Success semantics.

How come we are so successful, unless we are hooked up right to the world? A good

question, and one that suggests a way of thinking of our hook-up to the world. Success

semantics is the result of that suggestion. It is the view that a theory of *success in action*

is a possible basis for a theory of *representation*, or a theory of *content* or *intentionality*

(throughout this paper I shall use these interchangeably). At its most simple we can think

of representation in terms of disquotation, as in the famous "Fido" – Fido relationship.

Then the idea is that the disquotation of representation is explained or illuminated or even

analyzed by the disquotation of explanation, where whatever is represented explains

something about the person representing it. And what it explains is primarily the success

of the actions that the person bases upon the representing.

The view is an heir to the pragmatist tradition. At the most general level, the idea

is that we get our way, or flourish, or fulfill our desires or our needs because we get

things right about the world. The contents of our sentences are then whatever it is that we

get right.

The ancestor of success semantics, as of so much else, is Frank Ramsey, who

wrote that it is right to talk of a chicken's belief that a certain sort of caterpillar is

poisonous if the chicken's actions were such as to be useful if, and only if, the caterpillars

were actually poisonous. "Thus any set of actions for whose utility p is a necessary and

sufficient condition might be called a belief that p, and so would be true if p, i.e. if they

are useful".[1] Ramsey did not develop the idea, and it may even be doubted whether his

chicken was thought of in representative terms at all. Perhaps it was a primitive precursor

---

[1] F. P. Ramsey, 'Facts and Propositions', p. 144.

of a representing agent. But the idea is too tempting to let lie. It was later picked up and

paraphrased, by Jamie Whyte, in terms of the truth condition of a belief being that

condition that guarantees the success of desires based on that belief.[2]

But the idea that the truth-condition of a belief would be whatever *guarantees*

success in action based on the belief meets trouble, because nothing at all guarantees such

success. X's belief may have the truth condition that Cambridge is NNE of London, and

X may act on that belief, yoking it to his other belief that the way to travel NNE out of

London is to take the first departure from Paddington. X lands up in Bristol, failing in his

desire to get to Cambridge.[3] This suggests that no fact guarantees success in action,

because even when an agent apprehends a fact correctly, there may be an indefinite

amount of other rubbish in her head, waiting to misdirect action based upon it. This is the

familiar holism of the mental. And there may also be an indefinite number of things not

wrong with the agent, as in this example, but wrong with the environment: unknown and

unthought-of obstacles waiting to trip her up.

Recently, Pascal Engel and Jerome Dokic have attempted to protect the Ramsey-

White view against these difficulties.[4] Their idea is to bring into view the whole range of

actual and possible desires that might join with a given belief, and to suggest that 'true

beliefs are those that lead to successful action whatever the underlying motivating

desires'. They quote approvingly Ruth Millikan, who says of percepts of the world:

> The same percept of the world may be used to guide any of very many and diverse
> activities, practical or theoretical. What stays the same is that the percept must
> correspond to environmental configurations in accordance with the same correspondence
> rules for each of these activities. For example, if the position of the chair in the room

---

[2] Jamie Whyte, 'Success Semantics', *Analysis*, 50, 1990, p. 147.
[3] For non-English audiences: trains from Paddington go westward.
[4] Jerome Dokic and Pascal Engel, *Truth and Success*, Routledge, 2002, p. 46.

> does not correspond, so, to my visual representation of its position, that will hinder me equally in my attempts to avoid the chair when passing through the room, to move the chair, to sit in it, to remove the cat from it, to make judgements about it, and so on.

However, although Millikan is right that a false belief (here, a percept that does not represent a position correctly) stands ready to *wreck* an indefinite number of projects, it does not follow that a true percept similarly stands ready to *guarantee* the success of an indefinite number of projects, whatever the underlying desires. 'Guarantee' remains too strong. Taken strictly, the Paddington case alone falsifies the view that my saying about Cambridge, 'it's NNE of London' represents the fact that it does. For with this desire it failed to guarantee success, yet the formula requires that a truth condition guarantee success whatever the underlying desires.

