

Whither Experimental Semantics? *

Michael DEVITT

Received: 25.8.2011

Final version: 6.11.2011

BIBLID [0495-4548 (2012) 27: 73; pp. 5-36]

ABSTRACT: The main goal of the paper is to propose a methodology for the theory of reference in which experiments feature prominently. These experiments should primarily test linguistic usage rather than the folk's referential intuitions. The proposed methodology urges the use of: (A) philosophers' referential intuitions, both informally and, occasionally, scientifically gathered; (B) the corpus, both informally and scientifically gathered; (C) elicited production; and, occasionally, (D) folk's referential intuitions. The most novel part of this is (C) and that is where most of the experimental work should be. The secondary goal of the paper is to defend my earlier paper "Experimental Semantics" from the criticisms of Machery, Mallon, Nichols, and Stich in "If Folk Intuitions Vary, Then What?" They charge that I have seriously misunderstood their goal in "Semantics, Cross-Cultural Style" and that many of my arguments are "largely irrelevant". I argue that these charges are baseless.

Keywords: methodology; experiments; reference; linguistic intuitions; linguistic usage; corpus; Kripke; Machery; Mallon; Nichols; Stich.

1. Introduction

A group of philosophers, Edouard Machery, Ron Mallon, Shaun Nichols, and Stephen Stich ("MMNS") have looked critically at the *modus operandi* of the philosophy of language, the consulting of intuitions. They have noted that the intuitions consulted are nearly always those of philosophers.¹ They have wondered about the appropriateness of this methodology. Shouldn't the intuitions of the folk be considered too? Here they saw a problem. In light of some recent work in psychology by Richard Nisbett and colleagues (e.g., Nisbett et al. 2001), they predicted that there would be differences in the referential intuitions of East Asians ("EAs") and Westerners (Ws) about Gödel and Jonah cases like those used in Saul Kripke's famous refutation of description theories of names (1980): EAs would be more likely to have descriptivist intuitions, Ws, causalist ones. They conducted some experiments on undergraduates in Rutgers and Hong Kong that confirmed this prediction in the Gödel cases but not in the Jonah ones (Machery et al. 2004). This raised a dire thought in their minds "about the nature of the philosophical enterprise of developing a theory of reference" (2004, B1; see also Mallon et al. 2009).

The standard intuition-driven methodology of philosophy of language, indeed of "armchair" philosophy generally, has long bothered me (1994; 1996, ch. 2; Devitt and Sterelny 1999, 285-7). And it bothers others. Thus, Jaakko Hintikka remarks: "One

* I am indebted to the following for comments on a draft of this paper: Ellen Fridland, James Genone, Steven Gross, Justyna Grudzińska, Nathaniel Hansen, Edouard Machery, Genoveva Martí, Jennifer Nado, Gary Ostertag, David Pereplyotchik, Francesco Pupa, Georges Rey.

¹ Max Deutsch (2009) is strangely critical of this account of the philosophical methodology.



searches the literature in vain for a serious attempt to provide” a justification for the appeal to intuitions (1999, 130). In a similar vein, Timothy Williamson remarks: “there is no agreed or even popular account of how intuition works, no accepted explanation of the hoped-for correlation between our having an intuition that P and its being true that P.” He describes this as “a methodological scandal” (2007, 215). So I think we should welcome the critical stance that MMNS and other “experimental philosophers” have taken toward this standard methodology. I have found the challenge that MMNS have posed to what goes on in the philosophy of language particularly stimulating, as this paper attests. This having been said, I think that they are off on the wrong track.

My paper, “Experimental Semantics” (2011b),² accepts that the results in the Gödel cases seem to provide puzzling evidence against Kripke’s refutation but presents a three-step argument aimed to greatly diminish the significance of the results. In particular, I dismiss MMNS’s dire thought that the results raise questions about the enterprise of developing a theory of reference. And, right as I think they are to question the standard intuition-driven methodology – see, most recently, my “The Role of Intuitions” (2012) – they are wrong to respond by testing the folk’s referential intuitions. The appropriate response to their legitimate concern about armchair philosophy is not to move in more armchairs.³ So I do not endorse Machery and Stich’s vision for the philosophy of language:

[T]hat philosophers of language should emulate linguists, who are increasingly replacing the traditional informal reliance on their own and their colleagues’ intuitions with systematic experimental surveys of ordinary speakers’ intuitions. (2012, 495)

I have a different vision. My main goal in this paper is to present that vision. I shall propose a methodology for the theory of reference that does not abandon the role of philosophers’ intuitions but nonetheless gives a prominent place to experiments. However, the primary role for “experimental semantics” is testing linguistic usage rather than the folk’s referential intuitions. I shall start developing the case for this methodology quite early in pursuing my secondary goal.

In “If Folk Intuitions Vary, Then What?” (2012),⁴ MMNS respond to two critical reactions to their original paper: my paper, and one by Jonathan Ichikawa, Ishani Maitra, and Brian Weatherson (2011).⁵ My secondary goal is to defend my three-step argument from MMNS’s disappointing response. Apart for some constructive remarks about my second step, now called “the Expertise Defense”, the response is quite inadequate. Yet it concedes nothing. And it includes ill-based charges of misunderstanding and irrelevance.

My pursuit of this secondary goal begins in section 3 and continues on through the rest of the paper. However, my main goal about methodology is the dominant con-

² All unidentified references to my work are to this paper.

³ I owe this nice remark to Genoveva Martí who thinks she heard or read it somewhere but can’t recall the circumstances.

⁴ All unidentified references to MMNS are to this paper.

⁵ Martí 2009 is another critical reaction, to which Machery et al. 2009 is a response. Martí also emphasizes the importance of testing linguistic usage.

cern of sections 5 and 6. In section 2, before pursuing either of these goals, I want to raise some doubts about MMNS's initial prediction of cultural variation in referential intuitions. I think it unlikely that the puzzling variation MMNS have discovered shows us anything interesting about the use of intuitions in theorizing about reference. I suspect that the variation is a red herring.

Readers who are interested in the main goal's positive methodological proposal but not much in the secondary goal's disagreements with MMNS might do well to skim sections 2 to 4 and focus on sections 5 and 6.

2. *Is Cultural Variation a Red Herring?*

We should start by noting that there is one popular view of the source of linguistic intuitions that would predict *no* variation due to culture. This is the view, held by many linguists and to be discussed in some detail later (section 5), that a speaker's intuitions about her language are derived from her underlying competence in the language; on this view, the intuitions are, as I like to say (2006a, b), "the voice of competence" ("VoC"). So assuming, as we surely should initially (at least), that Ws and EAs have same competence with names, they should have the same intuitions about them. So if MMNS held VoC it would be surprising that they should predict cultural variations in these intuitions. Why would they think that the findings of Nisbett and colleagues were even *relevant* to those intuitions? *Do* MMNS hold VoC? It is hard to say because they are coy about the source of intuitions. Still the signs are that VoC is their view. Thus, what else could underlie the just-quoted passage about emulating linguists? A belief in VoC seems to be the only available explanation of the linguists' special emphasis on testing ordinary speakers' intuitions, as we shall see. If Machery and Stich don't hold VoC, why would they want to emulate the linguists? We shall see other signs of a belief in VoC later. Even if they are agnostic about VoC, that doctrine is so widespread that one wonders why it did not play a role in their predictions.

But, suppose that MMNS thought VoC false, as I will argue it is (sec. 5), should the recent work in psychology have led them to predict cultural variation between EAs and Ws in referential intuitions? It is hard to see why it should.

MMNS's general inspiration from psychology is summed up in a passage they quote (2004, B5) that distinguishes two styles of thought: the "holistic" style of EAs and the "analytic" style of Ws. Yet it is far from clear how the quoted explanation of these different styles, or the more detailed one in the paper that is its source (Nisbett et al. 2001), would yield any relevant prediction about referential intuitions. And Genevieve Martí (forthcoming) points out that one of the ways that this psychological literature distinguishes the styles of thought is as follows: EAs' thinking is socially oriented, Ws', individualistic. Yet, as she emphasizes citing Kripke, the causal theory seems to make reference socially determined, the description theory, individualistically determined. So she wonders why MMNS did not make the opposite prediction. So too does Gary Ostertag (forthcoming), on similar grounds.

My wonderings are a bit different, prompted by the passage in which MMNS are most specific about the reasoning that led to their prediction:

One range of findings is particularly significant for our project. The cross-cultural work indicates that EAs are more inclined than Ws to make categorical judgments on the basis of similarity; Ws, on the other hand, are more disposed to focus on causation in describing the world and classifying things (Norenzayan, Smith, and Kim 2002; Watanabe 1998, 1999). (Machery et al. 2004, B5)

One can see why this might seem to provide a basis for MMNS's prediction. But I have two problems. First, I can find no support in the paper, Norenzayan et al. 2002, for this claimed difference between EAs and Ws over causation-based judgments. Indeed, causality is hardly mentioned in the paper! Where MMNS talk of Ws favoring *causation*-based judgments, the paper talks of them favoring *rule*-based ones. But perhaps the paper otherwise supports MMNS's prediction? My second problem is that I can't see how it does. Ws' rule-based thinking does not seem to yield the prediction of a causalist intuition about reference. And it is not even clear to me why EAs' similarity-based thinking yields a descriptivist intuition rather than a causalist one. I'll spare the reader my reasoning.

Perhaps Martí, Ostertag, and I have missed the significance of this psychological work. In any case, it would be good to have a more careful spelling out than MMNS have provided so far of the bearing of this work on referential intuitions.

Of course, MMNS have something impressive going for them: they made a prediction and it was confirmed! But if recent psychology does not support their prediction, we would be left with no explanation of the variation that MMNS discovered in intuitions about Gödel cases. The variation would be a genuine anomaly. Given this, and the discussion to follow in section 4, I suspect that the variation is a red herring for theorizing about reference.

I begin my secondary goal of defending my three-step argument from MMNS's response.

3. *Are MMNS Misunderstood?*

In their opening remarks, speaking of their critics, MMNS claim:

[T]hey have quite seriously misunderstood our critical aim in Machery et al. 2004 and Mallon et al. 2009, perhaps because of one or two insufficiently careful sentences in these two articles. As a result, many of their arguments and just about all of their careful scholarship about Kripke's *Naming and Necessity* are largely irrelevant to the project we were pursuing...most of Devitt's and Ichikawa and colleagues' criticisms fail to address our concerns. (p. 2)

Later they claim that we fail to address their argument (p. 6).

This is very surprising. For, whatever their *aim*, I shall demonstrate that I addressed what they *claimed* and *argued*. And the considerations I adduced are very relevant to these claims and arguments. (I shall leave Ichikawa and colleagues to speak for themselves.) It is puzzling that MMNS feel so misunderstood. They have badly misdescribed the dialectical situation.

I start with the charge of serious misunderstanding which has two parts: I miss their real goal; I wrongly take them to have another goal. Sections 4 to 6 will address the charge of irrelevance.

