Denotation and Discovery

Gary Ebbs

PRACTICAL JUDGMENTS OF SAMENESS OF DENOTATION

Speakers of the same natural language typically take each other's words at face value. If you show me a ring and say, "This ring is gold," I'll take you to have asserted that the ring is gold, without first thinking about whether this way of taking your words is justified. When I take your word 'gold' at face value in this way, I take it to be true of the ring just in case my word 'gold' is true of the ring, hence just in case the ring is gold. This way of taking your word is like a judgment in that I may revise it in light of new information, but unlike a judgment in that it is unreflective and may never come up for review. I call it a practical judgment of sameness of denotation.

Our practice of taking other English speaker's words at face value extends across time, from moment to moment and, in some cases, for centuries. For example, if we learn that in 1650 a jeweler showed John Locke a ring and said, "This ring is gold," we'll take the jeweler to have asserted that the ring is gold. When we take the jeweler's word 'gold' at face value in this way, we in effect take it to be true of an object $x$ just in case our word 'gold' is true of $x$, hence just in case $x$ is gold. We thereby make what I call a practical judgment of sameness of denotation across time.¹

We express such judgments when we take ourselves to have made a discovery or to agree or disagree with other speakers. For instance, we take for granted that when chemists first accepted the sentence "Gold is the element with atomic number 79," they did not thereby introduce a new denotation for the word 'gold'.² Trusting our practical judgments of sameness of denotation for 'gold', and taking for granted that gold is the element with atomic number 79, we can agree with a jeweler who in 1650 showed John Locke a
ring and said, "This ring is gold," only if the ring was (a bit of) the element with atomic number 79. We realize that perhaps the ring wasn't gold, even if it passed all of the jeweler's tests for being gold. In this practical way we acknowledge that truth is independent of belief.

There is a diverse group of philosophers, including David Chalmers, Michael Dummett, Frank Jackson, John McDowell, Christopher Peacocke, and (sometimes) Hilary Putnam, who think that if no theory of what determines the denotations of our words justifies (or shows that we are entitled to) our practical judgments of sameness of denotation, then our impressions that we can make discoveries, that we can agree or disagree, and that truth is independent of belief, are illusory. Against this, I will emphasize that we trust our practical judgments of sameness of denotation more than any theory of what determines the denotations of our words. I will argue that the denotations of our words are not settled by a priori linguistic rules or by causal-historical facts, but that our practical judgments of sameness of denotation are of a piece with our pursuit of truth, so we should not reject them. Instead, we should build these judgments into a deflationary account of denotation and truth.

**Analyticity**

To begin with, consider the relationship between our practical judgments of sameness of denotation and the thesis that some of our sentences are "analytic," in the sense that we cannot abandon them without changing the subject. One version of this thesis is that we make some of our sentences true by agreeing on how they are to be evaluated. The main problem with this version of the thesis is that it conflates truth and belief. As Frege observed in a different connection, "Being true is different from being taken to be true, whether by one or many or everybody, and in no case is to be reduced to it. There is no contradiction in something's being true which everybody takes to be false" (1964, p. 13).

One way to support the thesis that some of our sentences are analytic without conflating truth and belief is to derive the thesis from a description of how we evaluate sentences. To see how this might be done, consider W. V. Quine's account of the deviant logician's predicament. Against the idea that deviant logicians may "reject the law of non-contradiction and accept an occasional sentence and its negation both as true," Quine argues as follows:

[They] think they're talking about negation, "¬", 'not'; but surely the notation ceased to be recognizable as negation when they took to regarding some conjunctions of the form 'p -P' as true, and stopped regarding such sentences as implying all others. Here, evidently, is the deviant logician's predicament: when he tries to deny the doctrine he only changes the subject. (1986, p. 81)

The moral is that even though truth is not up to us, for some words, including '¬', 'not', we can agree on criteria that settle whether or not a speaker is using them to talk about the same subjects that we talk about when we use them.

This is the truism behind what I call methodological analyticity—the idea that even though truth is not up to us, there are sentences we cannot reject without changing the subject. The least problematic version of the idea, due to Rudolf Carnap, makes sense only for sentences of an artificial language system. If we accept Quine's textbook explanations of '¬', sentences of the form '¬(p -P)' come close to being "analytic" in Carnap's strict sense of that troublesome word.

The sort of methodological analyticity I want to discuss is also supposed to be a feature of natural language sentences, such as 'Bachelors are unmarried adult males'. The idea is that we tacitly agree on criteria that settle whether or not a speaker is using 'bachelor' to talk about bachelors, 'adult' to talk about adults, and so on. We tacitly agree, for instance, that no one can reject 'Bachelors are unmarried adult males' without changing the subject. Moreover, the criteria on which we tacitly agree are in principle obvious to us—we can tell by reflecting on our own usage of the terms whether or not an explicit statement of the criteria is correct. Natural language sentences that are analytic in the sense can play a methodological role in our inquiries that is analogous to the more strictly defined methodological role of analytic sentences in Carnap's artificial constructed language systems. That is why I call this sort of analyticity methodological.

One might be inclined to dismiss methodological analyticity (even the pure form of it that is restricted to artificial languages) with the claim that for any sentence we accept, we can imagine that it's false. This claim may seem to follow immediately from Frege's distinction between truth and belief. But that distinction has no direct bearing on whether we can imagine that a sentence we now accept is false. What would it be, for instance, to imagine that a sentence of the form '¬(p -P)' is false?

