

Against Arguments from Reference*

RON MALLON

University of Utah

EDOUARD MACHERY

University of Pittsburgh

SHAUN NICHOLS

University of Arizona

STEPHEN STICH

Rutgers University

It is common in various quarters of philosophy to derive philosophically significant conclusions from theories of reference. In this paper, we argue that philosophers should give up on such ‘arguments from reference.’ Intuitions play a central role in establishing theories of reference, and recent cross-cultural work suggests that intuitions about reference vary across cultures and between individuals within a culture (Machery et al. 2004). We argue that accommodating this variation within a theory of reference undermines arguments from reference.

Interest in theories of reference is not limited to the philosophy of language. In fact, assumptions about theories of reference figure crucially in nearly every corner of philosophy, including the philosophy of mind, the philosophy of science, the philosophy of race, and meta-ethics, and it is widely agreed that identifying a correct theory would have far-reaching philosophical implications. In what follows, we focus on arguments that derive philosophically significant conclusions from the assumption of one or another theory of reference—what we call ‘arguments from reference.’ We review a recent empirical challenge to the

* We would like to thank Elizabeth Harman, Jonathan Weinberg, and Wayne Wu for helpful comments on previous drafts.

project of finding a correct theory of reference (Machery et al. 2004). At the core of that challenge are data that suggest strong variation in the intuitions used to find the correct theory of reference. We consider several ways that theorists of reference might accommodate this variation in intuitions about reference, and we show that arguments from reference are undermined on all of these options.

Here is how we proceed. In Section 1, we consider the structure of arguments from reference, and review a number of projects in several areas of philosophy that employ such arguments. Then, in Section 2, we show that while intuitions about reference are central in the philosophy of language for finding the correct theory of reference, the recent empirical work of Machery and colleagues suggests that intuitions about reference vary both within and across cultures. In Section 3, we take the variation suggested by this empirical work for granted and consider its implications for arguments from reference. We conclude that arguments from reference ought to be relinquished.

1. Arguments from Reference

Arguments from reference are common in projects throughout philosophy. These arguments can be analyzed into three stages. In the first, philosophers implicitly or explicitly adopt a substantive theory of the reference of a term t (or of a class of terms T , such as theoretical terms).¹ In the second stage, they claim that the reference of t or of members of T has some specific properties. For instance, in some arguments from reference, philosophers argue that the reference relation obtains or fails to obtain—that is, that t refers or fails to refer. Or, in other arguments from reference, they argue that the reference of t has changed. Finally, a philosophically significant conclusion is drawn. These conclusions include metaphysical conclusions—conclusions to the effect that the referent of t exists or does not exist (e.g. Stich 1983; Zack 1993)—and epistemological conclusions—conclusions about the nature of our knowledge about the referent of t (e.g. Boyd 1983, 1988; Kitcher 1993).

Where can we find such arguments from reference? Everywhere in philosophy, it would seem. We begin by reviewing how arguments from reference play a key role in the philosophy of mind, then we suggest

¹ We use the term “substantive” to rule out deflationary accounts of reference such as those suggested by Field (1986, 1994) and Horwich (1990).

they play a similarly important role in other areas including the philosophy of science, social theory, and metaethics.

1.1. The Philosophy of Mind: The Debate over Eliminative Materialism

Eliminativists in the philosophy of mind (Churchland 1981; Stich 1983) defend the surprising claim that the propositional attitudes like beliefs and desires that figure in the explanations of behavior offered by folk psychology literally do not exist. Eliminativists argue that these propositional attitudes are posits of a folk theory of mind that is spectacularly false in light of the emerging sciences of the mind and brain. They conclude that the posits of this theory—beliefs and desires—don't exist.

Consider how this argument fits the three steps of the arguments from reference.

(Step 1) Assumption of a substantive theory of reference: Eliminativists propose that mental state terms like 'belief' and 'desire' are defined by their role in a folk theory, namely the folk theory of mind. They assume that if these terms have referents, they must be entities that satisfy (or come close to satisfying) the relevant definitions. That is, eliminativists assume some version of a descriptivist theory of reference for mental state terms like 'belief' and 'desire.' While such theories may take a variety of forms, they typically agree on the following points:

- D1. Competent speakers associate a description with a term *t*. This description specifies a set of properties.

- D2. An object is the referent of *t* if and only if it uniquely or best satisfies the description associated with it.

In the absence of an entity that satisfies the description (or at least comes close), the term is empty.

(Step 2) Claim about reference: Eliminativists claim that the emerging scientific facts suggest that nothing satisfies the descriptions folk psychology associates with 'belief' and 'desire.' Thus, 'belief' and 'desire' do not refer.

(Step 3) Philosophically significant conclusion: Eliminativists conclude that, since 'belief' and 'desire' do not refer, beliefs and desires do not exist.

How does this eliminativist argument fare? As William Lycan (1988) has pointed out, the eliminativist conclusion follows from the falsity of folk psychology only on the assumption of some descriptivist theory of

reference (in step 1).² But descriptivist theories of reference have been sharply contested by causal-historical theories of reference, such as those defended by Kripke (1972/1980) and Putnam (1975). Like descriptivist theories, causal-historical theories of reference may take a variety of forms. However, they typically agree on the following points:

- C1. A term *t* is introduced into a linguistic community for the purpose of referring to a particular thing (e.g. a person or a property). The term continues to refer to that thing as long as its uses are linked to the thing via an appropriate causal chain of successive users: every user of the term acquired it from another user, who acquired it in turn from someone else, and so on, back to the first user who introduced the term.
- C2. Speakers may associate descriptions with terms. But after the term is introduced, the associated description does not play any role in the fixation of the referent. The referent may entirely fail to satisfy the description.

If, like Lycan, one adopts some causal-historical theory of reference, the eliminativist conclusion does not follow. Here is how Lycan makes the point:

I am entirely willing to give up fairly large chunks of our commonsensical or platitudinous theory of belief or desire (or of almost anything else) and decide that we were just wrong about a lot of things, without drawing the inference that we are no longer talking about belief or desire. To put the matter crudely, I incline away from Lewis's Carnapian and/or Rylean cluster theory of reference of theoretical terms, and toward Putnam's (1975) causal-historical theory. (Lycan 1988, 31–32)

So, by assuming a different theory of reference than the eliminativist, Lycan draws the opposite conclusion, viz. that beliefs and desires *do* exist.