One kind of defense is that the theory works only for ideal agents, meaning ones who never believe anything false, and are (vividly) aware of any obstacle that may wait to trip them up. For such agents, perhaps a fact does guarantee the success of action based on a representation of it. But these are agents that have to be described in the first place in terms of representational successes—indeed massive, unqualified representational successes—so there will be a lurking circularity in approach the issue by restricting the relevant agents in this way.

I believe we have to recognize that a true belief will certainly *aid* an indefinite number of possible projects, but it can do this while guaranteeing none, and realistically, we must still expect a failure rate, depending, as already indicated, on what else is in the agent's head, and how cooperative the environment proves to be.

In this paper I try to show that we can do better, without in any way departing from the spirit of Ramsey's position.

Success semantics might be wedded to a theory of biological function, as in some versions of teleo-semantics. But I regard that an option, and one that we should not take. Suppose we want to say that a particular brain-state in a frog represents the proximity of a fly (and not of any old small black thing). We could work in terms of some version of this:

(A) The brain-state represents the proximity of a fly = it is an adaptation, and evolution selected it because it is triggered by flies.

Or we could work with some version of this:

(B) The brain state represents the proximity of a fly = the state is involved in the genesis of action (behaviour) by the frog, and that behaviour is typically successful because of the proximity of flies.

I shall argue in favour of approaches taking the second option. The second approach is equally a success-based approach, but it confines itself to the present, or at least the extended present in which we we can talk of what typically causes what. It is by getting flies now that the frog flourishes. Evolution can stay in the background. It provides, no doubt, the correct explanation of the emergence of the system at all. But it leaves the content of any element in the system, such as the brain-state, to hinge upon the kinds of behaviour in which it is implicated, and the kinds of function this behaviour has. This is an advantage. David Papineau, for instance, once defended the other choice, talking of a

desire's 'real satisfaction condition as that effect which it is the desire's biological purpose to produce'. But this cannot generally be right, since many beliefs and desires have contents that are too late for any evolutionary process to have selected for them, and hence for any notion of biological purpose to apply. If Amanda wants a mobile phone, we cannot talk of this being an adaption, nor of the 'evolutionary purpose' of her desire. Evolution has had no chance to act selectively on people who do or do not desire mobile phones. Of course, Amanda may be instancing some evolutionary successful and much older strategy, such as acting like the rest of her group, or acting so as to attract a mate. But those are poor bricks out of which to build a theory of content, since desires, or beliefs, with almost any content can be seen in the same light. If Amanda can be said to instance such a strategy here, she presumably equally instances it when she wants a tattoo or a Britney Spears record. But these are different desires.

There are many choices for a theory of content. It can take language as primary, or something else. It can take the individual as primary, or the language-sharing group or community. It can help itself to notions such as action, or it can regard them as too heavily involved with intentionality to be part of any explanatory project. It can be heavily marinaded in normativity, or it can try to make do without. It can be wedded to some kind of project called naturalism, or it can turn its back on any such motivation. It can take sentences, or sub-sentential parts, or larger units such as theories as primary. It can be presented as a kind of realism about the mental, or it can come in the spirit of an intentional stance or interpretative manual. It can put up with, or celebrate, indeterminacies and underdeterminations, or it can insist upon facts of the matter. My

approach remains neutral on these issues for as long as possible. How long that is, remains to be seen.

What is the problem of content, or of intentionality or representation? It is usually expressed in terms of the mind's relationship with external things and states of affairs: things and states of affairs that may or may not exist, or may be relatively near to the subject, or that may be far, in space and time. It is the fact that we can think of distant things, past and future things, or even just imagined things. Sometimes the problem is put as the problem that these thoughts, identified as they are in terms of things different from ourselves, cannot supervene upon brain states of our own. That by itself need not strike us as much of a difficulty. Many facts about ourselves would not be facts but for the relations we have with other things. But we like to be able to say what those relations are, and therein lies the problem.

The Theory
Any theory of mind that takes our representational or intentional capacities as something to be explained seems likely to work in terms of some kind of distinction between vehicle and content, and that is what I shall do. The vehicle of representation, or what Ramsey called 'the subjective factor' is usually thought of in terms of the sentence, identified by features *other* than those intrinsically connected with meaning. So it is contingent whether a sentence has the content that it does. This standard approach need not preclude a wider theory, according to which there might be or actually are other kinds of representational vehicles. For example, there may be non-linguistic vehicles, or we might want to work towards a theory in which the whole person represents things, without there being anything as it were smaller, to count as a specific vehicle at all. We come to say

something about such extensions in due course, but for the moment it will do no harm to think in terms of sentences as paradigm representational vehicles.