A better argument against methodological analyticity is that we may at one time feel confident that we could not reject a given statement without changing the subject, but later realize that we were wrong. This point is usually attributed to Quine, but Hilary Putnam was the first to present examples that make it convincing. He observed, for instance, that in the eighteenth century, scientists had no idea how their theory that physical space is Euclidean could be false—no idea how one could reject the statement that physical space is Euclidean without changing the subject. After much theoretical work in mathematics and physics, scientists replaced that earlier theory of space with the theory that physical space is non-Euclidean. When scientists came to believe that physical space is non-Euclidean, they took themselves not to have changed the topic, but to have discovered that space is non-Euclidean.
To accept this description of the case, we must trust the later scientists’ practical judgments of sameness of denotation for the phrase ‘physical space’ more than we trust the earlier scientists’ speculation that one cannot reject the statement that physical space is Euclidean without changing the subject. The later scientists might be wrong about physical space—perhaps it’s Euclidean after all. But we take them to have discovered that physical space is non-Euclidean, and so we accept their practical judgments of sameness of denotation for ‘physical space’. These practical judgments are embodied in their use of that phrase to express what they take to be a discovery about physical space. Their use of that phrase links it to earlier uses of the phrase, and those earlier uses of the phrase are linked to even earlier uses of it. Taken together, these uses of ‘physical space’ constitute a transtemporal chain of practical judgments of sameness of denotation for ‘physical space’. In a similar way, every inquiry brings with it some chain or other of practical judgments of sameness of denotation across time. In this sense, our practical judgments of sameness of denotation are of a piece with our pursuit of truth.

One might think that this reasoning only shows that we can be radically wrong about our own tacit criteria for applying our terms, not that methodological analyticity is incorrect. As I defined it above, however, methodological analyticity implies that we can tell just by reflecting on our own current usage of a term whether or not a given explicit statement of how it should be applied is correct. Putnam’s counterexamples show that we can’t tell just by reflecting on our own current usage of a term whether or not a given explicit statement about how it should be applied will survive a conflict with our practical judgments of sameness of denotation. In this way, Putnam’s counterexamples undermine methodological analyticity.

**PRIMARY INTENSIONS**

One might grant that we trust our practical judgments of sameness of denotation, but think that if we cannot also justify them, then our impression that we can make discoveries is illusory. It is natural to think that for every word used by a given speaker, there are linguistic rules that determine the denotation of that word. If there are such rules, then, for instance, my practical judgment that another speaker’s word ‘gold’ has the same denotation as my word ‘gold’ is correct if and only if the denotation determined by the linguistic rules for her word ‘gold’ is the same as the denotation determined by the linguistic rules for my word ‘gold’. Can we identify linguistic rules that we can use to justify our practical judgments of sameness of denotation?

We cannot use disquotational rules to justify these judgments. A disquotational rule for applying a given word—a rule such as ‘gold’ denotes an ob-

ject x if and only if x is gold—tells us nothing about the conditions under which a practical judgment of sameness of denotation for that word is true. A justification of such judgments therefore requires more than a disquotational specification of rules for applying our words.

Assuming that for every word w used by a given speaker S there are rules that determine the denotation of w, it is tempting to think that those rules are settled by S’s beliefs about how w should be applied. Inspired by this thought, David Chalmers has recently proposed a theory of intensions (or concepts) that is meant in part to justify our practical judgments of sameness of denotation across time. He defines the primary intension of a word as a special sort of function from (agent-centered) worlds to extensions: in a given (agent-centered) world, the primary intension of a word picks out what the extension of the word would be if that (agent-centered) world turned out to be actual (Chalmers 1996, p. 57). To grasp the primary intension of ‘water’, for instance, we must grasp a function that yields the set of all and only portions of water as value if the actual (agent-centered) world has water in its rivers, lakes, and oceans, but yields the set of all and only portions of twin-water as value if the actual (agent-centered) world has twin-water in its rivers, lakes, and oceans.

What is distinctive of a primary intension, according to Chalmers, is that our grasp of it is independent of all our empirical beliefs. He argues that there must be a primary intension for any word that we can use to express a discovery. If we are to express a discovery about water that is based on our examination of a given sample of what we take to be water, he reasons, we must be able to say why it counts as a sample of water by appealing to rules that we can grasp without going through any empirical investigation or presupposing any empirical beliefs (p. 62).

This reasoning moves from a truism to a substantive epistemological claim. The truism is that we take ourselves to express a discovery about water by using the term ‘water’ only if we take for granted that the denotation of ‘water’ does not change as a result of our supposed discovery. The substantive epistemological claim is that we are entitled to take ourselves to express a discovery about water by using the term ‘water’ only if we can justify our practical judgments of sameness of denotation for ‘water’ by appealing to rules that we can grasp without going through any empirical investigation or presupposing any empirical beliefs.

A primary intension is well-suited to this justificatory role, according to Chalmers. “The intension specifies how reference depends on the way the external world turns out, so does not itself depend on the way the external world turns out” (p. 57). By reasoning about “what our words would refer to if the actual world turned out in various ways,” Chalmers thinks, we can simultaneously see that our words have primary intensions and discover what they are.
The main problem with this proposal is that what we actually say when we find ourselves in a previously imagined situation almost always trumps our earlier speculations about what we would say if we were to find ourselves in that situation. What we actually say when we find ourselves in a previously imagined situation reflects our best current judgment of what is true in that situation. When we are actually in the previously imagined situation, our best judgment of what is true in that situation brings with it practical judgments of sameness of denotation. If those practical judgments of sameness of denotation conflict with earlier speculations, then so much the worse for those speculations. A scientist in the eighteenth century might have confidently predicted that even if it turns out that there is some mathematically consistent non-Euclidean geometry, it can’t be the description of what he calls physical space. But the actual history of our practical judgments of sameness of denotation for ‘physical space’ yields the opposite conclusion that when scientists came to accept that physical space is non-Euclidean, they did not thereby change the subject by tacitly defining a new denotation for the phrase ‘physical space’. As I emphasized above, the later scientists’ practical judgments of sameness of denotation are not a piece with their inquiry into the shape of physical space. We take the later scientists’ practical judgments of sameness of denotation to trump the earlier scientists’ speculations because we are confident that the later scientists have discovered that physical space is non-Euclidean.

