

Content, Context and Explanation

Dennis W. Stampe

Meaning is this sort of property: when you know what the meaning of a thing is, you have an *understanding* of that thing. (Not all properties are like that: for instance, size isn't: one's knowing the size of a thing isn't necessarily understanding the thing.) Now generally, and roughly, to understand is to be able to explain. To understand an event, what it might mean, is to know what it might be the consequence of, or what consequences it might have; it is to know how to explain the event itself or to explain other events with it. To understand a sentence is, perhaps, to be equipped to make sense, in that way, of utterances of it. That, in any case, is the perspective of this discussion. Understanding an utterance, in the sense of recognizing its meaning, is not a matter of grasping some especially curious object, a "meaning" hovering between the utterance and the world; it's a matter instead of recognizing how the event might be explained.

Of course there are more reasons one might utter a sentence than the sentence has meanings. So to understand a sentence – i.e., to identify its meaning – would be to identify a *certain* thing that might explain a person's uttering it: what would explain it, perhaps, if certain particular conditions were to hold. Where, say, the fact that *p* would be the only possible cause of the utterance of the sentence, supposing those restricting conditions hold, then the unique meaning of the sentence is *p*.¹ Where these conditions do not obtain, it's as likely that something else, say *q*, caused its utterance; and it's not as likely that the statement is true. The meaning of the statement would still be *p*, not *q*, if the meaning of the statement is the state of affairs that would, under the *favoured* conditions, probably hold; the probable cause, that is, were this gauged just from the holding of those conditions. (For instance, I say "It's raining"; the fact that would most probably cause this utterance, were conditions of a certain sort to hold, is: that it is raining. And that's why *that's* what the sentence means, even if what actually caused me to utter the sentence is not that fact. Indeed, the example-giving conditions that do hold, right at the moment, reduce the probability of that being the actual cause, but that doesn't alter the meaning of the sentence.)

To identify the content of a representation is to identify a *particular* causal hypothesis by means of which the representation *might* be explained. The

hypothesis in question is of course not the only hypothesis by means of which it might be explained. It needn't even be one of the several hypotheses which *do* explain it in point of fact; it needn't provide a correct explanation. But it is – unless the representation in question is ambiguous – nonetheless, some single hypothesis, among those that might explain it. Its uniqueness, then, must be that it that alone meets certain constraints. Plainly one of these constraints is *not* that the hypothesis should be true. Truth is generally a good thing in an hypothesis, but not always: it is evidently not what we want from the one that captures what a representation “means,” what it represents *as* being the case. Neither do we want the hypothesis that is most probably true – most probable, that is, all things considered. The meaning-specifying causal hypothesis is the one most likely to be true, most probable, relative to just *certain* factors. It is not necessarily the one most likely to be true relative to all factors, or all things considered.

The fallen mercury in a thermometer might be caused either by the temperature having fallen, or by the nurse having shaken it down. Some have wondered why we don't just read the thermometer that way: it says that *either* the temperature's fallen *or* someone's shaken the mercury down. We say the needle's pointing to “E” in the gas-gauge means that the gas tank is empty, although it also reads empty if you park on a hill. So why not say that the needle says that either we're out of gas or the car is parked on a hill? After all, this disjunction has a better chance of being true than the first clause of it does. So why doesn't the disjunction express the right reading? The short answer to this question is that in reading a thermometer we are not trying to cite the actual cause of the mercury's position in the tube, nor even the causal hypothesis *most likely* to be true. That is just not what we're doing. *We are*, evidently, looking for the hypothesis that would most likely explain the mercury's position on certain assumptions – here, an assumption about what has or hasn't been done with the thermometer, and the assumption that the instrument is working right. (That is, we are looking for the causal explanation that certain assumptions, or facts, would make most likely.)

Now one may ask why we want the one sort of hypothesis and not the other?² Why *are* we interested in the hypothesis that would most likely explain the position of the mercury on the particular assumption that it hasn't been shaken down? Obviously it's because of what we do with the instrument; and that depends in part on what the instrument can do, and in part on what we want from it. What we want from it is to know the temperature, and a thermometer can tell us that. If we read it as saying that either the temperature has dropped or someone has shaken it down, then we would not *be* telling the temperature from it. We read it in such a way that it *may* tell us the temperature; that is, so that if we believe what it says, then, if the relevant assumptions are correct, we'll probably be right, and right about what the temperature is. We wear watches to tell us the time, and what the watch says is that it's four o'clock, not that ^{it's} either it's four o'clock or else its battery is dead. It's not clear what good such an instrument would be. Certainly, if that's what it said it would not tell the time. The right reading is not just the hypothesis which maximizes our chances of knowing *some-*

thing – whatever it might be – if we believe what the instrument (as interpreted) says.

