Intuitions in Philosophical Semantics

Abstract

We argue that the term “intuition”, as it is used in metaphilosophy, is ambiguous between at least four different senses. In philosophy of language, the relevant “intuitions” are either the outputs of our competence to interpret and produce linguistic expressions, or the speakers’ or hearers’ own reports of these outputs. The semantic facts that philosophers of language are interested in are determined by the outputs of our competence. Hence, philosophers of language should be interested in investigating these, and they do this by testing what we would say or understand in hypothetical communication situations. In the final section of the paper we suggest methods to investigate these outputs that are independent of the subject’s reporting them, and hence might be used as an alternative to the standard use of hypothetical cases.

1. Introduction

Contemporary metaphilosophy, i.e. the study of philosophical methodology, typically assumes that (at least analytic) philosophy is widely committed to the “method of cases”.[1] This method consists in eliciting an intuitive judgment as a response to a described hypothetical case. The judgment is then taken to be evidence for or against a philosophical theory. It is assumed that this methodology is widely used in all areas of analytic philosophy

1 This assumption is shared by rationalist defenders of the method of cases, such as George Bealer (1996), their experimental critics, such as Edouard Machery (Machery et al., 2004), and Jonathan Weinberg (2007; Weinberg et al., 2001), as well as deflationists about intuitive judgments, such as Timothy Williamson (2007). See (Cappelen, 2012) for a critical discussion of this assumption.

(and possibly beyond), and operates everywhere under the same constraints. Thus it can be critically evaluated from a rather general point of view.[2]

It should be noted from the start that this assumption is already *prima facie* implausible. In order to critically evaluate philosophical methodology, one needs to answer at least the following questions: Firstly, what is the nature of the subject matter of philosophical inquiry? (Is philosophy after analytic or conceptual truths, or after metaphysically necessary truths, or maybe after highly general empirical truths?[3]) Secondly, what is our epistemic access to that subject matter like? (Can we know about it *a priori*, by introspection, by a special faculty of intuition, by observation?) That different answers to these questions might yield different methodological recommendations should be obvious; e.g. if philosophy is after highly general empirical truths and we have no reliable a priori access to these, then it would seem inappropriate to use an exclusively a priori methodology.

However, not only will it be very hard to find two philosophers who would agree in their answers to these two questions for any given area of philosophical inquiry, it is certainly implausible that both questions should *have* uniform answers for *all* areas of analytic philosophy that are discussed in contemporary metaphilosophical debates (such as ethics, theories of reference, the analysis of knowledge, etc.). Consequently, some authors have recently pointed out that, if metaphilosophical inquiry hopes to arrive at any interesting insights, it should rather focus on the methodology of particular debates. (Author; Author; Cappelen, 2012).

In this paper we will argue that the assumption that intuitions have a certain sort of evidential role in all areas of analytic philosophy has led the metaphilosophical discussion

2 For an example, see Joshua Alexander's discussion of the role of intuitions in philosophy and the x-phi criticism of it in (Alexander, 2012).

3 This list doesn't exhaust the options, of course. Perhaps philosophy isn't a cognitive enterprise, but rather a normative one.

astray when discussing the role of intuitions in philosophy of language. In particular, the debate between Edouard Machery, Michael Devitt and Genoveva Martí over the relevance of certain lay speaker intuitions for the evaluation of theories of reference seems to suffer from a mistaken conception of the role of intuitions and function of thought experiments in philosophical semantics.

In the next section we will first show how the discussion of the so-called "expertise defence" between Edouard Machery and Michael Devitt presupposes a certain picture about the role of intuitions in philosophy: that they are judgments that are used as evidence for or against philosophical theories. In Section Three we will argue that one should distinguish at least two ideal[4] types of roles that intuitions can (and presumably do) play in philosophy. We will call these 'Evidential Role' and 'Constitutive Role'.[5] We will argue in Section Four that intuitions in philosophy of language (in particular those that are of concern in discussing theories of reference) should best be understood along the lines of the Constitutive Role model. We will also argue that, *pace* Devitt, this model does not presuppose any problematic assumptions about the psychological realization of linguistic competence. In Section Five, we will analyze what the interesting methodological issues are that should be addressed in a metaphilosophical analysis of the role of the method of cases in philosophy of language, and how philosophy of language could improve its methodology. Section Six will conclude with a discussion of how variation in (the relevant) intuitions could be accommodated, if it were to be found.

4 We use the notion of an "ideal type" in Max Weber's sense. As is argued in (Author) there are mixed cases too.

5 This distinction is similar to the distinction between constitutive and non-constitutive groundings for evidential relations, drawn in (Goldman, 2007).

2. The Experimentalist Challenge and the Expertise Defense

The "Experimentalist Challenge" is an attack against a supposedly widespread method in philosophy, the "method of cases". Here is how Edouard Machery explains what he takes the method of cases to be in philosophy of language:

> Lay people and philosophers alike have intuitions about what words of various kinds – such as proper names – refer to in actual and possible situations. For instance, people have the intuition that in the actual world "Barack Hussein Obama" refers to Barack Obama. Philosophers use these intuitions to assess theories of reference. A theory of reference is undermined if it entails that in an actual or a possible situation a proper name (or a natural kind term, etc.) refers to what people judge is not its correct reference, and it is supported if it entails that in an actual or a possible situation a proper name (or a natural kind term, etc.) refers to what people judge is its correct reference [...]
>
> Why would intuitions about reference play such a role? Presumably, for the same reason as ordinary judgments are often taken to provide evidence for particular facts. If I judge of an object that it is a chair, my judgment that it is a chair is evidence that it is a chair because I am reliable at sorting chairs from things that are not chairs.
>
> (Machery 2011, 38-39)

The method of cases presents first an actual or hypothetical case, about which the theory under investigation implies a certain verdict (for example about what a certain expression would refer to in this situation). This verdict is then compared with how philosophers intuitively judge the case. As Machery explains it, the reason for thinking that theories should

be tested against those judgments is the (implicit) assumption that people are reliable detectors of the property one is interested in (in this case, the relational property of referring to a certain object or kind).

Experimental philosophers like Machery see themselves as challenging this method on empirical grounds. They presented vignettes, modelled after thought experiments used (for example) in philosophy of language, and reported that lay speakers show considerable inter- as well as intracultural variation in their intuitive judgments about these cases (Machery et al. 2004), which should cast some doubt on the assumption that we are reliable detectors of the property in question.

The expertise defence against this challenge, championed by Kirk Ludwig (2007) and Michael Devitt (2011, 2011a, 2012, 2012a etc.), agrees with Machery on the role of intuitions in philosophy, but protests that the variation in intuitions of lay speakers does not immediately undermine the evidential value of the judgments of philosophers.[6] Philosophers are experts about the properties in question, hence their intuitive judgments should be considered more reliable than the judgments of ordinary folk. According to Devitt this is for two different reasons. Firstly, experts (for example in philosophy of language, or in linguistics) know, allegedly, more about linguistic reality from observation, and secondly, they have in general better theories than ordinary folk (Devitt, 2011).

The success or failure of the expertise defence would then hinge on whether it is in fact *true* that the intuitions of experts are immune from (or less prone to) the kind of variation found in the intuitive judgments of lay speakers, and whether their intuitions are in general more reliable than those of ordinary speakers. Both Machery (2011) and Devitt (2011, 2012a) suggest ways of empirically studying this question; we will discuss this in section 5.

6 In this paper we mainly concentrate on the expertise defence to motivate our discussion of the role of intuitions in philosophical semantics. Both authors have also offered other arguments against the Experimentalist Challenge.