So consider a subject S. S gets about the world, and we suppose that some of her actions are successful. She achieves what she desires. And suppose some of her actions are based upon a vehicle V. It is not going to be easy to say exactly what that means, but at a first pass it may mean that it is because of an event, whereby V becomes *salient* in her overall psychological state. Some writers like to think in terms of a sentence, such as V may be, entering S's "belief box". Without being so literal, or geometric, we can use that as a model, again if only for the purpose of approaching a wider theory. A slightly more realistic version for humans might be that S gets into a state in which, were she to be asked why she is doing what she is, the answer would contain V as an ineliminable ingredient. As for what distinguishes one's belief box from one's entertainment box, containing vehicles of content which we entertain without believing, we should look to functionalism. We should concentrate upon *force,* meaning that a belief differs from a mere entertainment of a thought precisely in that beliefs are, as it were, in gear. They are playing a role in the machinery of agency. So for the moment we are to think of an event, which I shall call the tokening of a vehicle, precipitating a vehicle into that machinery.

In order to come at the idea of V bearing a content (being a representation, having intentionality) we think in terms of explanation. What explains S's success as she acts upon the belief expressed by "the University Library is over there"? In the typical or paradigm case, she is successful because the university library is over there. She is not, on the other hand typically successful because of the whereabouts of Heffers or Trinity College or Grantchester. Why were S and R successful in meeting this afternoon?

Because they exchanged tokens "let's meet at the university library" and the university library was where they then went. Once more, it is their going to the university library, not their going past Heffers or through Trinity that explains their successful tryst. Why was S successful in her shopping? Because she said "can I have some haddock?" and haddock was what she got. The properties of neighbouring halibut and cod are not typically relevant to the success the actions based upon that tokening.

We could at this point go directly for an attempt to describe the representative content of the whole vehicle, the sentence. We might try something like this:

A vehicle V has the content p if and only if behaviour based on V is typically successful, when it is, because p.

However, difficulties lie down that direct road. I have in mind difficulties connected with the utility of false belief, which can accrue in various ways. Consider, for instance the vehicle "I am popular". Psychologists say that this is a useful thing to get into your belief box. It promotes your ability to get on well, even if it is false (maybe, especially if it is false). So it will not be true that behaviour based on tokening this sentence is typically successful, when it is, because the subject is popular. Yet this is the content of the sentence.

It would be possible to try to handle this kind of example as Engel and Dokic do, by bringing into view the variety of possible desires that might accompany the tokening. Then, while a false content explains occasional success, only true content could explain a general pattern of success across all these possible applications. I think this may work, although it takes us some distance from our actual evidence. We have no general access to the requisite patterns. We have to invent scenarios in which the tokening of "I am

8

popular" conspires with other desires to generate a whole pattern of actions, most of which are unsuccessful if, but only if, it is false.

A different range of problems comes into view if we think of approximations. Behaviour based on tokening the standard expression of the Boyle-Charles gas law (Pressure times volume = constant times temperature) is typically successful, when it is, because of the truth of van der Waal's equation.[5] But the sentence does not express the same thing as that equation. Here expanding our gaze to take in possible but non-actual desires does not seem likely to help, since it will always be true that it is the more complex relationship between the magnitudes involved that explains success in relying upon the simpler relationship.

In addition, there will be sentences that are too seldom tokened for there to be a typical way in which behaviour based on them is successful, let alone an explanation of any such pattern of success in terms of their truth. All in all then a direct approach looks unlikely to give us what we want.

If we want to stay more closely to the evidence, the remedy must be to go compositional. We do not want to ignore the structure of the representing sentence. So let us look at reference first, and try what I shall call the fundamental schema:

(FS) Suppose the presence of 'a' is a feature of a vehicle 'a…'. Then 'a' refers to a if and only if actual and possible actions based upon the vehicle 'a….' are typically successful, when they are, at least partly because of something about a.