3.1 Missed Goal?

MMNS describe their real goal as follows: “Our goal is to challenge the way philosophers of language go about determining what the right theory of reference is” (p. 3). Compare this with what I identify at the beginning of “Experimental Semantics” as one of MMNS’s “striking claims about theories of reference”, a claim that I just called their “dire thought”:

- (IV) The fact of these cultural differences [between EAs and Ws] “raises questions about the nature of the philosophical enterprise of developing a theory of reference” (Machery et al. 2004, B1); it points to “significant philosophical conclusions” (p. B8). (p. 419)

MMNS support their present claim about their goal by quoting this very passage from p. B1! So where’s my misunderstanding? There may be one. I saw this passage as so dire because I took it to be not simply about the standard intuition-driven *methodology* for developing a theory of reference but, more boldly, about the very *nature* of the referential enterprise altogether. I took this passage, and the paper’s final discussion (pp. B8-B9), to be raising doubts – as Stich has before (1996, 37-51; 2009, 199) and Machery and Stich are about to (2012, sec. 4) – about the very task of explaining reference. In any case, if this *is* a misunderstanding, then, first, it is clearly not one that MMNS have in mind: they never mention it and, as we shall see in a moment, their complaint is that we critics took their goal to be *less* bold, not more bold, than it actually was. Second, it is of little significance to our disagreement: I certainly took them to be challenging the intuition-driven methodology *at least*. And the first two of the three steps in my attempt to diminish the significance of the cross-cultural experiment, presented as reasons for rejecting (IV), count against that challenge. (The third step counts only against the bolder claim about the very task of explaining reference.) So, no serious misunderstanding here. Misunderstanding seems rather more on the other foot.

Later on MMNS claim, on two separate occasions, that I grant their “most important conclusion”. Now if this were so, then I must have *really* misunderstood them since I present myself as disagreeing with them. The first thing to note is that their statements of their “most important conclusion” on the two occasions are quite different! One statement is indeed a conclusion of their original paper, but I don’t grant it. The other statement, which I do grant, seems not to be present at all in that paper. I’ll discuss them in order.

One statement is, in effect, of their just-stated goal: “in light of [...] the variation in intuitions about reference, the standard, intuition-based method for determining the correct theory of reference should be revised” (p. 12). I am alleged to grant this because I proposed a way of investigating reference without using those intuitions (on which more in section 6). As already noted (sec. 1), I have indeed long thought that we need to revise our methodology. But I do not grant that we need to do it “in light of [...] the variation in intuitions”. So I do not grant their conclusion.

The other statement is in the following passage, after quotes from my paper:

[W]e wanted to raise much more general concerns about philosophers' implicit endorsement "of the view that the semantic task simply is the systematization of our ordinary intuitions about meaning, reference, and the like" (p. 424). Devitt concedes that, on this common understanding, our findings do indeed "raise questions about the philosophical enterprise of developing a theory of reference" striking "at the very subject matter of semantics" (p. 424). *And this is exactly what we wanted to show.* Our findings (if borne out by further work) suggest that the standard view of semantics is in deep trouble. So Devitt *agrees* with the most important conclusion we wanted to establish. (pp. 8-9)

So here MMNS take the most important goal of their original paper to be challenging the semantic *metatheory* according to which the *very task* of semantics is to study semantic intuitions. But challenging this metatheory is *very different* from challenging the ubiquitous semantic *method* of consulting intuitions. We have already seen two quotes supporting the latter view of their goals. Here is another: "what we have been really concerned with is the method...: the use of intuitions about reference to identify or justify the right theory of reference" (p. 4). Of course, the two challenges are related. If the metatheory were right it would both justify and require the method. So challenging the method challenges the metatheory. But challenging the metatheory does not challenge the method. In any case, it is obvious that challenging the metatheory is *not* their most important challenge. Indeed, so far as I can see, the metatheory gets no mention at all in the original paper.⁶ Their most important challenge is to the method. It looks as if I understand their goals better than they do!

And MMNS *should* think that using their experimental findings to challenge the method is more important than using it to challenge the metatheory, for two reasons. (1) Even if we are right to suppose that many who follow the method implicitly endorse the metatheory,⁷ doubtless many don't. It is mostly hard to tell what philosophers implicitly believe about the nature of the semantic task. And one can certainly follow the method without justifying it by the metatheory or, indeed, without justifying it at all. (2) There are good reasons, aside from any experimental findings, for thinking that the metatheory is deeply misguided; or so I have argued (1996; 2012).

3.2 *Wrong Goal?*

So much for the charge that I fail to see what their real goal is. What about the charge that I wrongly take them to have another goal? One of the "striking claims about the theory of reference" that I claim to be *implicit* in MMNS's original paper is:

(III) These results raise serious doubts about Kripke's refutation, which relies solely on the intuitions of Westerners. (p. 419)

⁶ They do briefly mention the Chomsky-inspired view that the semantic task is to give an account of "the implicit theory that underlies ordinary uses of names" (2004, B9). This is a different view of the task from the above metatheory but, as I note, it "may seem to be a reason" for the metatheory (p. 424n).

⁷ What I actually say is that "philosophers...implicitly seem...to endorse" the metatheory (p. 424). I have recently been more cautious (2012, sec. 2).

MMNS now claim that “our goal has never been to challenge Kripke’s argument against descriptivism. Our project has a different and, we dare say, broader target” (p. 3; see also p. 13). This is disingenuous. The broader target is the already-discussed intuition-based methodology. Now whether or not challenging Kripke was their “goal”, their whole case against the broader target is made using Kripke’s argument as an example of what is wrong with the philosophical method. Suffice it to say, any doubts MMNS have about the way philosophers of language in general “go about determining what the right theory of reference is” *must*, willy-nilly, carry over to their favorite example of such a procedure, Kripke’s argument. Claim (III) is indeed implicit in their earlier paper, just as I said. And, importantly, if they have failed to raise serious doubts about Kripke’s argument, as I argued they have, then they have failed to cast doubt on the philosophical methodology.

So much for the charge of misunderstanding. Turn now to the charge that many of my arguments fail to address their concerns and are “largely irrelevant”. As already noted, my argument proceeds in three steps. The first step points out that intuitions about Gödel cases are not of much significance in the intuitive support for Kripke’s refutation. The second step argues that we have good reason for preferring the intuitions of philosophers about Gödel cases to those of the folk. The third step argues that we should be seeking experimental support for theories of reference by testing usage not intuitions. I shall take them in turn, demonstrating the relevance of each. And I shall assess MMNS’s responses. For, although they think many of my arguments are irrelevant, they respond to them.

4. First Step: the Importance of Other Intuitions

(III) and (IV) are two of the four “striking claims about theories of reference” that I attribute to MMNS. Here are the other two:

- (I) Philosophical views about reference “are assessed by consulting one’s intuitions about the reference of terms in hypothetical situations” (Machery et al. 2004, B1).
- (II) [The cases of Gödel and Jonah] are “central” to Kripke’s refutation (p. B1). (p. 418)

MMNS describe these cases as “some of the most influential thought experiments in the philosophy of reference” (2004, B8). And, of course, intuitions about those cases are the ones that they test in their cross-cultural study.

4.1 The Argument

My first step in diminishing the significance of MMNS’s findings is to reject (I) and (II). I point out that although it is certainly the case that *some* intuitions used by Kripke to support his refutation are about reference in hypothetical situations – for example, the Gödel case – many intuitions he uses are not of this sort. I argue, first (pp. 420-1), that the most important intuitions Kripke uses, in what Kim Sterelny and I (1999) call “Ignorance and Error” arguments, are about humdrum *actual* not hypothetical situations, particularly not *fanciful* hypothetical ones. These are intuitions about ‘Cicero’,

‘Catiline’, ‘Feynman’, ‘Einstein’, ‘Peano’, and ‘Columbus’. I argue, second (pp. 421-3), that key intuitions used in the “Unwanted Necessity” and “Lost Rigidity” arguments are not semantic ones about reference but metaphysical ones about *modal properties*; for example, that although Hesperus might not have been the planet that is seen in the evening it is not the case that it might not have been Hesperus. I conclude:

Machery et al. leave untouched the referential intuitions about humdrum actual cases and the metaphysical ones about modal properties. These intuitions together are massively more important to the refutation than the intuitions about Jonah and Gödel cases that Machery et al tested. And whereas their test of Gödel intuitions counts against the refutation, *their test of Jonah intuitions confirmed the refutation*. Yet, it seems to me, the Jonah intuitions are more trustworthy than the Gödel ones because the Jonah case is less fanciful than the Gödel one. (p. 423)

Now I take it as obvious that this argument, including its “careful scholarship about Kripke’s *Naming and Necessity*”, is very relevant to (III): if sound, it should diminish doubts about Kripke’s refutation that may arise from MMNS’s findings about Gödel cases.⁸ So, it should diminish doubts about philosophical methodology. So the argument is very relevant to MMNS’s goal, contrary to what they claim.

4.2 MMNS’s Response

Despite their irrelevance allegation, MMNS do respond to some of it. First, instead of (I), they now claim that “philosophers... typically appeal to the intuitions of competent speakers about the reference of proper names (or other kinds of words) in *actual* and possible cases” (p. 3; emphasis added; see also Machery 2012, p. 38). The change to include actual cases – the cases I stressed in my criticism – is welcome.⁹ But what about my other criticism of (I)? I point out that the intuitions that are brought to bear on theories of reference are not all referential ones; some are metaphysical ones about modal properties. They do not respond to this criticism.¹⁰

⁸ A further reason for doubting the significance of these findings about Gödel cases is that they concern the names of *authors*, which seem to have a double life. I mentioned this problem in a note (p. 428, n. 9). In *Designation* (1981, 157-60), responding to Evans’ (1973) nice example of ‘Ibn Kahn’, I argued that the names of authors – my example was ‘Shakespeare’ – have a double life, sometimes as a regular causal name, sometimes as an attributive/descriptonal name like ‘Jack the Ripper’: thus it wouldn’t matter to the truth value of a critical assessment of some “work of Shakespeare” if the work was actually written by Bacon. So these names are tricky, seemingly covered by two conventions.

⁹ They do not acknowledge the change. Perhaps they had (I) in mind in admitting there were “one or two insufficiently careful sentences” in earlier papers (p. 2).

¹⁰ Perhaps the explanation of this is to be found in MMNS’s claim that they are “not concerned with ‘the strong theory-of-meaning construal of descriptivism’” but only with a weaker construal of descriptivism as simply a theory of reference (p. 4). According to strong descriptivism the description that determines the reference also expresses the meaning. MMNS’s lack of interest in strong theories is odd since, so far as I know, all *actual* descriptivist theories before Kripke were strong ones. In any case, modal intuitions, which count only against strong descriptivism, have to be taken into account in assessing the relative importance of Gödel intuitions to Kripke’s refutation of all those actual theories. (I should note also that their claim that I “correctly insist that the Gödel case is directed only at the weak form of descriptivism” is doubly wrong: I don’t insist on it and it isn’t correct.)

Now consider (II). Because of the important role of referential intuitions about humdrum actual cases, and metaphysical intuitions, I argue that the Gödel and Jonah cases are far from being “central” to Kripke’s refutation as MMNS claim.

