

# XIV\*—SHOULD WE TRUST OUR INTUITIONS? DEFLATIONARY ACCOUNTS OF THE ANALYTIC DATA

by Eric Margolis and Stephen Laurence<sup>1</sup>

**ABSTRACT** At least since W. V. O. Quine's famous critique of the analytic/synthetic distinction, philosophers have been deeply divided over whether there are any analytic truths. One line of thought suggests that the simple fact that people have 'intuitions of analyticity' might provide an independent argument for analyticities. If defenders of analyticity can explain these intuitions and opponents cannot, then perhaps there are analyticities after all. We argue that opponents of analyticity have some unexpected resources for explaining these intuitions and that, accordingly, the argument from intuition fails.

## I

**I***ntroduction.* In spite of W. V. O. Quine's famous attack on the notion of analyticity, people do have intuitions that they find natural to describe as 'intuitions of analyticity', that is, intuitions that seem to reflect truths of meaning. Whether or not there actually are any analyticities, it certainly does seem analytic that bachelors are unmarried and that chairs are furniture. In this paper, we consider the role of such intuitions as evidence for analyticity. Do these intuitions provide good grounds for believing that there really are analytic truths?

In an influential response to Quine, H.P. Grice and P.F. Strawson (1956) argue that they do. Grice and Strawson note that such intuitions are widespread and relatively stable and thereby amount to *prima facie* evidence for a real distinction, even if it isn't one that is antecedently well understood (142–3):

Is there a ... presumption in favour of the [analytic/synthetic] distinction's existence? *Prima facie*, it must be admitted that there is ... We can appeal, that is, to the fact that those who use the terms

1. This paper is fully collaborative; the order of the authors' names is arbitrary. We would like to thank Christopher Hookway, Philip Robbins and Jennifer Saul for helpful discussion of this material. We would also like to thank the AHRB and Rice University for their generous support of this research.

\*Meeting of the Aristotelian Society, held in Senate House, University of London, on Monday, 9th June, 2003 at 4.15 p. m.

‘analytic’ and ‘synthetic’ do to a very considerable extent agree in the applications they make of them ... This agreement extends not only to cases which they have been *taught* so to characterize, but to new cases.

Grice and Strawson also point out that the analytic/synthetic distinction finds support in the difference between judgments that are accompanied by a sense of certainty and those that seem to go further, apparently turning on the very meanings of the terms involved. They illustrate the difference by having us imagine someone asserting the following two sentences:<sup>2</sup>

- (1) My neighbour’s three-year-old child understands Russell’s Theory of Types.
- (2) My neighbour’s three-year-old child is an adult.

They remark that ‘it would be appropriate in the first case to say that we did not believe him and in the second case to say that we did not understand him’ (151).

These different strands in Grice and Strawson’s critique form the basis of an explanatory argument for analyticity. The thought is that our intuitions shouldn’t be dismissed out of hand. Moreover, if they can’t be accommodated without postulating real analyticities as their source, then perhaps this just shows that Quine was wrong after all. Georges Rey draws a conclusion along these lines (1993: 90):

Quine and his followers ... have no satisfactory explanation of ... either people’s projections of terms to novel actual and possible cases, or of their claims about constitutiveness, necessity, synonymy, or analyticity itself. Consequently, they have no serious rival to offer to the Traditional Explanation, which remains therefore at least an open empirical option.

Though in many ways a Quinean himself, Rey embraces analyticity on explanatory grounds of this sort. But even analyticity’s most ardent critics recognize that our intuitions of analyticity have to be addressed. For instance, Jerry Fodor writes (1998: 71–2):

I want to concentrate on the argument that the very fact that we *have* intuitions of analyticity makes a formidable case for there

2. For the purposes of this paper, it doesn’t matter what sorts of things are the proper bearers of analyticity (propositions, sentences, statements, etc.). Throughout we simply follow expository convenience.

being intrinsic conceptual connections. I'm sympathetic to the tactics of this argument. First blush, it surely does seem plausible that *bachelors are unmarried* is a different kind of truth from, as it might be, *it often rains in January*; and it's not implausible, again first blush, that the difference is that the first truth, but not the second, is purely conceptual. I agree, in short, that assuming that they can't be otherwise accounted for, the standard intuitions offer respectable evidence for there being cases of intrinsic conceptual connectedness.

For someone like Fodor, the task isn't to explain our intuitions of analyticity, but rather to explain them away.

We will call an account that takes the intuitions to reflect real analyticities a *face value account* (in that it takes the intuitions at face value) and an explanation that doesn't invoke real analyticities a *deflationary account*. In this paper, we propose to explore the prospects for deflationary accounts of people's intuitions of analyticity. Following some preliminaries we begin by examining past deflationary proposals. We argue that none of these is satisfactory. But a closer look at face value theories themselves suggests a natural deflationary approach that has all of the advantages of a face value theory.

This, of course, does not settle the question regarding the existence of analyticities, and for the purposes of this paper, we wish to remain neutral about the ultimate fate of the analytic/synthetic distinction. Still, the topic here does bear on this larger issue. For if it turns out that face value theories have no advantage over deflationary accounts, this removes one important form of argument for analyticity and suggests that perhaps our intuitions of analyticity shouldn't be trusted.

## II

*The Analytic Data.* Let's begin with a brief look at some of the intuitions that are at stake between face value and deflationary theories. It's important to remember that these are just intuitions and that the question of whether they are trustworthy is a matter of dispute.

The best examples of the intuitions in question concern the stock examples in philosophical discussions of analyticity. These are trivial claims and inferences that it is tempting to characterize in terms of the notion of analyticity:

- (3) Someone is a bachelor just in case he is an unmarried man.
- (4) A vixen is a female fox.
- (5) Chairs are furniture.

