Experimental Philosophy and the Theory of Reference

MAX DEUTSCH

Abstract: It is argued on a variety of grounds that recent results in 'experimental philosophy of language', which appear to show that there are significant cross-cultural differences in intuitions about the reference of proper names, do not pose a threat to a more traditional mode of philosophizing about reference. Some of these same grounds justify a complaint about experimental philosophy as a whole.

1. Introduction

The 'experimental philosophers', Ron Mallon, Edouard Machery, Shaun Nichols, and Stephen Stich (2004 and forthcoming) claim to have empirical evidence, in the form of survey data, that East Asians' intuitions about the reference of proper names differ significantly from those of Westerners. Mallon *et al.*[1] think that, since philosophers of language, and Saul Kripke in particular, argue in a way that assumes that intuitions about reference are universally shared, the empirical data they have gathered pose a serious problem for traditional theorizing about reference.

In what follows, I argue, firstly, that Mallon *et al.* misunderstand the way in which Kripke, for example, argues for and against the claims about reference of interest to him. As I will show, nothing in Kripke's famous argument against the *descriptivist theory of reference* for proper names hinges on assuming anything about peoples' intuitions. Secondly, I argue that Mallon *et al.*'s claim that there is significant cross-cultural variability in intuitions about the reference of proper names is not shown, nor even suggested, by the results of their studies. Mallon *et al.*'s experiments on intuitions are seriously flawed in a way in which anyone who is cognizant of the distinction between *semantics* and *pragmatics* can easily recognize.

This pair of criticisms of Mallon *et al.*'s experimental philosophy of language can be leveled at a good deal of other work being done in the burgeoning experimental philosophy movement. In much of this work, it is incorrectly assumed that

Thanks to John Collins, Harry Deutsch, Nick Georgalis, Mike Veber, and Gene Witmer for helpful comments on an earlier draft, and to an anonymous referee (this journal) whose comments led to a number of significant changes. I am especially grateful to Ron Mallon, who carefully read an earlier draft, and offered excellent advice on the paper's content and style.

Address for correspondence: 311 Main Building, Department of Philosophy, The University of Hong Kong, Pokfulam Road, Pokfulam, Hong Kong, China.
Email: medeutsc@hkusua.hku.hk

[1] Machery is lead author of the 2004 paper, while Mallon is the lead author of the forthcoming paper. Since most of the quotations come (with kind permission from its authors) from the forthcoming paper, I have chosen to refer to the collective as 'Mallon *et al*'.

Mind & Language, Vol. 24, No. 4 September 2009, pp. 445–466.

more traditional philosophers adopt methods that require uniformity in intuitions. And many experimental studies of philosophical intuitions exhibit a deleterious insensitivity to the semantics/pragmatics distinction. In the penultimate section of the paper (section 6), I criticize experimental philosophy as whole on precisely these grounds.

2. Kripke's Method

Kripke's main target in *Naming and Necessity* is a descriptivist theory of *meaning* for proper names, according to which the semantic content of a name is identical with the semantic content of the definite descriptions users of the name associate with it. Kripke offers direct arguments against this descriptivist theory of meaning, but he also objects to it indirectly by criticizing the theory of reference it entails. *D* encapsulates the theory of reference that is a consequence of the descriptivist theory of meaning:

> *D*: An ordinary proper name, *n*, as used by a given speaker, *S*, refers to the object that is the denotation of some/most/all of the definite descriptions *S* associates with *n*.

To show that *D* is false, Kripke simply describes counterexamples—cases in which a name, as used by a given speaker, does not refer to the denotation of the definite description(s) the speaker associates with the name. Here is one such case, one of Kripke's own (Kripke, 1972/1980, pp. 83–84): Imagine that Gödel did not prove the incompleteness of arithmetic but that some other man, Schmidt, did. Gödel stole the proof from Schmidt and published it under his own name. But now imagine a speaker who uses 'Gödel', but associates just a single description with it, namely 'the prover of incompleteness'. To whom does this speaker's uses of 'Gödel' refer, Gödel or Schmidt? The answer, Kripke says, is Gödel, not Schmidt. If Kripke is right, *D* is false.

The method Kripke employs in arguing against *D*—the method of proposing a counterexample to a generalization—is commonplace. It is used in philosophy but also in logic, mathematics, science, history, sociology, and even in postmodernist literary criticism. It is not a method confined to the philosophy of language, and there is nothing at all surprising or suspicious about it. One might think that Kripke is wrong about his case. Even so, the way he attempts to use it against *D* is utterly routine. Rational enquirers everywhere attempt to falsify generalizations with counterexamples.

Of course, Mallon *et al.* cannot have any reason to be suspicious of the method of presenting counterexamples to generalizations. However, in their 2004 and forthcoming, they suggest that there is something deeply amiss in Kripke's argument against *D*. They claim that Kripke's argument relies on a method peculiar to philosophy, and they go on to claim that this method is bankrupt, saying they have empirical evidence, in the form of survey data regarding peoples' *intuitions* about reference, to prove it.

These claims strike me as odd. Kripke's method seems to me no different in principle from the method someone might use in arguing against the generalization that, for example, *all mushrooms are edible*, namely by pointing to a poisonous variety of mushroom. And how, exactly, are peoples' *intuitions* about reference supposed to be relevant? There is no explicit appeal to intuitions either in my brief rendition of Kripke's argument, or in Kripke's original presentation in *Naming and Necessity*. In short, there seems to be nothing special, methodologically speaking, about Kripke's argument against *D*. It is an entirely ordinary use of a counterexample to object to a generalization. Nevertheless, we should examine Mallon *et al.*'s claims more closely. What they ultimately reveal, I think, are the strange misconceptions of at least one branch of the experimental philosophy movement.[2]

3. The Method of Cases

Mallon *et al.* claim that Kripke, and virtually every other philosopher of language working on the theory of reference, implicitly assumes a methodological claim that Mallon *et al.* label *the method of cases*:

> *The method of cases*:
> The correct theory of reference for a class of terms *T* is the theory which is best supported by the intuitions competent users of *T* have about the reference of members of *T* across actual and possible cases (Mallon *et al.*, forthcoming).

Mallon *et al.* then argue that the results of the surveys they have conducted show that there is no theory of reference that is 'best supported' by the intuitions of competent speakers. Their surveys show (or fairly strongly suggest), they say, that competent 'East Asian' English speakers tend to have 'descriptivist intuitions' compatible with *D*, while competent 'Western' English speakers tend to have 'Kripkean intuitions' incompatible with *D*. I will return in section 5 to these polls and their alleged significance for theories of reference. For now, I want to focus on Mallon *et al.*'s claim that Kripke and other philosophers of language implicitly assume the method of cases. What is their evidence for this?

[2] Experimental philosophy is a new movement whose practitioners design and run surveys meant to reveal peoples' intuitions about a variety of hypothetical cases and thought experiments that play a central role in philosophy. Mallon *et al.* have conducted several such surveys meant to probe peoples' intuitions about the reference of proper names. They claim that the results of these surveys seriously challenge Kripke's method and the method of philosophers of language more generally. I disagree with these claims, as the arguments in the main text will make clear. I note, however, that there are different kinds of experimental philosophy, not all of which take such a critical stance toward philosophical method. (See Nadelhoffer and Nahmias, 2007 for a useful sorting of the various branches of experimental philosophy.)