This example shows that statements we can’t imagine giving up without changing the subject are not thereby guaranteed to be true. But one might think that to accept Chalmers’s claim that some of our words have primary intensions that we can know a priori, we need only suppose that some of the statements that we can’t imagine giving up without changing the subject actually are true. The trouble is that Chalmers’s primary intensions are supposed to “back a priori truths”—statements that are “true no matter how the actual world turns out” (p. 59). Hence to accept Chalmers’s claim that some of our words have primary intensions that we can know a priori, it is not enough to suppose that some of the statements that we can’t imagine giving up without changing the subject are true.8

Like Chalmers, Frank Jackson tries to defend the inference from “we don’t understand how we could give up statement S without changing the subject” to “we could not give up statement S without changing the subject.” “Surely it is possible to change the subject,” Jackson reasons, “and how else could one do it other than by abandoning what is most central to defining one’s subject? Would a better way of changing the subject be to abandon what is less central?” (Jackson 1998, p. 38). The mistake here is to suppose that our best current judgment about what counts as changing the subject is immune to future revisions. It is a truism that if we want to change the subject, we must rely on our understanding of what is most central to defining it. But this truism does not establish that our current understanding of what is most central to defining our subject cannot be revised without changing the subject. This claim is discredited by many actual cases in which we were once confident that we could not revise a given statement without changing the subject, but discovered later that we were wrong.

Both Jackson and Chalmers sometimes claim to be able to accommodate such discoveries (Chalmers 1996, pp. 55–6; Jackson 1998, pp. 46–55).9 They do not seem to realize that to accommodate such discoveries is to concede that our practical judgments of sameness of denotations are more trustworthy than our explications of our own primary intensions. This concession undermines their thesis that our understanding of our own primary intensions justifies our practical judgments of sameness of denotation across time.10

THE CAUSAL-HISTORICAL THEORY OF DENOTATION

You may be convinced by these arguments against Chalmers and Jackson but hope to find a different sort of justification of our practical judgments of sameness of denotation. Saul Kripke and Hilary Putnam devised the causal-historical theory of denotation to explain and justify the practical judgments of sameness of denotation that led them to reject the description theory of proper names and natural kind terms. In outline, the causal-historical theory is that the denotation of a name or a kind term is initially determined by an “ostensive definition” that may partly rely on causal connections to samples or things that the name or term is to denote. Once the denotation of a word is established in this way, speakers of the language can be credited with using that word if and only if they are connected by an “appropriate” causal chain with other speakers who use the word, and they are “minimally competent” in its use. “Appropriate” and “minimally competent” are place-holders for specifications of the causal chains and competencies that explain and support our practical judgments of sameness of denotation.

There is no consensus about how to specify these supposed chains and competencies, and many philosophers concede that there are kind terms (such as names of biological species) that pose challenges for the causal-historical theory. Nevertheless, many philosophers still think that by citing causal-historical facts we can explain why the denotations of our word ‘gold’, for instance, did not change as a result of our discovery that gold is the element with atomic number 79. Even in this paradigm case, however, as I shall now try to show, we cannot justify our practical judgments of sameness of denotation by citing causal-historical facts.
A THOUGHT EXPERIMENT\textsuperscript{11}

The historical background for my argument is that platinum was not discovered until the mid-eighteenth century, when chemists called it "white gold" because of its striking similarities to what they previously called gold.\textsuperscript{12} Platinum has a higher melting point than gold. But like gold, platinum dissolves in aqua regia, which was named for its ability to dissolve gold.\textsuperscript{13} In 1650, a chemist applying this "acid test" to a sample of platinum might have concluded that it should be called gold.\textsuperscript{14} We now know that platinum and gold are different elements: platinum is the element with atomic number 78, and gold is the element with atomic number 79.

With this in mind, suppose that there is a Twin Earth that is indistinguishable from Earth up until 1651, when large deposits of platinum are uncovered in Twin South Africa, and that once it is established by Twin Earth chemists that the newly uncovered metal dissolves in aqua regia, members of the Twin English-speaking community call it 'gold,' treating it in the same way we treat gold: the platinum is mined as gold, hammered (and later melted) together with gold to produce coins and bars that are valued by Twin Earthlings just as we value gold. Everyone on Twin Earth trusts the Twin Earth chemists' judgment that the newly uncovered metal is properly called 'gold'.

Suppose also that on Twin Earth chemistry develops in almost exactly the same way in which it develops on Earth, except that when Twin Earth chemists investigate what they call 'gold', they conclude that there are two kinds of 'gold'—their word 'gold' denotes x if and only if x is (a bit of) the element with atomic number 78 or x is (a bit of) the element with atomic number 79.

Recall that Twin Earth is just like Earth with a slightly different future after platinum is first uncovered in Twin South Africa in 1651. To see the possibility of this Twin Earth scenario, it is enough to imagine a few accidental differences between the two communities that allow for the uncovering of large amounts of platinum on Twin Earth.\textsuperscript{15}

The crucial point is that just as members of our English-speaking community take for granted that the denotation of the English word 'gold' did not change as a result of the discovery that it denotes x if and only if x is (a bit of) the element with atomic number 79, so members of the Twin English-speaking linguistic community take for granted that the denotation of their Twin English word 'gold' did not change as a result of their discovery that it denotes x if and only if x is (a bit of) the element with atomic number 78 or x is (a bit of) the element with atomic number 79. Members of the two communities have different beliefs about what their word form 'gold' denotes, and they take these beliefs for granted even when they are evaluating utterances made by using gold in 1650. For instance, suppose that in 1650 John Locke and his twin on Twin Earth both uttered the words 'There are huge deposits of gold in those hills', with Locke indicating South African hills and Twin Locke indicating the corresponding Twin South African hills, both of which contain platinum but no gold. We take Locke's word 'gold' to be true of an object x just in case x is gold, whereas our contemporaries on Twin Earth take Twin Locke's word 'gold' to be true of an object x just in case (as we would say it) x is either gold or platinum. We conclude that Locke's utterance is false, and our contemporaries on Twin Earth conclude that Twin Locke's utterance is true.\textsuperscript{16}