We believe that this debate is on the wrong track. Although we would welcome a certain kind of empirical investigation in philosophical semantics (which we will briefly discuss in our final section), we believe that it is a waste of time and resources to test the reliability of the method of cases as described above, as it gives a highly misleading picture of the role of intuitions in philosophical semantics. The use of intuitions in philosophy of language does not depend on the assumption that we philosophers (or anybody else, for that matter) have a mysterious but reliable capacity to intuitively judge whether a certain described relation is the reference relation. And we are also not better than lay people at dealing with thought experiments because we possess the better theories (although we might well be better in dealing with them for other reasons). In fact we agree with the experimental philosophers this much; if analytic philosophy of language were committed to such assumptions, then it would be in trouble. Fortunately, it is not.[7]

3. The constitutive and the evidential roles of "intuitions"

In order to explain why this discussion between Machery and Devitt seems to us to be on the wrong track, we would like to explicate the distinction between an *evidential* and a *constitutive* role for intuitions in theorizing about a certain subject matter. As will become clear later, the distinction somewhat oversimplifies things, but we will add the necessary subtleties later.

7 By the way, we do not believe that all areas of analytic philosophy use intuitions in such unproblematic ways. We take contemporary metaphysics to be an area where intuitions have a problematic epistemic status, although we will not argue for this claim here. Likewise, if Devitt at least gives a correct account of what *he* is up to in philosophy of language, then his project *is* in trouble.

3.1 Physics vs. Folk Physics

To illustrate the distinction we are after, consider two areas of inquiry: *physics* and *folk physics*. If you are interested in physics, you are interested in how the physical world works. For example, you might want to know or learn about the general principles that determine what the trajectory of a stone will be, and which forces will be acting upon it, when you throw it into the air at a certain angle and a certain velocity. One *could* learn something about this on the basis of intuitions: one could just ask people what their intuitive, spontaneous judgments are about how the stone would fly and which forces would be acting upon it. These intuitions might be more or less accurate about how stones really fly, and about which forces really act upon it. Perhaps they will be somewhat reliable with respect to the trajectory of the stone, since we human beings are attuned to the physical behaviour of middle-sized dry goods, owing to our frequent interaction with them, such that some of our intuitive expectations about the behaviour of physical objects is even innate and can be found already in infants.[8] On the other hand, we might suspect that our ordinary intuitions about the forces acting upon the stone on its way are very unreliable; intuition probing is not likely to get us very far in understanding how the physical world works. Still, if one were to insist on using intuitions to study *physics*, perhaps it would be more fruitful to focus on expert intuitions: it might well be that people who have studied physics have more reliable intuitions about the behaviour of, and the forces acting upon, physical objects.

Let us contrast that with the study of *folk* physics, i.e. the study of how human beings intuitively expect the physical world to behave. In this case, it does not make much sense to ask to what extent *these* intuitions might provide reliable evidence for the study of folk physics. It might make sense to ask whether all human beings share the same folk physics, but it won't make sense to worry whether the intuitions of one subgroup are better evidence

8 Cf. (Baillargeon, 2004).

for folk physics than another: folk physics, as we conceive it here, is nothing but the systematization[9] of the intuitions we have about the physical world. Folk physics is *constituted* by these intuitions.

One might object here that we are conflating two issues, namely the question of what intuitive expectations human beings have with respect to the behaviour of the physical world around them (implicit in, and revealed by how their behaviour is attuned to the physical world), and the question of what kinds of intuitive *judgments* they make about physics, and that it is only the former, but not the latter, that folk physics studies. Thus, even if there is a sense in which folk physics aims at systematizing intuitions, in the sense that it aims at systematizing those intuitive *expectations*, its aim is not to systematize intuitive *judgments*. That is correct; as we will show in the next section, we should distinguish not just two, but four different notions of 'intuition'[10], often conflated in the literature.

3.2 Four Different Notions of Intuition

In this section we will distinguish between four different things that might be (and have been) called "intuitions" in the debate concerning the role of intuitions in the study of language. For example, some philosophers suggest that we distinguish between metalinguistic intuitions and linguistic intuitions, and that it should be the latter (if any) that matter for the study of language (Marti 2009). Other philosophers argue that when discussing the role of intuitions in the study of language, we should not consider intuitions to be judgments but instead consider

9 We are using 'systematization' in lieu of a better word. We don't consider systematization here to be revisionary and a form of normative regimentation, constructing order amongst our intuitions which isn't initially there. Thus, on our understanding, folk physics is descriptive (even if it allows for a competence/performance-distinction).

10 We use single quotation marks for mentioning, and double quotation marks for using a quoted expression.

the relevant data to be something prior to the judgment (such as, for example, linguistic *seemings*, c.f. Textor 2009).

We believe that there is some truth in both views, but that the distinctions that these authors draw (between metalinguistic and linguistic intuitions, and between linguistic seemings and linguistic judgments) do not fully clear up the confusions underlying the debate. For example, even if the relevant data for the study of language is linguistic seemings, it is still, of course, *judgments* resulting from these seemings that are presented as evidence, and even if we should systematize linguistic intuitions rather than meta-linguistic intuitions, these judgments are judgments about semantic properties of linguistic items, and hence prima facie meta-linguistic.

We want to suggest a different way of distinguishing different notions of "intuition", and hope to get clearer on exactly which of these so-called "intuitions", if any, can be used as evidence for various theoretical pursuits, and what the relationships between the different notions are. We will summarize our discussion in five general observations, which we will return to in later sections.

We should emphasize that nothing hinges on whether it is best to call each or any of the four different things we distinguish "intuition". Perhaps it would be better for the debate to eliminate that term. In any case, one can understand our definitions as purely stipulative ones, which use 'intuition' in lieu of a better word.

Consider the case of folk physics again. As noted above, some human beings, for example baseball players, have an impressive capacity to predict the trajectory of, for example, baseballs. Some of this interesting ability is acquired through training, some of it is perhaps innate. The outputs of this ability—particular expectations that the baseball-player has, about the trajectory of a particular ball—are intuitive in the sense that the baseball player forms these expectations in a spontaneous way without a prior conscious reasoning process.

In psychological literature it is this immediately formed expectation that is sometimes called "intuition". For example, Reed et al. characterize the skill of catching a ball as follows:

> [The skill] gives rise to a phenomenal sense of intuition. If you know how to catch and a ball is thrown towards you, you get an immediate feeling that you should run backwards or forwards to catch it. You do not do any conscious computation. (Reed et al. 2010, 64)

Let us call these immediately formed expectations that are realized as the feeling that one should move forward or backward in order to catch a ball "first-level intuitions"[11]. For all we know, these expectations are not proposition-like belief states, let alone propositional judgments. Nonetheless, they are intuitive in the ordinary sense of the word, and we will follow Reed et al. in calling them intuitions.

Suppose a baseball player sees a ball moving through the air, and we ask him where he expects the ball to land. The player *might* be able to accurately report what his first-level intuition is. But this report and the intuitive expectation reported are not the same, and should not be conflated. Someone might have accurate first-level intuitions, but be very inaccurate in recognizing and reporting them: he or she might be extremely skilled at throwing a ball and at moving appropriately to catch a ball, but very bad at verbally reporting what he or she is doing. In fact, our intuitive capacity to catch a ball, or to throw it in an intended direction is a complex sensorimotor skill and difficult (perhaps impossible) to access via introspection.[12] As Reed et al. found out, even if we ask someone right after catching a ball what her experiences were (for example, how their angle of gaze changed when following the ball, which is the relevant clue to catch a ball), the reports are highly unreliable (Reed et al. 2010).

11 In this paper we will call "first-level intuitions" any outputs of an intuitive capacity that are considered constitutive of the capacity by an inquiry that has that capacity as its subject matter.

12 Besides, we can already master such sensorimotor skills before the acquisition of any language.

Now consider a theoretician, interested in the impressive sensorimotor skills of baseball players, developing a first sketchy theory of how they manage to catch balls and throw them in the intended direction. Being at the early stage of theory development, the theoretician might begin with his or her intuitive understanding of what is going on in the baseball player and predict on the basis of that understanding (perhaps informed by observing what baseball players do) where the baseball player will expect a ball to land, if he is confronted with certain sensory input about the ball's trajectory. Let us call the intuitions of the theoretician "second-level intuitions"[13]. They aren't intuitions about the trajectory of the ball, but rather intuitions about the sensorimotor capacities of baseball players. Again we should distinguish these intuitions from the scientist's report of them. Since the second-level intuitions are theoretical, or in any case more theoretical than the first-level intuitions of the baseball player, perhaps they are generally less problematic to recognize and adequately report. Nonetheless, the intuitions, and the reports thereof, are different things, and should not be conflated.