Philosophers of language implicitly assume the method of cases, Mallon *et al.* say, because they must, at least implicitly, have an answer to the 'preliminary, methodological question: How do we know which theory of reference is correct?' (Mallon *et al.*, forthcoming). Mallon *et al.* lament the fact that 'philosophers of language have rarely addressed this methodological issue explicitly', but insist that it is 'clear from the specific arguments for and against specific theories of reference' that philosophers of language implicitly assume correctness for a theory of reference to consist in what the method of cases says it consist in, namely compatibility with the intuitions of competent speakers (Mallon *et al.*, forthcoming).

But what is it about these specific arguments concerning specific theories of reference that convince Mallon *et al.* that the intuitions of competent speakers matter so crucially? Like any other theory, a theory of reference is true only if it makes true predictions. But the predictions of a theory of reference concern terms and their referents, not competent speakers and their intuitions. For example, *D* predicts that, in Kripke's fiction, the relevant speaker's uses of 'Gödel' refer to Schmidt, not Gödel. If the prediction is false, so is the theory, but the theory makes no predictions at all concerning who will intuit what. Hence, in presenting the Gödel case, Kripke does not, and need not, make any claims about competent speakers' intuitions. He need only say, as he does, that the speaker's uses of 'Gödel', in the case he describes, do not refer to Schmidt, contrary to the prediction about the case implied by *D*. Mallon *et al.* see an implicit appeal to intuitions in Kripke's presentation. But, in fact, Kripke need not have any beliefs at all, implicit or otherwise, about what competent speakers might intuit about the Gödel case, and if he did have such beliefs, they might directly contradict the claim made by the method of cases. On considering the matter, Kripke might well admit that competent speakers will disagree with him about the case. (Descriptivists tend to be competent speakers, after all.) but insist that whether 'Gödel' refers to Schmidt is not to be settled by polling competent speakers. Such a view would be perfectly consistent with the way in which Kripke argues against *D*.

Mallon *et al.* offer another example, besides Kripke on the Gödel case, of a philosopher of language arguing in a way that supposedly betrays adherence to the method of cases: Gareth Evans (1973) on his famous Madagascar case. Evans uses the case to challenge so-called 'causal-historical' theories of reference for proper names. A simple causal-historical theory has the consequence that 'Madagascar' refers to a portion of the African mainland, since, when the name was first introduced, it did refer to a portion of the African mainland, and there is a causal-historical chain linking past uses of 'Madagascar' to current ones. But, as Evans points out, the trouble for the simple causal-historical theory is that current uses of 'Madagascar' do not refer to a portion of the African mainland. They refer instead to the large island off Africa's eastern coast.

Now where in any of this is the implicit appeal to the method of cases? Perhaps Mallon *et al.* would object to my describing Evans as *pointing out* that 'Madagascar' refers to the island. Mallon *et al.* view philosophers of language as an exceedingly cautious lot; according to them, 'according to Evans (1973), *people have the intuition* that nowadays the proper name 'Madagascar' refers to the large island near the south of Africa' (Mallon *et al.* forthcoming; emphasis added). But Evans' view is not that 'people have the intuition' that 'Madagascar' refers to the island. His view is instead that 'Madagascar' refers to the island. It is a view about a name and that name's referent. It is not a view, not even indirectly or implicitly, about what people intuit about the name and its referent.

Furthermore, the view is plainly correct; 'Madagascar' refers to the island, not to a portion of the African mainland. Mallon *et al.* appear to believe that Evans, if only he had reflected on his own methodology, would have had to retract the claim that 'Madagascar' refers to the island, and would have had to patiently await the results of an opinion poll concerning competent speakers' intuitions about the referent of 'Madagascar'. But that is preposterous. A philosopher of language such as Evans, just as easily as anyone else, could have simply checked his world atlas and seen that 'Madagascar' refers to the island. Facts such as the fact that 'Madagascar' refers to the island are *data* for theories of reference. Facts such as the fact that competent speakers intuit that the island is the referent of 'Madagascar' are data for a psychological theory, one that does not have any clear bearing on a theory of reference.

More evidence that Mallon *et al.* are wrong to attribute the method of cases to philosophers of language derives from the fact that plenty of philosophers of language *explicitly deny* that we know which semantic theory is correct is by testing to see whether competent speakers' semantic intuitions accord with the theory. Consider, for example, the semantic view of proper names known as *Millianism* or *Naïve Russellianism*. According to the Millian/Naïve Russellian, the meaning of a name is just its referent. A consequence of this view is that coreferential names are intersubstitutable in all sentential contexts (except quotational contexts) preserving truth. For example, the Millian/Naïve Russellian will say that, given that 'Superman' and 'Clark Kent' corefer, and given also that 'Lois Lane believes that Superman can fly' is true, it follows that 'Lois Lane believes that Clark Kent can fly' is true as well. As Millians/Naïve Russellians are well aware, however, competent speakers of English familiar with the Superman stories will intuit that 'Lois Lane believes that Clark Kent can fly' is *false*. Given this fact about competent speakers' intuitions, if Millians/Naïve Russellians subscribe to something similar to the method of cases as their standard of correctness for a semantic theory of names, they ought to abandon their Millianism/Naïve Russellianism. But they do not abandon it. They argue instead that the intuitions of competent speakers are not the final arbiter of correctness for a semantic theory and attempt to 'explain away' the contrary intuitions in a way that leaves their semantic view unscathed.

Or consider the view that Russell's Theory of Descriptions gets the semantics for English definite descriptions right. On Russell's theory, the features of something that is not an *F* are never relevant to the truth of a sentence of the form, 'The *F* is *G*'. However, as the defenders of Russell's Theory well know, there are cases in which someone uses a definite description, 'the *F*', intending to refer to something that is not an *F*. And they also know that, about such cases, many competent speakers will have intuitions that appear to conflict with predictions of Russell's Theory. Bob says, 'The man in the corner drinking a martini is happy tonight', intending to refer to Sal, who is in the corner, but happens to be drinking sparkling water, not a martini. On Russell's Theory, what Bob says is true if and only if there is exactly one man in the corner *drinking a martini* and that man is happy. But many competent speakers will intuit that what Bob says is true if *Sal* is happy, even when they are aware that Sal is not drinking a martini. Defenders of Russell's Theory do not immediately throw in the towel. Instead, like the Millians/Naïve Russellians, they argue that the intuitions of competent speakers are not to be trusted in these cases. In other words, they explicitly deny that something akin to the method of cases, applied to semantic theories of descriptions, is 'how we know' which theory of descriptions is correct.