These claims do not explicitly invoke the notion of analyticity. Nonetheless, to someone who has been introduced to the notion, these certainly seem like clear cut examples of analytic truths.<sup>3</sup> For example, (3)'s truth does seem to turn on little more than the meaning of 'bachelor', and it would seem that anyone who knows what 'bachelor' means should be in a position to know that (3) is true. Our judgments in these sorts of cases can be expressed in the more explicit claims (6)–(8).

- (6) It is analytic that *someone is a bachelor just in case he is an unmarried man*.
- (7) It is analytic that *a vixen is a female fox*.
- (8) It is analytic that *chairs are furniture*.

And just as with Grice and Strawson's (1) and (2), in all these cases—(3)–(8)—our intuitions contrast strikingly with our intuitions about such claims as (9)–(11):

- (9) Bachelors generally don't wear rings on their ring finger.
- (10) Vixen have fur.
- (11) Chairs are generally smaller than tables.

Purported analyticities don't all have the same intuitive force. For example, all of the following have at one time or another been claimed to be analytic, but they do not support equally strong intuitions to that effect, and none of them seems as clear as (6)–(8):

- (12) It is analytic that  $2 + 2 = 4$ .
- (13) It is analytic that *everything is self-identical*.
- (14) It is analytic that *space is Euclidean*.
- (15) It is analytic that *knowledge is justified true belief*.
- (16) It is analytic that *beliefs are not located in space*.
- (17) It is analytic that *pain is having a disposition to cry out, wince, etc.*

One difference between (12)–(17) and (6)–(8) is the philosophical interest of the terms involved. Not much turns on whether

3. That is, to the extent that there are any.

it's analytic that bachelors are unmarried. On the other hand, it matters a lot whether it's analytic that  $2 + 2 = 4$ . For then one might be in a position to begin explaining both the nature of mathematical truth and the nature of mathematical knowledge. Noting this difference, Hilary Putnam describes cases like (3)–(8), which have an air of triviality, as cases that involve 'one-criterion concepts'. Though denying the existence of more substantive analyticities like (12)–(17), Putnam endorses the existence of analyticities based on one-criterion concepts, remarking that while 'they [bake] no philosophical bread and [wash] no philosophical windows', they participate in analyticities all the same (1962a, p. 36). At this point, we will leave open the question of whether one should follow Putnam in endorsing a face value theory for such claims.<sup>4</sup> However, Putnam's contrast highlights another important feature of purported analyticities in that people's intuitions regarding their status is often variable. A given claim may seem analytic at one time, but later, in the context of a new and unanticipated theoretical development, it may come to be viewed as not even being true.<sup>5</sup> To illustrate this, consider (18):

(18) Cats are animals.

In the absence of arguments to the contrary, we suspect that many people have an intuition that (18) is analytic, though perhaps not so strong an intuition as with (3)–(5). And yet people's intuitions are often far less secure once they hear Putnam's (1962b) thought experiment in which the things we've been calling 'cats' are discovered to be Martian-controlled robots. Some would describe the situation as the discovery that cats aren't animals, others that there aren't any cats, and still others hover between the two.

In spite of their gradedness and potential variability, intuitions of analyticity such as those associated with (3)–(8) and (12)–(18)

4. We should also note that we aren't committed to there being any hard and fast distinction between trivial and substantial analyticities. Indeed, just as with (18) below, stories can be told to call into question the triviality of (5) or even (3). Car seats and ski lifts are arguably chairs, but are they furniture (Hampton, 1982)? Is the Pope or a gay man in a committed long term relationship really a bachelor (Fillmore, 1982, Lakoff, 1987)? For further doubts, see Lormand (1996) and Giaquinto (1996).

5. This seems to have been the fate of (14)–(17).

represent the sorts of intuitions that need to be accommodated by any minimally adequate account. But there may be other phenomena that need to be accounted for as well. Georges Rey stresses the fact that we are able to judge whether actual and hypothetical cases satisfy a given concept, and that we are able to formulate prospective analyses of concepts and appreciate their merits. He takes this to be a related source of evidence for analyticity. For instance, when people acknowledge the force of a Gettier example to the proposal that knowledge is justified true belief, they seem to appeal to intuitions concerning the application of the concept KNOWLEDGE. Their intuitions in such cases, which we might call *Socratic intuitions*, take the form:

- (19) Situation *S* is (or is not) an instance of concept *C*.

In this case:

- (20) [Gettier case *G*] is not an instance of KNOWLEDGE.

Rey calls the sum of these various phenomena—the intuitions of analyticity and the Socratic intuitions—the *analytic data* and goes on to suggest that the best explanation of the analytic data may be a traditional explanation that invokes ‘a process of reason alone whose justification does not depend upon experience’ (1993: 84).

Rey isn’t the only one who draws a connection between the analytic data and analyticity. For example, Alvin Goldman and Joel Pust write that it is ‘almost a matter of definition’ that when people are fully informed about a situation, their intuitive response to whether a concept applies in that situation is accurate, reflecting the concept’s constitutive structure (1998: 188–9). In short, in certain philosophical circles there is a new wave of optimism about the potential explanatory value of analyticities of a fairly traditional sort—and a corresponding sense that deflationists have an undischarged explanatory burden to provide a satisfying account of the analytic data. We’ll return to this case for analyticity and to face value theories in general in Section IV. Before doing that, however, we want to examine the prospects for deflationary theories.

### III

*Deflationary Theories.* A deflationary theory is one that aims to account for the analytic data without invoking real analyticities.

In this section, we'll look at the main deflationary theories that are currently available. These include some suggestions of Quine's, a theory of Jerry Fodor's, and a theory that we ourselves have advanced elsewhere. We begin with Quine.

Quine does not seem especially troubled by the analytic data and accordingly he doesn't expend a great deal of effort trying to explain the data away. Nonetheless, he does offer a few remarks that are suggestive of why it might seem that there are analyticities. If these have any merit, they may go some way toward undermining arguments for analyticity simply based on intuition. Quine's suggestions include (i) the idea that intuitions of analyticity are confused with what may seem obvious, (ii) the idea that intuitions of analyticity are the result of a special feature of the acquisition of certain concepts, and (iii) the idea that intuitions of analyticity are a function of how particular beliefs figure in the total network of beliefs.