---

[5] The more complex equation that corrects for the finite volume of gas molecules, and the attraction between them, which are ignored in the Boyle-Charles Law.

Here we imagine a sentence containing a name. Actions are sometimes based upon it. When they are successful, this is typically at least partly because of something about some object. And that is the object that is referred to in the sentence.

At this point we may wonder why 'success' is allowed to muscle its way to the front. After all, 'a' may represent something and then actions based upon a tokening that includes it would typically *fail*, when they do, at least partly because of something about whatever it represents. The university library being far from a mile away would explain why I failed to get there, acting on the tokening 'the university library is about a mile away'. Perhaps 'action semantics' would be a better title than 'success semantics'. I think this point is right, as far as it goes. But I also think that sucesss in action is the fundamental notion: like Davidson and Wittgenstein I incline to think that failure only exists against a background of success. It is only because of our successes that the representational powers we have are adaptive, and exist in the first place. So I shall retain the title, while remembering that it is the place of representation in the overall life of an agent that is the focus.

For any actual term, there will of course be a huge variety of possible sentences in which it may occur. So the pattern of success illustrated here for any one particular sentence can be enormously bolstered by thinking of other sentences alike in containing the term 'a'—enough so that the credentials of the object a as the focus, the uniquely invariant explanans of a huge variety of doings, will be abundantly established.

Some might worry that the 'something about a' introduces something suspicious and unscientific, such as surreptitious mention of facts.[6] But that is just an artefact of the generality of the proposal. The fundamental schema collects together a pattern of

---

[6] This objection was urged upon me by Gary Kemp.

explanation, and as is often the case, to generalize in this way we need mention of truth or fact. But in any particular case, the explaining is done without aanything suspicious of the kind. Why was John's action based on tokening 'the university library is about a mile away' successful? Because the university library is about a mile away. In other words, the introduction of a vocabulary of fact or truth is necessary for theorists generalizing about the phenomenon. But it does not indicate any mysterious residue in the phenomena themselves.

Before expanding this, and confronting objections, we should notice a few points. Some are obvious enough, but others deserve separate mention.

(1) There is a disquotation, in that the name, or other feature of the vehicle mentioned is used in the sentence following 'because'. It is in this sense that the schema reproduces the idea of representation as a vehicle-world relationship.

(2) There is a 'typically' qualification, corresponding to the fact that we want to tiptoe past deviant causal chains and the like. We gain our point of entry by thinking of the typical or normal case, just as we gain our point of entry to a causal theory of, say, vision, by thinking of the typical or normal case. There may follow a choice about how seriously we should look on deviant cases.

(3) There is an attempt to accommodate the thought that sometimes we refer to things which never enter into our action plans by expanding the explanatory range to actual and possible actions. I do not plan much on the basis of a tokening of 'Henry VIII', but were I to do so, or to have done so, and

generated a class of successful actions, that would be because Henry VIII was
one way or another.

(4) The actions specified may or may not be those of the agent. A baby may say
'biccy', and the fact that the word signifies biscuits is certified by the fact that
other peoples' actual or possible actions based upon his saying are typically
successful if, but only if, they involve biscuits. Things go well just when
biscuits are provided by someone else, in response. If the baby is at a stage
where this is not so—for instance, the success rate is the same if cereal is
provided—then reference is not so specific. Of course "biccy" in the baby's
mouth may express desire rather than belief. But that is an advantage of the
account: representation is different from force.

(5) There is no necessary restriction to linguistic vehicles. The schema would
apply, for instance, to maps. If a feature of my map of Cambridge is that the
presence of a big picture of a tower, here, then that can refer to the *position of
the university library* if actions based upon that feature of the map are
typically successful, when they are, because of *the position of the university
library.* Similarly, if we are moderately realistic about mental imagery, there
is no principled objection to features of the imagery carrying intentional
content. If we reify perceptions (I do not claim that we should, or even that it
is permissible to do so), then elements of perceptions can carry content,
referring to libraries, features, distances, and so on.