A DILEMMA FOR THE CAUSAL-HISTORICAL THEORY OF DENOTATION

Let's see if we can give a causal-historical explanation of our entrenched practical judgment that the denotation of the English word 'gold' did not change since 1650. Suppose that in 1650 members of both linguistic communities affirmed the following "ostensive definition":

\begin{enumerate}
  \item \textbf{x} is gold if and only if for most things y that I and other speakers in my linguistic community have on other occasions called gold, x is (a bit of) the same substance as y.
  \item The question is whether any such ostensive definition in 1650 actually determined that gold is true of x if and only if x is (a bit of) the element with atomic number 79. The answer is "no."
  \item To see why, note first that to explain our practical judgment that the denotation of the English word gold did not change since 1650 by appealing to (A), we must assume that:
  \begin{enumerate}
    \item For all \(x\) and \(y\), if \(x\) and \(y\) are gold, then \(x\) is (a bit of) the same substance as \(y\).
  \end{enumerate}
  \item is true in English and Twin English. Even if we stipulate that (B) is true in these languages, we have no good reason to believe that \(x\) is (a bit of) the same substance as \(y\) is true in Twin English of the ordered pair \(<x, y>\) only if \(x\) has the same atomic number as \(y\). The social (especially economic) practices in which the application of gold to gold or platinum is embedded in the Twin Earth community strongly suggest that if \(x\) is gold and \(y\) is platinum, for instance, then \('x and y are gold' and \('x\) is (a bit of) the same substance as \(y\) are true in Twin English of \(<x, y>\), even though \(x\) does not have the same atomic number as \(y\). One might try to rule this out by stipulating that:
  \begin{enumerate}
    \item For all \(x\) and \(y\), if \(x\) is (a bit of) the same substance as \(y\), then \(x\) has the same atomic number as \(y\).
  \end{enumerate}
\end{enumerate}
is true in both in English and Twin English. One problem with this strategy is that in 1650 no one was in a position to formulate (C), because analytical chemistry had not yet been developed. The crucial problem, however, is that even if we stipulate that (C) is true in English and Twin English, (B) may be false in Twin English. One may “ostensively define” ‘gold’ by affirming (A) without thereby guaranteeing that (B) is true.

One might think that to give a causal-historical justification of our practical judgment that the denotation of the English word ‘gold’ did not change since 1650, we need not suppose that an affirmation of (A) in 1650 could have guaranteed that (B) is true in English. One might think it is enough to suppose that (A) and (B) are true in English. But this is not enough. Even if (A) and (B) are true in English, as we all suppose, that does not show that our practical judgments of sameness of denotation across time for ‘gold’ are correct and the practical judgments of sameness of denotation across time for ‘gold’ in the Twin Earth community are incorrect.

This problem with Kripke-style ostensive “definitions” is similar to the problem with Chalmers’s theory of primary intensions. A Kripke-style ostensive definition of the denotation of a term can be part of a theory of what determines the denotation of the term, hence not just another entrenched belief that we express by using the term, only if we know that the ostensive definition could not turn out to be false. But we can know this only if, as Chalmers assumes, we can discover the rules for correctly applying our kind terms simply by reflecting on what we would say if the world turned out in various ways, without presupposing any empirical beliefs. As I emphasized earlier, however, what we actually say when we find ourselves in a previously imagined situation almost always trumps our earlier speculations about what we would say if we were to find ourselves in that situation. What we say when we find ourselves in a previously imagined situation reflects our best current judgment of what is true in that situation. When we are in the previously imagined situation, our best judgment of what is true in that situation brings with it practical judgments of sameness of denotation. If these practical judgments of sameness of denotation conflict with earlier speculations, we scrap the speculations, not the practical judgments of sameness of denotation. In our pursuit of truth, we trust our practical judgments of sameness of denotation across time, so we have no guarantee that a given “ostensive definition” of the denotation of a term is analytic, in the sense that we cannot revise it without changing the subject.

The causal-historical theory of denotation therefore faces a dilemma: the more informative the supposed ostensive definitions are, the more likely it is that they will later be revised without changing the subject; but the less informative they are, the less likely it is that there is only one way of correctly applying them. I conclude that affirmations of (A) in 1650 do not rule out either the Earthlings’ or Twin Earthlings’ discoveries about what their word ‘gold’ denotes and therefore do not justify the practical judgments of sameness of denotation for ‘gold’ in either linguistic community.

**DISPOSITIONS**

Suppose that Locke and Twin Locke would have accepted that gold is (a bit of) the element with atomic number 79 if they had been presented with the evidence we now have for this conclusion. One might think that an affirmation of (A) in 1650 rules out the Twin Earthlings’ practical judgments of sameness of denotation for ‘gold’ if it is supplemented with a counterfactual of this kind.19 The problem with this strategy is that whether or not an individual would accept or reject certain sentences may depend on the order in which he is presented with evidence that supports those sentences.20 It is plausible to suppose that Locke and Twin Locke would have affirmed the sentence ‘x is gold if and only if x is (a bit of) the element with atomic number 79’ if they had been presented with the same evidence that later English speakers encountered, in the same order in which they actually encountered it. But it is equally plausible to suppose that Locke and Twin Locke would have affirmed the sentence ‘x is gold if and only if either x is the element with atomic number 79 or x is the element with atomic number 78’ if they had been presented with the same evidence that later English speakers encountered, in the same order in which they actually encountered it. We have no independent grounds for saying that one of these presentations of the evidence is correct and the other is incorrect, and so an appeal to dispositions cannot show that our community’s practical judgments of sameness of denotation for ‘gold’ are correct and theirs are incorrect.21

**METHODOLOGICAL INTERLUDE**

Some causal theorists feel that the denotation relation is mysterious unless it is explained in causal-historical terms, so they are ready to reject our practical judgments of sameness of denotation if no causal-historical theory of denotation supports them. Various technical alternatives are open to someone who wishes to reject our practical judgments of sameness of denotation across time.22 For present purposes, however, the question is why we should feel driven to reject these practical judgments.

As I noted earlier, Saul Kripke and Hilary Putnam devised the causal-historical theory of denotation to explain and justify the practical judgments of sameness of denotation that led them to reject the description theory of proper names and natural kind terms. It ought to seem puzzling that some philosophers who were persuaded by Kripke and Putnam to
reject description theories of denotation because these theories conflict with our practical judgments of sameness of denotation are now inclined to reject those same practical judgments if they can’t be justified by citing causal-historical facts. One can’t have it both ways. No one who rejects description theories because they conflict with our practical judgments of sameness of denotation is in a position to reject these judgments just because they conflict with causal-historical theories of denotation. 23

This should be enough to convince us to reject causal-historical theories of denotation. But there is a deeper and more illuminating reason for rejecting such theories: they conflict with practical judgments of sameness of denotation that are (in the sense I explained earlier) of a piece with our pursuit of truth.