Quine's first suggestion, the idea that the seeming analytic is just the obvious, takes its foothold in claims involving simple logical truths and the most trivial of purported analyticities. The suggestion is that these may seem to be analytic but that one might just as well maintain that they are obvious. For example, in discussing the status of the logical truth that everything is self-identical, Quine remarks, 'We can say that it depends for its truth on traits of the language (specifically on the usage of "="), and not on traits of its subject matter; but we can also say, alternatively, that it depends on an obvious trait, viz., self identity, of its subject matter, viz., everything' (1954: 113). Clearly, Quine doesn't take his alternative explanation too seriously. His point is simply that it is just as good as the more loaded explanation in terms of analyticity.

Be this as it may, the problem with this account is that obviousness doesn't generally correlate with the analytic data. While it might be true that many purported analyticities are obvious, it's certainly not true that any claim that is obvious appears to be analytic in return. It's obvious that the sky is blue, and that bricks are hard, but no one holds that claims like these are analytic. So obviousness in and of itself can't explain why we have the intuition that certain claims are analytic.

Quine's second suggestion concerns the origins of our beliefs, how we come to learn, for example, that all bachelors are unmarried. Here the idea is that the process of learning the relevant

parts of English gives rise to the mistaken sense of analyticity. Quine makes a variety of remarks along these lines. One is to argue that we learn a word like 'bachelor' by 'learning appropriate associations of words with words'. As it is introduced to a child, it's connected with the phrase 'unmarried man', but not much else. 'Sever its tie with "unmarried man" and you leave it no very evident social determination, hence no utility in communication' (Quine, 1960: 56). In contrast, a word like 'momentum' may be introduced in the context of a phrase like 'mass times velocity', but it is embedded in a larger network of relations, allowing for subsequent changes that don't appear to be meaning changes. In a related vein, Quine characterizes language learning as a conditioning process where people learn to assent to certain sentences. Then he is able to suggest that analyticity can be recast in terms of the learning process, where a sentence that seems analytic is one where 'everybody learns that it is true by learning its words' (Quine, 1973: 79). Presumably, this means that we learn to assent to 'bachelors are unmarried men' as we learn the word 'bachelor', and that we learn the word 'bachelor' by being conditioned to link it with 'unmarried man'.

There are a number of difficulties with this account. One is the very naïve view of language learning that Quine endorses and his general tendency to work within a behaviourist framework. We won't go into the problems with either of these here, though we take it that both have been completely discredited.<sup>6</sup> A second difficulty is perhaps more relevant. This is that, even in cases where words are reliably associated with certain information in learning, such information is often easy to abandon later on. This point is nicely illustrated by Kripke's (1972) critical discussion of the description theory of proper names (see Rey, 1993: 87). For instance, when people first hear about Columbus, they are probably taught that he was the discoverer of America. And yet, they have no difficulty entertaining the possibility that this is just a mistaken view. Perhaps someone else beat Columbus to it, maybe a Viking sailor (not to mention the ancestors of America's indigenous populations).

6. See Chomsky (1959) on behaviourism. For a recent non-behaviourist account of word learning that nicely illustrates the complexity of the task that children face, see Bloom (2000).

Quine's third suggestion is that a belief's being central to one's overall system of belief, or its being held tenaciously, leads to the belief's seeming analytic. Quine is certainly right that simple mathematical and logical beliefs are at, or near, the very centre of anyone's system of belief and that, if it is possible to give them up at all, doing so would require massive revision, threatening the structure of the entire system. Though the mechanisms are a bit obscure, one can imagine that this might give rise to the sense that these beliefs are analytic. Something of this sort may also explain the temptation to think of the statements of certain natural laws as being analytic. They may be so central to the articulation of large amounts of theory that they seem to be definitive of the very concepts in which they are couched.

Promising as this may be, centrality can't be the whole story. Many cases that seem analytic just aren't sufficiently enmeshed in everything else we believe. For example, giving up the belief that bachelors are unmarried may have some impact on other beliefs, but it hardly permeates every aspect of our world view. It's at this point that Quine might rely on tenacity (i.e., understood independently of centrality). But that won't work either. Notice that there are many beliefs that people hold tenaciously that don't seem analytic in the least. Take any Moorean belief, that is, a belief that is so basic and commonsensical that any argument against it only ends up calling its own presuppositions into question. Consider, for instance, the belief in other minds, the belief that our individual experiences aren't just dreams, or the belief that the world has existed for more than a minute. All of these are beliefs that would be held tenaciously by anyone who entertained them and yet they don't seem the least bit analytic (see Rey, 1993: 88). Tenacity doesn't guarantee seeming analyticity, and seeming analyticity doesn't guarantee centrality.

Still, it might be that *some* of our intuitions of analyticity do in fact reflect centrality. This might not seem like much, but as Fodor (1998) points out, since deflationary theories are driven by the idea that our intuitions of analyticity are *faulty*, deflationists needn't suppose there is a single source of error that accounts for all of the mistaken intuitions. As Fodor sees it, Quine's suggestion regarding centrality is fine so far as it goes. What's missing is an explanation that can handle our intuitions regarding BACHELOR and the like. Fodor offers a theory to deal with

cases of just this sort. He presents his theory against the background of an information-based semantics wherein concepts have their content in virtue of nomic links to the properties they express. However, we needn't be concerned with Fodor's background assumptions. What matters is Fodor's core idea, which is that a deflationary spin can be given to Putnam's 'one-criterion' concepts.

