concepts rather than the ordinary concepts?

Turn now to the case of existence. In the contemporary metaontological debate, it is common to assume that there are alternative notions of existence that can be employed. Given the underlying Quinean assumption that existence is expressed by so-called existential quantification, this amounts to saying that there are alternative unrestricted existential quantifiers that can be employed. For example, according to the form of ontological realism defended in Sider (2009, 2011), ontological questions should be conceived of as questions about ‘what there is’, in the properly fundamental, or joint-carving, sense of ‘what there is’. Even if, in the ordinary sense of ‘there is’, there are tables and chairs, it can still be that in the more interesting, deeper sense of ‘there is’, there are no tables and chairs.

Let two languages be ontologically different if they employ different unrestricted quantifiers such that for some predicate F, “∃xFx” comes out true in one and untrue in the other, due to the differences in meaning between their quantifiers. It seems that if there can be alternative notions of existence as outlined, then there can be ontologically different languages; and Sider seems explicitly to commit himself to the possibility of ontologically different languages. But can there be ontologically different languages as characterized? Here is one reason to think not. Take a language supposedly ontologically different from ordinary language. For some type of thing F, we say truly, using our language, “there are Fs” but in this other language “there are no Fs” is true, and the reason, supposedly, has to do with the languages having a different (unrestricted existential) quantifier. But why exactly should we accept that ‘F’ in the alternative language means F? We would have such reason if in other sentences of the two languages ‘F’ behaved the same in the two languages. But suppose that was so; e.g. that the same atomic sentences involving ‘F’ were true in the two languages. Then given that there are true atomic sentences “F(a)” in our language there are also true atomic sentences “F(a)” in that alternative language. But then how can we regard the supposed unrestricted existential quantifier of that other language as such? Assuming that “there are no Fs” is true in that language, it doesn’t satisfy the standard introduction rule (or even a version restricted to atomic sentences). So suppose then that languages differ also in what atomic sentences are true; specifically there are no true atomic sentences of the form “F(a)” in the alternative language. Then the question becomes more pressing: why suppose that ‘F’ of that language means F? (The problem is not just what it takes for a speaker to use ‘F’ of that language to mean F, but it is about the content of the claim that it means F in the first place.)
All that is raised is a challenge – is there an answer to the question asked at the end of the previous paragraph? And even if the challenge in fact cannot be met, that does not show that there cannot be different quantifiers. What it would show is that differences in quantifiers go hand in hand with large-scale systematic differences of other kinds. One cannot, so to speak, selectively engineer the quantifier.22

As a third example, in addition to truth and existence, consider the ‘thinnest’ normative words – words like ‘good’, ‘right’, ‘ought’. These ‘thin’ (general, non-specific) normative words contrast with thick normative words like ‘courageous’, ‘brutal’, ‘lewd’, etc.23 There are reasons to think the strategy of conceptual engineering has certain limits here too. Consider the following scenario:

There is a linguistic community speaking a language much like English, except for the following differences (and whatever differences are directly entailed). While their words ‘good’, ‘right’ and ‘ought’ have the same evaluative and normative roles as our words ‘good’, ‘right’ and ‘ought’ have, their words aren’t coextensive with our ‘good’, ‘right’ and ‘ought’. So even if they are exactly right about what is ‘good’ and ‘right’ and what ‘ought’ to be done, in their sense, and they seek to promote and to do what is ‘good’ and ‘right’ and what ‘ought’ to be done in their sense, they do not seek to promote what is good and right and what ought to be done.

Is this sort of scenario possible? That is, can there be words non-coextensive with our ‘good’, ‘right’, etc. but with the same evaluative and normative roles? If not, then a certain kind of conceptual engineering isn’t possible: that of finding alternative concepts with the same normative roles as our actual concepts but different in extension. And I have elsewhere given an argument to the effect that this sort of scenario isn’t in fact possible. I will here relatively quickly rehearse the argument. (For the full argument see my (2012a).) All I want to do here is to outline the case for another conceptual fixed point.

If the sort of scenario sketched is indeed possible, then, a first thought may be, there is a question of which normative concepts are the best ones, the ones we employ or the ones employed by a community like this. And once that thought arises, a certain skeptical thought suggests itself: maybe our concepts aren’t the ones that ought to be used? (And how can we even know which concepts ought to be used?)

But there is something that seems very peculiar about these thoughts. We state them using ‘ought’ – our ‘ought’. And, while there certainly could be complications, it is reasonable to assume

22 See further Eklund (2009) and (2013).
23 The thin words I am focusing on are, as is often pointed out, context-sensitive, but I will be concerned with certain kinds of uses of them, the thinnest uses, as it were.
that “the concepts that ought to be used are ‘good’, ‘ought’, etc.” is a true sentence of our language even while “the concepts that ought* to be used are ‘good*’, ‘ought*’, etc.” is a true sentence of their language (where I use asterisks to indicate the counterpart expressions they use). In other words when we try to ask the supposed further questions that arise, we end up asking questions that are answered in our favor, and when they try to ask the supposed further questions that arise they end up asking questions that are answered in their favor. This means that we and they fail to ask the question sought. That question threatens to be ineffable.24

In general, if the sort of scenario sketched is possible, then either there is this further ineffable question or there is nothing more to it than what ought to be done, what ought* to be done, etc. One can naturally be hesitant about embracing the idea of ineffable questions. But, importantly, the other possibility seems strange too. Suppose there is a given action that would be deemed right* but not right. Suppose, for example, that there concepts are utilitarian and ours are deontological, and what’s deemed right* is sacrificing someone innocent for the greater good of the majority. As they are about to sacrifice this someone, we think this is something they ought not to do. We say “Stop it! You ought not do that!”. They say “Yes. We know. We ought not do that. But we ought* to do this!”. We persist, “But you ought to do what you ought to do!” . They say “Yes again. But also, we ought* to do what we ought* to do”. And if there is no further question there – some further fact that we and they disagree about – there is no disagreement over facts here, as they go about sacrificing the innocent. Certain antirealists are happy to embrace such a conclusion, but the conclusion is antithetical to the spirit of realism.

Someone who accepts the assumption that the scenario sketched is possible seems then to be forced either to accept ineffable questions or to embrace a conclusion antithetical to the spirit of realism. Assuming both these alternatives unacceptable, the conclusion is that the scenario sketched isn’t possible: there cannot be these non-coextensive concepts with the same normative role. There is a certain sense in which our thinnest normative concepts are irreplaceable. There cannot be other concepts, non-coextensive with these ones but with the same normative roles.

6. Concluding remarks

Perhaps a brief recapitulation is in order, especially since a number of importantly different themes have been brought up in the discussion. I began by discussing skepticism about intuitions, arguing that competence intuitions escape criticisms due to experimental philosophers relatively unscathed. It would have been nice to go on and defend a picture of philosophy as relying on competence

24 The complication alluded to is that if there can be these alternative notions of ‘right’, ‘good’, etc., then maybe there can be some alternative notions that are, as it were, self-undermining, maybe there can be a concept ought** such that “the concept that ought** to be used isn’t ‘ought***’ is a true sentence of the language of which it is part. But so long as some alternative notions are not self-undermining the problem arises.
intuitions, but I argued for a conception of philosophy as conceptual engineering, and given such a conception any role for competence intuitions is at best limited. Then, in the last section, I noted some limitations to the extent to which philosophy can be seen as conceptual engineering. Obviously some big questions in the vicinity remain entirely unresolved, such as what the proper methodology for conceptual engineering is, and also, separately, how we have the knowledge we seem to have in cases where the conception of philosophy as conceptual engineering is not straightforwardly applicable.

REFERENCES