A DEFLEXIONARY ALTERNATIVE 24

I have argued that we cannot justify our practical judgments of sameness of denotation by appealing to a priori rules or by citing causal-historical facts. I also explained why I think we should not reject our practical judgments of sameness of denotation just because we find that we are unable to justify them. I propose that we focus instead on describing our practice of disquoting our own words and taking other speakers’ words at face value.

A disquotation specification of the denotation of our term ‘gold’ tells us nothing about what determines the denotation of ‘gold’ or how to find out whether or not something is gold. What then is the point of having such a thin account of denotation? One good reason for constructing a disquotation account of denotation is to define a truth predicate that we can use to specify the laws of logic schematically for regimented sentences of our own language. 25 A regimented first-order fragment of English, for example, may include such sentences as ‘(This ring is gold) → (This ring is gold)’, ‘∀x(Fx → (x is gold))’, and ‘∃x∀y(x loves y) → ∀y∃x(x loves y)’—instances, respectively, of the schemata ‘p → p’, ‘∀x(Fx → Fx)’, and ‘∃x∀yGxy → ∀y∃xGxy’. To specify logical laws schematically, we can say, for example, “every sentence of the form ‘p → p’ is true,” “every sentence of the form ‘∀x(Fx → Fx)’ is true,” and “every sentence of the form ‘∃x∀yGxy → ∀y∃xGxy’ is true.”

To make such generalizations precise, we need to know which sentences of our regimented language are instances of a given logical schema, and what it means to say of one of these sentences that it is true. The first need is met by well-established syntactical criteria for admissible substitutions of regimented English sentences and predicates for schematic letters, 26 and the second need is met by a Tarski-style truth predicate that is defined recursively, using clauses that specify conditions under which English predicates are satisfied by sequences of objects.

If our metalanguage contains the object language, our accounts of satisfaction and denotation can be disquotional. Suppose that all the variables of the object language are numbered sequentially, and let the i-th variable in this sequence be called var(i). A sequence of objects is a function from positive integers to objects; for any such sequence s, let s_i be the i-th object in s.

If the metalanguage contains the object language, we can say, for example, that for every sequence s, s satisfies ‘gold’ followed by var(i) if and only if s_i is gold. 27 Similarly, we can say that for every sequence s, ‘gold’ followed by var(i) denotes s_i if and only if s_i is gold. 28 More generally, we can accept the results of applying the following disquotional patterns to any of our one-place predicates:

(S) For every sequence s, s satisfies ‘_____’ followed by var(i) if and only if s_i is _____.

(D) For every sequence s, ‘_____’ followed by var(i) denotes s_i if and only if s_i is _____.

Each speaker who understands these patterns can apply them to his own words. For instance, if I affirm the results of writing ‘gold’ in the blanks of (D), I assert that (my predicate) gold followed by var(i) denotes an object s_i if and only if s_i is (a bit of) gold.

To describe our practice of disquoting our own words and taking other speakers’ words at face value, it is enough to trust our practice of taking other speakers’ words at face value and to find applications of the disquotional patterns (S) and (D) to our own words both obvious and useful. 29 For instance, if I affirm the result of writing my word ‘gold’ in the blanks of (D), I can see that when I take another English speaker’s word ‘gold’ at face value, I in effect take for granted that her word ‘gold’ denotes something s_i just in case s_i is gold, and so her word ‘gold’ has the same denotation as my word ‘gold’, This result of combining my practice of taking other English speakers’ words at face value with applications of (D) to my own words is a practical judgment of sameness of denotation.

We can describe our practical judgments of sameness of denotation in this way without justifying them. I propose that we trust these judgments unless we have some special reason to doubt or revise them.

But what counts as a special reason? Here is an example. I was surprised to read recently that there is a group of astronomers who now say, "Pluto is not a planet, it’s just a big asteroid." At first I took the astronomers to be claiming that Pluto is not a planet. After I learned more about why these astronomers say, “Pluto is not a planet,” however, I started to suspect that their sentence ‘Pluto is not a planet’ amounts to a stipulative new definition for their term ‘planet’, so that when they say, “Pluto is not a planet,” they are not
claiming that Pluto is not a planet.\textsuperscript{30} I have since learned that many astronomers have come to the same conclusion: the startling sentence ‘Pluto is not a planet’ does not express a discovery but introduces a new denotation for the term ‘planet’.\textsuperscript{31} I also learned that some of the astronomers who were reported in the popular media to have said, “Pluto is not a planet,” really meant to say, “Pluto is not a major planet.”\textsuperscript{32} In short, I was wrong to take the first, unqualified version of the claim at face value.\textsuperscript{33}

There are no similar reasons to doubt or revise the practical judgments of sameness of denotation across time in either linguistic community of my thought experiment. I propose that we trust these practical judgments and accept that the denotation of ‘gold’ has not changed since 1650 in either English or Twin English—chemists in both linguistic communities discovered the underlying chemical properties of what they respectively call ‘gold’.\textsuperscript{34}

It may seem strange to suppose that the Twin English word ‘gold’ that was used on Twin Earth in 1650 is true of x if and only if x is either gold or platinum, even though in 1650 that word was not yet actually applied to platinum. In contrast, it does not seem strange to suppose that the English word ‘gold’ that was used on Earth in 1650 is true of x if and only if x is gold. But I have argued that nothing rules out the Twin Earthlings’ practical judgments of sameness of denotation for their word ‘gold’. The only salient difference is that our practical judgments of sameness of denotation for our word ‘gold’ feel familiar to us and theirs don’t. We should not mistake this feeling of familiarity for evidence that we are right and they are wrong.

To put our feeling of familiarity in perspective, it helps to note that there are terms whose use on Earth resembles the use of ‘gold’ on Twin Earth. Our word ‘jade’ is true of both jadeite and nephrite. The Chinese character called 翡翠 that we translate as ‘jade’ was actually applied only to nephrite until the eighteenth century, when the Chinese first encountered jadeite and started carving it. The mineralogical differences between nephrite and jadeite were discovered in 1863, a century after the Chinese practice of applying 翡翠 to both nephrite and jadeite became entrenched (Hansford 1968, pp. 26–29). They (and we) take for granted that the denotation of 翡翠 did not change when it was applied to jadeite. Similarly, Twin English speakers take for granted that the denotation of their term ‘gold’ did not change when it was applied to the element with atomic number 78. We are therefore in no position to dismiss their practical judgments of sameness of denotation for ‘gold’ just because these judgments seem strange to us.