4. Intuitions as Evidence?

I suspect that Mallon *et al.* would agree that the predictions of a theory of reference concern terms and their referents as opposed competent speakers and their intuitions, and hence would agree that the facts directly relevant to determining whether a theory of reference is correct are certain semantic facts, not psychological facts about intuitions. Mallon *et al.* would likely claim, however, that philosophers of language argue in a fashion that treats the latter sort of fact as at least indirectly relevant, by treating such facts as providing *evidence* for or against the predictions of a theory of reference. Regarding the Gödel case, for example, Mallon *et al.* would say that Kripke is appealing to the *intuitiveness* of the judgment that 'Gödel' does not refer to Schmidt in order to *justify* that judgment. Charitably interpreted, Mallon *et al.*'s attempt to saddle philosophers of language with the method of cases is simply their way of insisting that philosophers of language treat intuitions as *a source of evidence* for judgments about cases involving the reference of proper names.[3]

I won't argue here that philosophers of language never treat intuitions as evidence. There are a variety of metaphilosophies that treat intuitions as evidence and perhaps some philosophers of language subscribe to one or another of these.[4] I also won't argue that the intuitiveness of p is not evidence of any kind for p, though this is something that I happen to believe. What I will argue is that the cogency of the very arguments on which Mallon *et al.* choose to focus, namely Kripke's argument against descriptivism, and Evans' argument against the causal-historical theory, does

[3] Interpreting Mallon *et al.* in this way requires charity because one might reject the view that philosophers of language assume the method of cases without rejecting the view that they sometimes treat intuitions as evidence. For example, perhaps Millians treat competent speakers' anti-Millian intuitions as a piece of evidence against their view. Clearly, however, they would, and do, deny that the best semantic theory of names (Millianism, according to them) is the theory best supported by competent speakers' intuitions.

[4] See Bealer, 1999; Goldman and Pust, 1999; and Pust, 2000.

not depend on treating intuitions as a source of evidence. Even if intuitions are evidence, they *need not* be treated as such in the case of these two arguments.

Evans' argument *begins* with the clearly true claim that 'Madagascar' refers to the island, not the mainland. An appeal to intuition to justify the claim would be quite strange, and entirely unnecessary in any case. Various facts about how people use 'Madagascar' (including the fact that some people have labeled the island 'Madagascar' in atlases) are of course relevant to justifying the claim, but that is very different from facts about peoples' intuitions about the referent of 'Madagascar' being relevant. If some group or culture were to intuit that 'Madagascar' does not refer to the island, they would simply be mistaken.[5]

Since it depends on consideration of a counterfactual scenario, it may seem that Kripke's judgment regarding the Gödel case must be justified by intuition. I think this appearance results from confusing what justifies the judgment with the judgment's causal source. The causal source of Kripke's judgment about the Gödel case is intuition; this much is fairly clear. Kripke does not literally see or otherwise perceive that 'Gödel' does not refer to Schmidt in his fiction, and he presumably arrives at the judgment in the spontaneous, noninferential way characteristic of intuiting that something is so. However, in my view, the judgment's intuitiveness is not evidence that the judgment is true. Furthermore, Kripke himself does not say or suggest that he takes the judgment's intuitiveness as evidence for its truth.

In his own discussion of the case, Kripke, after spinning the tale of Gödel and Schmidt, and using 'we' to refer to those of us who, in the story, associate just 'the man who discovered the incompleteness of arithmetic' with 'Gödel', says that, on descriptivism, 'since the man who discovered the incompleteness of arithmetic is in fact Schmidt, we, when we talk about "Gödel", are in fact always referring to Schmidt' (Kripke, 1980, p. 83). Immediately following this comment, Kripke says, 'But it seems to me that we are not. We simply are not' (Kripke, 1980, p. 84). He does not say that it is *intuitive* that we are not talking about Schmidt; he says straight out, and emphatically, that we are not talking about Schmidt. Of course, the Gödel case is an intuitive counterexample to descriptivism for many readers of *Naming and Necessity*, but this is a logically inessential feature of the case. Kripke's argument against descriptivism succeeds if the Gödel case is a *genuine* counterexample. Whether it is an intuitive counterexample is not clearly relevant, and there is nothing in Kripke's presentation of the case that would lead one to believe that Kripke *thinks* it is relevant.

In fact, at various places, and in a variety of ways, Kripke argues that his judgment about the Gödel case is correct, but none of these arguments makes any appeal to intuitions or the intuitiveness of any principle or proposition. Here are three such arguments:

(a) Kripke points out that the imaginary Gödel-case has real life analogues. All that many of us 'know' about Peano is that he was the discoverer of certain

5 The Madagascar case shows very clearly that we have plenty of knowledge relevant to assessing various theories of reference that is not intuitive in character.

axioms concerning the natural numbers. But it turns out that Dedekind discovered those axioms. If descriptivism is true, many of us have been referring all along to Dedekind with our uses of 'Peano'. But we have not been referring to Dedekind with those uses. We have been referring instead to Peano, *mis*attributing to him the discovery of the axioms. This is not simply a further putative counterexample; it strengthens the claim that the Gödel-case is a counterexample by showing us that the way in which we *ought* to judge, with respect to the imaginary Gödel-case, should line up with the way in which we do in fact, and correctly, judge about the real-life Peano case. (See Kripke, 1972/1980, pp. 84–85.)

(b) Kripke argues that the view that 'Gödel' refers to Schmidt—the prediction made by descriptivism concerning the Gödel-case—suggests a more general view to the effect that *one can never be mistaken* in uttering a sentence of the form '*n* is the *F*', when 'the *F*' denotes, and is a definite description one associates with '*n*', a proper name. But one *can* be mistaken in uttering 'Peano is the discoverer of the axioms', even if one associates 'the discoverer of the axioms' with 'Peano'. The falsity of this general view is evidence that Kripke is right in claiming that 'Gödel' does not refer to Schmidt, in the Gödel-case. (See Kripke, 1972/1980, pp. 85n, 87.)

(c) Kripke argues for an alternative account of the way in which 'Gödel' refers (the causal-historical account) which explains, Kripke thinks, why 'Gödel' refers to Gödel in the Gödel-case. The existence of a satisfying general theory of reference that predicts that 'Gödel' refers to Gödel in the Gödel-case counts in favor of the view that 'Gödel' refers to Gödel in the case. (See Kripke, 1972/1980, pp. 91–93.)

In their forthcoming paper, although they do not quite rule the possibility out, Mallon *et al.* express puzzlement at the thought that there may exist evidence for a theory of reference that is independent of the intuitions of competent speakers, and they claim that, even if such evidence exists, arguing for a theory of reference on the basis of it would be at odds with the 'dominant tradition' in the philosophy of language of appealing to intuitions as evidence (Mallon *et al.*, forthcoming). The arguments (a)-(c) belie these views. The real puzzle is why Mallon *et al.* should be so puzzled. Like other philosophers, philosophers of language *argue* for their theories, and it is relatively rare for these arguments to bottom out in a simple appeal to what is or is not intuitive. Kripke is supposed to be the primary representative of the 'dominant tradition' of appealing to intuitions as evidence, but as arguments (a)-(c) show, Kripke's judgment about the Gödel case is justified by argument, not intuition.

Mallon *et al.* rush to ascribe a method to Kripke and other philosophers of language that gives significant evidential weight to intuitions because they think they have empirical data that shows, for example, that Kripke's own intuitive judgment about the Gödel case is not everyone's intuitive judgment about the case. I have argued that Kripke's method does not in fact, and need not in principle, assume anything at all about the evidential status of intuitions. Empirically determining who intuits what concerning the Gödel case won't help settle the question of whether it is a genuine counterexample to descriptivism. Still, it would be an interesting psychological discovery, if East Asian English speakers were found to have referential intuitions that differ significantly from Westerners'. Do Mallon *et al.* have any evidence that this is so? In the next section I argue that they do not.