The short answer to the disjunction question is important, if only because it is obvious; it is important not to lose sight of it. But it is less obvious how it, or some generalization of it, might answer the presumed parallel question about *mental* representations.³ Moreover, it cannot satisfy us if we wish to press a causal account of representation into the services of semantic “naturalism” of a more ambitious sort. Here two such theses might be distinguished. The less ambitious naturalism is the thesis that an adequate account of the semantic character of language or thought will employ only modes of analysis and explanation certified in other, paradigm regions of natural science. (This is a first, fairly modest contention for one who would “naturalize” semantics. It involves no immediate promise to reduce anything to anything.) The more ambitious semantic naturalism – reductive naturalism – is the thesis that properties and relations to which adequate semantic theories (linguistic and psychological) must advert, may be identified with properties and relations which themselves involve no semantic or intentional properties. (In this vein, for instance, the primitive intentional relation between the track of the deer and the hoof it is an impression *of* may be identified with the causal relation that holds between the depression in the mud and the hoof that caused the depression.)

Our assignments of content often turn on our epistemic aims and purposes: what we want or aim to know. But this fact does not help us, if we are in pursuit of a reductive naturalism, not just in regard to such artifacts as speech and thermometers, but in regard to the semantic properties of mental states as well, such as belief. For references to our cognitive purposes obviously involve references to our intentions, and to desired states of knowledge and belief, and all of these are themselves intentional states, possessed of content; and if our account of content attribution depends on the identity of such aims, that account will not have worked its way out of the intentional circle.

On the other hand, if we are ever to work ourselves out of the thicket of intentional concepts in which we think, we mustn't turn back every time it seems we might be going in a circle. For instance: we are trying to frame an account of contents, or, of content attribution; but it should be kept in mind, I think, that content *attribution* is, after all, a human activity, and it is bound therefore to be defined by human purposes, cognitive and practical. But even granted that we interpret representations, and discern and create semantic properties, in accordance with our purposes, what matters is that *what* we attribute to them, the properties themselves, should be entirely real, natural, properties, properties that representations do really have, in their own right. It is not required, of course, that the semantic vocabulary should somehow well up from nature inevitably, but only that in its application we are describing the real properties of the representations in question. And again, when we require that these be properties that representations possess “in their own right,” this cannot be taken to mean properties that they would possess even if there were no speech or thought or interpretation, if these are

the events and processes causally responsible for the properties in question.

This essay will touch on several questions that trouble our dreams of a semantic naturalism of the second, reductive kind: questions about content and object determination, the role therein of functions and well-functioning mechanisms, and the treatment of quantification. The discussion will approach these through an exploration of naturalism of the first, (for lack of a better term) methodological, and relatively uncontroversial kind. In particular, I wish first to explore the extent to which the content of a representation may be specified simply in terms of a certain hypothesis about its explanation, the relevant hypothesis in turn specified as one that satisfies certain normal desiderata in explanatory hypotheses, of a kind known elsewhere in natural science. I suggested, earlier, that the correct interpretation of a representation specifies "the" single thing that would explain it (or result from it⁴), under certain conditions; and that this must mean that there's a uniquely *satisfactory* hypothesis: exactly one hypothesis acceptable for a certain explanatory purpose. What I wish to explore, then, is the extent to which some familiar and normal constraints on the acceptability of explanations, rather than anything peculiar to semantics, might suffice to identify this unique hypothesis. To get further down the road toward a causal theory of representation, we need to articulate these constraints that narrow the acceptable hypotheses down, perhaps even to a single one. So far as I can presently see, there should be no more doubt about the rationale of our content-assigning practices than there is about our explanatory practices elsewhere in our efforts to understand phenomena, and phenomena that are as natural as they come.

Why is it, in interpreting representations, that we fasten on the particular hypothesis we do, hedged with just such provisos and not other ones? Does nature determine that we must? It is not the hypothesis most likely to be true, all things considered, or in the actual conditions, that gives the right interpretation, but the hypothesis that certain conditions, were they to hold, would make most probable. But why just those conditions and not others? If it isn't probable truth, what *is* the peculiar virtue of that hypothesis?

A large part of the answer, it seems to me, is that it is the peculiar explanatory *power* of the favored hypothesis. (I take it that the explanatory power of an hypothesis is something that can be assessed independently of any assessment of the actual truth of the hypothesis, or the probability of its being true. The measure of explanatory power is, to put it crudely, *how much* would be explained by the hypothesis *if* it were true.) And the explanatory power in question is specific: we are interested in an hypothesis with the power to explain *certain* things, beyond just the representation immediately in question – and, to explain them provided certain particular conditions hold, or under certain constraints.

For instance, we require a hypothesis to explain not just the representation itself (or its effects), but also other tokens of the same type. (Here a principle of parsimony in explanation becomes a Principle of Interpretative Uniformity.) This constraint plays a role in the determination of the object represented, in view of (a) true and false tokens with the same content,

representing the same thing; and (b) tokens with differing contents representing the same thing. The thing that is represented (the *object* of representation) may be distinguished from the *content* of a representation – the content being, what is represented as being the case regarding that object.⁵ Now, it is of events of certain types that we ask “what might explain this?” and we may require the answer to do for all the variant tokens of the type. It can do so, if the variation is wide enough, by abstracting to some determinable property of the hypothetical cause. This is ordinary stuff in explanation: “What might explain this?” we ask, referring to some instance of a puzzling phenomenon, and we want an explanation that covers any *such* event. What explains the orbit of Earth in a way that will also explain the orbits of other planets? We may abstract to a determinable property of the cause of one token of the type – e.g., the magnitudes (whatever they may be) of the gravitational and of the planet’s centrifugal forces. What explains the ten rings visible in this tree stump? And that that one shows six rings? What’s required is an hypothesis that will explain the first in a way that also explains the second: *the age* of the respective trees will explain the number of rings in its cross-section, and it’s the age of the tree that the number of rings represents. We pay a price for this generality or simplicity. We give up a certain determinateness; we have to give up the categorical proposition and make do with hypotheticals, or take on a burden of provisos, “ideally”s, “normally”s and the like.