POSSIBLE PASTS

Imagine a world in which the English-speaking linguistic community ceased to exist in 1651. It might seem that in 1650 in such a world, the denotation of the English word ‘gold’ must be indeterminate, because the use of the English word ‘gold’, described independently of its denotation, is compatible with both of the denotations described in the gold-platinum thought experiment. It might therefore seem that to accept the practical judgments of sameness of denotation for gold in both linguistic communities of the gold-platinum thought experiment, one must also accept the strange idea that the denotation of a word w at time t can be determined by the use of w at some time after t.\textsuperscript{35}

This reasoning presupposes that the use of a word determines its denotation. In my view, however, the use of a word never determines its denotation, for the simple reason that even our most deeply entrenched beliefs about how our words are correctly applied may be false.\textsuperscript{36} There is therefore no obstacle to making sense of the claim that in the world described in the previous paragraph, in 1650 the English word ‘gold’ denotes x if and only if x is a bit of the element with atomic number 79.

As Kripke explained in Kripke (1980), possible worlds are not discovered, they are stipulated. When we describe a world w in which the English-speaking community ceased to exist in 1651, we stipulate that we are talking about the English word ‘gold’. If the English word ‘gold’ denotes x if and only if x is a bit of the element with atomic number 79, as we believe, then it follows from our description of world w that in w in 1650, the English word ‘gold’ denotes x if and only if x is a bit of the element with atomic number 79. Similarly, in a world in which the Twin English-speaking linguistic community ceased to exist in 1651, in 1650 the Twin English word ‘gold’ denotes x if and only if either x is a bit of the element with atomic number 78 or x is a bit of the element with atomic number 79.

These possible pasts will seem puzzling only if we assume that the use of a word, described independently of its denotation, determines its denotation. I recommend that we reject this assumption and trust our practical judgments of sameness of denotation across time.

ARE WE MAKING THINGS UP AS WE GO?

But if we cannot justify our practical judgments of sameness of denotation, are we always just deciding how to apply our words? One might worry that, like “locomotives . . . which unroll their tracks before them as they move through a terrain,”\textsuperscript{37} we are just making things up as we go.

This worry may seem compelling, but I think it is confused. Our understanding of sameness of denotation is rooted in our practice of taking other speaker’s words at face value. When we try to suspend all our practical judgments of sameness of denotation, the very idea of sameness of denotation seems to vanish into thin air.\textsuperscript{38} We understand what it is to decide
how to apply our words only by contrast with cases in which we apply them without reflection, as a matter of course, hence without deciding how to apply them. We understand what it is to make up new applications for our words only by contrast with cases in which we apply them without reflection, as a matter of course, hence without making up new applications for them. We can't doubt all our practical judgments of sameness of denotation at once, so we can't make sense of the worry that we just are making things up as we go.39

I noted at the start that we make practical judgments of sameness of denotation when we take ourselves to have made a discovery or to agree or disagree with other speakers. Many philosophers believe that we can trust these aspects of our linguistic practices only if we can justify our practical judgments of sameness of denotation. Against this I have emphasized that we trust these practical judgments more than any proposed philosophical justification of them. Our practical judgments of sameness of denotation are of a piece with our pursuit of truth, in the sense that our confidence that we have made a discovery always rests partly on our confidence in some chain or other of practical judgments of sameness of denotation across time. But our confidence that we have made a discovery typically trumps any conflicting prior speculations about the denotations of our words. A surprising result of these reflections is that the very trust in our practical judgments of sameness of denotation that initially led many to embrace ambitious justificatory projects ultimately undermines those projects and supports a new kind of deflationism about denotation and truth.40

To support this deflationism I have highlighted the methodological role of our practical judgments of sameness of denotation. I have tried to persuade you that we should evaluate any proposed account or theory of truth by comparing our confidence in it with our trust in practical judgments of sameness of denotation. This comparison convinces me, for the reasons I sketched above, that we should trust our practical judgments of sameness of denotation and doubt that there is a substantive theory that justifies them.

ACKNOWLEDGMENTS

This chapter is an expanded version of the paper I presented at the Central Division meetings of the APA in May 2000, with Mark Wilson as commentator, and at the University of Wisconsin, Milwaukee, in October 1999. For challenging and helpful comments on various versions of the chapter, I thank Hugh Chandler, Brie Gertler, Richard Heck, Mark Kaplan, John Koethe, Michael Liston, Patricia Marino, Fred Schmitt, Bill Taschek, Joan Weiner, and Mark Wilson.

NOTES

1. These two opening paragraphs are modifications of the two opening paragraphs of Ebbs (2000).

2. In Kripke (1980), on page 138, Saul Kripke stresses that scientific discoveries do not constitute a change of denotation.

3. For instance, in McDowell (1984), John McDowell argues that unless we possess some account of how it is possible for our words to conform to independent, objective, patterns of application, we cannot avoid the conclusion that our impression that we can make judgments is illusory. In "On Truth" (first published in 1983, reprinted in Putnam 1994a), Hilary Putnam uses a similar argument against W. V. Quine's deflationary view of truth. See also Chalmers (1995); Dummett (1978b, pp. 420-40; Jackson (1998); McDowell (1994); Peacocke (1999); and essays 13, 15, and 17 in Putnam (1994a).

4. This version of the thesis goes hand in hand with what Paul Horwich calls "the strategy of implicit definition," according to which "terms may be provided with their meanings by the assertion of statements containing them" in such a way that some of the asserted statements could not be abandoned without changing the denotations of the terms they contain. Horwich rejects this position, for reasons he explains in chapter 6 of Horwich (1998a). I am sympathetic with Horwich's objections, but I think he does not expose the deepest problem with the strategy of implicit definition—that it ignores the diachronic dimension of our pursuit of truth. My criticisms in this chapter of the methodological version of analyticity and of Chalmers's theory of primary intensions highlight the diachronic dimension of our pursuit of truth and indicate how I would argue against the stronger and even less plausible thesis that we must some of our sentences true by agreeing on how they are to be evaluated.