5. The Surveys

Mallon *et al.* claim to have uncovered strong evidence that there is cross-cultural variability in intuitions about reference by running two opinion surveys on small groups of undergraduate students from the College of Charleston (18 students in the first survey), Rutgers (31 students in the second survey), and the University of Hong Kong (26 students in the first survey, and 40 students in the second survey). The surveys presented the students with cases modeled on the Gödel case, as well as on another of Kripke's anti-descriptivist cases—the Jonah case. The results, as Mallon *et al.* describe them, were that 'Western' students from the College of Charleston and Rutgers were more likely to have 'Kripkean intuitions' incompatible with *D* than were 'East Asian' students from the University of Hong Kong, who were more likely to have 'descriptivist intuitions' compatible with *D*.[6]

Even granting that philosophers of language assume the method of cases, or at least treat intuitions as a significant source of evidence for semantic theories, there is a serious difficulty with the view that the survey results have some bearing on a theory of reference. The difficulty is that it is not clear from the results that there really are any cross-cultural differences in referential intuitions in the first place. This difficulty emerges when we examine the wording of the vignettes Mallon *et al.* used in their surveys. Here is one of them:

> Suppose that John has learned in college that Gödel is the man who proved an important mathematical theorem, called the incompleteness of arithmetic. John is quite good at mathematics and he can give an accurate statement of the

[6] It is not especially clear why Mallon *et al.* think that the intuitions of undergraduate students might reveal something important about reference. The typical undergraduate student has no special training in semantics and so should not be expected to possess any intuitive insight about the nature of reference. For insight on the theory of reference, it makes much better sense to turn to smart, well-trained philosophers of language, such as Kripke and Evans. The undergraduates Mallon *et al.* surveyed are competent speakers of English (English is the language of instruction at HKU), that's true; but why suppose that mere competence suffices for intuitive insight? If one is competent in English, one can express one's beliefs and desires in English and be understood by other speakers of English. But being able to achieve these communicative goals is a far cry from knowing, even implicitly, how the reference of one's terms is secured. Knowing how the reference of one's terms is secured takes hard thinking and detailed semantic analysis and theorizing. There is no reason to think that every competent speaker is suited to this task, and certainly no reason to think that mere competence makes them suited to it.

> incompleteness theorem, which he attributes to Gödel as the discoverer. But this is the only thing that he has heard about Gödel. Now suppose that Gödel was not the author of this theorem. A man called 'Schmidt' whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and claimed credit for the work, which was thereafter attributed to Gödel. Thus he has been known as the man who proved the incompleteness of arithmetic. Most people who have heard the name 'Gödel' are like John; the claim that Gödel discovered the incompleteness theorem is the only thing they have ever heard about Gödel. When John uses the name 'Gödel', is he talking about:
>
> (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, p. B6).

The difficulty is that the question at the end of the vignette, 'When John uses the name "Gödel," is he talking about (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?', is *ambiguous*. One can use a name to 'talk about' x by using a name that does in fact refer to x in one's language, or one can use a name to 'talk about' x by using a name, which may or may not refer to x in one's language, but which one uses with the *intention* of referring to x. So the question at the end of Mallon *et al*.'s vignette (henceforth, the 'vignette question') can be taken as asking:

(Q1): To whom does *John intend to refer* when he uses 'Gödel'?

Or else it can be taken as asking:

(Q2): To whom does *the name, 'Gödel', refer* when John uses it?

More generally, questions about who or what a speaker is 'talking about' in using a term, or questions about who or what a speaker 'refers to' in using a term, are ambiguous questions that can be interpreted as questions about the *semantic reference* of the term, i.e. the object assigned as referent by the conventions of the language to the term, or else as questions about the *speaker's reference* of the term, i.e. the object to which the speaker intends to refer in using the term.[7]

---

[7] To my ear, the vignette question, 'When John uses the name "Gödel," is he talking about (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?', nearly forces a speaker's reference interpretation; it is a question about what *John* is doing with the name—making speaker's reference to the man (Schmidt) who actually discovered the proof—not a question about what the name *itself* is doing, which is, on a Kripkean causal-historical theory, semantically referring to the man who stole the proof.

Suppose I have two neighbors, Smith and Jones. I walk out onto my porch, and seeing one of my neighbors who I take to be Jones across the street raking some leaves, I say to myself, 'Jones is raking'. But I'm mistaken. It's Smith, not Jones, who is raking. The speaker's reference of my use of 'Jones' is, on this occasion, Smith, since I mean to be speaking of the person I see across the street raking leaves, and that person happens to be Smith. But the semantic reference of 'Jones' is Jones. That is why it is correct to describe me as having made a *mistake* in saying what I did. Despite the fact the person to whom I intend to refer is raking, the name I used semantically refers to Jones, and Jones, we may suppose, is napping, not raking.

Illustrative examples such as these, and the speaker's reference/semantic reference distinction itself, are familiar to every philosopher of language and indeed to most philosophers regardless of specialization. In fact, a famous paper of Kripke's (1977) is titled 'Speaker's Reference and Semantic Reference', and in it he discusses the distinction as it applies to definite descriptions as well as to proper names. (The Smith/Jones example from the preceding paragraph is a shortened version of one of Kripke's own examples.) As Kripke emphasizes, the distinction is closely related to the distinction between *semantic meaning* and *pragmatic meaning*, which is roughly the distinction between the proposition(s) conventionally encoded by a sentence and the proposition(s) a speaker intends to communicate in uttering the sentence. Understanding and appreciating this latter distinction has long been recognized as fundamental to understanding the nature of language and communication. Paul Grice's (1989) pioneering work on the related distinction between 'what is said' by an utterance and what are merely pragmatic 'implicatures' of the utterance has been enormously influential and continues to resonate not just in the philosophy of language but in many other areas of philosophy (and linguistics) besides. There is certainly no area of philosophical semantics that has not been profoundly affected by Grice's work and by the recognition of a real distinction between semantics and pragmatics, theories of reference and meaning for proper names being no exception. Given all of this, and given that the distinction between speaker's reference and semantic reference has such a clear bearing on Mallon *et al.*'s surveys, it is quite strange that there is no discussion of the distinction or its relevance to their work in either of their papers.

In any case, the fact that the vignette question can be interpreted as either (Q1), which asks for the speaker's reference of John's uses of 'Gödel', or (Q2), which asks for the semantic reference of those uses, casts severe doubt on Mallon *et al.*'s claim that the polls' results show that there are cross-cultural differences in referential intuitions. Given the ambiguity of the vignette question, it may be that some of their respondents were answering (Q1), while some were answering (Q2). If so, Mallon *et al.* cannot claim that their results show that Western and East Asian intuitions about the Gödel case *conflict*. They have no right, even, to another claim of theirs, which is that significant minorities in the Western and East Asian groups have intuitions that conflict with the majorities in those groups. The apparent conflict, in each case, might be explained by *consistent* answers to *different* questions: (Q1) and (Q2).