Of course the simple descriptivist and causal-historical theories we sketch here do not exhaust options for specifying a substantive theory

² In this article, we take for granted that if 'belief' and 'desire' refer descriptively and if the folk theory of mind is massively erroneous, then beliefs and desires do not exist. However, it is worth noting that this inference has been contested. Bishop and Stich (1998) have argued that one needs an additional premise to get from the claim that 'belief' does not refer to the desired conclusion that beliefs do not exist. Moreover, they claim that it is not clear how any of the premises that might fill the gap could be defended. We argue that even if Bishop and Stich's (1998) challenge were met, arguments from reference *still* wouldn't work.

of reference. But the moral we want to draw is quite general: depending on the substantive theory of reference one assumes about a term or a class of terms, one can draw different metaphysical conclusions.

1.2. Other Arguments from Reference

Arguments from reference have played an important role in metaphysical debates in the philosophy of mind, but their influence is much more widespread—reaching into almost every corner of philosophy. In the remainder of this section, we note three other important debates that look to hang on arguments from reference.

Consider first the debate in the philosophy of science over scientific realism. Several influential philosophers of science have defended the surprising claim that there is no scientific progress, thereby denying a key component of scientific realism (e.g. Feyerabend 1962; Kuhn 1970). They assume that theoretical terms like ‘mass’ or ‘energy’ are defined by their role in scientific theories. Because the role of these terms is fundamentally modified during scientific revolutions, such as Einstein’s revolution in physics, these philosophers of science conclude that the reference of theoretical terms changes during scientific revolutions. They conclude that there is no scientific progress.

But this conclusion follows from claims about scientific change only on the assumption of some descriptivist theory of the reference of theoretical terms (in step 1). If the causal-historical theory is the correct theory of the reference of theoretical terms (e.g. Putnam 1975), the inference from scientific change to the non-existence of scientific progress is blocked. Thus, Richard Boyd notes the crucial role causal-historical theories of reference have played in the defense of scientific realism against various stripes of anti-realism:

The anti-realist consequences which Kuhn (and Hanson) derived from descriptivist conceptions led to the articulation by realists of alternative theories of reference. Characteristically, these theories followed the lead of Kripke (... 1972) ... and Putnam (... 1975...) ... Each of them advocated a “causal” theory of reference ... It is by now pretty well accepted that some departure from analytic descriptivism, involving some causal elements, is a crucial component of a realist approach to scientific knowledge. (2002, sect 4.1)

While the moves and countermoves here can get very complex, the general point is quite simple: some important arguments regarding scientific realism and anti-realism are arguments from reference—they derive significant metaphysical and epistemological conclusions from a specific theory of reference.

Theories of reference have also played a key role in contemporary debates over the reality of race, with a number of theorists explicitly or implicitly adverting to such theories.³ Like eliminativists about propositional attitudes, race skeptics argue that race does not exist by denying that there is anything that satisfies the beliefs ordinary people hold about race (e.g. Appiah 1995; Zack 1993). As Robin Andreasen (2000, S661ff) notes, however, this conclusion tacitly depends on the endorsement of some descriptivist theory of reference. Andreasen goes on to point out that other theories of reference, like causal-historical theories, allow that the referent of a term may not satisfy much of the description common sense associates with the term, and she uses this strategy to defend her own account of race against the challenge of the race-eliminativists. Again we see that different metaphysical conclusions depend crucially upon the particular assumptions made about the theory of reference.

Finally, consider an example of the role that theories of reference have played in recent ethical theorizing. Robert Merrihew Adams (1979) and Richard Boyd (1988) have proposed that our ethical knowledge is similar to our scientific knowledge in some key respects: our knowledge about the nature of the good is a posteriori, it can progress, and our commonsensical ethical beliefs might turn out to be prejudices. They derive these surprising conclusions from the extension of the causal-historical theory of reference from scientific terms to moral terms. If moral terms, for instance ‘good,’ refer in a causal-historical manner, then it is an empirical question what the referents of these moral terms, for instance, the property *good*, are. Thus, a significant epistemological position in ethics is derived from a specific theory about the reference of a class of terms—i.e., moral terms.

2. Semantics, Cross-Cultural Style

2.1. *Finding the Correct Theory of Reference: The Method of Cases*

The arguments sketched in the previous section all hinge on what the correct theory of reference is. But how do we know which theory of reference is correct? Unfortunately, philosophers of language have rarely addressed this methodological issue explicitly. However, it is clear from the arguments for and against specific theories of reference that the correct theory of reference for a term (or for a class of terms such as proper names) is commonly thought to be constrained by our intuitions about the reference of this term (or about the reference of

³ E.g., Andreasen 1998, 2000; Appiah 1996; Glasgow 2003; Kitcher 1999; Mills 1998; Zack 1993, 2002; for a critical discussion, see Mallon 2006.

the members of a given class of terms) in actual and fictional cases. For instance, according to Evans (1973), people have the intuition that nowadays the proper name 'Madagascar' refers to the large island near the south of Africa, even when they learn that the term was historically used to refer to a region on the mainland of Africa.

We propose that to find the correct theory of reference, philosophers of language are committed to using what is sometimes called 'the method of cases':

The method of cases: The correct theory of reference for a class of terms T is the theory which is best supported by the intuitions competent users of T have about the reference of members of T across actual and possible cases.

The method of cases has played a crucial role in the challenge posed to traditional descriptivist theories of reference by the causal-historical theories championed by Kripke and others. Indeed, Kripke's masterstroke was to propose a number of cases that elicited widely shared intuitions that were inconsistent with traditional descriptivist theories (Kripke 1972/1980).⁴

It will be useful to briefly review one of Kripke's most widely discussed cases involving the reference of proper names. In this case ('the Gödel case'), Kripke imagines a scenario in which a name is widely associated with a description that is false of the original bearer of that name a , but true of some other person, b . Because descriptivist theories of reference hold that a term refers to the thing that (uniquely or best) satisfies the description associated with the term, a descriptivist theory of reference would seem to hold that the name in Kripke's example refers to b , the satisfier of the description. But, Kripke maintains, this is just wrong.