5. Carnap makes this explicit in the first sentence of his paper "Quine on Analyticity" (Creath 1990, pp. 427-32).


7. Chalmers presupposes that for all we know without empirical investigation, we may actually be in any world in which all of our subjective experiences are the same as they are in the actual world. This presupposition is incoherent, in my view, for reasons I explain in Ebbs (1996, 2001), but I do not question it here.

8. One can define a function $F$ from words and (agent-centered) worlds to extensions so that for any ordered pair of words and agent-centered worlds, the value of $F$ for that ordered pair is the extension of the word as used by the agent in that agent-centered world. Suppose we know a priori that for each pair of words and agents, there exists such a function. It does not follow, as Chalmers seems to assume, that we can know a priori what the value of the function is for the world we are actually in. For a similar criticism of a related position, see Stalnaker (1990).

9. Jackson also sometimes agrees with Lewis (1994) that we should disregard practical judgments of sameness of denotation that conflict with our explications of our primary intensions. See Jackson (1998, p. 38112). In this mood, Jackson explicitly concedes that the theory of primary intensions conflicts with a large number of our practical judgments of sameness of denotation.
10. My argument against Chalmers and Jackson suggests that "the syntactic construction of a quantity name may not reveal its actual ties to other quantities adequately," as Mark Wilson has stressed. See Wilson (1993 pp. 53-94); the quoted passage is on p. 82. I agree with Wilson that we may **discover** that many of our terms have what he calls honorable intensions. Unlike Chalmers's hypothetical primary intensions, **honorable** intensions are not known a priori.

11. I also use this thought experiment in Ebbs (2000).


13. See the entries for "gold" and "platinum" in the New Columbia Encyclopedia (1975).

14. Grosland reports that in 1752, a Swedish chemist named Scheffer concluded that the close similarity of (what we now call) platinum to gold justifies the claim that (what we now call) platinum is white gold (1962, p. 97). Grosland also points out that "the distinct nature of new substances was not always easy to demonstrate by elementary analytical methods and the skeptics could always maintain that any apparent discovery was really a substance previously known." (p. 98).

15. This thought experiment is similar in structure to the Druid thought experiment that Mark Wilson presents in Wilson (1982):

A B-52 full of regular American types landed on their uncharted island and the Druids exclaimed, "Lo, a great silver bird faieth from the sky." . . . [After this event] . . . the extension of the predicate is a bird for the cosmopolitan Druidese is something like the set of flying devices (including animal varieties) . . . [But] . . . If the hapless aviators had crashed in the jungle unseen and were discovered by the Druids six months later as they camped discontentedly around the bomber's hulk, their Druid rescuers would have proclaimed, "Lo, a great silver lleit in the jungle." . . . In this alternative linguistic community airplanes are no longer feared: are not held to be 'birds.' . . . Which extension should be assigned to 'bird' in cosmopolitan Druidese thus depends upon the history of the introduction of B-52's to the island. (pp. 549-50)

A similar thought experiment is briefly sketched by Daniel Dennett on page 312 of Dennett (1987).

16. This way of illustrating the odd consequences of the first thought experiment is adapted from a similar thought experiment presented by Keith Donnellan in Donnellan (1983, p. 103).

17. The similarity is not accidental. Chalmers's theory of primary intensions is a development of Gareth Evans's interpretation (in Evans 1985) of Kripke's claim in Naming and Necessity that there are contingent a priori truths. A model theoretic framework for Evans's idea of contingent a priori truths is sketched in Davies and Humberstone (1980). A number of philosophers, including David Chalmers, take for granted that Kripke-style ostensive definitions of natural kind terms must be contingent a priori, even if the discoveries we make about the natural kinds thus defined are not a priori. Chalmers (following Davies and Humberstone) credits Stalnaker (1978) with helping to develop the formal framework for the idea of a primary intension. But Stalnaker himself doubts that we can know primary intensions a priori. See Block and Stalnaker (1999). For an earlier criticism of an epistemological assumption similar to Chalmers's, see Stalnaker (1990).

18. Similarly, we have no guarantee that a "recognition capacity" that we associate with a given term determines the denotation of the term, in the sense proposed by Jessica Brown in Brown (1998). On any nonsemantic account of a "recognition capacity," the members of the Earth and Twin Earth linguistic communities in my thought experiment associate the same "recognition capacity" with the term 'gold' in 1650. Yet they later characterize the denotation of 'gold' differently and revise the "recognition capacities" they respectively associate with the word form 'gold,' without taking themselves to be changing the subject. If we take our practical judgments of sameness of denotation across time as our best guide to when we have changed the subject and when we haven't, then my thought experiment undermines Brown's theory of what determines the denotation of natural kind terms.

19. Hilary Putnam sometimes seems tempted by this idea. See, for example, "The Meaning of 'Meaning,'" reprinted in Putnam (1975); and Putnam (1988, pp. 30-37).

20. This formulation of the challenge my thought experiment poses for dispositional theories is due to Bill Robinson. The point was already implicit in Mark Wilson's druid thought experiment (see footnote 15).

21. The observations in this section suggest an argument against Paul Horwich's use theory of meaning. According to that theory, the meaning of a term is constituted by its possession of a certain "acceptance property" that can be specified independently of its meaning or denotation. Candidates for such properties are nonsemantic facts about a speaker's linguistic behavior, in particular, facts about which sentences the speaker is disposed to accept under various circumstances. For instance, according to Horwich, "the acceptance property that governs the speaker's overall use of 'and' is (roughly) his tendency to accept 'p and q' if and only if he accepts both 'p' and 'q'" (1998a, p. 45). Moreover, according to Horwich, two words express the same concept if they have the same basic acceptance property (p. 46), and any two words that express the same concept must have the same denotation (p. 69). Horwich is therefore committed to saying that the term 'gold' expresses the same concept and has the same denotation in 1650 in both of the linguistic communities of my thought experiment. This aspect of his use theory clearly conflicts with our confidence that we have not changed the denotation of our term 'gold' since 1650.

22. Of those alternatives, Harry Field's method of defining partial denotation, presented in Field (1973), is the best known and, in my view, the most promising.