Such considerations work to place the proper object of the representational state at a certain distance down the causal chain from the representation being interpreted. This bears on the question, pressing in any causal theory: if the thing represented is something that might cause the representation, why does the intentional arrow stop at one link in the causal chain and not one nearer, or further back? Let it be a perception, as we say, “of a grapefruit.” Why of “a grapefruit” and not of the retinal image? Because the perception is of the same type as, say, the tactile impression of the grapefruit, and so what would cause it (and any other representations of that same type) cannot be a retinal image; candidates must apply from further down the causal chain. Proximal causes won’t do. If we ask the awful question, Why do our states of mind give representation to objects outside our minds?, I suppose (unoriginally) that this is the answer. It is because our states of mind are of types all the variant tokens of which could be given a unitary explanation only by an hypothesis that would attribute them to a state of affairs outside our minds.

We often inquire after “the” explanation of a phenomenon, of whatever kind, as if there were a single explanation of it. When we do so, we conceive that phenomenon in certain ways, in terms of certain of its properties instead of others. If we didn’t do so, there would be no one explanation. Sometimes this is explicit: what explains the *size* of the footprint? Or what explains its shape? Sometimes it’s a property of a property: what explains the size of this footprint *being bigger than* the size of that one. The importance of this platitude comes out when we consider the role of contrast, where the relational properties of events define our inquiries.

Why is the corn in this field three feet high? Because there's a drought. Or, because it's been irrigated. Both explanations can be right. The first explains why the corn is three feet high instead of the normal six. The second why it's twice as high as the corn across the road. (This does not mean there's something arbitrary about explanation or causation, of a sort to unsettle our "naturalism". There are just two different phenomena, both solid facts, here being explained in two different ways. What we want an explanation of, of course, depends on our interests; but the existence of that phenomenon does not depend on our interests; and what the explanation of it *is* doesn't either.⁶)

We often want an explanation of an event that will serve further explanatory purposes. We might want to understand other events that the one in question might cause, and therefore to understand the one in question in such a way that it would do so. We have ulterior explanatory purposes, in asking for the explanation of an event. For the causal explanation of this event (the primary explanandum) might in turn explain certain other things (secondary explananda), which we might wish to understand. For instance, the explanation of the height of this corn might explain (if such a thing were to happen) why the farmer with three foot corn goes bankrupt while the one with withered corn survived. *This* one invested foolishly in irrigating a hopeless crop; the one across the road wisely cut his losses. The hypothesis that explains the height of the corn also explains the bankruptcy. Such ulterior explanatory purposes impose a tight pragmatic constraint on the acceptability of hypotheses. If you have an eye to understanding what's going to happen to this farmer, when you ask why his corn is so high, you will need the irrigation hypothesis, not just the drought hypothesis. The requirement that a satisfactory explanation should serve such ulterior explanatory purposes is one ordinary factor that delimits the acceptability of hypotheses.

This example, incidentally, might itself be construed as a matter of the reading of signs. For the relative height of the corn may be a *representation* of one thing or another. It may represent both the amount of rainfall a field has received (which you could tell from it, if you subtracted the effects of irrigation); and it represents, at the same time, the amount of money spent on irrigation (which you could tell by comparing it to an adjacent unirrigated field). It may *in fact* represent both these things, since its height is in fact less than the height of ordinary corn in ordinary years, and greater than the height of the corn across the road. The facts are there, even if what we may care to read from them depends, as of course it does, on what we may want to try to understand. And so the way we read it may depend on what we think the height of the corn might explain, as well as what might explain it.⁷

Of course, there is nothing in the nature of the phenomenon that determines that one and not the other is the actual message of the corn. The situation is somewhat different, however, in other cases, like the bimetallic thermometer which is also a circuit-closer in a thermostat. Here, there is a possible explanation of the primary explanandum (the change in the curvature of the thermometric strip) which is also the possible explanation of a secondary explanandum (the temperature's rising back to the set point).

Ability

That explanation is that the temperature has dropped, and that explanation explains both the bending of the bimetallic strip *and* the effects of its bending – which are, its closing the circuit to the heater and raising the temperature. That explanation explains what the thermometric strip is doing in this system.⁸ We identify the content of the representation in terms of that explanation of it that would explain the effects of the representation, when the whole mechanism is functioning properly. Perhaps similarly the content of a belief is given by that hypothesis that would explain the belief, and its effects, under certain conditions: and these conditions might be characterized as those under which the presumed mechanism responsible for the production of the belief, and its effects, is functioning properly.