Of course, there is an interesting relation between the first-level intuitions and the second-level intuitions. The second-level intuitions (and the respective second-level judgments which report those intuitions) are *about* the first-level intuitions. The theory our theoretician is developing, a theory of folk physics, is a systematization of the first-level intuitions. If you want to know what the state of the art in the reconstruction of folk physics is, you should ask a theoretician about his or her judgment (and it might presumably be better to ask him for considered judgment than for his intuitive judgment). Since this will inform you about the *state of the art* of the reconstruction of folk physics, it will also indirectly inform you about folk physics. Indeed, the better the reconstruction of folk physics currently is, the better that indirect evidence will be. However, if you want to study folk physics

13 Given an inquiry into an intuitive capacity, and first-level intuitions that are considered to be constitutive of the subject matter of that inquiry, we will call "second-level intuitions" the outputs of the intuitive capacity of a person undertaking that inquiry to form expectations or beliefs about the subject matter of the inquiry.

directly, rather than studying what experts currently *believe* about folk physics, you should be interested in first-level intuitions, because these intuitions *constitute* folk physics. This can be summed up as follows:

> *Observation 1. If one is interested in developing a second-level theory, then there is no sense in which first-level intuitions are "more or less accurate" evidence for that theory; instead, the problem of accuracy of first-level intuitions for a second-level theory concerns the reliability of reports of first-level intuitions.*

Combine this with the distinction we made earlier between folk physics and physics. If you are interested in how baseballs fly, the best way to study this is by studying baseballs, rather than the first or the second-level intuitions we talked about. The second-level intuitions aren't about baseballs but about *intuitions about* baseballs, and the first-level intuitions are only interesting to the extent that you have independent reason to think that our intuitive expectations about the physical behaviour of baseballs are accurate. Let's call theorizing about physics itself "intuition-independent"[14].

It might well be that one's first-level physical intuitions can become more accurate as a result of intuition-independent theorizing. Perhaps if we know more about the physics of baseballs, we get better in intuitively predicting how they fly. However, no amount of *second-level* theorizing is likely to make our first-level intuitions any more accurate: second-level theorizing is about intuitions, not baseballs. To sum up:

14 This is meant to remind the reader that these theories are neither about intuitions, nor about the underlying psychological processes that lead to them. Of course, in the development of physical theories, intuitions of various sorts play a role.

Observation 2. Typically, second-level theorizing doesn't improve the accuracy of first-level intuitions. At best, second-level expertise might improve the accuracy of reporting first-level intuitions.

Now, let's look at another example, folk *psychology*. Again, human beings have an impressive capacity to form expectations about and explanations for the behaviour of other human beings. For example, typically when one human being says to another human being 'Let's meet today at three in the afternoon in front of the town hall', one can find both of them at around three in the afternoon in front of the local town hall. What explains this? The usual explanation is that human beings intuitively ascribe mental states such as beliefs, intentions, desires, etc. to each other in forming expectations about behaviour. Because human beings do this in very similar (perhaps even uniform) ways, they are quite successful in coordinating their behaviour.

Again, this doesn't automatically mean that human beings are very good at *recognizing* and accurately *reporting* those intuitive ascriptions of mental states. Again, we *might* well be able to ascribe such states before having even learned the words 'belief', 'intention', or 'desire'.[15] Notice though that unlike in the case of folk physics discussed before, talking about the mental states we ascribe to other people is something we do on a regular basis: in fact, that is what human beings seem to be talking about most of the time (cf. Nagel, 2012). It is also easier to report these first-level intuitions, since they are not merely sub-personal sensorimotor skills, but available to consciousness. Thus, assuming that we are

15 Note that this involves at least two sources of error. On the one hand we might make a conceptual mistake and think that the mental state we assume somebody else (or ourselves) to be in falls into the wrong category (perhaps by refraining from classifying a mental state as an instance of knowledge when it's not a case of certainty), on the other hand, we might just be bad at verbally reporting our mental state ascriptions (for whatever reason).

quite reliable in reporting our first-level intuitions in the case of folk psychology is rather plausible. Thus,

> *Observation 3. How accurate we are in reporting our first-level intuitions is plausibly related to whether we have conscious access to them, and to how central the relevant capacity is in our discourse.*

As in the case of our folk physicist, we might take a second-level perspective on the matter, that of a folk psychologist, and consider what those mental states are that human beings ascribe to each other. (And again we should distinguish eventual second-level intuitions from our reports of them.) Let us consider an example. Suppose there is a theoretician (a folk psychologist) who intuitively thinks that we have privileged access to our own mental states, and that the reports we give, of our own mental states, are generally accurate. The states that get ascribed as, for example, belief states are states that we have privileged access to and that are thus authoritatively reportable on the basis of introspection.

Let's further assume that our theoretician is confronted with the case of Berta. Berta honestly describes herself as believing that women aren't less qualified than men for higher management positions. However, if asked to rank CVs according to the assumed qualification level for a higher management position, Berta nevertheless consistently ranks CVs she assumes to be from male applicants higher than those from female applicants. Now our theoretician is asked whether Berta believes that women are less qualified than men.

Notice that our theoretician can arrive at an answer via two different routes; he can consider the case of Berta and use his intuitive capacity to ascribe mental states to others to find out which mental state he'd ascribe to Berta in such a situation (by simply reporting what mental state he *does* ascribe to hypothetical Berta). Doing this, let's say, he judges that Berta

does believe that women are less qualified than men. This would be an instance of what we called a first-level intuition report above.

He can also consider the case as an instance of his intuitive theory of the ascription conditions of belief states, and perhaps judge on *that* basis that Berta does *not* believe that women are less qualified than men (since she is reporting the opposite). This judgment is based on the theoretician's second-level intuitions about folk psychology: it is a second-level intuition report. In this case, our theoretician's theory makes a predictive mistake about his own first-order intuitions: his intuitive theory of belief ascriptions does not agree with his own intuitive belief ascriptions.

Both judgments, then, if they are accurate, report an intuition. Notice further that the formulation of the judgment that our theoretician comes up with does not tell us whether it is a report of a first-level or a second-level intuition: both are naturally reported with a sentence such as 'Berta believes/does not believe that women are less qualified than men'. Notice finally that the amount of reflection entering into the eventual judgment does not make the difference either; our theoretician might reflect on the case very carefully and still arrive at his judgment about Berta on the basis of his intuitive capacity to ascribe mental states to other human beings (reporting a first-level intuition), and he might likewise not need to go through any explicit reasoning at all when judging the case as an instance of his intuitive theory of the ascription conditions of belief states (reporting a second-level intuition). Thus,

> *Observation 4. There are always at least four things that might be considered "intuitions"; what we called first and second-level intuitions, and their respective reports.*
>
> *Observation 5. "Intuitive judgments" can be ambiguous between first-level and second-level intuition reports. Whether a given judgment is a first-level or second-*

level intuition report cannot (at least not in all cases) be read off its formulation, but depends on the cognitive genesis of the state that is being reported.

Is there also an intuition-independent point of view, as in the case of physics and folk physics, discussed above? As in the case of physics we might wonder whether our capacity to predict and explain the behaviour of other human beings is generally very reliable, or whether its above-mentioned success is limited to specific circumstances. Some critics of folk psychology argue that folk psychology isn't generally reliable. Thus here, the intuition-independent theory might be a psychological theory, explaining and predicting human behaviour in other ways than folk psychology does (just as our physical theory would explain and predict the behaviour of flying stones differently than folk physics does). The last observation is important. First-level intuitions might play different roles, depending on the theoretical perspective we take. Are we considering them from the point of view of an intuition-independent theory, or from the perspective of a second-level theory?