(6) The schema speaks directly to *compositionality*. That is, its point of entry is a
feature of a sentence, correlating with a feature of the world. Thinking

intuitively, if we imagine an atomic sentence 'Fa', if one clause certifies the contribution made by the presence of 'a', another clause would certify the contribution made by the presence of 'F', in the obvious way (the presence of 'F' would refer to a property F if and only if actual and possible actions based upon the vehicle '…F' are typically successful, when they are, at least partly because of something about the property F.) This compositionality is put into the shop window, as it were, because it is *kinds* of vehicles that correlate with *kinds* of explanations of success, and the 'sub-sentential components', or as I prefer to call them features of vehicles, identify the kinds. Of course, our standard model would be the presence of words in a sentence, but other features of vehicles could easily have content on this methodology. I believe this aspect is actually true to Ramsey, who introduced the chicken example only as a preface to considering other sorts of belief, that is, ones expressed in composite linguistic vehicles.

(7) Although the point of entry is sub-sentential, there is no conflict with Frege's insistence on the priority of the overall vehicle or sentence. That priority can be maintained provided the idea of action being *based on* a vehicle requires a whole or unified vehicle: something like a sentence. But in this sense, a picture could *serve as* a sentence.

(8) Similarly there is no necessary conflict with a holistic view of language, for two reasons. Firstly, a feature may only be able to gain content, given this explanation, if it occurs in many different sentences, differently successful in different ways, provided there is a unifying thread in the explanation of those

successes. This will be the various properties some thing has (if it refers to a thing) or the varying instantiations or lacks of them that some property has (if it refers to a property). And secondly it remains possible, for all that the schema implies, that an action can never be regarded as being based upon a single vehicle, but only on any given vehicle in conjunction with others.

(9) The notion of 'the' explanation may worry some. In the simple, point-of-entry case, we imagine something like this. A person tokens a succession of vehicles 'the university library is over there', 'the university library is a mile away', 'the university library contains books' and so forth. He performs acts based upon these tokenings and is successful in some typical ways. Then the idea is that there is no 'total explanation' of the success of the first that fails to include the position of the university library relative to the subject, or of the second that neglects its distance from the subject, or of the third that neglects the fact that it contains books. Although in particular cases we might choose to emphasize something else, these facts will merely have been suppressed. They would need to be cited in a full story. Equally, we might wish to stress the differential or contrastive nature of explanation, in order to avoid the outcome that we are always referring to the presence of oxygen or the continuation of the gravitational field—things that are background general conditions of success.

(10)   The explanations in question need not be causal. Reference need not be confined to items that are causally anterior to the tokening of the vehicle. Actions successfully based on tokens such as 'tomorrow will be wet';

'tomorrow will bring the examination', and so on, may typically be successful

because the day after their tokening is wet, or does bring the examination. In

that case the day after the tokening is in good standing as the content or

reference introduced by 'tomorrow' as an element of the vehicle. Abstract

objects can be referents, insofar as (say) nineteen being one thing or another is

the explanation of the success of action based on the vehicle 'nineteen…'.

Reference to complexes such as aggregates or species and kinds clearly

follows on seamlessly. Actions based on the vehicle 'crowds are dangerous'

are typically successful, when they are, because crowds are dangerous, and

similarly for sheep being tasty or water being wet. Indeed it is the very

promiscuity of explanation, and its Protean capacity for covering all kinds of

topics, that largely explains the failure of causal theories and other attempted

naturalistic reductions of semantic notions (there is a comparison here with

similar frustrations in defining knowledge in other terms).

(11)   On the other hand, we could draw back at some putative cases of

reference. Can one refer to non-actual possible worlds, for example? That will

depend on how we can explain the success of actions based on putative

mention of them. If such explanations can cite the way possible worlds are, as

the explanans of this success, then reference is saved. But if this is not so, then

the referential credentials of the terms is put into question. I myself have grave

doubts whether useful explanations of our propensity to modality can take this

shape. If those doubts are well-founded, then by the fundamental schema there

is no such thing as reference to possible worlds. The schema does not solve our problems here, but it helps to pinpoint them.