The mention of functions seems to cause some consternation. There are worries about whether it is Nature itself or we and our purposes that determine what these functions are; and naive Darwinism is decried, as if it is being assumed that the mechanisms of belief-determination are the best ones possible, mechanisms ideally attuned to the truth. I cannot myself quite see how it matters what determines the function of a system; nor that any Panglossian assumptions are involved. No one is saying, surely, that the reason the belief that *p* is, under optimal conditions, caused by the fact that *p*, is that the mechanism of belief fixation happens to be a truth-tracking mechanism – as if this is something that had to be arranged, if not by a providential God, then by natural selection; as if otherwise the forces of darkness might have endowed us with falsity-tracking belief fixers.⁹ The idea is instead that the reason the belief that *p* is, under optimal conditions, caused by the fact that *p*, is that the belief wouldn't *be* "that *p*" – that wouldn't *be* its content – were it not the case that the fact that *p* would cause it, under optimal conditions. This is not some coincidence that needs an explanation, either by natural selection or divine design. It is simply an attempt to say what it is for something to be the content of a representational state. There is no particular *reason* it is so, any more than there is a reason why the contents of a bottle fit so neatly into the bottle: the contents just *are* what's in the bottle. It's not that some agency, natural or otherwise, arranged it so. The content of a belief, likewise, just *is* what goes into it, given what will come out of it.¹⁰

What comprises "the" explanation of a phenomenon depends upon what properties of it are to be explained thereby; and these properties may be ones that underlie the "active" powers of the state in the production of further events. These events include *actions*, whether of a person, or of a thermostat. The principal question, then, that elicits the content of a state of mind is: What might explain this character's state of mind in such a way that his state of mind might explain his actions, and more specifically, explain certain *properties* of his actions? You might want to know what to expect next. There are *syndromes* associated with kinds of mental states, and kinds of illocutionary acts, and the content-assignments are required to penetrate through the representation to explain its typical effects. As in medicine, where various phenomena occur together – headache, nausea, mood changes – there is perhaps one cause of them all; so, in belief, there may be

vocalizations, desire-satisfying bodily movements, mood-changes, and one cause, maybe, of the whole syndrome. What explains only a single symptom will not suffice to *satisfactorily* explain even that one.

Understanding any sort of thing may require a hypothesis having explanatory power of that kind; and "the" relevant explanation may be constrained to possess such power, whether it's representations or mental states or agricultural affairs that concern us. This is of peculiar importance in explanation by reasons. Our subject is in a state of belief, and we wish to find its content. Under some conditions, or with some probability, beliefs, together with desires, cause actions that satisfy those desires. (This is the familiar "syndrome" of practical reasoning: beliefs, desires, and acts that satisfy those desires run together; there is a way of understanding, explaining, beliefs that may explain the whole pattern.) Suppose the belief in question, and the desire to eat a peach, cause the subject to eat the object in his hand. What might the content of that belief be? Well, suppose that eating the object in his hand has satisfied the desire that caused it (i.e., to eat a peach). Now suppose the content of the belief is the state of affairs that would explain not just the belief, but the effects thereof - specifically, explain its effect (his eating the thing) being such as to satisfy the desire that was its partial cause. Whatever would explain *that* would *be* (on the present view) the content of the belief; or rather, would meet one qualification. What would explain it is the fact that the thing in his hand was a peach. That's what he believed; the content of his belief, was that the thing in his hand was a peach. Notice how the hypothesis is constrained by the requirement that it have this sort of explanatory power. If we didn't require the hypothesis explaining the belief also to explain the desire-satisfying feature of the ensuing action, there are many more hypotheses that would suffice.¹¹

To put it in the language of reasons, the content of a belief, *p*, must itself comprise a *reason* for the agent to do what it might cause the agent to do, given some desire. But the fact that *p* wouldn't be a reason to do anything, unless it would make the action satisfy that desire. The fact that there's a peach in his hand will make his eating what's in his hand satisfy his desire to eat a peach. If it wouldn't do so, his believing what he does would not comprise any reason to eat the thing in his hand. Notice, for instance, that the fact that it *looks* to him as if it's a peach would not do so. That would not explain why eating the object satisfies the desire to eat a peach, which is different from the desire to eat what looks like a peach.

So far I have relied on the desire-satisfying effects we take beliefs to have, and we can it seems explain our semantic descriptions of beliefs in terms of certain hypothetical explanations of the beliefs that explain the desire-satisfying properties of the acts caused by those beliefs. This of course still leaves us deep in the intentional circle, since the satisfaction of a desire is a *semantic* property of it, closely analogous to the truth of a belief. So are the prospects for *reductive* naturalism in any way brightened by this discussion?

Perhaps. Perhaps we are closer to the edge of the circle than we were. For we may look one step further down the causal road, and include in the syndrome that determines content assignments, further regular

effects – effects of the satisfaction of desire – that extend right out of the intentional circle. These effects might include the enhancement of the well-being of the organism; or would-be effects that would be conducive to its well-being; or the satisfaction of needs.¹² Or more simply, we may consider the eating of a peach itself, that event itself, leaving aside its capacity to satisfy the desire that caused it. There is nothing semantic or intentional about this effect: ingesting a peach is a wholly natural event. The belief might be caused by the fact, among others, that there is a peach in the person's hand; and it might also be explained by the fact that the person is hallucinating; that is, the belief is such that either might cause it. But now the hypothesis that would explain the belief in terms of the fact that there is a peach in hand would explain this *effect* of the belief, the eating of the peach, whereas the fact that the person is hallucinating a peach would *not* explain that fact. If the peach eating is an effect of the belief, the former hypothesis has greater explanatory power, *ceteris paribus*, than the latter. One question was, why not attribute to the belief the disjunctive content: either it's a peach or a peach-hallucination? The suggested answer is that that would be to describe the belief in terms of an hypothesis of *lesser* power, one incapable of explaining a certain secondary explanandum, here a nonintentional effect of the belief, the actual eating of a peach.