One-criterion concepts are supposed to be ones for which there is only one way of telling whether they apply. For example, BACHELOR is supposed to be a one-criterion concept because, by hypothesis, there is only one way of telling whether something is a bachelor (viz., check whether it is an unmarried man), while ENERGY isn't because there are numerous ways of telling whether the concept energy applies (evidenced by the many laws in which energy figures). Putnam argues that this distinction vindicates a commitment to a limited set of analyticities<sup>7</sup>—just the ones that Quine's notion of centrality misses. Of course, Fodor will have none of this. He wants to defend a deflationary account. His opposition to Putnam's face value account is that it can only work pending a principled way of individuating criteria. But, Fodor suggests, Putnam doesn't have a principled way of doing this. For instance, one may wonder whether BACHELOR has one criterion or two. Are UNMARRIED MAN and NOT MARRIED MAN the same criterion or different ones? As Fodor points out, that depends on whether the two are synonymous. Unfortunately, synonymy, like analyticity, is among the small circle of interdefinable notions that Quine's critique calls into question. 'So, it looks as though Putnam's construal of *analytic connection* in terms of *one-criterion concept* leaves us back where we started' (Fodor, 1998: 82).

Fodor doesn't mind criticizing Putnam's account for being unprincipled since Fodor thinks his own view requires only a purely epistemic reading of a concept's having one criterion (Fodor, 1998: 82):

Notice that since what I'm aiming for is not an account of the individuation of meanings, but just a diagnosis of some faulty

7. Putnam links one-criterion-ness to a claim of semantic constitutiveness by maintaining that something is in the extension of a one-criterion concept just in case it satisfies the associated criterion.

intuitions, telling my story doesn't presuppose a prior or principled account of the individuation of criteria. Unlike Putnam, I can make do with what I imagine everyone will grant: that for some concepts there are, *de facto*, lots of ways of telling that they apply and for other concepts there are, *de facto*, very few.

So while one-criterionhood may not be a mark of real analyticity, it can nonetheless be used to provide an account of some of our intuitions of analyticity. According to Fodor, the difference between one-criterion concepts and other concepts (such that one-criterion concepts give rise to intuitions of analyticity) is that the former are thought to have very few, and perhaps only one, way of telling that they apply that is independent of the other ways that might be used to tell whether they apply. When someone actually intuits this special (albeit nonanalytic) relation between concepts, this gives rise to a corresponding intuition of analyticity (Fodor, 1998: 83):

Suppose you think the only epistemic route from the concept *C* to the property that it expresses depends on drawing inferences that involve the concept *C\**. Then you will find it intuitively plausible that the relation between *C* and *C\** is conceptual; specifically, that you can't have *C* unless you also have *C\**. And the more you think that it is counterfactual supporting that the only epistemic route from *C* to the property it expresses depends on drawing inferences that involve the concept *C\**, the stronger your intuition that *C* and *C\** are conceptually connected will be.

Though Fodor's exposition is not terribly detailed, he does work through a few examples that are supposed to illustrate the structure of his account. He points out that there are many ways of telling whether *DOG* applies to something without having to deploy *ANIMAL*, and that it's even clearer that there are many ways of telling whether *WATER* applies to something without deploying *H<sub>2</sub>O*. 'But offhand, I can't imagine how I might determine whether John is a bachelor except by determining that he's male and un- (*viz.* not) married' (Fodor, 1998: 84).

One difficulty in evaluating this account is Fodor's insistence that he doesn't inherit the burden of saying precisely how to individuate criteria. Notice how he is happy to treat *UNMARRIED MALE* and *NOT MARRIED MALE* equivalently when it comes to advancing his own theory, even though he takes it to be the decisive blow against Putnam's that Putnam can't specify whether

these are distinct or not. It is hard to see why this issue is any less problematic for Fodor's position than Putnam's. After all, Fodor's theory of the analytic data turns precisely on how many criteria are associated with a concept. If that can't be pinned down, then he is in no position to be so confident about its practical implications. What's more, if Fodor can be allowed a little looseness in the individuation of criteria, one wonders why the same generosity can't be extended to Putnam. And if it can, then why not simply opt for Putnam's face value theory instead?

These are good questions, but we won't press them further. That's because we think there is a more serious objection to Fodor's theory. Notice that he maintains that, to the extent that  $\text{DOG} \rightarrow \text{ANIMAL}$  doesn't seem to be a good example of an analyticity, this is because there are multiple ways of telling whether dog applies without having to deploy ANIMAL. The example he cites is that we often tell whether something is a dog by how it looks. If it has a doggyish appearance, then it's very likely that it is a dog. Moreover, if that's right then there will be many other ways to tell whether something is a dog which also don't depend on employing the concept ANIMAL. One could focus on how it sounds, on the shape of its footprints, on what people call it, and so on. Notice that all of these are useful tests only because people know certain contingent information about dogs. The problem for Fodor is that if beliefs about contingent facts about dogs can be used to formulate numerous tests for whether something is a dog, *the same can be done for bachelors*. One might look to see whether the man is wearing a ring, for example, or whether he lives alone and spends lots of time out drinking at clubs, or whether, if you ask him if he's a bachelor, he says that he is. Of course, none of these tests is perfect. But neither is it a perfect test of whether something is a dog that it looks like one. In short, the problem with Fodor's deflationary theory is that his loose reading of 'criteria' allows for the purportedly one-criterion concepts to be associated with many criteria. But in that case, a concept like BACHELOR isn't in fact a one-criterion concept, and Fodor's explanation cannot account for the intuitions of analyticity associated with using a concept like BACHELOR.

Things are not looking terribly promising for deflationary theories at this point. Before turning to face value theories we want to look at one last deflationary account. This is an account that

we proposed ourselves in an earlier work. This theory takes as its starting point Quine's suggestion that the seeming analytic is just the obvious. Recall that the main problem with that view was that the obvious doesn't invariably give rise to a sense of analyticity. One way of dealing with this problem is to focus the claim more. Along these lines, the view we suggested took intuitions of analyticity to reflect *entrenched beliefs about constitutive conditions*, specifically, ones that people take to be intuitively or pretheoretically obvious (Laurence and Margolis, 1999). On this account, not just any obvious fact should seem analytic—only the ones that concern constitutive conditions. For this reason, the theory has no trouble accounting for the fact that neither the inference COLUMBUS → DISCOVERER OF AMERICA nor Moorean facts seem analytic. The theory also has little difficulty handling scientific discoveries concerning constitutive properties, for instance, the discovery that water is H<sub>2</sub>O. These don't seem analytic because, while they may be entrenched in the belief system, no one believes that such constitutive claims are intuitively or pretheoretically obvious. Finally, one of the nice features of this account is that it allows for graded intuitions concerning the purported analytic phenomena. Whether something is obvious is itself a graded matter, so it's no surprise that not all seeming analyticities are entirely on a par.