Suppose that Gödel was not in fact the author of [Gödel's] theorem. A man called 'Schmidt'... actually did the work in question. His friend Gödel somehow got hold of the manuscript and it was thereafter attributed to Gödel. On the [descriptivist] view ... when our ordinary man uses the name 'Gödel,' he really means to refer to Schmidt, because Schmidt is the unique person satisfying the description 'the man who discovered the incompleteness of arithmetic'. ... But it seems we are not. We simply are not. (Kripke 1972/1980, 83–84).

⁴ In his reply to Kripke, Evans (1973) also relies on intuitions about the reference of proper names, such as 'Madagascar.' Putnam (1973, 1975) relies on intuitions about the reference of natural kind terms such as 'gold.' See also Schwartz (1978, 1980) and Devitt (1981).

In contrast, a causal-historical theory of the reference of proper names is consistent with the intuition that the name continues to refer to its original bearer *a*, because *a* is the person causally-historically linked with contemporary uses of the name. Many contemporary descriptivists allow that these intuitions have falsified traditional forms of descriptivism and try to accommodate these intuitions within more sophisticated descriptivist theories (Evans 1973; Jackson 1998b).

A plausible justification for the method of cases might be the assumption that language users have an implicit theory of reference that produces intuitions about reference. The project for reference theorists can then be conceived by analogy with the Chomskyan project in linguistics. Philosophers of language use people's intuitions about reference to reconstruct the implicit theory that is part of each speaker's cognitive endowment (Segal 2001).

Despite agreement on the method of cases and agreement on many of the intuitions about the cases used in the philosophical literature, a consensus on the correct theory of reference remains elusive.⁵ As this paper goes to press, intuitions about cases continue to underdetermine the selection of a correct theory of reference. However, our argument is based not on the diversity of theories that may be constructed around the same set of intuitions, but on the possibility of variation in the intuitions themselves. Philosophers interested in reference are well aware that intuitions might differ for different classes of terms (e.g. natural kind terms, names, artifact terms, etc.). For instance, cases involving natural kind terms might elicit causal-historical intuitions, while cases involving artifact terms might elicit descriptivist intuitions (e.g. Schwartz 1978, 1980). Because intuitions might differ for different classes of terms, philosophers interested in reference are willing to allow different accounts of reference for different classes of terms (e.g. Devitt and Sterelny 1999). However, as we have previously noted (Machery et al. 2004), the possibility of diverse intuitions about the same cases, for instance about Kripke's Gödel case, plays little role in the contemporary search for a theory of reference. Indeed, contemporary participants in semantic debates seem to assume that the relevant intuitions about cases are more or less universal, and that exceptions can be explained away. As to why we should believe this, little is said. Just what happens to the search for a theory of reference (and the arguments that depend on it) if this assumption is mistaken, is a subject we turn to below. First, however, we consider a fledgling empirical

⁵ Recanati 1993; Abbott 1997, 1999, 2002; Jackson 1998a, b; Devitt and Sterelny 1999; Geurts 1997, 2002; Garcia-Carpintero 2000; Segal 2001; Soames 2002; Reimer 2002, 2004; Jeshion 2004.

program that casts doubt on this assumption, suggesting systematic diversity in intuitions about reference.

2.2. *Cultural Variation in Intuitions about Reference*

Recent work in cultural psychology and empirical philosophy has suggested the existence of real and systematic differences in philosophical intuitions. In an important series of experiments, Richard Nisbett and colleagues found large and systematic differences between people in East Asian cultures and people in Western cultures on a number of basic cognitive processes including perception, attention and memory. This burgeoning research program has also discovered group differences in describing, predicting and explaining events, in categorization of objects and in belief revision in the face of new arguments and evidence (for review, see Nisbett and Miyamoto 2001; Nisbett 2003; Nisbett and Miyamoto 2005). These findings suggest a dramatic role for culture in shaping human cognition. Inspired by this research program, Weinberg et al. (2001) decided to explore cultural differences in intuitions about cases drawn from philosophical epistemology. These cases were designed to elicit intuitions about the appropriate application of the concept of knowledge, and Weinberg et al. found that there are indeed systematic cross-cultural differences in epistemic intuitions.

The success of Nisbett's research program and Weinberg et al.'s results suggested that other philosophical intuitions, including intuitions about reference, might also admit of systematic cultural differences. In an earlier paper (Machery et al. 2004), we set out to explore this possibility. We began by noting that existing cross-cultural work suggests that East-Asians' categorization judgments depend heavily on similarity while Westerners are more inclined to focus on causation in classification (Watanabe 1998, 1999; Norenzayan et al. 2002), and we hypothesized that this emphasis on causation might make Westerners more likely to rely on causation in linking terms with their referents, favoring the sort of intuitions that Kripke used in support of his causal-historical theory. In fact, this is just what we found.

We constructed a set of vignettes suggested by Kripke's Gödel case, discussed above (Kripke 1972/1980, 93–92). The vignettes were presented in English to American and Chinese subjects.⁶ One of the vignettes was closely modeled on Kripke's own Gödel case (see Machery et al. 2004 for more details on the experiment):

⁶ The Chinese subjects were students at the University of Hong Kong where the language of instruction is English; all participants were fluent speakers of English.

Suppose that John has learned in college that Gödel is the man who proved an important mathematical theorem, called the incompleteness of arithmetic. John is quite good at mathematics and he can give an accurate statement of the incompleteness theorem, which he attributes to Gödel as the discoverer. But this is the only thing that he has heard about Gödel. Now suppose that Gödel was not the author of this theorem. A man called “Schmidt” whose body was found in Vienna under mysterious circumstances many years ago, actually did the work in question. His friend Gödel somehow got hold of the manuscript and claimed credit for the work, which was thereafter attributed to Gödel. Thus he has been known as the man who proved the incompleteness of arithmetic. Most people who have heard the name ‘Gödel’ are like John; the claim that Gödel discovered the incompleteness theorem is the only thing they have ever heard about Gödel. When John uses the name ‘Gödel,’ is he talking about:

(A) the person who really discovered the incompleteness of arithmetic?

or

(B) the person who got hold of the manuscript and claimed credit for the work?