(12)   The fundamental schema fleshes out the thought that are doings are successful because we are hooked up rightly to the world. This kind of formula may ring alarm bells, implying to some people an 'Archimedean point' or God's eye view whereby we Stand Above and Behind our own theories and applaud them for their real contact with Elements of the World. But this fear, whatever it amounts to, is in any event groundless. The explanation of our success that we give, when we cite the university library being one way or another, is not the offspring of some transcendental, Archimedean viewpoint. It is an explanation from within. It is no more mysterious than the way the university library blocks the view or costs money. These are things the library does, and there are others, and these are amongst them. The way it is in various dimensions sometimes explains the success of human doings, based on tokenings. (I should say myself that this 'deflationist' stance also explains what force there is to the 'no miracles' argument for realism about scientific theories. It is not that there is a metatheory, called realism, required to explain this success. It is that just as science explains pressures and temperatures, so it explains the successes of actions based on reference to those pressures and temperatures, and so on for the other theoretical elements of science).

(13)   Many people hold that representation is somehow essentially normative, and that this normative dimension is, fatally, missing from naturalized

accounts such as that of the fundamental schema (FS). This is difficult terrain, but at least *standards* for normative assessment are closely implicated by FS. For the notion of success is at the heart of the analysis. Where there is success there is also the chance of failure, either in a subject's state, or in the way his signals are taken by others. It will not require any other source to give us all that is needed from a notion of correctness or incorrectness in representation.

(14)   The disquotation in (FS) is one that *we* give. It is one we give when describing our own representations. So there is a sense in which, if it is the last word, we cannot stand outside our own skins—perhaps there is even a sense in which the early Carnap and others were right, that semantics is a very limited enterprise. But this does not mean that the proposal achieves nothing, or nothing more than a strictly modest or quietistic disquotational semantics does. It does not leave us with starkly irreducible notions like reference or predication, backed up, for all it tells us, by noetic rays. On the contrary, it naturalizes these notions by seeing them as applying to relational features of things we say to ourselves, or pictures or anything else we give ourselves, responsible for our success as agents acting in a surrounding world.

And so to difficulties. Some are easy to cope with, but others less so. The hardest, I believe, is voiced by Papineau. Papineau talks of Ramsey's different suggestion, criticized above. But the present proposal is just as vulnerable to the objection. Papineau complains:

> It explains truth for beliefs, only by assuming the notion of satisfaction, for desires. Yet
>
> satisfaction is as much a representational notion as truth, and so ought itself to be
>
> explained by an adequate philosophical theory of representation. [7]

So, for instance, consider our agent who wants a particular book, believes that the book is in the university library, and that the university library is in some direction from where he stands. Suppose all goes well. We can say that his success is explained by the book being in the university library, and the university library being where he expected. But his success is identified in terms of getting what he wanted, and that requires content or intentionality: he wanted a particular book, which he therefore had to represent to himself. If we cannot say that much about him, we have no reason, it seems, to talk of success at all. But to say that much requires some pre-existing representation, and that vitiates the proposal as a general account.

Should this objection silence us? It does not falsify the fundamental schema, but only suggests a limit to its utility. Yet how severe is this limit? If we were trying to give a reduction of *all* intentionality at a blow, it would be serious. But perhaps we do not have to claim any such ambition. It remains true that for any *particular* representative feature of a vehicle, we can use (FS) to give a truth-condition or account of its representative power. That account only works, it is true, by imagining the feature embedded in the psychology of an active, desiring agent. And it is true that when we turn to the fact of desire, other representative powers will be implicated. But these in turn can be explained by a reapplication of the schema. Suppose the book our agent desired was *Emma*, and suppose his desire was activated by the tokening of a representative vehicle: "I must read

---

[7] David Papineau *Philosophical Naturalism*, Oxford: Blackwell, 1993, pp. 70–1.

Emma". Then the fact that the term "Emma" represents *Emma* is given, according to (FS) by the fact that actual and possible actions based upon the vehicle 'Emma….' are typically successful, when they are, because of something about *Emma.* Notice that among these examples of success we can number the very occasion under discussion: the agent's success on this occasion arose because *Emma* was in the university library. Faced with this, it is not very clear how damaging Papineau's problem is. But in addition, we can approach it from a different angle.