Now the actual or formal property therein attributed to the belief, when we say it is the belief that there is a peach here, is the same property that makes it possible for an hallucination to produce it. So what are the causal powers of this property? It has the power to result in the ingestion of peaches. Or, in another circumstance – if an hallucination produces it – to yield another effect, the agent's taking a bite of air. We describe the property in terms of the former effect, not the latter. But that's all right. Being red is a natural property; but when we describe a substance as red we do so with reference to the effects it would yield under certain conditions – effects on human visual consciousness – notwithstanding the fact that under other conditions the same property would have other effects, as it would in the visual consciousness of something color-blind, like a dog. To focus on the fact that a representation could result from one thing, and could in turn cause another, is not to gainsay the fact that it could have had some alternative cause and effect. The fact upon which we focus remains, and it serves to describe a perfectly real property of the representation.

The content of a belief, *p*, must itself comprise a *reason* for the agent to do what it might cause the agent to do, given some desire: that is, the *fact* that *p* must comprise such a reason. This fact would explain the action, *via* the belief which contains that fact, if the agent is reasoning from that fact. Thus the belief has the content it does only because of its inferential or logical relations to other possible beliefs, desires and intentions, and actions. For the mechanism responsible for its production is an inferential mechanism. According to the present view, the content of a belief is identified by the hypothesis that would explain that state of mind, given the capacity of that state of mind to cause, in turn, certain secondary effects, other beliefs, intentions and actions, provided that that mechanism is functioning properly.

A plausible suggestion is that inference, or reasoning, has the general function of *maximizing* the range of true (or reliable) representations, generating new or revised representations the truth of which would be certain, or probable, while keeping the chances of error acceptably low. (Practical reasoning can also be brought under this characterization: if the desire that it be the case that *p* leads by practical inference to an action which would satisfy, or make that conative representation true, then rational action also serves to maximize truth among mental representations.) The maximizing function of reasoning is regulated, as we might think of it – enhanced and controlled – by that property of mental states that is reflected in the way the sentences expressing their contents are *quantified*, as the implicational powers of beliefs expressed by universally quantified sentences differ from the logical powers of existentially quantified beliefs. I wish to turn now to a question whether a causal approach to content-determination is competent to deal with such matters.

This thing is a peach; in view of that fact, I decide to eat the thing. The fact could explain the action if e.g. the general belief that peaches are edible may be attributed to me, as well as the desire to eat something. Let us focus on the general belief. Can the causal account of content determination explain why it should have the content it does?

The problem is this.¹³ It is proposed that the content of a belief is the fact that would cause that state of mind under certain conditions such that, were they to hold, the hypothesis of maximal explanatory power, relative to certain possible effects of the state, would be true. And if these effects include the act of eating something edible, then (given the belief that this is a peach, and the desire to eat) one fact that would do that is the fact that all peaches are edible. A credible causal account of content determination must explain, in a realistic way, how, under the favored set of conditions, the fact that all peaches are edible might have caused a state of belief – so that its content would *be* that all peaches are edible. If the account is to be realistic, the conditions in question must characterize the real mechanisms that determine our beliefs, taking us from one perceptual or belief state to another; this is the mechanism, presumably, of inference, and the conditions in question are those in which these mechanisms would be functioning properly. But what might cause us to believe that all peaches are edible? Not the “observation” of all peaches, i.e. each and every one of them; it is not that the edibility of each has been individually ascertained, or has impinged, causally, on the subject’s consciousness.

So far, the fact that all peaches are edible is a cause only at a distance from its alleged effect, the belief that all peaches are edible. The distance must be bridged by the machinations of inference, possibly inductive inference. Our thesis seems to require that if the inductive machinery of belief fixation has functioned properly, then the cause of the belief in question might be the very state of affairs that would make it true: all peaches being edible. But good inductive inferences to the conclusion that all *bs* are *f* need not proceed from a belief about each and every *b* to which the predicate *f* is extended, in the resulting judgement. Inductive mechanisms would afford no advantage

otherwise; the thinking organism that waits to observe every instance will not survive the wait.

An approach (Berkeleyian, perhaps, in spirit) that would be natural to pursue at this point – given the emphasis in the present discussion on the capacity of meaning-fixing hypotheses to explain the *effects* of the representation in question – is one which would favor the seemingly excessive hypothesis that the belief is caused by its truth condition because this one would explain the belief's effect on behavior: specifically, that the believer would be disposed by it (under certain conditions) to eat *any* peach. But I will instead go another route.

It may be noticed that an assumption of the problem, as stated here, is that the meaning of the formative “all” must be traced to the hypothetical cause of the very representation under discussion – e.g., of the utterance of that very sentence, “All peaches are edible,” or its counterpart in thought. One might suppose that a causal theorist is obliged to hold this:

[A] for every meaningful sentence of the language, there is a causal hypothesis about its generation under conditions of well-functioning that would postulate the truth-condition of the sentence as its cause.