In two separate studies using four different vignettes, we found that Americans were more likely than Chinese to give causal-historical responses. Thus, we found that probes modeled on Kripke’s Gödel case (including one that used Kripke’s own words) elicit culturally variable intuitions. As we had predicted, Chinese participants tended to have descriptivist intuitions, while Americans tended to have Kripkean intuitions.

It is important to note that we found significant intra-cultural differences as well. While for each vignette a majority of Americans gave causal-historical responses, in each case a sizable minority of the population (as high as 45% in one case) gave descriptivist responses. Similarly for the Chinese population, for each vignette, a majority of Chinese participants gave descriptivist responses, but in each case a sizable minority (in some cases over 30%) gave causal-historical responses.

2.3. Significance of these Findings

We have no illusions that our experiments are the final empirical word on the issue. This is a newly emerging type of research, and obviously it is too early to draw any definite conclusion about the variation of intuitions about reference.⁷ Nonetheless, our results already point toward some significant conclusions. Our findings suggest that some well-known semantic intuitions about proper names vary within and across cultures. If that conclusion is judged to be premature, at least our findings show that philosophers cannot simply *assume* that intuitions about reference are universal. Intra-cultural and cross-cultural variation in intuitions about reference is a live possibility.

Since intuitions about reference are used to support theories of reference, proponents of the arguments from reference should be eager to explore what consequences would follow for these arguments if intuitions about reference do indeed vary both within and across cultures.

3. Against Arguments from Reference

While the empirical results reviewed in Section 2 are still preliminary, they constitute a strong *prima facie* case that intuitions about reference used to construct theories of reference might vary from culture to culture and person to person. In the remainder of this paper, we will *assume* that such variation does exist, and we will explore its implications both for the theory of reference and for arguments from reference. We consider a number of ways that theorists of reference might accommodate this variation in intuitions about reference, but argue that none of them salvages arguments from reference.

3.1. Giving up on Substantive Theories of Reference

A first response to the diversity in intuitions about reference is to give up on the idea that the search for a substantive theory of reference is a viable enterprise.⁸ While such an abandonment might take a variety of forms, for present purposes we need only note that if there is no

⁷ Further empirical investigation should explore whether semantic intuitions about proper names elicited by other cases and semantic intuitions about other types of words, particularly natural kind terms, also vary across cultures. It is commonly assumed that the semantics of proper names and the semantics of natural kind terms are similar, in contrast with the semantics of other predicates such as artifact terms (e.g. Schwartz 1978; Devitt and Sterelny 1999). Furthermore, natural kind terms and proper names elicit similar intuitions about their reference from philosophers (Kripke 1972/1980; Putnam 1973, 1975).

⁸ This is a strategy that would be endorsed by deflationists like Field (1986, 1994) and Horwich (1990), albeit for other reasons.

correct substantive theory of reference, there can be no arguments from reference.

3.2. *Downplaying the Method of Cases*

A second response to variation in intuitions about reference is to downplay the role of intuitions in choosing a theory of reference in favor of other theoretical considerations.⁹ Philosophers could adopt a theory of reference on the grounds that it has some desirable philosophical consequences, independent of their intuitions about reference. For instance, a proponent of scientific realism might endorse a causal-historical theory of reference, on the grounds that it gives support to scientific realism.

But however well this approach to reference works, it is not our concern here.¹⁰ Arguments from reference *begin* with a theory of reference that is *independently* motivated, and *proceed* to philosophically significant conclusions. And the dominant way of independently motivating a theory of reference among philosophers of language is by appeal to intuitions about whether or how terms refer in various cases. Arguments that assume a theory of reference that is not independently motivated are not arguments from reference, as we use the term.

One might reply that besides the method of cases and the appeal to the philosophical consequences of theories of reference, some other considerations might be used to justify these theories. In reply, we first note that this move would involve breaking with the dominant tradition of employing the method of cases in the philosophy of language. More important, we have no idea what other considerations philosophers of language might appeal to. Thus, in the absence of concrete suggestions, we remain skeptical of the proposal to downplay the role of intuitions in choosing a theory of reference.

3.3. *Endorsing Referential Pluralism*

Yet another option for accommodating the variation in intuitions about reference in order to construct a theory of reference is to hold that differences in intuitions about the reference of a word *t* (or a class of words *T*) indicate that *t* (or every member of *T*) refers differently for different groups, just as different grammatical intuitions among members of different linguistic groups indicate different grammars for their languages.

⁹ Dennett (1996), Papineau (1996) and Laurence and Margolis (2003) have suggested intuitions are only one among several factors in choosing the correct theory of reference.

¹⁰ For doubts about this approach, see Mallon (2007).

In order to state this third option more precisely, we introduce the notion of an *intuition group*. An intuition group is simply a group of persons who share intuitions about a set of cases. Put in these terms, the apparent assumption of philosophers of language that intuitions about a type of case are universal amounts to the assumption that there is only one intuition group (or at least that there is only one for a given language). But the data we present above suggest that not everyone belongs to the same intuition group (because intuitions differ systematically), that intuition groups cross-cut language groups (because speakers of English can have systematically different intuitions), and that intuition groups cross-cut cultural groups (because significant minorities of both American and Chinese cultural groups have intuitions matching the majority of the other cultural group).

Like traditional philosophers of language, the referential pluralist holds that the correctness of a theory of reference for a class of terms is determined by the intuitions of the appropriate intuition group, but unlike them, the referential pluralist allows that there may be more than one intuition group. Referential pluralists thus hold the following:

The pluralist method of cases: The correct theory of reference for a class of terms *T* employed by members of intuition group *G* is the theory which is best supported by the intuitions that competent members of *G* have about the reference of members of *T* across actual and possible cases.