Let me be clear that the objection is not that the majority of East Asians Mallon *et al.* surveyed understood the vignette question as (Q1), while the majority of Westerners they surveyed understood it as (Q2). I suppose that if one were convinced of the truth of the causal historical theory of reference one might be tempted to view Mallon *et al.*'s data as showing that East Asians favor a speaker's reference interpretation of the vignette question while Westerners favor a semantic one. But one need not have any settled view about which theory of reference is correct in order to see that the ambiguity of the vignette question is problematic. The vignette question's ambiguity prevents us from determining whether or not the responses indicate *genuine disagreement* between East Asians and Westerners. It is not that it is obvious from the results that East Asians interpreted the vignette question as bearing on speaker's reference. It is rather that the results do not tell us *whether* they interpreted it that way or not. But if we do not know how the East Asian respondents interpreted the question, then we clearly cannot conclude, contrary to what Mallon *et al.* claim, that East Asians have referential intuitions that conflict with those of Westerners. And, of course, from their answers alone, we cannot determine whether the Western respondents were answering (Q1) or (Q2) either. Mallon *et al.* could safely conclude that East Asians' referential intuitions differ from Westerners' only if they had somehow ruled out the hypothesis that some of their respondents read the question as (Q1) while others read it as (Q2). However, since they have not ruled this hypothesis out, it could well be that the apparent conflict in their respondents' responses is merely apparent.

Mallon *et al.* are committed not only to there being genuine conflict between the referential intuitions of East Asians and Westerners but also to the claim that the referential intuitions of their respondents concerned the *semantic references* of the relevant names. When Mallon *et al.* claim that their results show that there is cross-cultural variability in referential intuitions, they presumably mean that their results show, with respect to the John/Gödel vignette quoted above, for example, that East Asians take the *semantic reference* of John's uses of 'Gödel' to be what *D* implies, i.e. the man who, in the vignette, really discovered incompleteness (Schmidt). However, given the ambiguity of the vignette question, how can Mallon *et al.* be so sure? What is their evidence that their East Asian respondents were interpreting the vignette question as (Q2) instead of (Q1)? I cannot myself see that there *could* be any evidence for this, since, if East Asians tend towards descriptivism, they will answer the vignette question with, '(A) the person who really discovered the incompleteness of arithmetic', regardless of whether they interpret it as (Q1) or (Q2). According to *D*, the semantic reference of John's uses of 'Gödel' *is* the man who really discovered incompleteness, but, in the vignette, that *same* man is arguably the speaker's reference of John's uses of 'Gödel'; John intends to be referring to the man who really discovered incompleteness when he uses 'Gödel'.[8]

It seems safe to suppose that some of Mallon *et al.*'s respondents' reactions were pragmatically driven intuitions about speaker's reference. At the very least, there is no reason to think that all of Mallon *et al.*'s respondents' reactions were semantically driven intuitions about semantic reference. And, even if some were, there is no way to tell which were and which were not. I conclude that, even if philosophers of language did assume something like the method of cases, they would have nothing to fear from the results of Mallon *et al.*'s surveys. Those results simply do not say whether East Asians tend to have intuitions about the semantic reference of proper names that differ from the intuitions had by Westerners about the semantic reference of proper names. Perhaps further surveys with univocal vignette questions would establish that they do. As far as I can see, this is a completely open empirical question.

6. Experimental Philosophy More Broadly

Mallon *et al.*'s critique is of a piece with a good deal of other work in experimental philosophy. Many experimental philosophers seek to challenge more traditional philosophy by showing, via surveys similar to those conducted by Mallon *et al.*, that philosophical intuitions generally, not just intuitions about reference, are culturally and otherwise variable. A background assumption in much of this work is that, as a group, philosophers accept something similar to the method of cases, but formulated more broadly, as an account of correctness for philosophical theories generally, not simply for theories of reference. Many experimental philosophers seem to believe, that is, that all or most philosophers assume, perhaps implicitly, that the correct philosophical theory of reference, knowledge, intentional action, moral responsibility, or any of the other traditional topics of philosophy, is the theory best supported by competent speakers' intuitions about actual and possible

[8] Kirk Ludwig (2007) also points out that Mallon *et al.*'s vignette question is ambiguous between a speaker's reference and semantic reference reading. But he goes on to complain that, if it is given a semantic reference reading, then, given the way the vignette is told, 'there is only one correct response to the [vignette] question' (Ludwig, 2007, p. 150). His support for this is that 'in the description of what John was told, the name 'Gödel' is used to refer not to the discoverer of the incompleteness of arithmetic, but to the person who stole the manuscript' (Ludwig, 2007, p. 150). In fact, however, the vignette says that John is told *only* that 'Gödel is the man who proved an important mathematical theorem, called the incompleteness of arithmetic' (Machery *et al.*, 2004, p. B6). It is not obvious that the use of 'Gödel', in this single thing that John is told, refers to the proof stealer. Other sentences in the vignette clearly do use 'Gödel' to refer to the proof stealer, but these are not part of the description of what *John* is told, they are instead part of what the reader of the vignette is told. Furthermore, even if 'Gödel' refers to the proof stealer in what John is told, there is no clear reason, independent of various competing theories of reference, why John's uses of 'Gödel' must semantically refer to that same person. So I think Mallon *et al.* are more careful in their phrasing of the vignette than Ludwig gives them credit for.

cases involving those topics. Or, if they would not put it quite that strongly, they believe at least that philosophers tend to treat competent speakers' intuitions about a given philosophical topic as a significant source of evidence for truths regarding that topic.[9]

Philosophy's assumption of a generalized version of the method of cases (or something slightly weaker) is allegedly revealed by the philosopher's appeal in his or her theorizing to 'what we would we would say' about cases. Would we say that the speaker's uses of 'Gödel' refer to Schmidt in the Gödel case? Would we say that a subject in a Gettier case knows? Would we say that agents in 'Frankfurt cases' are morally responsible for their actions? Would we say that a CEO who starts a new program solely to increase profits, but, in so doing, and with advance knowledge, harms the environment, has *intentionally* harmed the environment?[10] According to many experimental philosophers, asking and answering these sorts of questions is pretty much all there is to more traditional philosophizing about reference, knowledge, moral responsibility, and intentional action. And the main difficulty with the method, say the experimental philosophers, is that it typically involves completely unsupported empirical speculation. How in the world could a philosopher know 'what we would say', given that the 'we' refers to all of us competent speakers of English, if he or she never even bothers to *ask* us? Thus the surveys: If we really want to know what competent speakers would say about Gettier cases, for example, let's formulate the relevant vignettes, go to McDonalds, and start asking people. One cannot tell, 'from the armchair', what the good people dining at McDonalds will say about whether a subject in a Gettier case knows.