Papineau's problem will probably seem most intractable if we think synchronically. We might imagine the simultaneous tokening of two vehicles, V_B carrying the content of the belief, and V_D carrying the content of a desire. And we perplex ourselves because "success" underdetermines the identity of these two things together. "Success" could consist in the belief having one content, and the desire a related content, or the belief having a different content with an accommodating difference in desire, and so on indefinitely. Underdetermination stares us in the face.

But suppose we think a little more diachronically. We find out what baby wants by finding what brings peace. We could be wrong: baby may have wanted a biscuit, but be pacified by a rattle. But as the days go by, typical patterns emerge. If "biccy" reliably correlates with pacification by a biscuit, we get one entry into our lexicon. If when baby seems to want a biscuit we direct his attention successfully by saying where it is, and the words we use become part of baby's repertoire, then we take them to be representing wherever it is. And so it goes, entry by entry. But at the end of the process there is only one thing to think, sometimes, about what the emergent child believes and wants. And by then representative features of vehicles are available either to enter the function of

pushing and pulling, the "desire box", or the function of guiding the actions appropriate to the pushings and pullings, the "belief box".

This solves the epistemological problem. We play off macro behaviour and microstructure of vocabulary, and just as with a crossword puzzle, clue at a time, fallibly, but eventually uniquely, a solution emerges. But does it solve the metaphysical or ontological problem? Does it tell us what representation *is*, or how intentionality is *possible*? Does it, for instance, make room for misrepresentation?

I believe so. Consider misunderstanding first. Suppose subjects S and R want to meet, and S says 'let's meet in New York' and R hears 'let's meet in Newark'. They will fail to meet. S intended R to token something with one kind of power, and he tokened something with a different kind. Instead of directing him to New York, the event set him off toward Newark. It is an event which reliably does that, because there is a feature of the vehicle (which might be 'Newark is the place to go'), and actual and possible actions based upon the vehicle are typically successful, when they are, because Newark is the place to go. On this occasion, it is not, and action will fail.

With falsity we imagine an agent whose tokenings of 'a' and of 'F' generally slot into the fundamental schema so as to compel interpretations as referring to a, and to the property F, respectively. We suppose that the (syntactic) structure (or some other feature) of the vehicle ensures its indicative form. So the subject bases action on 'Fa', interpreted as a being F.[8] In other words, he acts on the belief that a is F. Unfortunately a is not F. So either the subject will be unsuccessful, or his success will not be explicable in the typical, disquotational fashion. He is not successful because a is F, but in spite of a not being F.

---

[8] Clearly, I am assuming that 'concatenation' in a simple atomic sentence as vehicle has the consequence that the vehicle represents whatever is referred to by the name term as having whichever property is represented by the predicate.

There is no principled difficulty about isolating such cases, and saying the right thing about them.

If the theory allows misrepresentation to fall nicely into place, it is difficult to see what remains of the idea that it is at best epistemologically adequate, but failing on some metaphysical or ontological ground. However, another general problem looms.

So far we have run representation and reference together, imagining that a theory of the latter, in terms of explanations of success, will automatically give us a theory of the former. But there are difficulties here, since obviously a subject can represent without referring. This is the problem of empty names, (or predicates that fail to pick out properties, although historically that seems not to have been so worrisome). Johnny represents Santa Claus to himself (and success may attend his actions based on this representation). But his tokenings do not refer to Santa Claus, or to anything, since the explanation of his success does not consist in Santa Claus being one way or another. Johnny himself may suppose the explanation of his success to be the doings of Santa Claus, but he is wrong about that.

Maybe the typical explanation of the success of Johnny's actions, such as hanging up a stocking, are the doings of Johnny's father. So can we avoid the result that his tokenings in fact refer to his father? We may not want to avoid it: it is no accident that when Johnny grows up, one way for his father to reveal the truth is to say 'I was Santa Claus all along', or 'it was me to whom you wrote all those letters'. If we want to avoid the interpretation, we can invoke several other features of the situation. Johnny's friends all suppose one person to be the common reference of the name, but no one person explains their successes equally. Johnny's conception of what Santa Claus is like is quite

at variance with what Johnny's father is like, and although that does not preclude reference, it at least counts against it. But there still remains the question: how are we to analyse the distinction between representation and reference that the case opens up?