The case in point, and other cases as well, should make us wary of such a commitment. But can it be avoided?

Something that might make one uneasy about this assumption in any case, is that it seems oblivious to the compositional character of sentences and presumably of their mental counterparts. That is, the fact that sentences are composed of constituents of fixed meaning, and that the meaning of the whole sentence is in some way built up out of the meaning of the constituents, has nowhere played a role in the account of content determination. But surely it should. What determines that a sentence has the meaning it does, is that its constituents have the meaning they do, and that they stand to one another in the syntactic relations they do. A causal theorist, then, should expect that what determines the content of a constituent will be the causal circumstances of that constituent, not necessarily in any given sentence, but, perhaps, in some particular sentence, or sentential context. But if so, then it may be that there are certain *definitive contexts* within which a formative occurs, such that the cause of its occurrence, in that context, under ideal conditions, fixes its meaning, not just in that context, but in every other context as well – including contexts in which the element occurs, also under ideal conditions, but in which the cause of that sentence would not be the state of affairs that would make it true.

We are focusing on the meaning of the formative “all”, and our question is, to put it bluntly, why is it, according to the causal theory, that “all” means *all* and not *some*? For given the observation of the edibility of a fair sample of peaches, the appropriate causal impact of *some* peaches is precisely what would suffice to cause the perfectly well-functioning inferential machinery to install “all” in the context “__peaches are edible.”

Of course the meaning of the adjective “all” is the same, regardless what

variety of inference might be responsible for its deployment: that is, if the thought "All peaches are edible" were to result, as it might, not from induction from particular cases, but from deduction from the thought that "All fruits are edible," its content would be exactly the same. Given that the conditions under which an inference is properly made differ from one kind of inference to another, it seems hopeless to try (as thesis A would require) to characterize the uniform semantic contribution of "all" with reference to some uniform causal prerequisite of well-made inferences. For if we generalize across all inferential contexts, our characterization of the background cognitive conditions, or fidelity conditions, will be nearly vacuous; so vacuous that it not be capable of constraining the relevant hypotheses in the way the theory requires.

The causal theory asserts that there is a unique causal hypothesis that fixes the meaning of any representation or significant constituent thereof. But it need not assert that that hypothesis might explain (in the same sort of way) every well-made occurrence of that representation. It may be, instead, that that hypothesis furnishes the favored explanation only in a particular, *definitive context*. The leading idea, then, takes this form: the meaning of an expression is given by the hypothesis that would, under conditions of well-functioning, explain its projection *in some particular context of its occurrence*.¹⁴ And the causal circumstances of its proper occurrence in that context can then fix the meaning of the element across *all* contexts, including those to which the meaning-fixing hypothesis would have no direct realistic application.

We might try a tentative account of the universal quantifier along these lines. It might be claimed that there is a certain context in which the meaning of "all" is fixed, namely a context in which only the fact that all the instances being referred to have the property being attributed to them, could adequately and realistically explain the formation of the belief. Thus there are, on my desk, some pencils. Now suppose I think, and say: *This one is yellow; that one is yellow; and that one is yellow; these pencils are yellow*. Here the pencils being referred to as "these" pencils, have been demonstrated one by one, and the reference of the plural "these" is anaphoric, to the referent of the "this", "that", and "that." ("These" is just the plural of the demonstrative "this.") Each one is observed to be yellow; one by one they causally impinge on my consciousness in the way reported. Now I pass, by some process, to the summary general statement, "These pencils are yellow." The universal quantifier is yet to be deployed. But now I may properly deploy it: "All these pencils are yellow." We may imagine a mechanism that produces this change in the representation, in accordance with certain syntactic and inferential procedures. This mechanism will be functioning properly in installing the quantifier "all" in this sentential context when and only when the fact that *all* the items *referred* to (or perceived) – notice, no fewer than all of them – are observed to have the property yellow. *Here*, only the fact that *all* the pencils referred to are yellow, then, would cause the representation "All these pencils are yellow," if the process that generates it is this so-called "perfect" induction, and the

mechanism generating it is functioning properly.¹⁵ In this restricted context, notice, there is no difficulty about how the fact that all of them are yellow could yield this result, since the yellowness of each of them may have a causal impact.

It's a delicate question, but I do not think my description of the context, or the process, has smuggled in any illicit or question-begging use of the universal quantifier. Notice that the process which yields the universally quantified judgement from the exhaustively observed cases should not be thought of as a process of deduction. For from "This, that, and that are yellow" one cannot validly deduce that "All these are yellow," without a universally quantified premise, "'These' refers to nothing other than the things referred to by 'this', 'that', and 'that' - i.e., those are *all* the ones being referred to." Neither is the process in question induction of the normal "ampliative" sort, where one passes substantively beyond what is given. In this process of "perfect" induction, the truth condition of the inferred statement does not differ from that from which it is inferred, even if their meaning, or what they say, does differ.