The problem with this line of thinking is that there are no good reasons to suppose that philosophers really do accept a generalized version of the method of cases as their account of correctness for philosophical theories. The reason cannot be merely that philosophers say, as they do, that this or that judgment about a case is 'intuitive' or represents 'what we would say' about the case. These expressions may be simply philosopher-speak for, 'this judgment is true', or perhaps, 'this judgment is pretty obviously true', or maybe even, 'this judgment should be accepted as true unless compelling reasons can be given for rejecting it'. In any case, there is no reason we must take the expressions as abbreviations of 'this is what most competent speakers would say about the case', and there is still less reason to

---

9 How significant? Presumably, the answer varies from one experimental philosopher to the next. However, if one thinks that statistically significant variability in intuitions about x would pose a serious challenge to philosophical theories of x, then one is assuming that philosophers treat intuitions about x as a very important—perhaps necessary—source of evidence for philosophical theories of x. I think many experimental philosophers take their more traditional brethren to believe that intuitions are an important or necessary source of evidence. In some cases, experimental philosophers seem to assume something stronger. The method of cases, for example, says that the correct theory is the theory best supported by intuitions. To attribute the method of cases to philosophers is to attribute to them the view that a theory is correct *only if* it's the theory best supported by intuitions.

10 This question comes from a survey Joshua Knobe (2003a and 2003b) used in some of his well-known work in experimental action theory.

think that philosophers either typically do or must accept that, if their pet theory is not best supported by the intuitions of competent speakers, then that theory is false or in serious jeopardy, evidentially speaking. My guess is that very few philosophers conceive of the truth or evidential basis of their views as determined by the intuitive judgments of competent speakers. Why should they? Competence in a language does not buy one insight into the nature of reference, knowledge, moral responsibility, intentional action, or any of the other traditional topics of philosophy.

This is not to say that there are no methodological questions about appeals to hypothetical cases in philosophy. On the contrary, there is an old question, itself one of the traditional questions of analytic metaphilosophy, concerning how one knows that a given philosophical judgment about a case is true. Suppose we philosophers know that a subject in a Gettier case does not know. How do we know this? The traditional answer is: By *thinking* about the case. Giving this answer appears to commit one to the existence of *a priori* knowledge.[11] And not just a priori knowledge of the content of our concepts or the meanings of our words—the knowledge in question appears to be knowledge about *knowledge itself*, not merely about the concept *knowledge*, or the meaning of the English word 'knowledge'.

Some philosophers are skeptical of a priori knowledge, and some who are not skeptical of a priori knowledge are skeptical that we can have it about anything other than the contents of our concepts or the meanings of our words. Perhaps, for the philosopher who is a skeptic about a priori knowledge, the existence of variability in philosophical intuitions spells trouble.[12] For example, supposing that there are cross-cultural differences in competent speakers' intuitions about Gettier cases, the philosopher who is a skeptic about a priori knowledge perhaps does need to offer some independent support for the claim, if he or she is inclined to make it, that a subject in a Gettier case does not know.[13] He or she cannot claim to know a priori, by thought alone, that a subject in a Gettier case fails to know, and if he or she is skeptical only of a priori knowledge that is not knowledge of concepts or word meanings, the situation is perhaps even worse, for what right does the philosopher have to say that his or her own intuitions reveal the contours of *the*, presumably shared, concept *knowledge*, or *the* meaning of 'knowledge'?

But what of non-skeptics? For them, there is no problem, for they can say that the philosophical method includes a significant a priori component. The non-skeptic will say that whether a subject in a Gettier case knows is just not something that

[11] One can know a priori that a subject in a Gettier case fails to know without this knowledge being immediately or non-inferentially derived from considering the case. Perhaps coming to know that a subject in a Gettier case fails to know takes considering a range of cases, and reflection on various epistemic principles. There is no need to say that one can 'just see' that such a subject fails to know.

[12] Only 'perhaps', since it is open to such a philosopher to argue that philosophers' judgments about conceptual contents or word meanings are more reliable than those of the folk.

[13] Weinberg *et al.* (2001) report the results of a study that shows (they allege) that there are cross-cultural differences in intuitions about Gettier cases.

empirical experimentation will settle for us; if we know that such a subject does not know, we do not know it a posteriori. In particular, we do not—cannot—know it by conducting a survey. Hence, for the philosopher who is not a skeptic about a priori knowledge, and who conceives of the philosophical method as including a significant a priori component, the results of the experimental philosophers' surveys are irrelevant. For the surveys to be relevant, the philosophical questions upon which they allegedly bear must be conceived as being answerable via a posteriori methods. But non-skeptical philosophers do not conceive of the questions this way.

It may be that many experimental philosophers doubt existence of a priori knowledge, and a good skeptical argument against the a priori really would challenge the method employed by a great many philosophers. But it is hard to imagine how the surveys of which experimental philosophers are so fond might bear on the issue of the existence a priori knowledge. If skepticism about the a priori is what drives the experimental philosopher's challenge to philosophy, experimental philosophers ought to stop conducting surveys, settle into their armchairs (if they haven't gotten rid of them yet), and fashion their case against the a priori.[14]

Many experimental philosophers seem to assume that the majority of more traditional philosophers are themselves skeptical of the a priori. As a result, experimental philosophers end up treating these other philosophers as though they are all, deep down, *ordinary language philosophers*. All philosophy, they suppose, is ordinary language philosophy, but dressed up in way that masks its true nature—it is ordinary language philosophy *in disguise*. This would explain why experimental philosophers think their surveys matter so much to philosophy and its methods. Since experimental philosophers think other philosophers care deeply about how ordinary people *talk* about reference, knowledge, morality, action, etc., they think philosophers ought to care about the results of their surveys. The surveys expose, in a way no a priori method could, how ordinary folk talk about the traditional topics of philosophy.[15]

Treating all philosophers as ordinary language philosophers is really a very strange thing to do, however. These days, most philosophers would cringe at the claim that 'all philosophical problems are problems of language', or the claim that philosophical problems can be solved by the 'linguistic analysis' of ordinary speech. But these are precisely the sorts of claims to which experimental philosophers think more

[14] Experimental philosophers might say that Quine has already done this job for them, but, firstly, not everyone has been convinced, and, secondly, if one *is* convinced, what need is there for surveys? If Quine's reasons for rejecting the a priori suffice for rejecting the a priori, then that is the end of the story. Any bit of philosophy that depends on a priori methods—and this, I think, is nearly all the bits—gets thrown out. Who cares who intuits what?

[15] Part of the problem with ordinary language philosophy was that, although its practitioners were supposed to care about the 'ordinary use' of a term, in practice they paid attention to only how they and a small group of their colleagues used the term. But there was no guarantee that they used the term in an ordinary way! If it were not dead already, the experimental philosophers' surveys might form the basis of mildly interesting critique of ordinary language philosophy.

traditional philosophers are committed. The claim that the correct philosophical theory of *x* is the theory best supported by competent speakers intuitions about actual and possible cases involving *x*—the generalized version of the method of cases—is merely a new way of expressing the now passé view that our ordinary talk about *x* reveals the truth about *x*.