This is the difference between the evidential role of intuitions and the constitutive role. On the evidential role model first-level intuitions are more or less reliably tracking an independent property or independent facts, and according to their degree of reliability they can be used as evidence for intuition-independent theories about those facts. But there is also a connection between intuition-independent theories and first-level intuitions in the other direction. Since experts with intuition-independent theories *might* have better first-level intuitions, their intuitions might also be better evidence for intuition-independent theories than those of ordinary folk. That is the idea the expertise defence against the experimentalist challenge tries to exploit.

On the constitutive role model, first-level intuitions constitute what second-level theories describe. Being more expert with second-level theories doesn't make your first-level

intuitions “better” evidence, because they aren’t evidence for second-level theories in the first place. Since there isn’t a set of independent facts that could be captured by an intuition-independent theory, there is also no other kind of expertise that could tutor or outperform first-level intuitions. The question of evidence arises with our access to those intuitions. How reliable are our first-person reports of our first-level intuitions? Being more expert with second-level theories might make your reporting of first-level intuitions more accurate. Perhaps you know better what to attend to and how to correctly use the technical vocabulary to report it. But this doesn’t affect the status of the intuitions reported.

4. Where to locate theories of reference?

The question of whether the expertise defence makes sense in a metaphilosophical defence of the use of intuitions in philosophical semantics depends, then, on whether the evidential role model is the right model for analysing what is studied in philosophical semantics. In other words, if we apply our distinction between levels and theories to philosophical semantics, what level are theories of reference properly located at?

4.1 Theories of reference are second-level theories

As several authors have argued (e.g. Martí, 2009; AUTHOR), first-level intuitions in philosophical semantics are our dispositional responses in our normal employment of linguistic competence. We have a capacity to intuitively interpret and produce utterances. This intuitive capacity, as any such capacity, can be studied from a second-level perspective.

As we have argued in detail elsewhere (AUTHOR) philosophical semantics and theories of reference should be understood as a systematization of those first-level intuitions, the outputs of our linguistic competence. The reason is simply that the way our terms refer

ultimately only supervenes on (actual and potential) usage. There are no independent additional facts over and above usage that an intuition-independent theory could be about. Hence, theories of reference are properly located at the second level: they are theories *about* our first-level intuitions, and (actual and potential) first-level intuitions are constitutive for reference.[16]

To be sure, the reference of an expression in an utterance of a speaker is not simply whatever the speaker intends or believes to be the referent of the term. As we know from the discussion of externalism about reference, *what an expression refers to* is not (always) internal to the speaker in that way. However, *what theory of reference is true of an expression* is, in our view, internal in this sense. To be more precise, how[17] a linguistic expression E in an utterance U by a speaker S refers and which theory of reference is true of E is determined by S's dispositional states to apply and interpret E in actual and hypothetical circumstances. These dispositional states include, for example, also her disposition to correct or change her

16 This also explains why thought experimentation is a more viable strategy for theorizing about reference than about (say) physics. Consider Devitt's favourite example of expertise (*e.g.* in Devitt, 2006, 104): a palaeontologist who is very skilled at distinguishing fossils of pig jawbones from other objects. In our terminology, she has highly reliable first-level intuitions about pig jawbones. Obviously, an intuition-independent theory of pig jawbones is available. Just as obviously, thought experimentation is not going to tell us much about what pig jawbones are. Suppose, however, that we are interested in finding out *how* our palaeontologist manages to distinguish the jawbones from the non-jawbones; which criteria is she using? (That is, suppose we are looking for a second-level theory of the palaeontologist's first-level intuitions.) Now things change completely: presenting the palaeontologist with imaginary cases (hypothetical white stones with various perceptible features) might tell us a *lot* about what kinds of factors are guiding her first-level intuitions. (But there is no reason to assume that thought experimentation is our *best* way of studying her first-level intuitions; cf. our observation 1 above.)

17 'How a linguistic expression E refers' is our shorthand for, roughly, 'whether external factors can play a role in determining what E refers to, and if so, which kinds of external factors: causal chains, or underlying essences, or the judgments of experts, or . . .'

usage when more information becomes available to her. It is such dispositional states that matter for how an expression refers in utterances of a speaker: there is nothing over and above them that could settle the semantic facts.

As the case of reference shows, for some first-level intuitions there simply might not be an intuition-independent theory that could outperform them, because the first-level intuitions simply don't track a property that exists over and above those first-level intuitions. In these cases there is only second-level systematizing of first-level intuitions and hence our observations 1-3 should apply.

The idea that we are dealing with second-level theories, that is, theories about first-level intuitions, seems to many to be a reasonable assumption in the case of syntax, but not in semantics. Henry Jackman, who introduced a similar distinction, calling intuitions about what a term refers to (our first-level intuitions) 'type-2 intuitions', and theories explaining why our terms refer to what they refer to (our second-level theories) 'type-3 theories', expresses the disanalogy thus:

> A semantic theory bears a 'normative' relation to intuitions of a sort that syntactic theory does not. Of course, syntactic intuitions can be mistaken, but such mistakes are explained in terms of performance errors that don't change the fact that the 'correct' intuitions can be understood as directly reflecting the underlying syntactic reality. Semantic intuitions, on the other hand, even when correct, are correct in virtue of their relations to other intuitions, not to some underlying concept, and because of this, it is easier to think of semantic intuitions (like ethical or epistemic intuitions) as mistaken. The process of systematizing them may be best understood as revisionary.
>
> (Jackman 2009, 167)

The reason why Jackman thinks that the case of semantics is different is that it is prototypes that underlie our semantic intuitions, and that semantics, however, is after *concepts* which can at best be *normatively regimented* on the basis of prototypes (Jackman 2009, 167).[18]

It is not clear why Jackman thinks that conceptual analysis is playing a role at all here. Either he seems to think that theories of reference are conceptual analyses of REFERENCE which would get some support with what he is saying later in the paper[19], but is obviously in tension with the idea that—what he calls—"type-2 intuitions" are intuitions about terms like 'Gödel' when uttered in specific situations. Or, less likely, he may not be distinguishing properly between theories of reference and theories of meaning, thinking that the theories in question are in the business of reconstructing the concept underlying our application of 'Gödel'.

As Jackman correctly notes earlier in his paper, we have learned from discussions about externalism that our terms do not necessarily refer to what we take them to refer to (Jackman 2009, 164). Thus, our intuitions about the reference of individual names taken out of context do not inform us much about the actual reference of those names and gives little or

18 A similar account of conceptual analysis is defended in (AUTHOR).

19 Later in the paper Jackman discusses how different systematizations could lead to pluralism in the concept of reference, such that it would be correct for Asians to say about us and our usage of 'Gödel' that the latter "refers" to Schmidt. But this seems to assume that what is at issue is a conceptual analysis of REFERENCE. No doubt, 'reference' is a technical term with a specific, regimented meaning in philosophy of language. But, as we argued elsewhere (AUTHOR), theories of reference are accounts of how terms of certain semantic types refer; there is no reason to assume that such theories would fall out of a conceptual analysis of a technical term like 'reference'. But if, on this latter conception, a second-level theory of our first-level intuitions about the reference of expressions would *normatively* influence our intuitions, then philosophers would end up speaking a language different from ordinary English, and it's not clear why philosophers should be interested in *that*. (Note that, on the other hand, it is attractive to have regimented technical concepts, even if they diverge from those expressed by terms in ordinary English. See AUTHOR for details.)

no valuable input to theorizing about reference. However, the intuitions here considered, namely our intuitive application and interpretation of names, given the complete relevant context of utterance (including, for example, facts about beliefs of the speaker, causal chains of name usage etc.) are---just like syntactic intuitions---only mistaken if a performance error occurs. The basis for *that* isn't a concept (or its psychological counterpart), but our competence with proper names in general. This competence, in turn, doesn't consist in always using names in the correct way, nor in always interpreting them correctly, but in being sensitive to the right kind of information and circumstances in one's usage and interpretation of the name.[20]

4.2 Do thought experiments in philosophy of language test first or second-level intuitions?

If the constitutive role model is adequate for philosophical semantics, and observations 1-3 are correct, *and* if the intuitions that are typically considered in philosophical semantics and the discussion of theories of reference were first-level intuitions, the expertise defence would be a non-starter. However, Michael Devitt and Genoveva Martí think it is not first-level intuitions, but second-level intuitions, that are tested in the discussion of theories of reference (Devitt 2011, Martí 2009, 2014). As we said above, second-level intuitions play merely an indirect evidential role for second-level theories. If Devitt and Martí were right, then the expertise defence would be back in the game: perhaps the second-level intuitions of experts (with second-level theories) are more accurate than those of laypersons.