MMNS are totally unmoved. Why? Their first reason is:

Evidence suggests that *some* intuitions about the reference of proper names vary within and across cultures – viz. the intuitions about the reference of “Gödel” in the Gödel case. This variation inductively suggests that *other* intuitions about reference are also likely to vary. (p. 3; see also Machery 2012, p. 41)

The variation does not inductively suggest this at all! Why suppose that a finding about intuitions in a fanciful hypothetical case will generalize to intuitions in humdrum actual cases (the only “other intuitions” that are relevant to my criticism)? I argue, in the spirit of an earlier Stich (1983, 62n), that intuitions about the former cases are much harder than those about the latter. If so, we should not expect this generalization to be true. In any case, it is an empirical question that can be easily settled by testing the folk’s referential intuitions in the humdrum actual cases. I encourage MMNS to do so. I predict that no significant cultural variation will be found.

MMNS’s second reason is that intuitions about actual cases will not be sufficient to choose between theories of reference: intuitions about Gödel-type cases will be essential to that choice in the end. They offer two considerations in support.

(i) They rightly point out that, however successful Kripke is against a range of description theories, others can be proposed; and that coming up with a “full-fledged” causal theory that moves beyond Kripke’s “picture” is hard, requiring choice between many options.¹¹ They claim, without argument, that “intuitions about possible cases are likely to be needed to determine what the correct theory of reference is” (p. 6). Removing Gödel from Kripke’s discussion would make hardly any difference to the effectiveness of Kripke’s argument. Why must it be any different for anyone else’s?

(ii) MMNS claim that intuitions about actual cases do not pose ignorance and error problems for two “more complex forms of descriptivism” that they propose. To settle the fate of these description theories we will need intuitions about the Gödel case “since [in that case] we can stipulate that everyone is mistaken” (p. 8). Their two proposals are: “the reference of a proper name is determined by the description conventionally associated with this proper name”; and, “by the description that experts in the relevant linguistic community associate with the name” (p. 8). Even if these proposals were promising so far as they go (which I don’t think they are), they don’t go far enough to be testable theories. For, they don’t connect the conventional associations or expert descriptions with the referential capacities of the competent but ignorant users of the name. Thus, I use the name ‘Bruce’ to refer to many people I could not identify. On any occasion it refers to one of those people in particular. How is one of the millions of sets of conventional associations or expert descriptions for the name ‘Bruce’ brought to bear to determine that my reference on that occasion is to that particular person? I bet that a descriptivist answer to this question will fall foul of Krip-

¹¹ I’d be the last to deny the latter having written a book (1981) that attempts just such a theory. MMNS rightly draw attention to the many appeals to intuition in that book (p. 7).

kean intuitions about actual cases: ignorance and error problems once again (see, for example, Devitt and Sterelny 1999, sec. 3.4).

Let us take stock. What is at issue is the critical significance of the cultural variation discovered in the Gödel cases for the philosophical method of using intuitions as evidence for theories of reference. My first step claims to diminish that significance by arguing that the discoveries about referential intuitions in a fanciful hypothetical case like Gödel do not reflect on, (a), referential intuitions about humdrum actual cases; nor on, (b), intuitions about modal properties. And (a) and (b) combined are massively more important to the theory of reference than intuitions about Gödel cases. In response, MMNS claim, without argument, that the cultural variation in Gödel-case intuitions is likely to generalize to actual-case intuitions. They claim further, without any adequate argument, that we will have to rely on Gödel-case intuitions in the end because actual-case ones will not suffice to choose a theory of reference. Both these claims are implausible, in my view. MMNS do not mention the role of modal intuitions as evidence. They failed to discover cultural variation in the Jonah-intuitions. There is no evidence of variation in the intuitions of philosophers. All in all, this strikes me as a very thin basis for the sort of general conclusion about cultural variation in intuitions that they need to cast serious doubts upon the philosophical method. The cultural variation in Gödel-cases is puzzling (sec. 2) but this is not the right response to it. I suspect that the variation is a red herring for theorizing about reference.

I put a lot of store by this first step because its theoretical burden is small and it stays very close to the indubitable historical facts of theorizing about reference. MMNS are on very weak ground here. Indeed, I could simply rest my case against the significance of their findings with this first step. However, criticizing MMNS is only the secondary goal of this paper not the main one. The main goal is to propose a methodology for the theory of reference with a prominent place for experimental semantics. The first step is little help with that. The second and third steps, in contrast, bear importantly on that methodology. The discussion of these steps in sections 5 and 6 is more to further that goal than the secondary one.

5. *Second Step: "The Expertise Defense"*

One response that a philosopher of language might make to MMNS's findings about folk intuitions is to claim that her own intuitions are superior to those of the folk. MMNS are scornful of this: it "smacks of narcissism in the extreme" (2004, B8-B9). My second step in diminishing the significance of the MMNS's findings is to argue that philosophers' intuitions should indeed be preferred to folk intuitions because philosophers are more expert on matters referential (pp. 424-9). Machery calls this, aptly, "the Expertise Defense" (2012, p. 37).¹² He and his team, MMNS, are dismissive of it. They find support in an important paper, "Are Philosophers Expert Intuiters?", by Jonathan Weinberg, Chad Gonnerman, Cameron Buckner, and Joshua Alexander (2010). These papers make some good points against the Expertise Defense.

¹² Machery also cites Ludwig 2007.

Machery begins his paper in this volume with the claim that the Expertise Defense is perhaps “the most influential complaint” against a body of work in experimental semantics that he has co-authored. By the end of his paper, the Expertise Defense seems to have become the *sole* basis for my critical view of the original MMNS paper, the most famous part of that body of work: I am alleged to have argued “that these empirical results can be dismissed because experts’ intuitions provide better evidence about reference than lay people’s” (p. 53). This is a misunderstanding. The Expertise Defense is just one of three steps in my argument against the striking conclusion that MMNS draw about “the philosophical enterprise of developing a theory of reference”. It is not even the most important step. I think that the considerations in my first step, just discussed, are much more decisive than the Expertise Defense against that striking conclusion. However, if we are thinking not of the significance of MMNS’s findings but of how the future should go in the philosophy of language, the main concern of Machery and Stich (2012), another of the papers in that body of work, and my main concern here, then I do think the Expertise Defense should loom large. For, the evidential weight we should attach to the intuitions of the philosophers and of the folk in theory construction hinges on the effectiveness of this defense. So I will devote considerable space to it.

Any assessment of the Expertise Defense is best guided by some theory of the *source* of intuitions about language. For our assessment of the likely reliability of linguistic intuitions, whether the folk’s or the philosophers’, should depend on where we take the intuitions to have come from. My support for the Expertise Defense does rest on a theory. But before considering that theory we should consider another theory that counts *against* the Defense.

5.1 “Voice of Competence” (“VoC”)

This is the “voice of competence” theory (“VoC”), popular in linguistics, that I mentioned in section 2. It not only counts against the Expertise Defense but also, as noted, against MMNS’s prediction of cultural variation in these intuitions. I suggested that although MMNS are coy on the source of intuitions, there are signs that they hold VoC. I mentioned one sign: the suggestion that philosophers should emulate linguists by testing folk referential intuitions. Another sign is that VoC seems to be an underlying presence in their dismissal of the Expertise Defense, soon to be discussed. If it is, then there is of course a tension between that dismissal and their prediction of cultural variation. And there is the problem that VoC is false; or so I have argued at length elsewhere (2006a, b, c). I shall summarize that argument in a moment.

First, we need to say a bit more about VoC. Noam Chomsky provides a nice statement of the doctrine:

[I]t seems reasonably clear, both in principle and in many specific cases, how unconscious knowledge issues in conscious knowledge [...] it follows by computations similar to straight deduction. (1986, 270)

The “unconscious knowledge” is the speaker’s knowledge of her language, her linguistic competence, residing in a module of the mind, the “language faculty”;¹³ the “conscious knowledge” is an intuition about the language. Carson Schütze gives another statement of VoC: “the assumption that grammaticality judgments result from interactions among primary language faculties of the mind” (1996, xi). I have described VoC as follows: linguistic competence, all on its own,

[P]rovides information about the linguistic facts [...]. So these judgments are not arrived at by the sort of empirical investigation that judgments about the world usually require. Rather, a speaker has a privileged access to facts about the language, facts captured by the intuitions, simply in virtue of being competent [...]. (2006a, 483-4; 2006b, 96)¹⁴

Competence not only plays the dominant role in *linguistic usage*, it also provides informational content to *metalinguistic intuitions*. Those intuitions are indeed, “noise” aside, the voice of competence. That is why they are reliable. And if competence really does produce them then we have no reason to prefer those of the linguists to those of ordinary competent speakers. Indeed, we should prefer those of the latter because the linguists will be prone to a sort of noise that lessens their credibility: theoretical bias.

For Chomsky and the linguists, VoC is a theory of the source of syntactic intuitions, particularly those about grammaticality/acceptability.¹⁵ Stich has suggested that philosophers of language might follow the lead of linguists in seeking a justification for the authoritative role given to referential intuitions (1996, 40).¹⁶ He suggests that philosophers may think that speakers derive their referential intuitions from a representation of referential principles.¹⁷ So, just as the true grammar that linguists seek to discover is already represented in the mind of every speaker, so too, according to this suggestion, is a true theory of reference. Referential intuitions, like syntactic ones, are the result of something like a deduction from a represented theory. Thus, speakers have access to linguistic facts simply in virtue of being competent.

5.2 Objections to VoC

Someone who took the usual linguistic view of grammatical intuitions might well be tempted by this analogous view of referential intuitions. So, it is interesting to note that Chomsky is not tempted. He expresses skepticism about “contemporary philosophy of language” and its practice of “exploring intuitions about the technical notions ‘denote’, ‘refer’, ‘true of’, etc.” He claims that:

¹³ Thus, Machery calls this “the modularist conception of intuitions” (2012, p. 41).

¹⁴ The evidence that VoC is the received Chomskian view of linguistic intuitions is overwhelming yet some strangely resist the attribution: Collins (2008a, 16-19); Fitzgerald (2010, sec. 3.4). I have responded (2010c, sec. 4). See Sprouse and Almeida (forthcoming) for recent further evidence.

¹⁵ Linguists tend to make too much of the distinction between intuitions about grammaticality and acceptability, in my view (2010c, sec. 1.3).

¹⁶ Hintikka (1999) and Williamson (2007) make similar suggestions.

¹⁷ Stich does not say whether or not he endorses the thought himself; that old coyness again.

[T]here can be no intuitions about these notions, just as there can be none about ‘angular velocity’ or ‘protein’. These are technical terms of philosophical discourse with a stipulated sense that has no counterpart in ordinary language. (1995, 24)

So Chomsky is skeptical about the use philosophers make of referential intuitions. But he is not, of course, similarly skeptical about the use linguists make of syntactic ones. Why the difference? If skepticism about semantic intuitions is appropriate, then surely just the same skepticism is appropriate about the syntactic ones, and for just the same reason. All the terms in linguistic theory are, in the relevant sense, technical and theory-laden. A few like ‘grammatical’ and ‘sentence’ have counterparts in ordinary language but so too do ‘denote’ and ‘refer’. Semantic and syntactic intuitions are on a par. Chomsky seems to have given a good objection to VoC altogether.

There are other good objections.