All told, we think that this last deflationary theory has much to be said for it. Still, we no longer think it is wholly satisfactory. The problem is that, while there may be a close correlation between beliefs about obvious constitutiveness and the seeming analytic, it remains somewhat obscure *why* the one would account for the other. The theory says, in effect, that the correlation between the two is accounted for by the fact that we are often confused about what we are really intuiting. We aren't intuiting real analyticities; instead, we are intuiting that certain constitutive claims are obvious. But this doesn't fully address the issue at hand. One still wants to know why we confuse such intuitions with intuitions of analyticity. Something further needs to be said. Much the same problem extends to Quine's and Fodor's suggestion regarding centrality. Even if some cases of the seeming analytic do involve beliefs that are at the very centre of anyone's system of belief, it remains to be explained why people would become confused about these beliefs, taking them to express claims about meaning relations.

In sum, none of the deflationary theories we have examined provides a fully satisfactory treatment of the basic intuitions of analyticity. Moreover, none has had anything to say about the Socratic intuitions that worry Rey, Goldman and Pust, and others. Face value theorists will see these gaps as further motivation for re-examining the prospects of a traditional account of the analytic data in terms of analyticity itself. It is time, then, to have a closer look at the face value approach.

#### IV

*A Return to Face Value Theories?* If there isn't an acceptable explanation of the seeming analytic that avoids postulating analyticities, then perhaps we should just take our intuitions at face value. Maybe they aren't confused; maybe the reason certain claims and inferences seem analytic is because they really are analytic. What's more, if Rey *et al.* are right, our Socratic intuitions may provide further support for face value theories. In that case, analyticity might turn out to have considerable explanatory appeal, despite whatever Quine may have said. In this section we want to take a closer look at face value theories and, in particular, address the question of how analyticities explain what they are said to explain on a face value account.

How *do* analyticities explain the seeming analytic? This may seem an odd question at first. The whole point of adopting a face value theory is to claim that our intuitions aren't confused, that they actually have it right. But even so, one should still ask how they get it right. Are we to believe we just have an inexplicable faculty of intuition that puts us directly in touch with the facts of analyticity? That seems fanciful, to say the least. Surely, a more promising approach is to maintain that psychological processes of some kind account for our intuitions. If so, we need to ask what these processes are like. This is an explanatory burden that face value theorists need to take more seriously; after all, the force of their charge that *deflationists* have no account of the analytic data depends on it. Ultimately, of course, this question belongs more properly to psychology than to philosophy. If a psychological process is at work, then it's up to psychology—the scientific study of the mind—to articulate its structure and character. But in the absence of detailed psychological study of

this question, we can still begin to think about it by considering possible models in broad outline. Here we want to consider two schematic models that seem representative of the options available to face value theorists.

The first includes a variety of models that can all be grouped together under the heading of *Theory Based Accounts*. What these accounts have in common is that they claim that intuitions of analyticity are generated in part by an implicit theory of meaning. The theory might be a naïve or 'folk' theory (akin to folk psychology), it might be universal or culture-specific, it might simply involve a few basic beliefs plus general intelligence. The suggestion is that, however the implicit theory is to be understood, it takes categorization judgments as input and generates, among other things, judgments that certain claims have a special status owing to the meanings of their constituent terms. We all use our words and concepts in accordance with our categorization dispositions. These form the basis for a huge range of judgments that we are able to reflect upon. In these reflections, whether we realize it or not, we may be appealing to the principles of an implicit theory, one that tells us that certain uses are meaning constitutive and that these are exhibited in certain patterns of inference.<sup>8</sup> Having detected the relevant patterns in our own categorization, we are then able to view them in the way that is specified by our implicit theory. In other words, these patterns come to strike us as being engendered by meaning and having a special status as a result of meaning. Once we acquire an explicit concept of analyticity, we take these intuitions to be intuitions of analyticity.

Now the existence of mechanisms generating categorization dispositions is something that shouldn't be controversial on any theory of the mind. Categorization is about as fundamental as a psychological phenomenon can get. The whole point of having concepts is to apply them, to form judgments about which things fall under them and which don't; it's only by engaging in

8. Just what these patterns are is something that a detailed account would need to specify. Since our main purpose here is to explore the possibilities for deflationary accounts, and not to develop a fully articulated face value account, we will not pursue this question in any detail. Still, one possibility, suggested by Paul Horwich's (1998) theory of meaning, is that the relevant patterns involve 'centrally explanatory uses'. Another, suggested by Christopher Peacocke's (1992) theory of concepts, is that they involve 'primitively compelling' inferences. Yet another is that they involve the sense that some categorization judgments simply can't be given up.

categorization that we are able to bring to bear the information we know about various types of things. What's more, psychologists have been studying the nature of categorization for some time now and have developed a number of theories that have a good deal of empirical support (Smith and Medin, 1981; Murphy, 2002). As for the implicit theory of meaning, though this is more controversial, it is not unimaginable that such a theory is psychologically real. People do find certain ideas about meaning (however inchoate) to be relatively natural. Perhaps these ideas are sufficiently consistent and well articulated to do the sort of psychological work that a face value theorist requires.