Their view seems to be this: they agree that first-level intuitions of lay speakers would be evidence for theories of reference, but second-level intuitions of lay speakers are not. If intuitions are tested in philosophical thought experiments, the intuitions triggered are the

20 For details, see (AUTHOR).

second-level intuitions of the semanticist. The semanticist is an expert with linguistic usage and has carefully observed his or her own usage, as well as the usage of other speakers, to arrive at an intuitive generalization about it. It is this judgment that is used as evidence in philosophical semantics.

Martí calls the second-level intuitions "metalinguistic" intuitions. How to tell whether an intuition is "metalinguistic"? The idea seems to be the following: if an intuition is reported using meta-linguistic vocabulary (such as 'refers to', 'is about', 'is true') that talks about the referential and semantic properties of expressions, then the intuition is metalinguistic. Here is Edouard Machery's attempt to interpret this distinction (we quote this in some length, because we think that the way in which Machery struggles to make sense of the distinction to some extent reveals what's wrong with it):

> To substantiate this criticism, Martí draws a distinction between two types of intuitions: metalinguistic intuitions and what we will call 'linguistic' intuitions. Meta-linguistic intuitions are judgements about the semantic properties of mentioned words (e.g. their reference), while, if we understand Martí correctly, linguistic intuitions are judgements about the individuals (substances, classes, etc.) described in the actual and possible cases used by philosophers of language. These judgements would be expressed by sentences using words rather than mentioning them. An example might clarify this distinction. In Kripke's Gödel case [...], the judgement that the proper name 'Gödel' refers to Gödel and not to Schmidt is a metalinguistic intuition, since it is about the reference of the proper name 'Gödel'; by contrast, the judgement that in this case Gödel should not have claimed credit for the incompleteness theorem is a linguistic intuition.
>
> (Machery et al., 2009, 689)

Thus, the *linguistic* intuition, which would be proper evidence for semantic theorizing, is the intuition that Gödel should not have claimed credit for the incompleteness theorem? This, obviously, can't be right. *That Gödel should not have claimed credit for the incompleteness theorem* is perhaps a moral intuition[21], but it is certainly not a linguistic intuition.[22]

But if *this* isn't a linguistic intuition, then what is? It seems that in order to be a *linguistic* intuition, the intuition would need to be *somehow* language related, but as soon as we have something explicitly language related, isn't the intuitive judgment automatically *meta*-linguistic? We believe that this failure to be able to explain what "linguistic" as opposed to "meta-linguistic" intuitions should be, points to an underlying confusion. In order to better understand where the confusion lies, we should reconsider Martí's initial objection to the study by Machery et al.

Machery et al had presented their test subjects with a vignette, retelling Kripke's Gödel/Schmidt case, featuring a hypothetical speaker, John, who only believes about Gödel that he proved the incompleteness of arithmetic (while according to the story that was in fact Schmidt's achievement). At the end of the story, they asked the following question:

When John uses the name 'Gödel,' is he talking about:

---

21 Depending on how 'should' is understood here, it could also be an instance of some other type of intuition, but certainly not a linguistic one.

22 As will hopefully become clear below, this isn't a linguistic intuition if judged by the same criterion which makes the judgment that the proper name 'Gödel' refers to Gödel and not to Schmidt a "metalinguistic intuition" in the quote above (i.e. by taking the statement expressing the judgment at face value and checking what it seems to be about). As we explain below, there is---also according to this criterion---a sense in which there *is* a linguistic intuition *encapsulated* in the judgment that Gödel should not have taken credit for the work: the linguistic intuition in this case is the speaker's intuitive choice to use the name 'Gödel', rather than some other expression, in reporting his or her moral intuition.

> *(A) the person who really discovered the incompleteness of arithmetic?*
>
> or
>
> *(B) the person who got hold of the manuscript and claimed credit for the work?*
>
> (Machery et al., 2004, B6)

Martí objected that the question asked at the end of the vignette triggered the wrong kind of intuition:

> Presenting the participants in the experiment with the Gödel/Schmidt scenario and asking them whether John talks about (i.e. refers) to one or another of two people, described under alternatives A and B, when he uses 'Gödel', invites the participants to reflect on how, in their opinion, reference is determined, and so the question MMNS pose tests opinions as regards which way of thinking, i.e. theorizing, about reference determination is the correct one.
>
> (Martí, 2009, 44)

Indeed, if a test-subject would, in answering this question, "reflect on how, in [her] opinion reference is determined", the judgment would, in our terminology, be a second-level judgment. However, although this is one way to arrive at an answer to the test question, it is not the only way. Alternatively, the test subject could arrive at an answer by considering a hypothetical utterance by John, and reporting how she would intuitively interpret that hypothetical utterance. She could, for example, imagine John saying something like 'Gödel was a brilliant mathematician' and see whether she would understand this as referring to the person who proved the incompleteness theorem or the guy who claimed credit for the work.

Martí has, of course, a valid point against the study by Machery et al.; the question at the end of the vignette *could* have been understood as asking for a second-level intuition. It *might* "invite the participants to reflect on how, in their opinion, reference is determined", and thus it *might* have triggered the wrong responses; second-level intuitions of lay speakers which are, arguably, not very good evidence for theories of reference. But this doesn't mean that the question at the end of the vignette *had* to be answered that way by a lay speaker.

This leads back to the points that we made earlier in discussing first and second-level intuitions about folk psychology and that we summarized in observations 4 and 5. The statement 'In John's utterance 'Gödel' refers to the guy who claimed credit for the work' can report a first-level intuition, a second-level intuition, or perhaps no intuition at all. What it reports isn't a matter of the terminology used in the statement—it's not a matter of whether the word 'reference' or cognate words occur in the sentence. What matters is which kind of considerations form the basis for the judgment. To see this, compare the statements

(a) This rose is red.

(b) 'This rose is red' is true.

(b) is meta-linguistic in the sense that it mentions a sentence and uses semantic vocabulary. It might be that your reasons for accepting or judging (b) are based on your (intuitive or reflected) views about semantics and about what makes sentences true, given that you believe that this rose is red. But you might just as well have judged (b) on the same basis as (a), viz. based on your belief about the colour of this rose, given your mastery of the predicate 'is true' and the use of quotation marks. The fact that an intuition report contains semantic vocabulary does not automatically indicate that the intuition reported was metalinguistic. In fact, we believe that the best way to understand the Gödel thought experiment is not to understand it

as intended to elicit a second-level intuition. We think that the original Gödel thought experiment as well as many other thought experiments in the philosophy of language (Author) ask for judgments made on the basis of our linguistic competence, *not* on the basis of our empirical generalizations, formed by observing linguistic behaviour of ourselves and others in the past (pace Devitt).[23] Hence it asks for a judgment based on first-level intuitions.