This suggestion about the universal quantifier would need serious work before anyone should seriously assert it. But one can, I think, assert this. What a causal theorist is committed to is *not* the view that [A] there exists for every meaningful sentence of the language, a causal hypothesis about its generation, which would postulate its truth condition as the ideal cause of its utterance. The commitment is more nearly to this claim:

- [A*] (1) There exists, for every meaningful sentence S of the language, an analysis into its significant constituents, such that for each such constituent C, there exists a sentence S* of which C is a constituent, and there is a causal hypothesis about the generation of S*, in some context X, which would postulate the truth condition of S* as its ideal cause; and, (2) this circumstance - i.e., that there is such a causal relation between S* and that condition - fixes the meaning of C across all the sentential contexts in which C may occur.

The second clause of this thesis will need support. The compositionality principle will impose a general presumption that formally indistinguishable elements bear the same meaning in distinct syntactic environments. But this presumption of interpretative uniformity will not be enough. It will not be enough for the causal theorist merely to find some context or other in which he can plausibly equate truth-conditions with ideal causes. It will have to be plausible, in addition, to hold that that context is one in which the meaning of the element might actually be determined for *all* contexts. It may be relevant if that context is one in which the meaning of the element might be *learned*. And one might argue that the learnability of languages imposes a further constraint on acceptable interpretive hypotheses.¹⁶

The meaning of elements will not differ, presumably, merely because the causal or epistemic contexts of their occurrence are different. But that presumption is defeasible, and it is a fair question when it is defeated.

Suppose a meaning for the element "all" is fixed in contexts of perfect induction, and then for the first time some genius uses the term in a context requiring ampliative induction. Its meaning would presumably be the same (it would still mean *all*, not *some*), but it's a good question why that is. I should think it must have to do with the fact that ampliative induction is the same sort of process as is perfect induction – a process of inference – both subserving the same function of maximizing true or satisfied representations. And the conditions or contexts in which this process of inference can effectively fulfill this function need not be confined to the meaning-definitive context.

The foregoing problem about the universal quantifier arises from an assumption which I have rejected, that no matter what mechanism is responsible for a representation's occurrence, the cause of its occurrence would be its truth condition, if the mechanism is functioning properly. It may be, instead, that its truth condition would be the cause of its occurrence only when the representation is produced by a certain particular one of several processes that might produce it, when that particular one is functioning properly. Of course, this possibility can be exploited only if that mechanism can be singled out in a way that does not beg crucial questions.

I wish now to turn to a line of criticism which turns quite explicitly on the assumption I have rejected. Jerry Fodor, a one time but fleeting friend of the teleological version of the causal theory, more recently finds it unsatisfying. One reason he gives, is that he sees no reason to believe that there is a single function that the mechanisms of belief fixation are designed to serve. "[Why] should we believe," he asks, "that the mechanisms of belief fixation are designed always to deliver truths?"¹⁷

But why we should have to believe this? What we are required to believe might instead be something more believable: namely that there is *one* mechanism with such a function – that is, one such that when *it* is functioning properly only a certain state of affairs would cause it to produce a given representation; that that state of affairs is (consequently) the content of that representation; and that this syndrome determines the content of other representations of the same type, even when they are produced by other mechanisms, with distinct conditions of proper functioning.

Maybe, Fodor says, there are mechanisms that do not function to deliver truths at all, but falsehoods instead. He imagines one that might serve to repress unbearable truths. But there is an actual such case to discuss. There actually are, after all, two distinct mechanisms that generate representations of the visual sort – visualizations, if you will – mechanisms presumably having distinct functions, and one of them, when it's functioning properly, produces misrepresentations. I refer to the dream-producing mechanism, as opposed to the mechanisms of visual perception. Now we can imagine with Freud that the mechanism that produces dreams does have a function. Let's suppose its function is to let off steam in certain ways, and that e.g., sexual frustration causes it to produce images of oneself in flight when it's working properly; this (presumably) will be a misrepresentation of what one is doing. So, in a dream one visualizes oneself flying. Now why is *that* the content of

the representation? Its content is supposed to be whatever state of affairs would cause the representation if the mechanism that produces it is functioning properly – and that mechanism is the dream generator. If this were what fixes contents, then one would be dreaming that one was sexually frustrated, not that one was flying. (Maybe you're dreaming you're flying *because* you're frustrated, but you aren't dreaming *that* you are frustrated.) So the content of the dream is not the function-satisfying cause of it.

But all this just shows that the dream-producing mechanism isn't the imagery-producing mechanism by which the content of dream imagery is fixed. The optical system is. In the recipe which says, to find the content look for the cause of the production of the representation when "the" mechanism that produces it is working right, the reference is to a *certain* mechanism. But which one? We will not get a unique explanatory hypothesis, nor thus a unique determinant of meaning, unless we can specify the relevant mechanism. Can we do no better than to say, circularly, "It's the one with reference to which content is fixed." Surely we can. We may look, most obviously, to the *fundamental* one, the one on which the successful functioning of the other is parasitic, or dependent, insofar as its functioning depends on the content of its representations, and their consequent effects. There will be, in fact, an "asymmetric dependency,"¹⁸ in those cases where one content assignment is correct. Indeed, surely this is *why* there is an asymmetric dependency between the false, or hallucinatory representation and the veridical one. That is, the capacity of the dream generator to play the functional role it does, would presumably be dependent on the capacity of the optical system to produce such representations *because* one is flying. Otherwise, presumably, the dream image wouldn't have the symbolizing capacity that it does have, in this "Freudian" explanation. (Dreaming about sexual frustration, instead of about flying, probably wouldn't help matters.)