What I call the “standard” version of VoC, implied by the Chomsky quote and by Stich’s suggestion, is based on the “representational thesis” that linguistic rules (and principles) are *represented* in the language faculty. Speakers are then thought to derive their intuitive judgments from these representations by a causal and rational process like a deduction. We are given no details of the causal-rational route from an unconscious representation of rules in the language faculty to a conscious judgment about linguistic facts in the central processor. And we need details to turn this sketch into a theory. Still, the idea of one sort of representation leading to another is familiar and so this standard explanation may seem promising. I produce several reasons for thinking it is not promising at all (2006a, 488-91, 503-5; 2006b, 100-3, 114-7). The most important objection is to the representational thesis that is the basis of the explanation. A major conclusion of *Ignorance of Language* (2006b) is that there is no significant evidence that linguistic rules *are* represented in the minds of speakers and, given what else we know, it is implausible to suppose that they are.

Despite the evidence that the standard version is the right way to interpret VoC, it is not certain that linguists really do see intuitions as having their source in *represented* rules. And that implausible representational thesis is certainly rejected by many Chomskian philosophers of linguistics (e.g., Smith 2006; Collins 2006, 2007, 2008a; Pietroski 2008; Slezak 2009). So, perhaps what I call the “nonstandard” version of VoC is the right interpretation: the intuitions are provided somehow by embodied but *unrepresented* rules (2006a, 482-6; 2006b, 96-8). But this version faces an apparently overwhelming objection: we do not have *any idea* how embodied but unrepresented rules might provide linguistic intuitions (2006a, 506-7; 2006b, 118). Not only do we lack the details needed for a plausible explanation, but attention to other similar systems gives good reason to suppose that the linguistic system does not provide these intuitions and so we *could never* have the details. The explanation would require a relatively direct cognitive path from the embodied rules of the language to beliefs about expressions of that language, a path that does not go via central-processor reflection on the data. What could that path be? Consider some other examples. It is very likely that rules that are embodied but not represented govern our swimming, bicycle riding, catching, typing, and thinking. Yet there does not seem to be any direct path from these rules to relevant beliefs. Why suppose that there is such a path for linguistic beliefs? Why suppose that we can have privi-

leged access to linguistic facts when we cannot to facts about these other activities? We do not have the beginnings of a positive answer to these questions and it seems unlikely that the future will bring answers.¹⁸

Since writing *Ignorance*, I have become aware of a body of developmental literature that provides persuasive empirical evidence against VoC.¹⁹ 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 behavior is not a direct reflection of linguistic competence” (1996, 95). It looks as if VoC is false.

These objections were aimed primarily at VoC as a theory of syntactic intuitions. They apply just as much to VoC as a theory of referential intuitions.

If VoC is not the right account of the source of our metalinguistic intuitions, what is? Michael McKinsey gives one answer. He thinks that it is “fairly clear” that “the principle that the meanings of words are knowable a priori...is taken for granted by most philosophers of language and by many linguists” (1987, 1). I think that he is probably right that this principle is taken for granted. But we can be confident that MMNS have no more time for the principle than I have (2011a). So let us set it aside.

5.3 *The Modest Theory of Intuitions*

In “Experimental Semantics” (pp. 425-7), I summarized another answer that I have given elsewhere (2006a, b, c; 2010b). I claim that intuitive judgments about language, like intuitive judgments in general, “are empirical theory-laden central-processor responses to phenomena, differing from many other such responses only in being fairly immediate and unreflective, based on little if any conscious reasoning” (2006a, 491; 2006b, 103).²⁰ Although a speaker’s competence in a language obviously gives her ready access to the *data* of that language, the data that the intuitions *are about*, it does not give her ready access to the *truth* about the data; the competence does not provide the informational content of the intuition. In this respect the view is sharply different from VoC. And it is sharply different in another respect: it is *modest*, making do with cognitive states and processes we were already committed to. So, following Mark Textor (2009), I shall call it “the Modest Theory”.

¹⁸ Georges Rey strongly disagrees (2006, 562-7). I shall respond to his proposal soon.

¹⁹ See particularly, Hakes (1980), Ryan and Ledger (1984), Bialystok and Ryan (1985), Bialystok (1986).

²⁰ How dependent is my view on a classical “central processor” model? Hardly at all. When I first introduce talk of a central processor in *Ignorance*, I mention the opposing “massive modularity” model and set it aside claiming that any close-to-plausible version of it would not affect my arguments (p. 94). Thus, suppose that the locus of personal-level, largely conscious, theorizing about area X is (mostly?) in an X-module rather than a central processor. One could remove “central processor” from my account of intuitions in favor of “X-module”. (Note that although the language faculty may be a locus of *subpersonal*-level theorizing about language, it is not of *personal*-level theorizing.)

The Modest Theory's view that intuitions about language are "theory-laden" is important to the Expertise Defense and so needs explaining.²¹ MMNS take the view to be that these intuitions "are the product of people's more or less inchoate empirical theories" (p. 9; see also Machery 2012, sec. 3.1). I put the view a bit differently. First, the view is *not* that these intuitions are theoretical judgments or the result of theorizing. Rather, the intuitions are mostly *the product of experiences* of the linguistic world. They are like "observation" judgments. As such, they are "theory-laden" in just the way that we commonly think observation judgments are. The antipositivist revolution in the philosophy of science, led by Thomas Kuhn and Paul Feyerabend, drew our attention to the way in which even the most straightforward judgments arising from observational experiences may depend on a background. We would not make the judgments if we did not hold certain beliefs or theories, some involving the concepts deployed in the judgments. We would not make the judgments if we did not have certain predispositions, some innate but many acquired in training, to respond selectively to experiences.²² There is need for some cautionary words about this theory ladenness.

(a) The power of the background to influence judgments should not be exaggerated. Thus a person observing the Müller-Lyre arrows will judge that one "looks longer" than the other even though she knows perfectly well that they are the same length. (b) The view is not that we consciously bring this background into play in a way that amounts to *theorizing* about the experience. Surely, we mostly don't. Nonetheless, the background plays a causal role in the judgment. (c) The view is not that we need to have done a deal of thinking about language before having linguistic intuitions: a thoroughly ignorant person may *learn* to have intuitions in an experimental situation (2006a, 502; 2006b, 114).²³ (d) Finally, the theory ladenness we are discussing is *epistemic*. It should not be confused with *semantic* theory-ladenness, the view that the meaning of an observation term is determined by the theory containing it. This "semantic holism", also part of the revolution, has little to be said for it in my view (1996, 87-135).

The Expertise Defense falls out of the Modest Theory, as Machery nicely explains (2012, sec. 2.2): we should prefer the linguistic intuitions of linguists and philosophers because they have the better background theory and training. But the Defense is not so easy, as we will soon see (5.5).

I take it as obvious that this argument for the Expertise Defense, like the argument in my first step, is not "largely irrelevant" to MMNS's project of challenging philosophy's intuition-based methodology. For, if sound, the Defense diminishes the significance of cultural variation in *folk* intuitions for the theory of reference. MMNS's charge of irrelevance is not looking well.

I emphasize immediately three things that are *not* consequences of the Expertise Defense. First, it is not a consequence that we can simply rest with the intuitions of a

²¹ I draw here on my 2006d and 2010b.

²² So "theory" in "theory-laden" has to be construed *very* broadly to cover not just theories proper but also these dispositions.

²³ I claim that this is the way to view intuitions of the ignorant in the ingenious "minimal pair" experiments (2006a, 499; 2006b, 110).

group of linguists or philosophers, even less, with the intuitions of one or two linguists or philosophers. These intuitions are open to test against linguistic reality, at least; see section 6. Second, it is not a consequence that the practice of gathering those intuitions informally is always appropriate. It would surely be better to gather them sometimes, at least, in a proper scientific way. Third, it is not a consequence that folk intuitions should never be sought nor that they provide no evidence (cf. Machery 2012, secs. 2.3 and 4.2). Indeed, they may sometimes provide evidence that is as good as, or even better than, that provided by the experts' intuitions. I shall say more about this (sec. 5.5).

However, it certainly is a consequence of the Expertise Defense that linguists and philosophers in the grip of different theories about some theoretically interesting cases may have different intuitions as a result: they may be *biased*. And false theories may lead to false intuitions:

Linguistic education should make a person a better indicator of linguistic reality just as biological education makes a person a better indicator of biological reality. Of course a person educated into a false theory may end up with distorted intuitions. But that is an unavoidable risk of epistemic life, in linguistics as everywhere else. We have no unsullied access to any reality. (2006a, 504; 2006b, 115)

Machery and Stich make much of the risk of bias, as we shall see (sec. 5.5).

5.4 MMNS's Response to the Modest Theory

What do MMNS have to say in response to the Modest Theory of linguistic intuitions? Nothing in the way of argument. However, they do appeal to authority:

As far as we know, there is not a single well known linguist who has endorsed Devitt's critique of the "voice of competence" account in print or embraced Devitt's alternative account. (p. 10)

As far as *I* know, MMNS are right, but I wonder why they think that this is worth saying.

First, linguists hardly ever discuss theories of intuitions *at all*, presumably feeling that they have better things to do, like constructing grammars. Indeed, they mostly seem to just *presuppose* VoC without even stating it explicitly. There seems to be little if any attention to the key epistemological question: *Why* are these metalinguistic intuitions good evidence? This is surprising given the importance attached to these intuitions as evidence in grammar construction. It is particularly surprising given the common concern about the evidential use of the informally gathered intuitions of linguists, particularly of just one or two linguists, rather than the use of intuitions gathered from ordinary speakers in a proper scientific way (Schütze 1996).²⁴

The lack of attention may stem partly from the received Chomskian "psychological conception" according to which the grammar for a language is about a cognitive system in the language faculty of its speakers. Assuming that the grammar is more or less

²⁴ Thus, consider Sorace and Keller (2005), Featherston (2007), and Myers (2009). The authors of these recent papers are among those cited by Machery and Stich as examples of people who "have not only criticized syntacticians' reliance on their own and their colleagues' intuitions, they also have put forward an alternative methodology: the careful survey of the intuitions of ordinary competent speakers" (2012, 497). Yet none of these papers raises the key epistemological question about these intuitions.

true, it follows from this conception that the grammar's rules (and principles) are embodied in a speaker's mind. A lot of work still has to be done to get VoC, of course: we need the details of *how* the embodied rules yield a speaker's metalinguistic intuitions. Still, it may be tempting to think that the embodied rules *must* be responsible for her intuitions, even *sans* details. Tempting or not, VoC does still need the details. Aside from that, this route to VoC faces a serious problem, in my view: the psychological conception is false. I have argued against it and in favor of a "linguistic conception" according to which, a grammar is about a nonpsychological realm of linguistic expressions, physical entities forming a symbolic or representational system (2003; 2006b, ch. 2; Devitt and Sterelny 1989).²⁵ It is then an open question whether competence in a language is constituted by the embodied rules of the language.

Whatever the reason for this lack of attention, as far as I know, no linguist has *argued* for VoC in recent times, nor responded to any of my arguments against VoC and for the Modest Theory. Philosophers of linguistics are a different story. Many have responded,²⁶ but they have not, I claim (2006d; 2008c; 2010b, c), succeeded in undermining my arguments.

This brings me to my second point. This issue is not to be settled by appeal to the authority of linguists,²⁷ or even to the authority of philosophers of linguistics; it is to be settled by argument. MMNS have offered none, nor even an alternative theory. Do they think VoC goes without saying?

5.5 Criticisms of the Expertise Defense

We are now in a position to consider the criticisms that MMNS and others make of the Expertise Defense. I start with two criticisms about bias that I don't think should worry us. The remaining three criticisms are more troubling.