4.3 What are the commitments of this view?

It is not clear to us how we could positively show that our view about how thought experiments in philosophy of language are intended is the right one. After all, we can only say that, as addressees, we understood many thought experiments in this way, while Devitt apparently always understood them in his way. We could point out that, given our reconstruction of which intuitions are involved in evaluating semantic theories, and the fact that first-level intuitions are constitutive for the kind of second-level theories that semantic theories are, it would make a lot of sense to understand thought experiments in our way: if we are right, then philosophers of language have not been massively misled in their armchair methodologies. To be sure, we do not consider such considerations conclusive.[24] However, in the absence of strong reasons against it, we think our view should be preferred, given that it gives a coherent overall view of why and how thought experiments can be relevant for theorising about reference. We will now show that the two main objections Devitt provides

23 Unless beliefs about such empirical generalizations were part of linguistic competence. As we will explain below, we only need to assume that there is such a thing as linguistic competence; we don't need to make any strong assumptions about its realization in the human mind/brain.

24 There is a trend in metaphilosophy to argue for the epistemic purpose of a certain methodology on the basis of what would make most sense given that we use that methodology. This form of argument strikes us as bizarre, since it puts methodologies first and the epistemic interests second. It's like, "hey, since we don't want to leave these comfy armchairs, what could we be interested in, so we can study it from here?"

against our kind of view fail. Perhaps that will make it easier to accept our interpretation of the thought experiments in question.

Devitt seems to have two related objections. One objection is that our view presupposes an empirically implausible connection between linguistic competence on the one hand, and reports of the outputs of that competence on the other. The second objection is that our view presupposes an empirically implausible account of the psychological realization of linguistic competence. Fortunately, we don't have to engage in the difficult empirical discussion of the second question. We will argue that our view simply isn't committed to either of these two empirical assumptions.

Let us consider the view that the judgments that are elicited by philosophy of language thought experiments (such as, in the example above, that in John's usage 'Gödel' refers to the person who got hold of the manuscript and claimed credit for the work) are simply expressions of the subject's linguistic competence. Devitt calls this the *Voice of Competence View*, VoC. As we explained above, we do believe that thought experiments in philosophy of language ask for reports of elicited first-level intuitions, and we also argued that a report of a first-level intuitive understanding of John's utterance can well have the form above, i.e. make use of expressions like 'refers', 'means that' etc.[25] Thus, in this sense, our view seems to fall under Devitt's characterization of VoC. But is our view also affected by Devitt's criticism of VoC? Here are his objections:

1. VoC presupposes that semantic rules are either represented in the mind/brain or otherwise embodied, but the former seems psychologically implausible, and it's not clear what the latter is supposed to mean (c.f. Devitt 2011, 17).

25 Remember that Devitt classifies metalinguistic judgments on the basis of their syntax, not on the basis of the source of the judgment.

2. VoC presupposes that there is a more or less direct path from linguistic competence to—what we call—first-level intuition reports.

> 2.1 It is unjustified to simply assume that there is such a path from abilities to reports about the outputs of those abilities, since there are obvious examples where such path is lacking (examples are touch typing and bicycle riding) (c.f. Devitt 2011, 17).
>
> 2.2 There is empirical evidence showing that "the ability to speak a language and the ability to have intuitions about the language are quite distinct, the former being acquired in early childhood, the latter, in middle childhood" (Devitt 2011, 18). Thus, VoC is false and the assumed link between linguistic competence and the judgments elicited in philosophy of language thought experiments is broken (c.f. Devitt, 2011, 17).

It is true that some Chomsky-inspired linguists and philosophers of language feel inclined to speculate about the psychological realization of linguistic competence, but we don't see any reason why we need to engage in such speculation. Semantic linguistic competence *in a language*, the way we understand it, is the competence to use and interpret expressions in correspondence with the prevalent way that speakers in the relevant linguistic community use and interpret those expressions. Normal human beings that grow up in linguistic communities acquire that competence at some point. It is undeniable that there is such a competence; what its psychological realization looks like, is for our purposes a secondary question (see AUTHOR). Thus, the objection in (1) doesn't apply. It's an interesting psycholinguistic question how semantic competence is psychologically realized, but VoC doesn't have to take a stance and speculate about that realization.

A similar reply should be made in response to objection 2.1. The objection was that VoC doesn't have a good account of how linguistic competence issues in metalinguistic

judgments. Moreover, there doesn't seem to be a good account forthcoming, because typically there simply is no competence to make reliable judgments about other competences we have, such as touch typing or bike riding. Indeed, we would have a hard time recognizing and reporting, for example, which finger we use to press, say, the letter 'l' when typing 'irrelevant'. But do we seem to have an equally hard time to recognize and to report, for example, what we understood the name 'Gödel' to be referring to in the utterance we just heard? Or how we meant a certain expression in an utterance we just produced, or which expressions we would use to describe a certain situation?[26]

It seems obvious to us that we don't have as great difficulties as in the touch-typing case, in reporting how we understood or meant something. In fact, this is a rather central part of our everyday communication. We spend a lot of time talking about what other people said and what we have meant in saying something, as we noted in observation 3 this makes it plausible that we have some competence in reporting these matters accurately.

Objection 2.2, in turn, claims that even though it is true that we make meta-linguistic judgments like these, there is empirical evidence that these are not in fact reporting the outputs of our linguistic competence, but are an altogether different competence. Here is Devitt again:

> Since writing *Ignorance*, I have become aware of a body of developmental literature that provides persuasive empirical evidence against VoC. The evidence suggests that the ability to speak a language and the ability to have intuitions about the language are quite distinct, the former being acquired in early childhood, the latter, in middle childhood. Carson Schütze ends a critical discussion of much of this evidence with the observation that "it is hard to dispute the general conclusion that metalinguistic

26 That meaning is consciously available to us, and thus in principle reportable, should arguably be taken as a fundamental fact that a theory of meaning has to accommodate (cf. Smith 2006).

> behavior is not a direct reflection of linguistic competence" (Schütze 1996, 95). It looks as if VoC is false.
>
> (Devitt, 2011, 18)

First of all, the reported empirical evidence holds perhaps for grammaticality judgments, and it is not clear that it carries over to semantics. But, anyway, it wouldn't be surprising if also the ability to report first-level *semantic* intuitions was a competence that developed later than the ability to speak and understand a language. After all, one might first need to acquire meta-linguistic concepts in order to be able to express reports of first-level intuitions. However, the fact that there are two competences, viz. the competence to speak and understand a language, and the competence to recognize and report outputs of that first competence does not conflict with VoC in any way. It would only conflict with it if we were assuming that the second, later competence was *unrelated* to the first. As Schütze correctly explains, there are two extreme views about metalinguistic judgment:

> There are two extreme positions one can take on the relation between these two processes [of judging a sentence as grammatical/ungrammatical and the ordinary processing of a sentence in conversation]: they might be identical, or they might be totally different. In the first case, some might argue (this is perhaps the null hypothesis) that the only difference between processing for judgment and processing for conversation is that in the former case the reply consists of a "yes" or "no" (or a numeric rating, or whatever), instead of a pragmatically related utterance. Obviously, the decision between the possible judgments has to come from somewhere, but on this view the processing of the sentence itself is identical. [...]

> At the other extreme, one might say that judging is nothing at all like understanding and involves none of the same cognitive mechanisms. If you are told you will have to judge a sentence, you route it to the sentence-judging processor in the mind, rather than the sentence-comprehending processor. These two modules are entirely separate and might differ in arbitrary ways. (If this were put forward as a serious proposal, one would have to address the question of how and why such a separation would come to exist in the mind.)
>
> (Schütze, 1996, 81-2)

VoC is not committed to the first extreme. Devitt seems to commit himself to the other extreme, however. He seems to think that our ability to make intuitive metalinguistic judgments develops from our observation of linguistic behavior (that of others and ourselves), and that these judgments are thus normal empirical judgments, not judgments based on the recognition and report of outputs of our linguistic competence:

> The ordinary competent speaker is surrounded by expressions that refer and so is in a good position to have well-based opinions about reference by reflecting on these expressions. This is not to say that she will reflect. Indeed, a totally uneducated person may reflect very little and hence have few if any intuitive judgments about her language. Still it is clear that the normal competent speaker with even a little education does reflect on linguistic reality just as she reflects on many other striking aspects of the world she lives in. And this education will usually provide her with the terms and concepts of folk semantics, at least. As a result she is likely to be able to judge in a fairly immediate and unreflective way what an expression refers to. Such intuitive opinions are empirical central-processor responses to linguistic phenomena.