The view no longer enjoys much currency partly because it is now widely recognized to be insensitive to precisely the distinction between semantics and pragmatics—the very distinction which, as we saw above, throws a wrench into Mallon *et al.*'s experimental philosophy of language. Part of Grice's motivation for developing his theory of pragmatic implicature was to diagnose what he believed to be a mistake in many ordinary language philosophers' argumentative strategy. According to Grice, ordinary language philosophers often incorrectly inferred *falsity or lack of truth-value* from *inappropriate usage*.[16] For example, from the fact that it would sound strange to describe morally neutral everyday actions such as eating one's breakfast as having been done *voluntarily*, Gilbert Ryle (1949) concluded that an action may be truly described as voluntary only if it ought not to have been done. And Norman Malcolm (1949), following Wittgenstein, objected to the use of 'knows' in cases in which there is 'no inquiry underway' concerning the truth of the relevant proposition. On Malcolm's view, the assertion that one knows that one has hands would be a 'misuse', and so, presumably, a misapplication, of 'knows', unless there is some genuine question about *whether* one has hands.

Grice reacted to this sort of move by pointing out that the use of a term may be inappropriate *even if it correctly applies*. In applying a term, a speaker may say something not informative enough, or too informative, or irrelevant (among other inappropriate things) without saying something untrue. Breakfast-eatings are not typically *in*voluntary, so if I describe mine as voluntary, I perhaps suggest that there was something *special* about it. Perhaps even, as Ryle would have it, that it ought not to have been done. But this suggestion is not part of 'what is said', in Grice's famous phrase; it is not part of my description's literal truth conditions. At best, it is a pragmatic implicature of what is said.

If Ryle and the rest of the ordinary language philosophers can mistake pragmatic implicatures for literal truth conditions, then so can your average competent speaker of English. Can and do: It is these days widely accepted among linguists and philosophers of language that 'seemingly semantic intuitions',[17] intuitions which strike their possessor as intuitions about a sentence's literal truth or falsity but are in fact intuitions about the merely pragmatic effects of uses of the sentence, are common. It is especially easy to mistake what a sentence is *typically used to communicate* for its literal content. Of course, sometimes a sentence is typically used to communicate its literal content. But, more often than is generally recognized, sentences are not used that way. Nonliterality is the rule, not the exception.

[16] Grice (1989) makes this point against ordinary language philosophy in the 'Prolegomena' to *Studies in the Way of Words*.

[17] The phrase is Kent Bach's (2002), from a paper of his of the same name.

A typical use of (1), for example, will communicate the thought that *not all children have been immunized*:

(1) Some children have been immunized.

A competent English speaker, correctly recognizing that a typical use of (1) communicates this thought, might mistakenly infer that (1) is true only if not all children have been immunized. But this really would be a mistake, since a circumstance in which every child is immunized is necessarily one in which (1) is true. That not all children have been immunized is an implicature of, not a literal truth condition on, a typical use of (1).

The relevance to experimental philosophy is plain: If competent speakers can and do mistake implicatures for truth conditions, then, when we poll them asking them to intuit whether some sentence, *S*, is true or false, we must somehow make sure this mistake is not being made. Are they intuiting that *S*'s literal truth conditions are satisfied or fail to be? Or are they intuiting instead that one of *S*'s implicatures is true or fails to be? The fact that many sentences are typically used to communicate something other than their literal truth conditions compounds the problem, for, in the case of such sentences, separating out their truth conditions from their implicatures can take hard work, and the difference between the two is sometimes, even for theorists armed with the truth-conditions/implicature distinction, 'hard to hear'.

It might be objected that whether experimental studies of peoples' intuitions run into trouble stemming from insensitivity to the truth conditions/implicature distinction can be determined only on a case-by-case basis, and that the burden of proof for someone wishing to criticize a given study on such grounds is surely on the critic. A plausible implicature must be specified and the claim that it is the source of some subset of the relevant intuitions must be justified. These complaints would have some force only if intuitions about cases were, by default, intuitions about the literal truth conditions of the sentences used in eliciting them. Nonliterality is rampant, however; a great many sentences are used, on some occasions, to mean something different from what they literally mean, and a good number of others are *typically* used to mean something different from what they literally mean. Given this, a competent speaker's judgment/intuition that *S* is false, for example, is evidence only for the disjunction: *the speaker believes of the semantic content of S that it is false* or *the speaker believes of (at least) one of the implicatures of (this use of) S that it is false*. In general, then, we should not take intuitions about cases to be intuitions about the truth-conditions of the sentences used in drawing those intuitions out. An always live alternative hypothesis is that they are intuitions about the pragmatic implicatures of those sentences.

To this general point, it may be added that some of the experimental philosophical surveys conducted to date have, to their detriment, failed to factor out the influence of pragmatic effects in recording their respondents' intuitive reactions to cases. We have already seen how the pragmatic notion of speaker's reference ruins the surveys on reference conducted by Mallon *et al*. But there are further cases. For example, Fred Adams and Annie Steadman (2004) have forcefully argued that the now notorious 'Knobe-effect' might be explained by appeal to pragmatic implicatures associated with the use of intentional language, as opposed, as Knobe (2005) and others imagine, to revealing something surprising about our concept of intentional action. Knobe (2003a) reports that people he has surveyed are far more likely to describe a 'side-effect' of an agent's performing some action—a causal result of the action the agent knows of but is not trying to bring about—as 'intentional' if the side-effect is widely regarded as bad. In Knobe's vignette, the agent is a CEO who opts to start a new program because it will increase profits. The CEO knows that starting the program will also harm the environment but professes 'not to care at all' about harming the environment. A large majority of people polled will say that the CEO in Knobe's vignette harmed the environment intentionally. However, if the side-effect is described as *help*, instead of harm, to the environment, and the CEO again claims not to care about the program's environmental effects, a large majority of respondents to the vignette thus modified will say that the CEO did *not* intentionally *help* the environment. Should we conclude, as Knobe (2005) does, that the (perceived) normative status of a side-effect matters to whether it was produced intentionally? A less exciting alternative is to understand the Knobe effect as a case of people mistaking implicatures for truth-conditions. As Adams and Steadman argue, it is plausible, given that we often excuse ourselves for the objectionable actions we perform by saying that they were performed unintentionally, that there is an implicature associated with most uses of sentences of the form, '*S A*-ed unintentionally', to the effect that *S* should not be blamed or held responsible for *A*-ing. Perhaps the majority of the respondents to the harm version of Knobe's vignette correctly recognized that the CEO deserved some blame for harming the environment, and then mistakenly inferred that the CEO harmed the environment intentionally. The inference would be a mistake because, as Adams and Steadman maintain, the CEO's being blameworthy is merely an implicature of, not a literal truth-condition on, the claim that the CEO unintentionally harmed the environment.[18]

---

[18] Some experimental philosophers have claimed that further experiments on the Knobe effect have refuted Adams and Steadman's implicature explanation. For example, Shaun Nichols and Joseph Ulatowski (2007) report that, in a study run by Nichols, people given the harm version of Knobe's vignette, and then asked to choose between (a) and (b) below, overwhelmingly opt for (a).

(a) The CEO intentionally harmed the environment and is responsible for it.
(b) The CEO did not intentionally harm the environment but is responsible for it.