(1) MMNS are very exercised about theoretical bias (p.11). Thus, Machery and Stich claim that "syntacticians' theoretical commitments risk influencing their intuitions, undermining the evidential role of these intuitions" (2012, 497). And they are quite right, of course. But that is just the sort of epistemic risk we *always* run in science. It is the price of learning anything: "the innocent eye is blind, the virgin mind is empty". Machery and Stich seem to hanker after the unobtainable: a bias-free world. Of course if VoC were correct then we could come close to escaping the epistemic risk in linguistics by consulting uneducated folk. Perhaps Machery and Stich hold VoC. But then they need to supply a good reason for doing so.

The Expertise Defense claims that we should, in general (see (3) below) prefer the intuitions of experts, despite the inescapable risk of theoretically biased intuitions. We

²⁵ This rejection has received a deal of criticism (some of it very harsh): Antony 2008; Collins 2007, 2008a, b; Dwyer and Pietroski 1996; Laurence 2003; Longworth 2009; Matthews 2006; Pietroski 2008; Rattan 2006; Rey 2006, 2008; Slezak 2009; Smith 2006. Devitt 2006d, 2008a, b, c, and 2009b are recent responses to some of these criticisms.

²⁶ Collins 2006, 2007, 2008; Rattan 2006; Rey 2006; Smith 2006; Mišćević 2006; Pietroski 2008; Slezak 2009; Textor 2009; Fitzgerald 2010.

²⁷ Indeed, I think that there is a deal of confusion among linguists about intuitions (2010c, secs. 2-3).

should not exaggerate the likely effect of that bias. The intuitive judgments that scientists, including linguists (see Sprouse and Almeida forthcoming), make about their domains tend to be in agreement. This is not surprising because the intuitions are not determined simply by theoretical background: they are determined largely, we hope, by experiences of the reality of that domain.

It is worth noting that philosophers who wanted to save description theories of names in the face of Kripke's arguments did not reject his referential intuitions, whether about humdrum cases or Gödel and Jonah cases, but rather tried to construct description theories that were compatible with those intuitions (see Devitt and Sterelny 1999, sec. 3.5 for discussion).

(2) Weinberg and colleagues emphasize another concern about bias. There is evidence that “non-truth-tracking factors” like “the order of presentation of the cases” lead to unreliability in folk intuitions. The concern then is:

[W]hat the purveyors of the expertise defense require is that philosophers' intuitions are sufficiently less susceptible to the kinds of unreliability that seem to afflict the folk intuitions studied by experimental philosophers. (2010, 333; original emphasis)

But the Expertise Defense does *not* require this. The Defense requires only that the philosophers' intuitions be better, in general (see (3) below), even if just as influenced by non-truth-tracking factors as the folk's. We assume that those factors do not *alone* determine the intuitions: background theory and linguistic reality play determining roles. So, the better the background, the better the intuitions. Similarly, the calls of a biased professional baseball umpire should be preferred to those of an equally biased fan.

We should, of course, always be on guard against bias of one sort or another in using philosophers' intuitions as evidence. There are some obvious ways to counter it. Thus, where the concern is theoretical bias, we can consider the intuitions of philosophers of various theoretical persuasions. This is in effect how Kripke's intuitions were informally tested. And we can look for other evidence. This can be found in usage; see section 6. It might even be found in folk intuitions. For, to repeat, it is not part of the Expertise Defense that folk intuitions should never be sought.²⁸

We move now to more serious worries about the Expertise Defense.

(3) MMNS think that the Defense lacks empirical support:

[W]hile Devitt simply assumes that the linguistic intuitions of linguists and philosophers of language will be more reliable than the intuitions of ordinary speakers, methodology-savvy syntacticians have begun to explore the issue empirically. As we read this growing body of literature, there is little in it to support the idea that the intuitions of linguists and philosophers of language are more reliable than those of ordinary speakers, and there is some reason to think that they may in fact be *less* reliable. (p.11)

First, it is not true that I *simply assume* that linguists and philosophers have “more reliable” intuitions. (1) The Modest Theory predicts that the intuitions of linguists and philosophers will be better in their respective linguistic domains: briefly, the better the

²⁸ One reason for seeking folk intuitions in linguistics, emphasized by MMNS (p.11), is a concern that the linguist's idiolect may differ from the folk's.

theory that intuitions are laden with, the better the intuitions. So my arguments for that theory and against its rival VoC support the prediction. (2) I offer some more empirical considerations in favor of the prediction from other domains and linguistics (2006a, 492-3, 499-500; 2006b, 104-5, 111).

Still, MMNS are right to demand more evidence that the prediction is right. But we need to be careful about the exact nature of the prediction.²⁹ Although it follows from the Modest Theory that we should prefer the intuitions of experts in the area in question, it does not follow from this that these experts will have more reliable intuitions than the folk about *every* fact in the area. It does not follow, for example, that the paleontologist who is better than the folk at identifying something as a pig's jawbone will also be better at identifying something as a skull. Perhaps educated folk would do just as well because they have *enough* expertise. What does follow from the Modest Theory is that the more expert a person is in an area, the better the person's background, the wider her range of reliable intuitions in the area. Turning to linguistics, the theory does not imply that linguists have more reliable intuitions about *every* linguistic fact:

[W]e can often be confident that such intuitions of normal educated speakers are right. We often have good reason to suppose that these core judgments of folk linguistics, partly reflecting "the linguistic wisdom of the ages", are good, though not of course infallible, evidence for linguistic theories. (2006a, 498-9; 2006b, 110)

We should prefer the linguists' intuitions particularly "when we get beyond the simple cases to theoretically interesting ones like 'The horse raced past the barn fell' and 'Who do you wanna kiss you this time?'" (2006a, 499; 2006b, 111).³⁰ The Modest Theory predicts that the more expert a person is in linguistics, the wider his range of reliable linguistic intuitions.

There is an obvious difficulty in testing this prediction: the theory provides no guidance as to what level of expertise is required to be a reliable intuiter about any particular sort of linguistic fact. Indeed, how could the theory? The level required is an independent empirical question.

These subtleties need to be kept in mind in assessing the interesting findings of Culbertson and Gross (2009). MMNS cite these findings as evidence against the Expertise Defense (p. 11; see also Machery 2012, sec. 3.2). From the perspective of the Modest Theory, the findings are indeed a little surprising. But given what that theory actually predicts, and issues arising from the vexed distinction between "acceptability" and "grammaticality" judgments, I argue that the bearing of these findings on the Expertise Defense is far from obvious (2010c; Gross and Culbertson 2011 is a response).

MMNS cite (p. 11) another experiment that is alleged to cast doubt on the Expertise Defense. It is another Gödel-case experiment (Machery 2012). The subjects were divided into three groups: Group 1, experts in semantics and the philosophy of language; Group 2, experts in discourse analysis, historical linguistics, and sociolinguistics; Group 3,

²⁹ I draw here on my discussion of Culbertson and Gross (2009) in Devitt (2010c).

³⁰ The experiments I cite at this point, Spencer (1973) and Gordon and Hendrick (1997), though open to criticism as Culbertson and Gross point out (2009, 727-8), illustrate the sort of non-simple cases where we might expect linguists' intuitions to differ from those of the folk.

comparably educated lay people. The first thing to note about the findings is that the intuitions of all groups were decisively Kripkean. However, group 1 was more Kripkean than 2 and 3, and 3 was more Kripkean than 2. The only *significant* result was the comparison of 1 and 2. Machery has a strange take on this result. Noting that groups 1 and 2 are both experts about language and yet their intuitions differ significantly, Machery concludes: “This inconsistent influence of expertise on intuitions about reference casts doubts on whether expertise really improves the reliability of these intuitions” (p. 50). Yet, of course, the two groups are experts on *different aspects* of language. What matters to the Expertise Defense is expertise about *reference*. We can confidently select Group 1 as likely experts about that. With Group 2, who knows? Their expertise in other aspects of language may well not transfer to reference. Consider the following passage, quoted by both Weinberg and colleagues (2010, 35) and MMNS (p. 10): there is “little transfer from high-level proficiency in one domain to proficiency in other domains—even when the domains seem, intuitively, very similar” (Feltovich et al. 2006, 47; emphasis added). Machery’s indecisive results cast no serious doubt on the Expertise Defense and may even give it some gentle support.

All in all, as far as I know, experimental work has not undermined the prediction that philosophers have a wider range of reliable referential intuitions than the folk.

But does experimental work *support* that prediction? MMNS have this to say about linguistic intuitions: “Experimental work on linguists’ and ordinary competent speakers’ intuitions has not shown that the former are more reliable than the latter” (p. 11). Maybe not. And, turning to referential intuitions, there does not seem to be experimental work supporting the analogous claim about philosophers and those intuitions. It would certainly be good to have some work. Here is a relatively easy way to get some.

We can probably assume that nearly all philosophers of language agree with Kripkean intuitions about Gödel cases; and Machery’s experiment, just discussed, supports that assumption. In contrast, MMNS have shown (let’s suppose) that EA folk do not agree. Furthermore, both EA and W folk “reveal considerable intra-cultural variation” in their intuitions (B8). So in Gödel cases we have a clear divergence between the philosophers and the folk. If we could now produce evidence that the philosophers are *right*, we would have shown that the philosophers’ intuitions are indeed more reliable here than the folk’s. We can hope to find this evidence in studies of usage (sec. 6), perhaps even in folk intuitions about humdrum cases, that count against description theories and in favor of Kripke’s causal “picture” of names.

I think that theories of reference need evidence from usage anyway (sec. 6). That evidence will be doubly helpful if it supports the Expertise Defense, thus justifying a preference for philosophers’ intuitions over the folk’s. And the evidence is needed also to address the following two concerns.

(4) Weinberg and colleagues (2010), in a discussion cited by MMNS (p. 10), draw attention to some literature on the development of expertise. This does not challenge the Modest Theory but does throw doubt on whether that theory really does support the Expertise Defense as I have claimed. It seems that there is:

[T]remendous diversity in the development of expertise according to the characteristics of the task and the learning environment. Some areas, such as meteorology and chess, have proved conducive to acquiring expertise; others, such as psychiatry, stock brokerage, and polygraph testing, have tended not to produce real expertise. (p. 334)

So, is philosophy of language in the meteorology group or the psychiatry group? My hunch clearly is that it is in the former group. But hunches are not enough. Evidence of the sort just outlined would be helpful.

Even if philosophy is in with meteorology, MMNS have another worry, following Weinberg and colleagues again:

[E]ven if it is the case that philosophers of language have a great deal of expertise about many aspects of natural language, it does not follow that their intuitions about thought-experiments concerned with reference are more reliable than those of other speakers. (p. 10)

But expertise in philosophy of language simply *is* an expertise about reference, meaning, truth conditions, and the like. According to the Expertise Defense *this* expertise is the background for referential intuitions *whenever they are formed*, whether in a thought experiment (Kripke's Gödel case) or not (Kripke's Einstein case).