> They have no special authority: although the speaker's competence gives her ready access to data it does not give her privileged access to the truth about the data.
>
> (Devitt, 2011a, 426)

In our terminology, Devitt thinks that these judgments are reports of second-level intuitions, not reports of first-level intuitions. But this is *entirely* implausible. Even if one agrees that it takes expertise to recognize and report first-level intuitions in meta-linguistic vocabulary, that doesn't make these reports something other than reports of first-level intuitions.

For the case of grammaticality judgments, Schütze quotes Birdsong's remark that "meta-linguistic data are like 25-cent hot dogs: they contain meat, but a lot of other ingredients, too. Some of these ingredients resist ready identification" (Schütze 1996, 54). VoC could live with that, if the "meat" is first-level intuition. What Devitt seems to think is that meta-linguistic data are like one dollar "not dogs"[27].

5. Expertise and Experiment

Let's now turn back to the discussion of the experimentalist challenge and the expertise reply to it. As we have already indicated, there seem to be some valid objections against the study by Machery et al. There is on the one hand, as we already mentioned, Genoveva Martí's objection that Machery et al. tested the wrong intuitions. On our account, the intuitions of interest are first-level intuitions. Martí seems to think that those weren't tested at all, but that the vignette only tested second-level, metalinguistic intuitions. Here we disagree. What is problematic, according to our account, is that the vignette *could* be understood as asking for a

27 "Not dogs" are vegetarian hot dogs, also sometimes called "pickle dogs", cf. Wikipedia (2012).

second-level intuition. Because first-level as well as second-level intuitions can be reported by syntactically metalinguistic sentences, and laypersons might not have understood the task in the vignette as a question about their first-level intuition.

As was pointed out and empirically shown by Sytsma and Livengood (2011), the vignette was also ambiguous as to the perspective the test subject was supposed to take. Should she answer the question from John's perspective, or from the narrator's perspective? Moreover, Sytsma and Livengood showed that, once that ambiguity is resolved, the variation in test subjects' responses is significantly reduced.

Both of these observations, in addition to the remarks we already made above, point to another expertise defence against the experimental challenge. Perhaps the intuition reports of philosophers of language are better than those of lay speakers, because philosophers of language much better understand what type of intuition, from what kind of perspective, is being asked for in a thought experiment. We also said above that it is one thing to have the relevant first-level intuitions, and another to be able to recognize and accurately report them. Since first-level intuitions of ordinary speakers are constitutive for reference, and since second-level theorizing can't make first-level intuitions "better", the first-level intuitions of expert philosophers aren't possibly "better" or more relevant than those of ordinary folk. However, not everyone needs to be very skilled in recognizing and reporting first-level intuitions, and there is reason to think that philosophers of language are better at this than the folk (cf. observation 2).

One of the reasons to think so is that reporting first-level intuitions in second-level terminology requires that this terminology is sufficiently entrenched. 'Reference', for example, is a technical term in philosophy of language. It has acquired a very specific meaning since the development of theories of reference and denotation at the beginning of the last century. Thus when philosophers of language report what a certain expression refers to,

given how they understood an utterance, they will presumably give more consistent answers than lay speakers who don't have the philosophers' regimented notion of reference.

Note that being competent with a certain technical vocabulary doesn't mean that judgments expressed in that terminology are therefore theory-laden. As we've argued elsewhere, 'reference' in philosophy of language is a functional kind term (AUTHOR). Philosophers can well agree on the functional characterization of reference, while disagreeing about which theory properly identifies the relation that satisfies that functional role (for a particular expression). Thus there is a crucial role for expertise after all.

In light of this result, one might now perhaps wonder whether our objection to the discussion between Machery and Devitt is merely verbal. Perhaps Devitt's expertise defence should just better be phrased in terms of the experts' privileged access to the right vocabulary for describing the kind of intuitions she shares with the layperson, and Machery should be understood as attacking *that* claim.

That the debate between Machery and Devitt doesn't turn on this, can be seen when we consider the argument Machery presents against Devitt's claim that experts about language should have more reliable intuitions about metalinguistic matters than the folk, because they have access to better theories. Machery reports data he takes to indicate that linguists working in socio-linguistics have different intuitions about reference than analytic philosophers of language, and argues on this basis that expertise with language does not improve one's intuitions (Machery, 2011, 50).

If we are right, this result shows nothing against our version of the expertise defence. It is to be expected that philosophers of language, who are familiar with what is at stake in theories of reference, will understand the vignette in the intended way, and will thus be more likely to give accurate reports of their first-level intuitions than linguists from other areas. Moreover, they are more likely to report those first-level intuitions that are *relevant* for

theories of reference, rather than some other intuitions. That's not because those other linguists' first-level intuitions are inferior, it's just that this methodology involves so many ambiguities that experts are more likely than non-experts to get the task right.

Experimental philosophers might reply that this is a defence strategy that merely attempts to turn a problematic theoretical bias into an epistemic virtue. We agree that this is a real worry. Thus we should try to find better, alternative evidence for theories of reference: evidence that gives us access to first-level intuitions, but not via the problematic and ambiguous metalinguistic statements. How could that be achieved?

Devitt, who also believes that first-level intuitions (he considers them to be facts about usage) are much more relevant for theories of reference than second-level intuitions, suggests the following methodology: instead of asking the test subjects how they understand an expression in a certain hypothetical utterance, one should instead confront the test subjects with actual utterances and then find out how they understood the utterance on the basis of what belief states they would ascribe to the speaker of the utterance. If we have reason to assume that our test subjects are themselves competent in using the vocabulary in question (a proper name '*a*', for example), their ascriptions should provide us with evidence of what is required for this name to refer to *a*:

> These attributions by the subjects are significant because we theorists can then reason as follows. If the subjects are right in the meanings they ascribe, then those meanings will all be about *a*: the attributions won't be right unless the thoughts and utterances of the [the speakers] co-refer with the subjects' '*a*' and the subjects' '*a*' refers to *a*. And we have good reason to suppose that the subjects will mostly be right in their attributions: first, we can expect the subjects to be, like the rest of us, generally successful in ascribing meanings to explain behavior and gain information about the

> world; second, given their indubitable competence with '*a*', there is no reason to suppose that this general success will not be exemplified in this particular context.
>
> (Devitt, 2011a, 430-31)

Although we think it is progress to test the comprehension of actual utterances rather than that of hypothetical utterances, we doubt that evidence from reports of belief ascriptions will straightforwardly inform us about what the test subjects took the speaker to be referring to: referential relations are just one among the many factors that can be relevant for our belief ascriptions. Here is a real life example to illustrate this point. One of us regularly teaches logic courses and asks in the final exam questions about the decidability of theoremhood in first-order predicate logic. As an answer to the question 'Can you prove in first-order predicate logic for any given well-formed formula whether it is a theorem? Please substantiate your answer.' he sometimes gets as an answer 'No you can't, due to a result proved by Alfred Tarski in 1931'. Since he knows what he told his students about Tarski, Gödel, incompleteness and decidability, the belief that the teacher ascribes to the student, is that the student mistakenly believes that Gödel's incompleteness theorem established that first-order predicate logic is undecidable. Note that the name 'Gödel' did not play a role in the student's answer, nor is it part of the correct answer to the question. If, on the other hand, asked what the student was literally referring to with the name 'Tarski', that teacher would certainly have said that the student actually referred with it to Tarski. In other words, belief ascriptions depend on myriads of contextual factors that have to do with why we are interested in ascribing those beliefs (in this case, the teacher is interested in what the student understood about metalogic, not in whether the student is able to remember names of famous logicians correctly). But if belief ascriptions are context dependent, the semanticist will have a hard time using these to distill how the test subjects understood the proper names used.