In any event, there is an alternative to Theory Based Accounts that some face value theorists may wish to explore. This second class of models we'll group together under the heading *Similarity Based Accounts*. As before, people are supposed to have categorization dispositions that govern the use of their concepts and terms. But instead of an implicit theory that generates intuitions of analyticity, there is a psychological mechanism embodying a similarity metric that classifies categorization judgments as being more or less alike. According to this classification, claims that are analytic constitute a natural similarity class. For example, *bachelors are unmarried* would be judged to be more like *vixens are female* than either would be judged to be like *grass is green*. This gives rise to the intuition that claims that are analytic constitute a special class. And again, once we acquire an explicit concept of analyticity, we take these intuitions to be intuitions of analyticity.

Clearly, face value theorists need something along these general lines; otherwise they'd be committed to a mysterious and inexplicable faculty of intuition. What's more, though our remarks only give an initial sketch of what a substantive face value theory might look like, they do give some credence to the face value theorist's claim to having an account of our intuitions of analyticity. Nor is it unreasonable to suppose that a plausible and more detailed face value account could be developed along such lines.

## V

*Deflationism Again.* If a face value account of our intuitions of analyticity can be developed along the lines of one of the models

suggested in the previous section, then the challenge of finding a comparable deflationist account becomes pressing. How might deflationists go about meeting this challenge? Surprisingly, the discussion of possible face value accounts in the previous section suggests a natural deflationary alternative that has all of the advantages of a face value theory: deflationists can simply mimic the face value accounts without committing themselves to the existence of any real analyticities. The models canvassed in the previous section are all *psychological* accounts of the genesis of intuitions; they make *no* essential appeal to the existence of analyticities. As a result, deflationists are free to co-opt those models more or less wholesale.

Consider, for example, the Theory Based Accounts. On these models, intuitions of analyticity are the product of an implicit theory which takes sets of categorization judgments as input and delivers as output intuitions of analyticity. Nowhere in these models is an appeal made to actual analyticities. So a deflationist can happily adopt such a model, while denying that there are any analyticities. The implicit theory may imply that certain claims are analytic, but who's to say that it's right? The theory doesn't have to be true in order to be explanatory; people would have the very same intuitions either way. In other words, *analyticity per se does no explanatory work in accounting for our intuitions of the seeming analytic.*

The moral here is that once face value theorists take seriously the need for a psychological account of intuitions, the intuitions can no longer motivate the postulation of analyticities. It may be that there are analyticities that correspond to certain of our intuitions, but the intuitions themselves can be explained without ever mentioning this fact. The upshot is that analyticity has to be argued for on other grounds altogether; intuitions don't provide the support that face value theorists claim for them.

Deflationist accounts that mimic face value theories don't have any of the problems that are characteristic of the previous deflationist accounts. Since mimicking accounts simply adopt the essential aspects of face value theories, they cannot be accused of being incomplete or overly sketchy; they will be just as complete and detailed as the face value theories on which they are based. Likewise, there is no longer any mystery as to why, on a deflationist approach, certain uses of a concept are connected

with meaning intuitions. That was a problem with the earlier deflationary models, where it was claimed that people simply confuse centrality or obvious constitutiveness with meaning constitutiveness. Now that mystery simply disappears. Deflationists can handle this worry in exactly the same way as face value theorists, for example, by claiming that people readily form meaning intuitions, not because they are confused, but because they are following the dictates of an internalized theory.

Finally, what about the gradedness of our intuitions of analyticity? Earlier we noted that a good explanation of the seeming analytic ought to accommodate the fact that our intuitions aren't all on a par. How well do deflationary accounts that mimic face value theories do in this respect? Certainly they do at least as well as face value theories. And in the end, they may actually do very well. Consider Theory Based Accounts. Since the implicit theory may be somewhat vaguely specified or underdeveloped, there may be significant scope for differing strengths to one's intuitions that given claims are analytic. The implicit theory may pick out certain patterns as clearly exhibiting meaning constitutive relations, while being a bit hazy when it comes to other patterns. Judgments that conform to these other patterns would then lack the certainty that goes with the clear cut cases.

The availability of deflationist accounts that mimic face value accounts answers the face value theorist's challenge, and thereby undermines the argument for analyticity based on intuitions of analyticity. Deflationists, however, are not limited to such accounts. Before closing this section we would like to briefly explore a deflationary Similarity Based Account, according to which so-called intuitions of analyticity are entirely misdescribed: they do not involve the seeming analytic at all.

Suppose that, prior to learning the explicit concept of analyticity, all we have is an intuitive sense that certain claims are more or less similar; there is no prior tendency to view the purportedly analytic ones in terms of the notion of meaning. When we learn the explicit concept of analyticity, this generally involves a theoretical gloss ('true in virtue of meaning alone'), and it nearly always involves being given a few paradigmatic examples (*bachelors are unmarried men, vixen are female foxes*). As we attempt to find an application for this new concept, asking ourselves whether this or that claim is analytic, there are then two

aspects of the newly introduced distinction that we can rely upon. One is the theoretical gloss on the distinction, the other is the paradigmatic examples. In such circumstances, it may well be the examples that end up doing most, or all, of the work in generating our intuitive sense of this distinction. We come to say that certain claims are analytic because they are similar to the paradigms we've been given. *But it may be that the similarity metric we employ has nothing to do with meaning or analyticity at all.* In employing this concept, we trust that the distinction we intuitively mark corresponds to the philosophical gloss that we are given when we are learning about analyticity. But our trust may be misplaced.

Though we did not emphasize it earlier, the similarity metric is really the heart of any Similarity Based Account; it determines what dimensions the concept generalizes along. If the similarity metric involved has nothing to do with meaning or analyticity, what dimensions does it generalize along? One possibility is that it involves the sorts of properties appealed to in the deflationary accounts discussed in Section III above. For example, it may be that the similarity metric concerns what's *an entrenched belief about constitutive conditions that people take to be intuitively or pretheoretically obvious*.<sup>9</sup> If something like this were true, then our intuitions wouldn't be intuitions of analyticity at all; they wouldn't even be intuitions that certain claims *seemed* analytic. Rather, they would be intuitions about what's obviously constitutive, misconstrued in light of the philosophical gloss that's associated with acquiring the concept of analyticity. On this view, the reason we confusedly suppose ourselves to be intuiting the analytic/synthetic distinction, is just that we have been *indoctrinated* by philosophy.