23. David Lewis is one proponent of a priori intensions who accepts this methodological point and embraces description theories of denotation. In Lewis (1994), he concedes that very often more than one denotation is compatible with the descriptions that we associate with a term. In such cases, Lewis recommends that we use Field's method of defining a partial denotation for the term (see Field 1973).

24. Some of the material in this section is adapted from Ebbs (2002).


26. See for instance Quine's syntactical criteria for substitution, presented in chapters 26 and 28 of Quine (1982).

27. The satisfaction clauses for predicates are needed to give inductive specifications of satisfaction conditions for sentences containing quantifiers. Suppose our
regimented language contains just negation (symbolized by ‘¬’), alternation (symbolized by ‘∨’), and a universal quantifier (symbolized by ‘∀’). (In this language there is no separate symbol for the existential quantifier; existential quantifications must be expressed in terms of negation and universal quantification. Other truth functional connectives, such as ‘→’ and ‘∨’ can be expressed in terms of ‘¬’ and ‘∨’.) Then the satisfaction clauses we need, in addition to those for the n simple predicates of the language, may be formulated as follows:

(n+1) For all sequences s and sentences S: s satisfies the negation of S if and only if s does not satisfy S.

(n+2) For all sequences s and sentences S and S': s satisfies the alternation of S with S' if and only if either s satisfies S or s satisfies S'.

(n+3) For all sequences s, sentences S, and numbers i: s satisfies the universal quantification of S with respect to var(i) if and only if every sequence s' that differs from s in at most the i-th place satisfies S.

Suppose that together with the satisfaction clauses for the n simple predicates of the language, these clauses inductively define satisfaction for all sentences of the language. Using this inductive definition of satisfaction, we can then define truth for this language as follows: a sentence of the language is true if and only if it is satisfied by all sequences. (The above satisfaction clauses are modeled on Quine's formulations in Quine 1986, chapter 3.)

28. Note that on this use of the term 'denote', a word does not denote its extension; the extension of a word is the set of objects that the word denotes. See Quine (1982, p. 94). Note also that to specify the denotation of a predicate it is necessary to identify objects as members of sequences; I do this here only to highlight the intimate connection between denotation and satisfaction.

29. I do not claim that applications of the disquotational patterns (S) and (D) yield analytic truths, or that the left and right sides of the resulting biconditionals are synonymous. If every deflationary view of truth entails such dubious claims, then the view of truth that I propose is not deflationary. Naturally I don't accept the antecedent of this conditional. But see Gupta (1995) for a defense of the antecedent and a criticism of the view that applications of the disquotational patterns (S) and (D) yield analytic truths.

30. Some astronomers maintain that Pluto is one of the many Kuiper Belt objects, and that if we had known of the existence of all of these objects when Pluto was first discovered, we would not have called Pluto a planet. The argument fails, however, for the same reason that the appeal to dispositions can't show that our practical judgments of sameness of denotation for 'gold' are correct, and those on Twin Earth in the gold-platinum thought experiment are incorrect (see §7 above). What we are disposed to accept depends in part on the order in which we are presented with the relevant evidence. There are no independent grounds for establishing that the actual order in which we were presented with the relevant evidence is the wrong order for settling whether or not Pluto is a planet.

31. For an argument that Pluto is a planet, see Stansberry (2001); for the opposing view, see Jewitt (2001).

32. For an explicit statement of this weaker claim, see Minor Planet Center (2001).

33. We can criticize particular practical judgments of sameness of denotation only against a background of practical judgments of sameness of denotation that we accept without justification. Such judgments cannot all come up for review at the same time. Hence the fact that our practical judgments of sameness of denotation are revisable does not imply that there are standards independent of all them relative to which they are justified or unjustified.

34. Their discoveries do not conflict because the denotations of their terms are different, even though the terms themselves are spelled in the same way. Although we might have been tempted at first to take the Twin English term 'gold' at face value, if we are convinced that it denotes an object x if and only if x is either gold or x is platinum, and we do not revise our judgment that platinum is not a type of gold, we will not take it at face value. If by some strange circumstance we were ever in a position to learn their language, we would be able say what the denotation of their term is disquotationally, but we still would distinguish their term 'gold' from ours, possibly by using subscripts that mark the difference.

Some superficially similar conflicts about the denotations of words cannot be resolved in this way. Suppose, for instance, that one linguistic community actually splits into two communities, C1 and C2, that become isolated from each other. Let T be an unambiguous term of the language used before the split, and assume that after the split

1. the use of T in C1 comes to differ from the use of T in C2,
2. the characterization of T’s denotation that is accepted in C1 is different from the characterization of T’s denotation that is accepted in C2, and
3. members of both communities take for granted that the denotation of their term T is the same as it was before the split.

If members of the two isolated communities later realize that their practical judgments of sameness of denotation across time for T lead back to a single unambiguous term of the language used before the split, they are not likely to relinquish their practical judgments of sameness of denotation across time for T. They will likely take themselves to disagree about the denotation of T, and so they will try to persuade each other of their views.

35. For a defense of the view that the denotation of a word w at time t can be determined by the use of w at some time after t, see Jackman (1999b).

36. See Ebbs (2000) for a detailed argument in defense of this claim.

37. This striking image is from Wilson (1982, p. 586). Wilson does not believe there is any real danger that we are just making things up as we go.

38. At one point in Kripke's exposition of Ludwig Wittgenstein's remarks on rule following, Kripke writes: "It seems that the entire idea of meaning vanishes into thin air" (1982, p. 22). In Ebbs (1997, chs. 1, 10), I reconstruct and criticize Kripke's skeptical argument.

39. This is an argument I first presented in Ebbs (2000).

40. In this short chapter, I have focused exclusively on explaining how we can incorporate our practical judgments of sameness of denotation within a deflationary view of truth. I have said nothing about how I would try to answer the many fundamental criticisms of deflationary views of truth that are now familiar in the literature.
on truth. My answers to some of these criticisms agree with the answers of Field (1994), Horwich (1998b), Leeds (1978), and Quine (1986). But my way of describing practical judgments of sameness of denotation leads me to disagree with Field, Horwich, Leeds, and Quine about some key points. For a sketch of my answers to some of the central objections to deflationary views of truth, see Ebbs (2002).
References