In the remainder of this section, we argue that referential pluralism is an implausible way of accommodating the variation in intuitions about reference.

3.3.1. *Are Referential Intuitions Evidence for Reference?*

Referential pluralism assumes that speakers' intuitions about reference provide evidence about reference. This assumption would be quite plausible if variation in intuitions about reference mapped onto variation in languages or dialects. Similarly, we are confident that intuitions about the grammaticality of sentences provide evidence about grammatical properties, because variation in these intuitions map onto variation in languages or in dialects. People who have different intuitions about the grammaticality of sentences tend to speak different languages or different dialects.¹¹ The same is true of other linguistic intuitions, such as intuitions about synonymy, antonymy or polysemy. Now,

¹¹ There are some tricky issues involved in identifying languages and dialects, but for present purposes they can safely be ignored.

consider the situation in which people who evidently speak the same dialect have different intuitions about the grammaticality of sentences. Plausibly, this would cast doubt on the assumption that intuitions about grammaticality provide reliable evidence about the grammatical properties of the dialect they speak. At the very least, syntacticians would be hard-pressed to find a justification for this assumption. Our data seem to show that two individuals can belong to two distinct intuition groups despite evidently speaking the same dialect (because they speak the same language, belong to the same culture and have much the same socio-economic status). Faced with this variation, it is very tempting to abandon the assumption that intuitions about reference provide evidence about reference all together. Instead, one might, for example, propose that a speaker's intuitions about reference are caused by a variety of factors that turn out to have nothing do with reference, including her culture and perhaps her philosophical commitments (Stich 1996, 85, fn. 35). But referential pluralism is committed to the method of cases, and so must make this assumption, despite the fact that it is an assumption that is in dire need of justification.

3.3.2. *From Referential Pluralism to Referential Relativism*

For argument's sake, let's grant that proponents of arguments from reference can justify the assumption that intuitions about reference provide evidence about reference, even though variation in intuitions about reference does not map onto variation in dialects or languages. Then, according to the referential pluralist, the correct theory of reference for interpreting a person's utterance is the theory of reference supported by the intuitions of the intuition group to which the person belongs.

How does referential pluralism affect arguments from reference? Consider the argument for the elimination of propositional attitudes.¹² Suppose that members of Group A have descriptivist intuitions about predicates such as 'belief,' while members of Group B have causal-historical intuitions. The referential pluralist concludes that predicates such as 'belief' refer differently when they are used by members of Group A and by members of Group B. They refer descriptively when used by members of Group A, and they refer in a causal-historical manner when used by members of Group B. Suppose also that the description associated with 'belief' is derived from a theory that is massively erroneous and thus that the description is satisfied by nothing. Because 'belief' refers descriptively when used by members of Group A

¹² The same line of argument applies, *mutatis mutandis*, to the other arguments from reference.

and because the description associated with ‘belief’ is satisfied by nothing, when a member of Group A says ‘Beliefs do not exist,’ the referential pluralist concludes that what this person says is true. However, because ‘belief’ refers causal-historically when used by a member of Group B, when a member of B says ‘Beliefs do exist,’ the referential pluralist concludes that what this person says is also true. But these two conclusions seem to flatly contradict one another. Surprisingly, referential pluralism seems to lead to contradictions.

The referential pluralist need not be daunted, however, for he can simply argue that there is no contradiction when a member of A says truly ‘Beliefs do not exist’ and when a member of B says, also truly, ‘Beliefs do exist.’ Consider the following situation. John and Jean are talking to each other by phone. John is in New York, while Jean is in Paris. It’s noon in New York and 6 pm in Paris. John says truly ‘It’s noon,’ while Jean says truly ‘It’s not noon.’ It’s raining in New York, but not in Paris. John says truly ‘It’s raining,’ while Jean says truly ‘It’s not raining.’ John and Jean are not contradicting each other, and it is clear to them that they are not. For the truth of what John and Jean say when they say ‘It’s noon’ and ‘It’s raining’ depends upon the contexts of use of the two sentences ‘It’s noon’ and ‘It’s raining.’ And the context of use is not the same for Jean’s utterances and for John’s utterances. The context of use for Jean’s utterance of ‘It’s noon’ and ‘It’s raining’ involves the weather and the time *in Paris* when the phone conversation takes place, while the context of use for John’s utterance of ‘It’s noon’ and ‘It’s raining’ involves the weather and the time *in New York* when the phone conversation takes place. Relativization to a context of use is extremely common in natural languages.

The referential pluralist might simply argue that a similar phenomenon is going on when our imaginary member of Group A says truly ‘Beliefs do not exist,’ while our imaginary member of Group B says, also truly, ‘Beliefs do exist.’ The context of use for our member of A’s utterance of ‘Beliefs do not exist’ and the context of use for our member of B’s utterance of ‘Beliefs do exist’ are not the same. When the member of A says ‘Beliefs do not exist,’ the context of use includes how terms such as ‘belief’ refer when they are used by members of Group A, which itself depends, according to the referential pluralist, on what kind of intuitions about reference members of Group A have. When a member of B says ‘Beliefs do exist,’ the context of use includes how terms such as ‘belief’ refer when they are used by members of Group B, which in turn depends on what kind of intuitions about reference members of Group B have. The truth of what a speaker says when she utters ‘Beliefs do exist’ or ‘Beliefs do not exist’

is relativized to the intuition group of this speaker. And similarly for every utterance. Like John and Jean, members of Group A and Group B do not contradict each other, for their utterances are evaluated according to the distinct contexts appropriate to their respective groups.