(5) Weinberg and colleagues have another related worry about the Expertise Defense. It starts with the following claim:

One of the most robust consensus findings of the study of expertise is that expert judgments can only become more reliable where experts are readily confronted with clear, reliable feedback on which to train. (2010, 340)

Weinberg and colleagues wonder "what, specifically, might that feedback be" in philosophy? They contrast philosophy unfavorably with other disciplines in this respect: "Philosophy rarely if ever...provides the same ample degree of well-established cases to provide the requisite training regimen" (p. 341). They conclude pointedly that philosophers are in the category

of the mildly self-deceived, and at a minimum, these considerations indicate that proponents of the expertise defense need to offer real, substantive scientific evidence that this is not so. (p. 342)

I think that they are right that we could do with that evidence. The evidence I have outlined would again be helpful. Nonetheless, we should not overlook that philosophers of language are confronted informally by language use that does provide feedback. Thus, my Kripkean intuitions about names have been confirmed, day in and day out for forty years, by observations of people using a name to refer successfully to an object that they are ignorant or wrong about. More on this in section 6.

5.6 Conclusions

I have proposed the Modest Theory according to which intuitions are theory-laden. This implies that we should prefer the intuitions of those with better theories and training, the experts. However, the theory provides no guidance as to what level of expertise is required to be a reliable intuiter about any particular sort of fact. When it comes to referential intuitions, I predict that the folk's intuitions may be as reliable as the philosophers' about reference in humdrum actual cases; the folk are likely expert enough for these cases. In fanciful hypothetical cases, I predict that the philosophers'

intuitions will be more reliable. Gödel cases are of that sort. That is the Expertise Defense against the significance of MMNS's findings.

One way to undermine this Defense is to reject the Modest Theory. MMNS implicitly do so, but without argument.

But is there any alternative to this theory of referential intuitions? Setting aside apriorism, the only alternative appears to be VoC. MMNS seem to favor this theory, and it does count against the Defense. According to VoC, the true theory of reference for names is represented, or otherwise embodied, in the minds of all competent speakers. Has anyone explicitly endorsed this view? In any case, I have argued that we have no reason to believe VoC and that it runs counter to the evidence that linguistic competence precedes metalinguistic competence in a child's development. Furthermore, if the theory of reference were embodied *without* being represented, we would have no idea how it *could* yield referential intuitions.

Still, the Expertise Defense faces a problem. The problem is not risk of bias, for that risk is a feature of epistemic life in general and can be guarded against. The problem is that the Modest Theory's prediction that the referential intuitions of philosophers of language should be preferred to those of the folk could do with empirical support. The need for this support becomes particularly pressing in light of the following concerns: that training in philosophy may not produce real expertise; in particular, the training may not provide the requisite reliable feedback. The place to look for this empirical support is in linguistic usage. We shall consider that in the next section.

The Expertise Defense bears on both goals of this paper. It bears on the secondary goal because, if it is right, it further undermines the significance of MMNS's findings of variation in the folk's referential intuitions. But it is more important to the main goal because, if it is right, our methodology should attach more weight to the intuitions of philosophers than to those of the folk in constructing theories of reference. This would give some legitimacy to the practice of consulting only the intuitions of philosophers. Should we discover that the Expertise Defense was wrong, which I don't expect, then there would be no basis for preferring the philosophers' intuitions over the folk's. Both sets of intuitions might still mostly be reliable enough to use as evidence but clearly this discovery would diminish the evidential value of philosophers' intuitions and increase the importance of other evidence.

One further thought on philosophers' intuitions. At present these are seldom gathered by a proper scientific survey. Nonetheless they are publicly aired in departments, conferences, and journals. Any that are not shared are likely to be challenged.³¹ So, for the most part, it is likely to be unnecessary to do a proper survey before using these intuitions as evidence.

The Expertise Defense raises interesting and controversial issues. Clearly its status needs to be settled.

Finally I move to the third step, the crux of my main goal to propose a methodology for the theory of reference in which experiments feature prominently.

³¹ Colin Phillips (forthcoming) makes a similar point about the informal use of linguists' intuitions; see also the discussion in Sprouse and Almeida (forthcoming).

6. *Third Step: Testing Usage*

I have talked of testing theories and intuitions about reference against “linguistic reality” and of doing this by checking “linguistic usage”. I think that this, rather than testing referential intuitions, should be the focus of experimental semantics. I shall finish this paper by saying more about this.

A language is a system of representations or symbols, governed by rules, that scientists sometimes posit in a species to explain its communicative behavior. Thus, languages have been posited for honey bees, for Gunnison’s prairie dogs, and, of course, for humans. The “linguistic reality” that concerns us here is the language in a particular community of humans. Each time members of the community speak the language, we have a piece of “linguistic usage”. What about reference? Referential relations are commonly assumed to be central to the nature of language. If this assumption is right, then theories of reference are central to our theory of a language.

If the referential assumption is not right – and Machery and Stich at least toy with the idea that it is not right (2012, sec. 4; see also Stich 1996, 37-51; 2009, 199) – then there is no place for theories of reference at all and hence no call to gather any evidence about reference. If one does not make the assumption then one needs, of course, to explain language’s relation to the world in terms other than reference. Thus Paul Horwich (1998, 2005) has a deflationary theory of reference and proposes a use theory of language.³² In any case the referential assumption is a presupposition of any methodology for theories of reference.

Previous sections have been concerned with the gathering what is, in effect, *indirect* evidence about nature of reference: consulting people’s *intuitions about reference relations*. Yet science does not generally proceed like this. For example, we don’t mostly do biology by consulting people’s intuitions about living things; we seek direct evidence about the living things themselves. Similarly, I suggest, we should do semantics by seeking *direct* evidence about *the reference relations themselves*. How do we do that?

6.1 *The Corpus*

Reference relations are manifested in usage. So one way to gather direct evidence is to look at the corpus of usage. An example of what we can learn from the corpus is provided, unwittingly (and ironically), by MMNS themselves in the very experiment we have been discussing. *Their own uses* of ‘Gödel’ in a vignette designed to *test* description and causal theories of reference, are *inconsistent with* what (standard) description theories would predict.³³ (As a result, the experiment is biased against description theories³⁴ and so flawed in design.) Their ‘Gödel’ vignette reads as follows:

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

³² I don’t think that this theory works (2002, 2011c).

³³ Similar remarks apply to James Genone and Tania Lombrozo’s use of the invented term ‘tyleritis’ in a vignette used in another piece of experimental semantics (forthcoming).

³⁴ Which, of course, adds to the puzzle that EAs were found to have descriptivist intuitions.

he can give an accurate statement of the 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. (Machery et al. 2004, B6)

This vignette contains eight uses of the name ‘Gödel’ (and one mention). Now consider the question: Who do these uses refer to? MMNS are the authors of this vignette and there can be no doubt that these philosophers are fully competent with the name ‘Gödel’. And the referent of this name out of the mouths of the fully competent is to the eminent logician who did in fact prove the incompleteness of arithmetic and spent many years at Princeton. So *that* is who MMNS’s eight uses of the name in the vignette refer to. But then their use of the name in the following passage disconfirms the description theory: “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.” For, if MMNS’s use of ‘Gödel’ refers to that eminent logician in virtue of their associating with it the description, ‘the prover of the incompleteness of arithmetic’, this passage is not something that MMNS would be disposed to say. They would not, in one and the same breath, both refer to Gödel and suppose away the basis of that reference. Similarly, according to the theory that the reference of ‘bachelor’ is determined partly by its association with ‘unmarried’, competent speakers would not be disposed to say: “Suppose that the bachelors in Iceland are married.” But here the description theory seems to survive because we would *not* be disposed to say this.³⁵

This is an example of how we can use the corpus to argue for/against a theory of reference. Of course there are difficulties in using the corpus in a scientific way. First, one has to note something in the linguistic phenomena that *is* evidence for/against some theory of reference. Then one has to have a record of it, which is problematic if it is spoken rather than written. And one may need to document quite a lot of information about the speaker and circumstances. So my recent claim (sec. 5.5) of forty years of observations in support of Kripkean intuitions does not qualify. Still, it does illustrate what a mass of evidence the corpus provides that *could* be mined scientifically. And it indicates the important role that the corpus plays as informal evidence about reference.³⁶

³⁵ I once said much the same about Kripke’s case of Jonah: “Note that it is *not possible*, according to the description theory, for an earlier scholar to speculate, or to find evidence, that Jonah was a certain ordinary man that he, the scholar, has tracked down: that Jonah was the subject of superstitious stories; and so forth. Such speculations and evidence cannot be *about Jonah* because they deny the descriptions on which our use of the name depends. (1981, 19; see also Devitt and Sterelny 1999, 56).

³⁶ I made a similar point about the evidential role of the corpus in linguistics as part of a response to the tendency in linguistics to exaggerate the role of speakers’ intuitive judgments (2006a, 486-7; 2006b, 98-9).

6.2 Elicited Production

Fortunately, we don't have to rely on the corpus for direct evidence in usage: we can *induce* usage from competent speakers in experimental situations. Consider this description of "the technique of *elicited production*" in linguistics:

This technique involves children in a game, typically one in which children pose questions to a puppet. The game orchestrates experimental situations that are designed to be uniquely felicitous for production of the target structure. In this way, children are called on to produce structures that might otherwise not appear in their spontaneous speech. (Thornton 1995, 140)

This technique is frequently used on children, partly because of the difficulty of getting helpful intuitions from them.³⁷ Clearly much direct evidence could be gathered in this way. However, contriving appropriate situations in an experiment is likely to be a laborious business.

I proposed an easier technique of elicited production for linguistics. Instead of constructing situations to see what people say and understand in those situations, "we can *describe* situations and ask people what they would say or understand in those situations" (2006a, 487; 2006b, 99). Note that this is quite different from the much-discussed earlier method of describing situations that include utterances and asking people to judge the linguistic properties of those utterances. The present method is not to prompt these *metalinguistic intuitions*, yielding indirect evidence, but to prompt *linguistic usage*, yielding direct evidence.³⁸ Such a method has surely often been employed informally: linguists and philosophers ask themselves, and sometimes ordinary speakers, what they would say or understand in various situations.³⁹ I don't know whether this technique has ever been used experimentally in linguistics and I am quite sure it has not been in the theory of reference. The method could provide a rich source of evidence. It seems to me to be the way forward in experimental semantics.

6.3 An Example of Elicited Production

How does one go about testing theories of reference by this method of inducing usage? In "Experimental Semantics" (pp. 429-432), drawing on an earlier discussion of methodology (1994, 1996), I proposed one way for testing a description theory of names (although I did not describe it, as I should have, as inducing usage). I shall not repeat the details but the basic idea is to use qualitative measures to evaluate how a character's association of descriptive information with a proper name influences what mental states people competent with the name will ascribe to the character. The Kripkean prediction is that it will make no difference to those people's readiness to use the name in their ascriptions whether or not the character associates with the name an

³⁷ See Crain et al. (2005, sec. 4) for a nice summary.

³⁸ The expression of a metalinguistic intuition is, of course, a piece of usage but its evidential role as such an expression is quite different from its role as simply a piece of usage (Devitt 2010c, sec. 1.2).

³⁹ I sum up my discussion of linguistic evidence: "the main evidence for grammars is not found in the intuitions of ordinary speakers but rather in a combination of the corpus, the evidence of what we would say and understand, and the intuitions of linguists" (2006b, 100).

identifying description of the name's bearer. Wesley Buckwalter and I have begun conducting experiments in just this way on English native speakers. Our preliminary results confirm the Kripkean prediction (2011).