Nichols and Ulatowski say this shows that Adams and Steadman's implicature explanation of the Knobe effect 'was not borne out' (Nichols and Ulatowski, 2007, p. 353). However, if the implicature explanation is correct, a majority opting for (a) is just what we ought to expect. In fact, if the implicature explanation is correct, (b) ought to strike speakers as *inconsistent*. The mistake speakers are making, according to the implicature explanation, is the mistake of taking 'not responsible' to be *implied by* 'not intentional', when, really, there is no implica*tion*, just an implica*ture*. But if they take it to be an implication, then no wonder they do not think (b) accurately describes the facts: The second conjunct of (b) negates what many speakers take to be an implica*tion* of (b)'s first conjunct.

Given the problems it poses for two central examples of experimental philosophy, it would not be surprising if the implicature/truth-conditions distinction posed equally severe problems for other examples of experimental philosophy. My guess is that many 'folk' reactions to philosophical cases are reactions to pragmatic implicatures instead of literal contents.[19] However, even if it were somehow possible to construct surveys that excluded intuitions about implicatures and revealed only intuitions about the semantic contents of the sentences used in describing philosophical cases, we should not lose sight of the fact that the data collected by such polls would still only be data about *what people believe* about the semantic contents of the relevant sentences. And there might be a gap between what people believe and what is so. As I have been emphasizing throughout the paper, competent English speakers, even large groups of such speakers, even every such speaker, might be wrong about the truth-values of English sentences such as, '"Gödel" refers to the man who stole the proof', or 'The CEO intentionally harmed the environment'. Getting them to understand the difference between a semantic content and an implicature is not going to make competent English speakers magically immune to erroneous judgments about the truth-values of the literal contents of the sentences they encounter. When the sentences describe puzzling philosophical cases, chances of a mistake shoot way up, whether the speaker is aware of the implicature/truth-conditions distinction or not.

7. Conclusion

Not all philosophy is ordinary language philosophy in disguise. Surveying competent speakers' intuitions is not, therefore, a means of empirically testing philosophical theories, or gathering evidence relevant to the truth of such theories. At best, intuition surveys will tell us whether competent speakers *believe* the semantic contents of the sentences used to describe philosophical cases. Competent speakers can get it wrong, however. Competence in a language does not bestow philosophical insight.

At worst, surveys of competent speakers will record judgments about merely pragmatic implicatures, instead of semantic contents, and so their results will not qualify as evidence that different groups of people have genuinely different intuitions

[19] Philosophers are not immune to confusing the truth or falsity of a pragmatic implicature for the truth or falsity of a literal content, but some of them, having been exposed to the general distinction between semantics and pragmatics, and so aware of the of the ever-present possibility of the confusion, are less prone to it. All parties to the debate over descriptivism, for example, know that the views they defend concern the semantic reference of proper names. The fact that 'Gödel' might be *used* to make speaker's reference to someone other than the semantic reference of 'Gödel' is not a fact that is likely to distort their semantic views. Of course, agreement that the relevant theory is a theory of semantic reference does not add up to agreement about which theory is correct.

about what counts as a case of referring to Gödel, or an intentionally produced side effect.

We saw that, in the case of Kripke's argument against descriptivism, it need not be assumed—and is not assumed by Kripke himself—that Kripke's judgment about the Gödel case is shared by all, or even most, competent speakers of English. What matters is whether the judgment is true. And whether one is justified in supposing that it is true will depend on the quality of the arguments for its truth, not on how many people intuit that it is true.

Something similar can be said of philosophical arguments more generally. Philosophers need not assume that their own intuitions about cases are universal. So surveys showing them that they are not universal are irrelevant. Majority opinion does not determine the truth, or constitute the primary source of evidence in philosophy, and despite appeals to 'what we would say' about cases, majority opinion has never been thought to play these roles in philosophical argument.

Experimental philosophers who use intuition surveys to criticize philosophy assume that once they have uncovered variability in intuitions it is up to the more traditional philosopher to scramble to explain this variability. This strikes me as backwards. It is rather the experimental philosopher who must explain why variability is a real problem. If the reply is simply that the philosophical method is inconsistent with variability in intuitions, what is the evidence that philosophers have actually adopted such a method? I have argued that philosophers of language need not be troubled by variability in referential intuitions, and that, in any case, Mallon *et al.* have not demonstrated that there is any genuine cross-cultural variability in such intuitions in the first place. I suspect that the fact that many speakers can be misled in their truth-value judgments by pragmatic implicatures is a real obstacle to establishing that different groups of competent speakers have differing philosophical intuitions, but the more important point is that, even if they did, it is very difficult to see how this could amount to a criticism of philosophy. Uniformity in philosophical intuitions would be the real shocker; variability is utterly unsurprising and inconsequential.

Department of Philosophy
The University of Hong Kong

References

Adams, F. and Steadman, A. 2004: Intentional action in ordinary language: core concept or pragmatic understanding? *Analysis*, 64, 173–81.

Bach, K. 2002: Seemingly semantic intuitions. In J. Keim Campbell, M. ORourke, and D. Shier (eds), *Meaning and Truth*, New York: Seven Bridges Press, 21–33.

Bealer, G. 1999: A Theory of the A Priori. *Philosophical Perspectives*, 13, Epistemology, pp. 29–57.

Evans, G. 1973: The causal theory of names. *Supplementary Proceedings of the Aristotelian Society*, 47, 187–208.

Goldman, A. and Pust, J. 1998: Philosophical Theory and Intuitional Evidence. In W. Ramsey and M. DePaul (eds), *Rethinking Intuition*. Lanham, MD: Rowman & Littlefield.

Grice, H. P. 1989: *Studies in the Way of Words*, Cambridge, MA: Harvard University Press.

Knobe, J. 2005: Theory of mind and moral cognition: exploring the connections. *Trends in Cognitive Sciences*, 9, 357–9.

Knobe, J. 2003a: Intentional action and side-effects in ordinary language. *Analysis*, 63, 190–93.

Knobe, J. 2003b: Intentional action in folk psychology: an experimental investigation. *Philosophical Psychology*, 16, 309–23.

Kripke, S. 1972/1980: *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Kripke, S. 1977: Speakers reference and semantic reference. In P. French, T. Uehling, and H. Wettstein (eds), *Midwest Studies in Philosophy vol. II: Studies in the Philosophy of Language*. Morris, MN: University of Minnesota, 255–76.

Ludwig, K. 2007: The epistemology of thought experiments: first person versus third person approaches. *Midwest Studies in Philosophy*, 31, 128–59.

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

Malcolm, N. 1949: Defending common sense. *The Philosophical Review*, 58, 201–20.

Mallon, R., Machery, E., Nichols, S., and Stich, S. forthcoming: Against arguments from reference. *Philosophy and Phenomenological Research*.

Nadelhoffer, T. and Nahmias, E. 2007: The past and future of experimental philosophy. *Philosophical Explorations*, 10, 123–49.

Nichols, S. and Ulatowski, J. 2007: Intuitions and individual differences: the Knobe effect revisited. *Mind & Language*, 22, 346–65.

Pust, J. 200: *Intuitions as Evidence*. New York: Routledge/Garland.

Ryle, G. 1949: *The Concept of Mind*. Chicago: The University of Chicago Press.

Weinberg, J., Nichols, S., and Stich, S. 2001: Normativity and epistemic intuitions. *Philosophical Topics*, 29, 429–59.