3.3.3. *Relativizing Assertions in Philosophy and Beyond*

All this has very far-reaching implications for how we carry on discourse, ordinary and philosophical. To see why, recall that while our data suggest systematic differences in philosophical intuitions between Hong Kong and U.S. students, they also suggest a high degree of *intra-cultural variation*. So, while our data suggest that culture has a systematic effect on the intuition group to which one belongs, they also indicate that intuition groups cross-cut cultural groups. Call this the *ubiquity of variation*. As we have seen, our referential pluralist argues that because individuals may belong to different intuition groups, we should relativize their utterances to the theories of reference appropriate to their intuition groups. And while this allows the referential pluralist to resolve apparent contradictions, when combined with the ubiquity of variation, referential pluralism entails that conclusions of the arguments within cultural groups must be similarly relativized. In philosophy, this means that arguments over the existence of beliefs (or the existence of races, the progress of science, the nature of our epistemic access to moral properties and so on for the conclusions of every other argument from reference) have to be relativized. Thus, suppose that Kuhn and Feyerabend were right that the role of theoretical terms is fundamentally changed by scientific revolutions. Philosopher A might truly say that ‘science progresses’ because she has causal-historical intuitions about theoretical terms and Philosopher B might truly say that ‘science does not progress’ because she has descriptivist intuitions, and according to the referential pluralist, the two philosophers *would not disagree*. Similar considerations apply to the case of philosophical agreement. When asserting or denying that beliefs exist or that science progresses (or the conclusion of any other argument from reference), philosophers A and B *would agree or disagree only if* they belonged to the same intuition group. If, for example, Churchland and Fodor both belonged to an intuition group whose intuitions about reference support a descriptivist theory of reference for mental state terms, then they might genuinely disagree about whether beliefs exist.

We take it that these conclusions are very surprising and would involve a very substantial revision of philosophical methodology. For they suggest that philosophical disagreement and agreement among

even speakers of the same language, who belong to the same culture, have the same socio-economic status, and even attended the same graduate program in philosophy may be illusory if the speakers have different intuitions about how terms refer in actual and fictional cases. Moreover, this conclusion is not limited to philosophical debate, but appears to extend to all discourse. And there is worse to come. For, as we now argue, when combined with an additional (we think very plausible) premise, referential pluralism has conclusions that are far more absurd than any we have suggested so far.

3.3.4. *The Uncertain Membership of Intuition Groups*

This additional premise is, simply, that we do not really know to which intuition groups any of us belongs. Before arguing for it, we point out that this premise, combined with referential pluralism and the ubiquity of variation, leads to the *prima facie* absurd conclusion that we do not really know, of any of our discourse, whether it agrees or disagrees with the discourse of anyone else. Consider how this applies to a philosophical case. Suppose that there are a number of different intuition groups and that it is unclear what intuition group philosophers A and B belong to. Then, if referential pluralism is correct, when philosopher A says 'Beliefs exist' and philosopher B says 'Beliefs do not exist,' it is unclear whether philosophers A and B disagree or whether they are speaking at cross-purposes. Thus, if it is unclear which intuition group we belong to, far from salvaging the arguments from reference, referential pluralism leads to the absurd conclusion that we simply have no idea when proponents of these arguments agree, when they disagree and when they talk past each other. And again, this conclusion is not limited to debates in philosophy, but rather extends quite generally to discourse about anything at all!

So, the question is pressing: Do we know which intuition group each of us belongs to? We think the answer is no: Our knowledge of the intuitions each of us has about the relevant cases is very far from complete. A series of considerations suggest why this is the case:

1. Every speaker must be considered because of the ubiquity of variation.

As we have already mentioned, intuitions even about the simple cases we have tested vary within cultural and language groups. It follows that one cannot assume that people who share a language and a culture are members of the same intuition group. This means every speaker must be considered individually!

2. Because lots of cases are relevant, intuition groups must be fine-grained.

Above we said an intuition group was a group of persons who shared intuitions about a set of cases. Since we are interested in theories of reference, the relevant intuitions are all those judgments about cases that decide between distinct theories of reference. For example, Machery et al. (2004) considered hypothetical cases that have been used to test whether descriptivist or causal-historical theories of reference better comport with our intuitions. But this is only the tip of the iceberg of relevant intuitions for there are many varieties of descriptivist theories, causal-historical theories, and blends of such theories that might be best supported by different sets of intuitions.

Moreover, the schematic theories of reference we have discussed must be accompanied by *auxiliary assumptions* in order to apply the theories, assumptions that themselves are typically justified in part by appeal to intuitions. For example, causal-historical theorists must decide exactly *what individual or thing is picked out by the historical introduction of a term, and whether any “switching” to another individual or thing has occurred along the way* (Evans 1973). And descriptivists must decide on *what the reference-fixing description is, and how closely a thing must satisfy it* in order to qualify as a referent.

But do intuitions that are needed to determine the right auxiliary assumptions really vary in ways similar to the intuitions tested by Machery et al. (2004)? While evidence is limited, there is every indication that the answer is yes. Consider, for example, the extensive literature on what knowledge is. Much of this literature employs the method of cases in order to elicit judgments about whether there is knowledge in various counterfactual cases in the hopes of arriving at an understanding of knowledge. In recent decades, a vast amount of work has been driven by responses to Edmund Gettier’s (1963) thought experiments. Famously, Gettier suggested that certain sorts of justified true beliefs (for example, those that were accidentally true) did not intuitively count as knowledge, and so the search has gone on for additional conditions that would allow a statement of the necessary and sufficient conditions of knowledge. Other work on knowledge has also employed the method of cases. Keith Lehrer (1990) employed the ‘Truetemp’ case to explore intuitions regarding the internalism/externalism debate—a debate about whether factors external to a subject’s introspective access are relevant to the application of epistemically normative concepts like *justification*.

As we mentioned above, Weinberg et al. have found variation in judgments about whether cases count as instances of knowledge. In particular, they found variation regarding *both* Gettier cases and

Truetemp cases, and they found it *both within and across cultures*. This suggests that the ubiquity of variation may well extend to judgments elicited to determine the right auxiliary assumptions to make. For example, these results suggest there may well be variation in judgments about what the right description to associate with the term ‘knowledge’ is.

Because it takes a great many assumptions to determine and apply a theory of reference, and because the intuitive judgments required to decide which of these assumptions is correct for a person likely vary within and across cultures, it follows that the sorts of intuition groups into which people must be divided in order to determine full-fledged theories of reference will be very fine-grained.

3. Numerous cases must be considered for each speaker.

It seems to follow from 1 and 2 that assigning every speaker to the correct intuition group would require a vast amount of careful work. Numerous cases, actual and fictional, would have to be considered for every speaker. Nothing of this kind has ever been attempted.