Surprisingly, then, deflationists have a number of interesting options. One is simply to adopt whatever account face value theorists have on offer. Deflationists can do this because face value theories that take seriously the need to explain our intuitions end up making no essential reference to analyticity. This first option

9. A different alternative might take the intuitive distinction to involve something like what seems clearly *a priori*. If this were true, then, ironically, our sense that something real is marked by the analytic/synthetic distinction would be inherited from the *a priori/a posteriori* distinction (or more precisely, the seeming clearly *a priori/a posteriori*).

alone shows that there is no independent argument for analyticity to be had based on considerations of intuitions of analyticity. But another more aggressive approach is to claim that we don't really have intuitions of analyticity in the first place (i.e., intuitions that certain claims seem analytic). The assumption that we do stems from a misplaced trust in the gloss that philosophers have put on our intuitive sense that such truths as *bachelors are unmarried* form a distinctive class.

## VI

*Socratic Intuitions.* We have shown that it is possible to account for intuitions of analyticity without there necessarily being any analyticities at the bottom of these intuitions. So while intuitions of analyticity may have been thought to provide good evidence for analyticity, it's evidence that can be safely put to the side; at least this much of the analytic data doesn't argue for analyticity. Still, there is the other half of the analytic data—the Socratic intuitions that Rey *et al.* emphasize—and one may wonder whether these, at least, continue to argue for analyticity. In this section, we take up this last issue.

Socratic intuitions are intuitions about whether a concept applies in various actual and hypothetical situations. They are the sorts of intuitions that philosophers appeal to when attempting to formulate and assess analyses of a given concept. Rey claims that opponents of analyticity have special problems in accounting for these intuitions, but he isn't alone; the sentiment is widespread (see, e.g., Jackson, 1998).

What account can deflationists give of Socratic intuitions? How can they explain our ability to judge whether a case falls under a given concept or our ability to formulate and assess prospective analyses of our concepts? The answer is that deflationists can provide rich accounts of all these abilities. Indeed, just as with intuitions of analyticity, deflationists can provide accounts with the same essential structure as the ones given by face value theorists.<sup>10</sup>

10. Though it's worth mentioning that deflationists may have a different view about the significance of conceptual analysis as a philosophical method. Many maintain that conceptual analysis is a bankrupt enterprise. But whether conceptual analysis is a worthwhile endeavour is one thing; how to explain the intuitions that often figure in the endeavour is another. The point in the text is that deflationists aren't at a disadvantage when it comes to the latter.

Consider what's involved in our ability to judge whether actual and hypothetical cases fall under a given concept. Notice that the ability just amounts to the ability to *categorize*. And as we mentioned earlier, categorization has been extensively studied by psychologists. One result of this important body of research is that categorization isn't a unitary phenomenon; rather, it involves a number of distinct yet related processes. Among these differing processes, some are rough and ready; they underwrite the quick and unreflective categorization that we spontaneously rely upon in everyday cognition. This sort of categorization is relatively well understood and is thought by many to involve similarity comparisons. On some models, the similarity of a target is measured against an exemplar stored in memory (i.e., an especially typical or representative instance). An affirmative answer is reached so long as the target is deemed sufficiently similar to the exemplar. For example, a common exemplar for the concept BIRD is a representation of a robin or sparrow; things that are judged to be similar enough to robins and sparrows are consequently judged to be birds. On other models, the similarity of a target is measured by matching features. A concept like BIRD would then be taken to be a complex item that has constituents such as HAS WINGS, HAS A BEAK, FLIES, etc. and whether something is deemed a bird would depend on whether it is judged to have a sufficient number of the corresponding properties (Smith and Medin, 1981; Murphy, 2002).

Rough and ready categorization undoubtedly plays a large role in our mental lives, but when it matters, people can also engage in more deliberate and thoughtful categorization. This sort of categorization can often be painfully slow and involve an enormous amount of reflection. And unlike the quicker forms of categorization, it isn't particularly well understood. In principle, thoughtful categorization may involve specific rules, implicit or explicit theories, and various represented information. Indeed, given the holistic nature of confirmation, it may implicate large chunks of one's background beliefs. Realistically, it's probably a mistake to speak of thoughtful categorization as a single process. At best, it's an all-things-considered determination of whether a given concept should be applied in a particular circumstance, a determination that might make use of any and all of one's cognitive resources.

A lot more could be said about categorization. But what matters for present purposes is that there is nothing to stop a deflationist from adopting any of these models in explaining categorization. So to the extent that anyone can explain how people judge that a concept applies in a given circumstance, deflationists can as well. Much the same goes for our ability to appreciate counterexamples to proposed analyses. What one does in appreciating a counterexample to an analysis of a concept *C* is consider the case described (usually a hypothetical case) and note a mismatch: either it satisfies the concept but not the analysis, or else it satisfies the analysis but not the concept. Yet whichever way it goes, the processes involved are again just categorization processes. So deflationists shouldn't be troubled. By appealing to psychology's best models of categorization, they can explain our ability to appreciate counterexamples.

Finally, consider our ability to formulate prospective analyses of concepts. What one usually does is start with a few examples that are thought to satisfy the concept to be analysed. Then one tries to come up with a provisional analysis, followed by counterexamples, modifications, more counterexamples, and so on. Arriving at good counterexamples and good provisional theories and modifications requires a great deal of creativity. Unfortunately, intellectual creativity isn't something for which psychology has any worked out and widely agreed upon models. But creativity isn't a particular problem for deflationists. The processes that underlie abductive and scientific reasoning are in the same boat, and there's little reason to think that these involve resources that are not available to deflationists.

In sum, deflationists have no more trouble than face value theorists in accounting for the fact that we are capable of generating Socratic intuitions. Socratic intuitions don't argue for face value theories any more than intuitions of analyticity do.