This proposal for investigating reference without using referential intuitions was the core of my third step in diminishing the significance of MMNS's experimental findings. We have already seen that the first two steps escape MMNS's charge of being "largely irrelevant" to their project. And so does this one if we take that project to have been the bold one of challenging the very nature of the referential enterprise. For, the third step shows that the task of explaining reference is still viable, whatever the concerns about the reliability of intuitions. So, far from *many* of my arguments failing to address their concerns, *none* did. I think that I had good reason to take their project to be the bold one (sec. 3.1). Still, perhaps the project was the less bold one of just challenging the intuition-based methodology of philosophy, as MMNS now say it was. Then my third step is indeed irrelevant, for it simply proposes a different methodology. Even so, one out of three does not make many.

My proposal is just one way to test theories of reference by the method of elicited production. The success of the method does not of course depend on this way.

6.4 MMNS's Response

MMNS make two brief criticisms of this proposal.⁴⁰ (1) They allege that it is "really a bait-and-switch":

We are promised that one can provide evidence about the nature of reference without appeal to intuitions, but, when Devitt describes his method in detail, it turns out that it does appeal to intuitions after all, just not intuitions *that are explicitly about the reference of terms*. Rather, the evidential basis for deciding between theories of reference is supposed to be mental states ascriptions – intuitive judgments about other people's mental states. (p. 12)

I didn't promise no appeal to intuitions. My proposal is to seek evidence in some linguistic usage, in thought ascriptions. These ascriptions in response to a situation, like any immediate "intuitive" assertion in response to any situation, might be described as "intuitions" about the situation: I'm easy about the term 'intuition' (2006a, 491; 2006b, 103). But if those ascriptions are described as intuitions then all the "observational judgments" of science should be so described. Intuitions are then unavoidable as evidence. What is at issue is not the evidential role of intuitions but of intuitions *about reference* in a theory of reference. MMNS have missed the point.

(2) They go on:

Furthermore, it is unclear that this alternative involves no appeal to intuitions about reference since ascribing beliefs or thoughts to others involves determining what they are thinking about. On many accounts, this amounts to determining what they are referring to. (p.12)

Suppose that in the vignette a character Frank says, "Cicero was an ancient Greek". Suppose next that the participant, Alice, as competent as anyone with 'Cicero', re-

⁴⁰ They also refer to Stich's criticisms (2009, 197-200) of a brief (ill-judged) earlier proposal of mine for testing a theory of reference (2009a, 49-50). But these criticisms are beside the point as that proposal, with the role it assigns to folk intuitions, is quite different from the present proposal.

sponds to prompts with, “Frank believes that Cicero was an ancient Greek”. *Simply on the strength of Alice’s competence with ‘Cicero’*, theorists can reason from Alice’s use of that name in her response that Frank’s use of it in the vignette probably referred to the famous orator Cicero. This counts against a description theory. The claim that the mental processes that cause Alice’s response must involve *thoughts about reference*, in particular the intuition that Frank was referring to Cicero, is theoretically controversial at least and in my view quite wrong. But the important point is that we need take no stand on this theoretical controversy to use Alice’s response as evidence. It counts as evidence without the theoretical burden of any views about referential intuitions, about their cognitive roles, epistemic value, or whatever. That’s why linguistic usage like Alice’s is such good and basic evidence.

Finally, I wonder why MMNS are so unsympathetic to the idea of looking to usage for evidence. They are clearly very influenced by the debate about methodology in linguistics. Though this debate is certainly dominated by attention to the role of intuitions as evidence – far too much so, in my view (2006a, 486-7; 2006b, 98-9) – the role of the corpus as evidence is usually acknowledged, even if only in passing: it is acknowledged for example in Featherston 2007 and Myers 2009. And the idea of getting evidence from elicited production has a respectable place in linguistics.⁴¹ So it is curious that, in theorizing about reference, MMNS focus all their energy on testing intuitions and never even mention looking to the corpus or testing usage.

6.5 *The Methodological Proposal*

My main goal in this paper was to propose a methodology for the theory of reference in which experiments feature prominently. I am now in a position to do so.

(a) *Philosophers’ Referential Intuitions*: The received methodology includes the informal testing of theories against philosophers’ intuitions. If the Expertise Defense is right, then those intuitions are likely to be reliable and so it is appropriate that they should have an evidential role. Doubtless it would sometimes be appropriate to survey these intuitions scientifically but, given how publicly these intuitions are typically aired, this may not be necessary often.⁴² Even if the Expertise Defense is right, the evidential value of these intuitions does not match that from (B) and (C) below. And, if the Defense should turn out to be wrong, the value of these intuitions would be greatly diminished although not eliminated.

(b) *The Corpus*: I suspect that the present methodology also includes much informal testing of theories against philosophers’ observations of language use, observations of the corpus; for example, observations of the ignorant successfully referring with a

⁴¹ All the items mentioned (as well as reaction time studies, eye tracking, and electromagnetic brain potentials) are mentioned by Manfred Krifka (forthcoming) in a helpful summary of the evidence that linguists use in semantics.

⁴² “Though it is relatively common to assume that formal experiments provide ‘better’ results than informal results, the current state of the field suggests that many of the perceived benefits of formal experiments ultimately disappear under closer empirical scrutiny.” (Sprouse and Almeida forthcoming)

name. We should add to this informal practice the systematic and scientific study of the corpus for evidence.

(c) *Elicited Production*: The most novel part of my proposal is that we should induce linguistic usage from competent speakers to test theories. This could be done by constructing situations to see what people say and understand in those situations, but it is far easier to describe situations and ask people what they say or understand about those situations. I have provided an example of how we might go about doing this.

(d) *Folk's Referential Intuitions*: Finally, there is surely some place for the scientific surveying of the folk's intuitions, pioneered by MMNS. However, I think that it is a relatively small place. There is a point to surveying folk rather than more expert philosophers only in the rather unlikely circumstance that there is no other way of guarding against theoretical bias (or if there is concern that the folk have a referentially different idiolect). Even if the Expertise Defense turned out to be wrong, I suspect that wherever it might otherwise seem appropriate to test folk intuitions it would be better to seek direct evidence in folk usage.

7. Conclusions

"Experimental Semantics" offered a three-step argument aimed to greatly diminish the significance of the findings in MMNS's original paper. In response they claim that I have seriously misunderstood their goal and that many of my arguments are "largely irrelevant". They concede nothing. The secondary goal of my paper has been to defend my argument from this disappointing response.

But the paper starts by raising some doubts about MMNS's prediction of cultural variation in referential intuitions. It is hard to see how the recent psychological work that is alleged to support this prediction really does support it (sec. 2).

In response to the charge of misunderstanding, I have argued that, except perhaps in one respect that MMNS do not have in mind, taking their goal to be a bit bolder than it was, I understood their goal well (sec. 3). As to the charge of irrelevance, my first two steps are obviously relevant to their goal and the third is relevant if their goal was the bolder one (secs. 4-6). Their charges seem baseless.

Despite the alleged irrelevance of my arguments, MMNS do respond to them. My first step argued that the referential intuitions about Gödel cases that MMNS tested are relatively insignificant in Kripke's argument against description theories: the combination of referential intuitions about humdrum cases and modal intuitions are massively more significant. In response, MMNS claim, implausibly and largely without argument, that the cultural variation in Gödel-case intuitions is likely to generalize to actual-case intuitions; and that we will have to rely on Gödel-case intuitions in the end because actual-case ones will not suffice to choose a theory of reference. MMNS are on such weak ground here that I could rest my case against the significance of their findings with this first step (sec. 4).

In light of this discussion and doubts that the psychological work predicts cultural variation in referential intuitions, I suspect that the variation in Gödel-case intuitions is a red herring.

My main goal in the paper is to propose a methodology for the theory of reference in which experiments feature prominently. The relative place of philosophers' referential intuitions in this methodology depends on the Expertise Defense. If the Defense is right, we should attach more weight to the intuitions of philosophers than to those of the folk in constructing theories of reference. And if it is right, it further undermines the significance of MMNS's findings of variation in the folk's referential intuitions. My second step argued that the Defense is right. MMNS and others have shown that this claim needs more empirical support (sec. 5).

Finally, I have argued for testing usage as central to the methodology. I suspect that theories are already being informally tested by observations of the corpus and recommend that we should add to this some systematic and scientific study of the corpus. My most novel proposal is that we should use the technique of elicited production to test usage: we should conduct experiments in which situations are described and people are prompted to see what they say and understand in those situations. Finally, I propose a small place for the scientific surveying of the folk's referential intuitions (sec. 6).

In sum, I am urging a methodology that uses: (A) philosophers' referential intuitions, both informally and, occasionally, scientifically gathered; (B) the corpus, both informally and scientifically gathered; (C) elicited production; and, occasionally, (D) folk's referential intuitions. The most novel part of this is (C) and that is where most of the experimental work should be.

The philosophical tradition is to base semantic theories on philosophers' intuitions. MMNS go along with the traditional focus on intuitions as evidence but add folk intuitions to the mix. This is the wrong way for experimental semantics to go. What is needed is a different evidential focus: the testing of usage.

REFERENCES

- Antony, L. 2008. MetaLinguistics: methodology and ontology in Devitt's *Ignorance of language*. *Australasian Journal of Philosophy* 86: 643-56.
- Barber, A., ed. 2003. *Epistemology of language*. Oxford: Oxford University Press.
- Bialystok, E. 1986. Factors in the growth of linguistic awareness. *Child Development* 57: 498-510.
- , and E. B. Ryan. 1985. A metacognitive framework for the development of first and second language skills. In *Metacognition, cognition, and human performance*, ed. D. L. Forrest-Pressley, G. E. MacKinnon, and T. G. Waller, 207-52. New York: Academic Press.
- Buckwalter, W. and M. Devitt. 2011. Testing theories of reference: Usage vs intuitions. Metro Experimental Research Group, City University of New York Graduate Center, invited presentation 11 February 2011.
- Chomsky, N. 1986. *Knowledge of language: Its nature, origin, and use*, New York: Praeger Publishers.
- . 1995. Language and nature, *Mind* 104: 1-61.
- Collins, J. 2006. Between a rock and a hard place: A dialogue on the philosophy and methodology of generative linguistics. *Croatian Journal of Philosophy* VI: 469-503.
- . 2007. Review of Devitt 2006b. *Mind* 116: 416-23.
- . 2008a. Knowledge of language redux. *Croatian Journal of Philosophy* VIII: 3-43.
- . 2008b. A note on conventions and unvoiced syntax. *Croatian Journal of Philosophy* VIII: 241-7.
- Crain, S., A. Gualmini, and P. Pietroski. 2005. Brass tacks in linguistic theory: Innate grammatical principles. In *The Innate Mind: Structure and Contents*, ed. S. Laurence, P. Carruthers, and S. Stich, 175-97. Oxford: Oxford University Press.