4. Explicit views are not a good guide to intuition group membership.

It remains open to a referential pluralist to insist that at least philosophers have considered a wide range of actual and fictional cases in deciding what their own intuitions are. And, as we noted above, some philosophers are explicitly committed to descriptivist or to causal-historical theories of reference for specific classes of words (e.g., proper names). It therefore seems open to the referential pluralist to contend that a philosopher’s explicit commitments are good evidence about the intuition group she belongs to. For instance, if a philosopher endorses a descriptivist theory of the reference of theoretical terms, this could be taken to be evidence that she has descriptivist intuitions about the reference of theoretical terms. If so, the referential pluralist could insist that, at least for philosophers, we can have some idea about whether they agree and disagree.

Call this assumption that a philosopher’s explicit views are a guide to her intuition group membership *the limpidity assumption*. We think the limpidity assumption is eminently questionable. Remember that prior to Kripke and Putnam, pretty much everyone was a descriptivist. But as soon as Kripke proposed his famous cases, including the Gödel case, many philosophers discovered that they had intuitions that were incompatible with descriptivist theories of reference and adopted some version of the causal-historical theory of reference. It is overwhelmingly plausible that the distribution of intuitions elicited by these cases would

have been much the same ten or fifty years before Kripke and Putnam. That is, had Anglo-American philosophers been asked about the Gödel case in 1931 rather than in 1971, many would have had causal-historical intuitions. But almost nobody was committed to the causal-historical theory of reference before Kripke and Putnam. They simply hadn't considered a relevant case.

Indeed, many philosophers of language are implicitly committed to denying the limpidity assumption. In debates about what the right theory of reference is, philosophers of language often contend that their opponents have failed to consider the cases that would have elicited intuitions inconsistent with the theory held by their opponents. That is, they contend that their opponents have ignored some of their own intuitions, by failing to consider the whole gamut of actual and fictional cases. Thus, philosophers of language take for granted that their opponents might hold a theory of reference that is not supported by their opponents' intuitions. This is tantamount to denying the limpidity assumption.

Because philosophers represent the most plausible candidates for persons whose explicit views about reference track their intuition group membership, the failure of the limpidity assumption for philosophers is very bad news for any suggestion that each of us knows which intuition group she belongs to. Furthermore, while the limpidity assumption is implausible for philosophers, it is of no use at all in considering discourse among ordinary people who typically have neither considered their intuitions about actual and fictional cases nor explicitly endorsed a theory of reference.

Together, these considerations strongly support the view that we simply do not know to which intuition group any of us belongs. And that completes our *reductio*, for since it is unclear which intuition group each of us belongs to, and because we may well belong to different groups (even if we share, e.g., a language, culture, and socio-economic group), referential pluralism leads to the absurd conclusion that no one knows when proponents of arguments from reference agree, when they disagree, and when they speak at cross-purposes. This conclusion is so bizarre that we ought to abandon the referential pluralism that leads to it.

4. Metaphysics without Arguments from Reference

Arguments from reference are a basic philosophical currency, used to establish philosophically significant conclusions in a variety of areas. But if we are right, arguments from reference have to be rejected, given the plausible (but, by no means, conclusively established) assumption

that the intuitions many take to be important in finding the correct theory of reference are themselves diverse. The three ways we have considered of accommodating the diversity in intuitions about reference in order to build a theory of reference undermine these arguments from reference.

- (i) If philosophers give up on substantive theories of reference, then, obviously, they ought to give up on arguments from reference, since arguments from reference begin with a substantive theory of reference.
- (ii) If philosophers endorse a theory of reference because it gives support to their metaphysical commitments, then they do not need arguments from reference. And absent concrete proposals for the justification of theories of reference on the basis of something other than the method of cases, we remain skeptical of this response to the variation in intuitions about reference.
- (iii) If philosophers endorse referential pluralism, then they must justify the assumption that intuitions about reference provide evidence about reference, although variation in these intuitions do not map onto variation in languages or dialects. Supposing this can be done, philosophers must accommodate a relativization of the conclusion of the arguments from reference to intuition groups that may cross cut languages, cultural groups, and so forth. This might be a way to accommodate the variation in intuitions about reference provided that we know which intuition group a person belongs to. However, we do not know. Without this knowledge referential pluralism leads to the absurd view that we do not know when people agree, when they disagree and when they speak at cross-purposes.

So philosophers must choose. They can abandon arguments from reference. Or they can hold on to the hope that, despite evidence to the contrary, variation in intuitions about reference does not really exist. One of these is clearly a safer bet.

References

- Abbott, B. 1997. A note on the nature of 'water.' *Mind* 106: 311–319.
— 1999. Water = H₂O. *Mind* 108: 145–148.

- 2002. Definiteness and proper names: Some bad news for the description theory. *Journal of Semantics* 19: 191–201.
- Adams, R. M. 1979. Divine command ethics modified again. *Journal of Religious Ethics* 7(1): 66–79.
- Andreasen, R. O. 1998. A new perspective on the race debate. *British Journal for the Philosophy of Science* 49: 199–225.
- 2000. Race: Biological reality or social construct? *Philosophy of Science* 67: S653–S666.
- Appiah, K. A. 1995. The Uncompleted Argument: Du Bois and the Illusion of Race. In L. A. Bell and D. Blumenfeld (eds.), *Overcoming Racism and Sexism*. Lanham, MD: Rowman and Littlefield, 59–78.
- 1996. Race, culture, identity: Misunderstood connections. In K. A. Appiah and A. Gutmann (eds.), *Color Conscious: The Political Morality of Race*. Princeton, NJ: Princeton University Press, 30–105.
- Bishop, M. and Stich, S. P. 1998. The flight to reference, or how not to make progress in the philosophy of science. *Philosophy of Science* 65: 33–49.
- Boyd, R. 1983. On the current status of the issue of scientific realism. *Erkenntnis* 19: 45–90.
- 1988. How to be a moral realist. In G. Sayre-McCord (ed.), *Essays on Moral Realism*. Ithaca: Cornell University Press.
- 2002. Scientific Realism. *The Stanford Encyclopedia of Philosophy (Summer 2002 Edition)*, Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/sum2002/entries/scientific-realism/>.
- 2003a. Finite beings, finite goods: The semantics, metaphysics and ethics of naturalist consequentialism, part I. *Philosophy and Phenomenological Research* 66: 505–553.
- 2003b. Finite beings, finite goods: The semantics, metaphysics and ethics of naturalist consequentialism, part II. *Philosophy and Phenomenological Research* 67: 24–47.
- Churchland, P. M. 1981. Eliminative materialism and the propositional attitudes. *Journal of Philosophy* 78: 67–90.
- Dennett, D. 1996. Cow-sharks, Magnets, and Swampman. *Mind & Language* 11(1): 76–77.
- Devitt, M. 1981. *Designation*. New York: Columbia University Press.
- and Sterelny, K. 1999. *Language and Reality: An Introduction to the Philosophy of Language* (2nd ed.). Cambridge, MA: MIT Press.
- Evans, G. 1973. The causal theory of names. *Supplementary Proceedings of the Aristotelian Society* 47: 187–208.