## VII

*Conclusion.* At one point Grice and Strawson suggest that Quine should not be taken as claiming that there is no analytic/synthetic distinction (143):

For, in general, if a pair of contrasting expressions are habitually and generally used in application to the same cases, *where these cases do not form a closed list*, this is sufficient for saying that there are kinds of cases to which the expressions apply; and nothing more is needed for them to mark a distinction.

The resulting sense in which there are ‘analyticities’ is, however, rather anaemic, as it is entirely consistent with deflationism. Deflationists can hold that there is a sense—some sense—in which *bachelors are unmarried* should be sorted into one group and *grass is green* into another. But the partitioning may do no more than divide the claims into ones that *seem* analytic and ones that don’t. This concession, if you want to call it that, is perfectly compatible with there being no analytic truths, much as the fact that the witch/non-witch distinction was used projectibly is compatible with there being no witches (Harman, 1967). What we have here is hardly the basis for a significant independent argument for the existence of analyticities. Seen in this light, it shouldn’t be so surprising that deflationists are able to mimic face value accounts of the genesis of intuitions of analyticity. They can mimic face value accounts of the genesis of intuitions of witch-hood as well.

But as we’ve pointed out, deflationists might also adopt a more aggressive stance. It may be that the distinction marked by our intuitive sense that some claims are analytic and others synthetic isn’t even about the *seeming* analytic and the *seeming* synthetic. The real source of our intuitions may be nothing more than a sense of what seems obviously constitutive or not—a distinction of little apparent philosophical interest. That we think of it otherwise, that we are tempted to describe it in terms of the analytic/synthetic distinction, may say more about the extent to which we’ve been indoctrinated by philosophy than about the nature of what we are actually intuiting. Grice and Strawson may have something like this in mind when they suggest on Quine’s behalf that the distinction might be ‘totally misunderstood by those who use the expressions, ... the stories they tell themselves about the difference are full of illusion’ (143).

In any case, the argument from the analytic data fails. The mere fact that we have intuitions of analyticity and Socratic intuitions does not provide grounds for believing that there really

are analyticities. To that extent, our intuitions should not be trusted.

*Eric Margolis*  
*Rice University*  
*Department of Philosophy—MS 14*  
*P. O. Box 1892*  
*Houston, TX 77251–1892*  
*USA*  
*Margolis@rice.edu*

*Stephen Laurence*  
*Department of Philosophy*  
*University of Sheffield*  
*Sheffield S10 2TN*  
*S.Laurence@sheffield.ac.uk*

#### REFERENCES

- Bloom, P., 2000, *How Children Learn the Meanings of Words* (Cambridge, Mass: MIT Press).
- Chomsky, N., 1959, 'Review of B. F. Skinner's *Verbal Behavior*', *Language*, 35: 26–58.
- Fillmore, C., 1982, 'Towards a Descriptive Framework for Spatial Deixis', in *Speech, Place, and Action*, R. Jarvella and W. Klein (eds) (London: Wiley).
- Fodor, J., 1998, *Concepts: Where Cognitive Science Went Wrong* (New York: Oxford University Press).
- Giaquinto, M., 1996, 'Non-Analytic Conceptual Knowledge', *Mind*, 105: 249–68.
- Goldman, A., and Pust, G., 1998, 'Philosophical Intuitions and Psychological Theory', in *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*, M. DePaul and W. Ramsey (eds) (Lanham, MD: Rowman and Littlefield).
- Grice, H. P., and Strawson, P. F., 1956, 'In Defense of a Dogma', *Philosophical Review*, 65: 141–58.
- Hampton, J., 1982, 'A Demonstration of Intransitivity in Natural Categories', *Cognition*, 12: 151–64.
- Harman, G., 1967, 'The Death of Meaning', in his *Reasoning, Meaning, and Mind*, (Oxford: Oxford University Press).
- Horwich, P., 1998, *Meaning* (Oxford: Oxford University Press).
- Jackson, F., 1998, *From Metaphysics to Ethics: A Defense of Conceptual Analysis* (New York: Oxford University Press).
- Lakoff, G., 1987, 'Cognitive Models and Prototype Theory' in *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, U. Neisser (ed.) (New York: Cambridge University Press).
- Lormand, E., 1996, 'How to Be a Meaning Holist', *Journal of Philosophy*, 93: 51–73.
- Kripke, S., 1972/1980, *Naming and Necessity* (Cambridge, Mass: Harvard University Press).

- Laurence, S., and Margolis, E., 1999, 'Concepts and Cognitive Science' in *Concepts: Core Readings*, E. Margolis and S. Laurence (eds) (Cambridge, Mass: MIT Press).
- Murphy, G., 2002, *The Big Book of Concepts* (Cambridge, Mass: MIT Press).
- Peacocke, C., 1992, *A Study of Concepts* (Cambridge, Mass: MIT Press).
- Putnam, H., 1962a, 'The Analytic and the Synthetic', in his *Mind, Language, and Reality: Philosophical Papers, Vol. 2* (New York: Cambridge University Press).
- Putnam, H., 1962b, 'It Ain't Necessarily So', in his *Mathematics, Matter, and Method: Philosophical Papers, Vol. 1* (New York: Cambridge University Press).
- Quine, W. V. O., 1954, 'Carnap and Logical Truth', in his *The Ways of Paradox and Other Essays*, revised edition (Cambridge, Mass: Harvard University Press).
- Quine, W. V. O., 1960, *Word and Object* (Cambridge, Mass: MIT Press).
- Quine, W. V. O. 1973, *Roots of Reference* (La Salle, Illinois: Open Court).
- Rey, G., 1993, 'The Unavailability of What We Mean: A Reply to Quine, Fodor and Lepore' in *Holism: A Consumer Update*, J. Fodor and E. Lepore (eds) (Amsterdam: Rodopi).
- Smith, E. and Medin, D., 1981, *Categories and Concepts* (Cambridge, Mass: Harvard University Press).