A much better way to test how expressions are semantically interpreted would be to test how subjects understand the reference of a certain expression without having to ask them anything at all. Fortunately, there are such ways. For example, psycholinguists have for some time now successfully used eye-tracking in the so-called virtual world paradigm, in order to test the resolution of referring expressions in the comprehension of utterances (cf. Tanenhaus and Spivey-Knowlton, 1996). In these experiments the subjects are looking at a static or moving visual scene, which contains, among some distractors, objects that are potential referents for the referring expressions in the tested utterances. The test subject then receives as a stimulus the utterance, and it is measured at what time after onset or offset of the expression in question, fixation probabilities for eye movements significantly favour which of the visible objects.

Since these eye movements are automatic, not under our voluntary control, and clearly correlated with language comprehension, this data provides evidence about how hearers semantically interpret referential expressions that is much more direct than that obtainable from surveys. There is no need to ask the test subjects afterwards how they interpreted the utterance semantically, but if one would like to know how reliable we are in reporting such first-level comprehension intuitions, one should test whether these reports agree with the eye-tracking data.

We believe that this is more likely to be a way forward in the experimental investigation of semantics than the use of vignettes modelled after thought experiments, or the indirect evidence one can hope to get from belief ascriptions. Perhaps there are other experimental methods that could be used, too. The main shortcoming of the existing empirical studies, we think, is that they have not distanced themselves far enough from the thought-experiment setups commonly used in philosophy. As we have explained above, thought experimentation about fictional (and often far-fetched) scenarios can be a perfectly

legitimate tool for a philosopher to investigate his or her own first-level intuitions, but when we study non-philosophers, who lack expertise in dealing with thought experiments, other kinds of methods may be in order.

6. If Intuitions Vary Then What?

We have argued that the discussion of the expertise defence in reaction to the study of Machery et al. is on the wrong track, since—insofar as intuitions are relevant for philosophical semantics—first-level intuitions matter and they don't improve from second-level theorizing. We also said, however, that expertise might be required for reliably reporting first-level intuitions, which might, after all, to some extent account for the variation found by Machery et al.

But what if the factors we cited as possibly explaining the variation, do not explain it? In the last section we suggested a methodology for testing first-level intuitions directly. What if such experiments reveal the same or even more variation in the comprehension of referential expressions? We believe that here is an important second way in which the debate between Machery et al. and their critics is on the wrong track. If there is variation in first-level intuitions, then this doesn't show that our intuition-based methodology is flawed. It can't show *this*, because first-level intuitions are *constitutive* for second-level theories. So, what conclusion should be drawn from variation in first-level intuitions for theories of reference?

We believe that this depends on the amount and systematicity of the variation. In Machery et al. (2004) the claim was that the variation in intuitions is correlated with cultural background. If there were such variation in first-level intuitions, the right thing to do would

be to propose different theories of reference for the different cultures. If the variation were found to be correlated with other factors (stress, drugs), the right thing to do would be to propose a competence/performance distinction. If, however, the variation were found to be wildly unsystematic and ultimately random, we'd agree with Mallon et al. (2009) that this should discourage theorizing about reference. The reason, however, would *not* be that we don't have enough good evidence for reference to continue; the reason would rather be that there simply isn't a phenomenon of semantic reference that could be studied at the level of shared languages and that could play a substantial part in the systematic explanation of events of successful communication. First-level intuitions *constitute* reference: if there were no pattern and no systematicity to the speakers' first-level intuitions, theorizing about them would have no value (at least as far as theories of reference and the explanation of linguistic communication are concerned). Whether our first-level intuitions *do* exhibit enough systematicity and pattern for the existence of semantic reference (and for the viability of theorizing about reference) is an empirical question, but at present there is no evidence whatsoever to think that our semantic first-level intuitions lack the necessary amount of systematicity.

References

Alexander, J. 2012: *Experimental Philosophy: An Introduction*, Cambridge: Polity.

Baillargeon, R. 2004: Infant's Physical World, *Current Directions in Psychological Science*, 13, 89-94.

Bealer, G. 1996: A priori knowledge and the scope of philosophy, *Philosophical Studies*, 81, 121-142.

Cappelen, H. 2012: *Philosophy without Intuitions*, Oxford: Oxford University Press.

Devitt, M. 2006: *Ignorance of Language*, Oxford: Clarendon Press.

Devitt, M. 2011: Whither Experimental Semantics? *Theoria*, 72, 5-36

Devitt, M. 2011a: Experimental semantics. *Philosophy and Phenomenological Research*, LXXXII, 418-35.

Devitt, M. 2012: The role of intuitions in the philosophy of language. In Gillian Russell and Delia Graff Fara (eds.), *Routledge companion to the philosophy of language*. New York: Routledge, 554-65.

Devitt, M. 2012a: Semantic Epistemology: Response to Machery. *Theoria*, 74, 229-233.

Goldman, A. 2007: Philosophical Intuitions: Their Target, their Source, and their Epistemic Status. *Grazer Philosophische Studien*, 74, 1-26.

Jackman, Henry. 2009. Semantic Intuitions, Conceptual Analysis, and Cross-Cultural Variation. *Philosophical Studies*, 146, 159-177.

Ludwig, Kirk. 2007. The epistemology of thought experiments: First-person approach vs. third-person approach. *Midwest Studies in Philosophy*, 31, 128-59.

Machery, E. R. Mallon, S. Nichols, and S. P. Stich 2004: Semantics, cross-cultural style. *Cognition*, 92, B1-B12.

Machery, E., C. Y. Oliviola, and M. de Blanc 2009: Linguistic and metalinguistic intuitions in the philosophy of language, *Analysis*, 69, 689-694.

Mallon, R., E. Machery, S. Nichols, and S. P. Stich 2011: Against Arguments from Reference. *Philosophy and Phenomenological Review*, 2009, 332-356.

Machery, E. 2011: Expertise and Intuitions about Reference. *Theoria*, 72, 37-54.

Martí, G. 2009: Against semantic multi-culturalism. *Analysis*, 69, 42-48.

Martí, G. 2014: Reference and Experimental Semantics. In E. Machery and E. O'Neill (eds.). *Current Controversies in Experimental Philosophy*. Routledge 2014, 17-26.

Nagel, Jennifer 2012: Intuitions and Experiments: A Defense of the Case Method in Epistemology, *Philosophy and Phenomenological Research*, LXXXV, 495-527.

Reed, N., McLeod, P., Dienes, Z. 2010: Implicit knowledge and motor skill: What people who know how to catch don't know. *Consciousness and Cognition* 19, 63-76.

Schütze, Carson T. 1996. *The empirical base of linguistics*. Chicago: The University of Chicago Press.

Smith, Barry C. 2006. Publicity, Externalism and Inner States. In Tomáš Marvan (ed.), *What Determines Content?: The Internalism/Externalism Dispute*. Cambridge Scholars Press.

Sytsma, Justin, and Jonathan Livengood. 2011. A new perspective concerning experiments on semantic intuitions. *Australasian Journal of Philosophy*, 89, 315-32.

Tanenhaus, M.K. and Spivey-Knowlton, M.J. 1996. Eye-tracking. In F. Grosjean & U. Frauenfelder (eds). *Language and Cognitive Processes: A guide to spoken word recognition paradigms*, 11, 583-588.

Textor, Mark. 2009. Devitt on the epistemic authority of linguistic intuitions. *Erkenntnis*, 71, 395-405.

Weinberg, Jonathan, Shaun Nichols and Stephen Stich. Normativity and epistemic intuitions. *Philosophical Topics*, 29, 429-60.

Weinberg, Jonathan. 2007. How to challenge intuitions empirically without risking scepticism. *Midwest Studies in Philosophy*, XXXI, 318-343.

Wikipedia 2012. http://en.wikipedia.org/wiki/Vegetarian_hot_dog.

Williamson, Timothy. 2007. *The philosophy of philosophy*. Malden, MA: Blackwell.