- Feyerabend, P. K. 1962. Explanation, reduction and empiricism. In H. Feigl and G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, vol. 3: *Scientific Explanation, Space, and Time*. Minneapolis: University of Minnesota Press, 28–97.
- Field, H. 1986. The deflationary concept of truth. In G. MacDonald and C. Wright (eds.), *Fact, Science, and Value*. Oxford: Blackwell, 55–117.
- 1994. Deflationist views of meaning and content. *Mind* 103: 249–85.
- Fodor, J. 1987. *Psychosemantics: the Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.
- Garcia-Carpintero, M. 2002. A presuppositional account of reference-fixing. *Journal of Philosophy* 97(3): 109–147.
- Geurts, B. 1997. Good news about the description theory of names. *Journal of Semantics* 14: 319–348.
- 2002. Bad news for anyone? A reply to Abbott. *Journal of Semantics* 19: 203–207.
- Glasgow, J. 2003. On the new biology of race. *The Journal of Philosophy* 100: 456–474.
- Horwich, P. 1990. *Truth*. Oxford: Blackwell.
- Jackson, F. 1998a. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford: Oxford University Press.
- 1998b. Reference and description revisited. In J. Tomberlin (ed.), *Philosophical Perspectives*, vol. 12: *Language, Mind, and Ontology*. Oxford: Blackwell, 201–218.
- Jeshion, R. 2004. Descriptive descriptive names. In A. Bezuidenhout and M. Reimer (eds.), 2004.
- Kitcher, P. 1993. *The Advancement of Science*. Oxford: Oxford University Press.
- 1999. Race, ethnicity, biology, culture. In Leonard Harris (ed.), *Racism*. New York: Humanity Books, 87–120.
- Kripke, S. 1972/1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Kuhn, T. S. 1970. *The Structure of Scientific Revolutions*, 2nd edition. Chicago: University of Chicago Press.
- Laurence, S. and Margolis, E. 2003. Concepts and conceptual analysis. *Philosophy and Phenomenological Research* 67: 253–282.
- Lycan, W. 1988. *Judgement and Justification*. Cambridge: Cambridge University Press.
- Machery, E., Mallon, R., Nichols, S. and Stich, S. 2004. Semantics cross-cultural style. *Cognition* 92: B1–B12.
- Mallon, R. 2006. Race: Normative, not metaphysical or semantic. *Ethics* 116(3): 525–551.

- 2007. Arguments from reference and the worry about dependence. *Midwest Studies in Philosophy* 31(1): 160–183.
- Mills, C. 1998. *Blackness Visible: Essays on Philosophy and Race*. Ithaca: Cornell University Press.
- Nisbett, R. E. 2003. *The Geography of Thought: How Asians and Westerners Think Differently ... and Why*. New York: Free Press.
- Nisbett, R. E. and Miyamoto, Y. 2005. The influence of culture: holistic versus analytic perception. *Trends in Cognitive Sciences* 9(10): 467–473.
- 2001. Culture and systems of thought: Holistic vs. analytic cognition. *Psychological Review* 108: 291–310.
- Norenzayan, A., Smith, E. and Kim, B. 2002. Cultural preferences for formal versus intuitive reasoning. *Cognitive Science* 26: 653–684.
- Papineau, D. 1996. Doubtful intuitions. *Mind & Language* 11(1): 130–2.
- Putnam, H. 1973. Explanation and reference. In G. Pearce and P. Maynard (eds.), *Conceptual Change*. Dordrecht: Reidel.
- 1975. The meaning of ‘meaning’. In H. Putnam, *Mind, Language and Reality*. Cambridge: Cambridge University Press.
- Recanati, F. 1993. *Direct Reference*. Oxford: Basil Blackwell.
- Reimer, M. 2002. Ordinary proper names. In G. Preyer and G. Peter (eds.), *On Logical Form and Language*, Oxford: Oxford University Press.
- 2004. Descriptive names. In A. Bezuidenhout and M. Reimer (eds.), 2004.
- Schwartz, S. P. 1978. Putnam on artifacts. *Philosophical Review* 87: 566–574.
- 1980. Natural kind terms. *Cognition* 7: 301–315.
- Segal, G. 2001. Two theories of names. *Mind & Language* 16(5): 547–563.
- Soames, S. 2002. *Beyond Necessity: the Unfinished Semantic Agenda of Naming and Necessity*. New York: Oxford University Press.
- Stich, S. P. 1983. *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press.
- 1996. *Deconstructing the Mind*. Oxford: Oxford University Press.
- Taylor, P. 2000. Appiah’s uncompleted argument: DuBois and the reality of race. *Social Theory and Practice* 26: 1 103–28.
- 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.

- 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., Nichols, S. and Stich, S. 2001. Normativity and Epistemic Intuitions. *Philosophical Topics* 29(1/2): 429–459.
- Zack, N. 1993. *Race and Mixed Race*. Philadelphia: Temple University Press.
- 2002. *Philosophy of Science and Race*. New York: Routledge.