- Culbertson, J., and S. Gross. 2009. Are linguists better subjects?. *British Journal for the Philosophy of Science* 60: 721-36.
- Deutsch, M. 2009. Experimental philosophy and the theory of reference. *Mind and Language* 24: 445-65.
- Devitt, M. 1981. *Designation*. New York: Columbia University Press.
- . 1994. The methodology of naturalistic semantics, *Journal of Philosophy* 91: 545-72.
- . 1996. *Coming to our senses: A naturalistic program for semantic localism*, Cambridge: Cambridge University Press.
- . 2002. Meaning and use. *Philosophy and Phenomenological Research* 65: 106-21.
- . 2003. Linguistics is not psychology. In Barber 2003a: 107-39.
- . 2006a. Intuitions in linguistics. *British Journal for the Philosophy of Science* 57: 481-513.
- . 2006b. *Ignorance of language*, Oxford: Clarendon Press.
- . 2006c. Intuitions. In *Ontology studies cuadernos de ontologia: Proceedings of VI international ontology congress San Sebastian, 2004*, ed. V. G. Pin, J. I. Galparaso, and G. Arrizabalaga, 169-76. San Sebastian: Universidad del País Vasco. Reprinted in Devitt 2010a.
- . 2006d. Defending *Ignorance of language*: Responses to the Dubrovnik papers. *Croatian Journal of Philosophy* VI: 571-606.
- . 2008a. Explanation and reality in Linguistics. *Croatian Journal of Philosophy* VIII: 203-31.
- . 2008b. A response to Collins' note on conventions and unvoiced syntax. *Croatian Journal of Philosophy* VIII: 249-55.
- . 2008c. Methodology in the philosophy of linguistics. *Australasian Journal of Philosophy* 86: 671-84.
- . 2009a. On determining what there isn't. In Murphy and Bishop 2009: 46-61. Reprinted in Devitt 2010a.
- . 2009b. Psychological conception, psychological reality: A response to Longworth and Slezak. *Croatian Journal of Philosophy* IX: 35-44.
- . 2010a. *Putting metaphysics first: Essays on metaphysics and epistemology*, Oxford: Oxford University Press.
- . 2010b. What 'intuitions' are linguistic evidence?. *Erkenntnis* 73: 251-64.
- . 2010c. Linguistic intuitions revisited, *British Journal for the Philosophy of Science* 61: 833-65.
- . 2011a. No place for the a priori. In *What place for the a priori?*, ed. Michael J. Shaffer and Michael L. Veber, 9-32. Chicago and La Salle: Open Court Publishing Company. Reprinted in Devitt 2010a.
- . 2011b. Experimental semantics. *Philosophy and Phenomenological Research* LXXXII: 418-35.
- . 2011c. Deference and the use theory. *ProtoSociology* 27: 196-211.
- . 2012. The role of intuitions in the philosophy of language. In *Routledge companion to the philosophy of language*, ed. Gillian Russell and Delia Graff Fara. New York: Routledge: 554-65.
- , and K. Sterelny. 1989. Linguistics: What's wrong with 'the Right View'. In *Philosophical Perspectives, 3: Philosophy of Mind and Action Theory, 1989*, ed. James E. Tomberlin, 497-531. Atascadero: Ridgeview Publishing Company.
- , and —. 1999. *Language and reality: An introduction to the philosophy of language*, 2nd edn. 1st edn 1987. Oxford: Blackwell Publishers.
- Dwyer, S. and P. Pietroski. 1996. Believing in language. *Philosophy of Science* 63: 338-73.
- Evans, G. 1973. The causal theory of names. *Proceedings of the Aristotelian Society*, Suppl. vol. 47:187-208.
- Feltovich, P. J., M. J. Prietula, and K. A. Ericsson. 2006. Studies in expertise from psychological perspectives. In *The Cambridge handbook of expertise and expert performance*, ed. K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman, 41-67. Cambridge: Cambridge University Press.
- Featherston, S. 2007. Data in generative grammar: The stick and the carrot. *Theoretical Linguistics* 33: 269-318.
- Fitzgerald, G. 2010. Linguistic intuitions. *British Journal for the Philosophy of Science* 61: 123-60.
- Genone, J. and T. Lombrozo. Forthcoming. Concept possession, experimental semantics, and hybrid theories of reference. *Philosophical Psychology*.
- Gordon, P. C. and R. Hendrick. 1997. Intuitive knowledge of linguistic co-reference. *Cognition* 62: 325-70.
- Gross, S. and J. Culbertson. 2011. Revisited linguistic intuitions. *British Journal for the Philosophy of Science* doi: 10.1093/bjps/axr009
- Hakes, David T. 1980. *The development of metalinguistic abilities in children*. Berlin: Springer-Verlag.
- Hintikka, J. 1999. The Emperor's new intuitions, *Journal of Philosophy* 96: 127-47.
- Horwich, Paul. 1998. *Meaning*. Oxford: Clarendon Press.
- . 2005. *Reflections on meaning*. Oxford: Clarendon Press.

- Ichikawa, J., I. Maitra, and B. Weatherson. 2011. In defense of a Kripkean dogma. *Philosophy and Phenomenological Research*: DOI 10.1111/j.1933-1592.2010.00478.x
- Krifka, Manfred. Forthcoming. Varieties of semantic evidence. In *Handbook of Semantics*, ed. Claudia Maienborn, Paul Portner, and Klaus von Heusinger. Berlin: Mouton de Gruyter.
- Kripke, S. A. 1980. *Naming and necessity*. Cambridge, MA: Harvard University Press.
- Laurence, S. 2003. Is linguistics a branch of psychology? In Barber 2003: 69-106.
- Longworth, G. 2009. Ignorance of linguistics. *Croatian Journal of Philosophy* IX: 21-34.
- Ludwig, K. 2007. The epistemology of thought experiments: First-person approach vs. third-person approach. *Midwest Studies in Philosophy* 31: 128-159.
- Machery, E. 2012. Expertise and intuitions about reference. *Theoria*, Vol. 27/1, N°73, 37-54.
- , Mallon, S. Nichols, and S. P. Stich. 2004. Semantics, cross-cultural style. *Cognition* 92: B1-B12.
- , —, —, and —. 2012. If Folk Intuitions Vary, Then What? *Philosophy and Phenomenological Research* LXXXII. DOI 10.1111/j.1933-1592.2011.0055.x
- , Olivola, C. Y., and de Blanc, M. 2009. Linguistic and metalinguistic intuitions in the philosophy of language. *Analysis* 69: 689-694.
- , and S. P. Stich. 2012. The role of experiments in the philosophy of language. In *Routledge companion to the philosophy of language*, ed. Gillian Russell and Delia Graff Fara. New York: Routledge: 495-512.
- McKinsey, M. 1987. Apriorism in the philosophy of language. *Philosophical Studies* 52: 1-32.
- Mallon, R., E. Machery, S. Nichols, and S. P. Stich. 2009. Against arguments from reference. *Philosophy and Phenomenological Research* 79: 332-56.
- Martí, G. 2009. Against semantic multi-culturalism. *Analysis* 69: 42-48.
- , Forthcoming. Empirical data and the theory of reference. In *Topics in contemporary philosophy, volume 10: Reference and referring*, ed. M. O'Rourke. Cambridge, MA: MIT Press.
- Matthews, R. J. 2006. Could competent speakers really be ignorant of their language? *Croatian Journal of Philosophy* VI: 457-67.
- Miščević, N. 2006. Intuitions: The discrete voice of competence. *Croatian Journal of Philosophy* VI: 523-48.
- Murphy, D. and M. A. Bishop, eds. 2009. *Stich and his critics*. Oxford: Wiley-Blackwell.
- Myers, J. 2009. Syntactic judgment experiments. *Language and Linguistics Compass* 3: 406-23.
- Nisbett, R. E., K. Peng, I. Choi, and A. Norenzayan. 2001. Culture and systems of thought: Holistic vs analytic cognition. *Psychological Review* 108: 291-310.
- Norenzayan, A., E. Smith, and B. Kim. 2002. Cultural preferences for formal versus intuitive reasoning. *Cognitive Science* 26: 653-684.
- Ostertag, G. Forthcoming. The 'Gödel' effect.
- Phillips, C. Forthcoming. Should we impeach armchair linguistics?. In *Japanese/Korean linguistics 17*, ed. S. Iwasaki. CSLI Publications.
- Pietroski, P. 2008. Think of the children. *Australasian Journal of Philosophy* 86: 657-69.
- Rattan, G. 2006. The knowledge in language. *Croatian Journal of Philosophy* VI: 505-21.
- Rey, G. 2006. Conventions, intuitions and linguistic inexistents: A reply to Devitt. *Croatian Journal of Philosophy* VI: 549-69.
- . 2008. In defense of folieism: Replies to critics. *Croatian Journal of Philosophy* VIII: 177-202.
- Ryan, E. B. and G. W. Ledger. 1984. Learning to attend to sentence structure: Links between metalinguistic development and reading. In *Language awareness and learning to read*, ed. John Downing and Renate Valtin, 149-71. New York: Springer-Verlag.
- Schütze, C. T. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.
- Slezak, P. 2009. Linguistic explanation and 'psychological reality'. *Croatian Journal of Philosophy* IX: 3-20.
- Smith, B. C. 2006. Why we still need knowledge of language. *Croatian Journal of Philosophy* VI: 431-56.
- Sorace, A. and F. Keller. 2005. Gradience in linguistic data. *Lingua* 115: 1497-524.
- Spencer, N. J. 1973. Differences between linguists and nonlinguists in intuitions of grammaticality-acceptability. *Journal of Psycholinguistic Research* 2: 83-98.
- Sprouse, J. and D. Almeida. Forthcoming. The role of experimental syntax in an integrated cognitive science of language. In *The Cambridge handbook of biolinguistics*, ed. Kleantes Grohmann and Cedric Boeckx. Cambridge: Cambridge University Press.

- Stich, S. P. 1983. *From folk psychology to cognitive science: The case against belief*. Cambridge, MA.: Bradford/MIT Press.
- . 1996. *Deconstructing the mind*. New York: Oxford University Press.
- . 2009. Reply to Devitt and Jackson. In Murphy and Bishop 2009: 190-202.
- Textor, M. 2009. Devitt on the epistemic authority of linguistic intuitions. *Erkenntnis* 71: 395-405.
- Thornton, R. 1995. Referentiality and wh-movement in child English: Juvenile *d-link*uency. *Language Acquisition* 4: 139-75.
- Watanabe, M. 1998. Styles of reasoning in Japan and the United States: logic of education in two cultures. Paper presented at the American Sociological Association annual meeting, San Francisco, CA.
- Watanabe, M. 1999. Styles of reasoning in Japan and the United States: logic of education in two cultures. Unpublished PhD thesis, Columbia University, New York.
- Weinberg, J. M., C. Gonnerman, C. Buckner, and J. Alexander. 2010. Are philosophers expert intuiters? *Philosophical Psychology* 23: 331-355.
- Williamson, T. 2007 *The philosophy of philosophy*. Oxford: Blackwell Publishing.

MICHAEL DEVITT is Distinguished Professor of Philosophy at the Graduate Center at CUNY. His research interests include the philosophy of linguistics and the philosophy of language. He is the author of *Designation* (Columbia, 1981), *Coming to Our Senses* (CUP, 1996), *Language and Reality* (with Kim Sterelny, MIT, 1999), *Ignorance of Language* (OUP 2006), and the editor (with Richard Hanley) of the *Blackwell Guide to the Philosophy of Language* (Blackwell, 2006). His most recent book is *Putting Metaphysics First* (OUP, 2010).

Address: The Graduate Center, City University of New York, 365 Fifth Ave., New York, NY 10016, USA.
E-mail: mdevitt@gc.cuny.edu