What we need to cope with this is the idea of Johnny's mind-set being appropriate to a Santa Claus world, although we do not inhabit such a world. We can do this if we use the notion of a dossier that Johnny associates with the tokening of 'Santa Claus': giving presents, visiting once a year, climbing down chimneys, and so forth. This dossier corresponds to beliefs Johnny has about what Santa Claus is like and what he does. The existential quantifications associated with those descriptions should not be problematic, for we have already suggested an approach to the general problem of misrepresentation. That he ties these quantifications together under the heading of Santa Claus gives us our understanding of Johnny's mind-set. It is inappropriate to the actual world, and its token 'Santa Claus' has no reference. But it does not show us a Johnny who is irrational or uninterpretable, and certainly not one whose tokenings fail of content, so that he fails to think at all. Johnny thinks, hopes, desires, is grateful just as if Santa Claus were a real person.

There is more to be said about fiction, and fantasy (I have imagined Johnny in the grip of a real mistake, not fantasizing about non-existents in full awareness that this is what he is doing). But I do not see that the phenomena will force any significant move away from (FS).

Less global difficulties may remain. One intriguing worry might be that the theory falls into a mirror-image of one problem that afflicts a causal theory. Causal theories of reference do not easily allow reference to the future. Is it possible that success based

theories do not easily allow reference to the past? For after all what causes success is a matter of what will be the case when the time to reap rewards comes, which will be the future. So, for instance, how can I refer to the present position of my car, when it is the future position of my car that will explain my success or failure as I walk in the direction in which I am prompted by some tokening?

There are two kinds of answer to this. One would point out that explanations cast a wide net, and we do not confine explanations to immediate or proximate explanations. True, it is whether the car will be in a place to which I walk that proximately determines my future success. But it is where it is now that explains where it will be (in the normal case, in which the car is stationary). The other kind of answer reminds us of the wide class of actual and possible actions. My actual actions, based on a tokening of the present position of the car, reap their rewards in the future. But *possible* actions based on the same tokening could have reaped their rewards now or in the past. Hence, there is a wider class of possible actions whose success would typically be, or have been explained in accordance with the fundamental schema, by facts about past or present objects.

I have talked throughout of tokenings as events, in which a vehicle is somehow summoned into an active area of the brain or mind: a belief box or desire box, implicated in the machinery of action. We may wish to point out that as well as episodic events like this, there are 'standing beliefs' or for that matter standing desires, which may seem to be implicated in action but with no event of this kind taking place. To accommodate this idea, I take it we can expand our conception of what it is for action to be based on vehicles. We might think of some vehicles exerting a standing pull. The words 'it is a bad idea to walk into a wall' do indeed not have to go through my mind for me to act daily on

the belief that it is a bad idea to walk into a wall. But the fact that it is *that* belief upon which I am acting has to lie somewhere. Presumably it lies in my being in a state both in which I am strongly disposed to avoid walls, and in which I am disposed to cite something like the belief mentioned as the rationale for my first disposition. In the habituated agent a tokening does not have to precede an action based on the belief that the tokening expresses. Similarly we can say that the batsman played the stroke as he did because he foresaw the flight of the ball. But we don't have to think of an antecedent mental picture with elements representing that flight. It is enough that afterwards he could produce such a picture, either mentally or on paper or in any other way.

I have talked in very simple terms of actions being based on tokenings, and some may be poised to object that this makes a mockery of the delicate space of reasons. Representational tokenings should not be thought of as pushing action in some hydraulic or mechanical way. Rather, they inject contents into the space of reasons, and whatever action emerges is only the resultant of operations within that space.

But it is a mistake to think that the simplicity of the conception is inconsistent with the complexity of our reasonings. The simplicity of the conception is supposed to take some of the mystery out of representation or intentionality. It does require a notion of basing an action upon a tokening (or background disposition to token). Such 'basing' may become more complex than any simple two-factor, desire-belief, model suggests. Actions may turn out almost never to be based upon one tokening at a time. Standing beliefs and desires complicate things indefinitely. But at the end of the day there is such a thing as basing action on belief, and expressing belief in vehicles, just as there is such a

thing as basing the direction of movement on a map. And this is all that is required to launch success semantics as a going concern.

Simon Blackburn.

May 21st 2002.