Now to play my earlier theme again, we are looking for the hypothesis that will explain the representation in a way that will enable us to explain certain effects of that representation. The "I'm flying" hypothesis wins over the "I'm frustrated" hypothesis because it has the explanatory power we're looking for, in which the image would explain why it is that having this dream releases the dreamer's sexual tension.

I shall sum up the main point. Suppose the schema of the meaning-fixing explanatory hypothesis is this:

Under conditions of well functioning of mechanism M, working in context X, its being the case that p would explain the occurrence of the representation R.

The possible explanans, p, specifies the meaning of R, in *every* context in which M produces R. I have pointed out that a proposal of this type does not imply that that explanans, p, is capable of *explaining* the occurrence of R in any context other than X. It may be that p could not possibly explain it in some other contexts. But this is not to say that the meaning-fixing hypothesis is irrelevant to the explanation of occurrences of R outside the definitive context X. I am inclined to think it cannot be irrelevant. But its relevance is

not that "p" specifies the explanation of R (under favorable conditions). Rather, its relevance is this. The reason the mechanism M will produce R in contexts other than X may be, that *in X*, the cause of R would be the fact that p. For instance, the reason that inference may lead us to the belief that *all* peaches are edible in a context requiring ampliative inference is that in the context permitting perfect inductive inference, that same mechanism, operating so as to fulfil its maximizing function, would yield the belief that *all the peaches one has eaten* were edible, only if the truth condition of that belief caused it to do so. And the reason that the mechanism producing visualizations will produce a representation of one flying in the *dreaming* "context", may be that what would explain the action of the mechanism producing such representations in the *perceptual* context would be, the fact that one *is* flying. When you are dreaming, of course, it is not the one's flying would explain the visualization of oneself flying; not likely. But what would explain it is that one's flying would explain it when you are seeing what you are doing. The unique explanatory power of the meaning-fixing hypothesis, then, is manifested not just in its capacity to identify the explanation of secondary explananda, but also in its capacity to explain other tokens of representations of the same type, appearing in other contexts of various kinds, produced by the same or other mechanisms belonging to such an integrated system.

/st

Notes

A version of this paper was presented to the Conference on Information Theoretic Epistemology and Semantics, held in Tepoztlán, Mexico, in August 1988. The paper has been revised in the light of those discussions, and in particular, of remarks of Dan Dennett, Dick Grandy, Hartry Field, and Brian Loar. I am also indebted to Melinda Hogan for helpful remarks.

- 1 This idea is proposed in "Show and Tell," in *Forms of Representation* eds B. Freed, et al., (North Holland, 1975), and in "Towards a Causal Theory of Representation," in French, Uehling, and Wettstein (eds), vol. 2, *Midwest Studies in Philosophy*, (Minneapolis, 1977), and it is discussed further in "Verificationism and a Causal Account of Meaning," *Synthese* 69, (1987).
- 2 But first, Why *shouldn't* we want the one we do? Why should we want the disjunctive one? Because it's more likely to be true? No, the question is: why should we want the one more likely to be true. Surely it depends on what we aim to do with the hypothesis. (Or is the thought that what is *really* more probable is what the symbol *really* means? Then "real" meaning becomes Grice's "natural" meaning. But why should we say that? If that's the thought, it's just wrong: what the symbol really means is *not* what is "really" more probable.) For the "disjunction-problem" see Fodor, e.g., "Semantics, Wisconsin Style" (*Synthese*, 1984) p. 240, where the problem is said to arise from this "key assumption . . . that, *ceteris paribus*, if the correlation of a symbol with a disjunction is better than its correlation with either disjunct, it is the disjunction, rather than either disjunct, that the symbol represents. This is a sort of 'principle of charity' built into causal theories of representations: 'so construe the content of a symbol that what it is taken to represent is what it correlates with *best*'." But this is no charity, and

it is not built into a causal theory of representation unless the theorist in charge is committing sabotage. Understanding an utterance is not making an inference to the best causal explanation of it, unconditionally, but rather making an inference to what would be the best causal explanation of it provided certain conditions hold.

- 3 *Perceptual* representations may be said to function, like utterances, to tell us things, or like pictures, to show us things; and we perhaps construe them accordingly. But a belief or thought is not so plausibly described in such terms; nor is it so plausible to say that a belief *is* interpreted, or construed, at all, by the subject whose belief it is. But beliefs do presumably play some functional role in the life of the subject – a belief, like a thermometer, does, or can do certain things; it can, in its way, inform our thinking, and decision making, and action, and we may plausibly take these capacities to define its content, as the capacity of thermometer to inform us about the temperature determines the content of its representations.
- 4 Throughout, I will ignore such cases for the sake of simplicity – cases such as the order, “You report at dawn.” Here the meaning or truth condition – that you report at dawn – is something that would, ideally, *result from* the utterance, and be explained *by* it, not something that would cause or explain the utterance. Similarly, the content of an intention is an act that would result from and be explained by the intention, not the other way around. Any formulations in what follows that fail to embrace these cases should be understood to be abbreviations for one that would.
- 5 The specification of the *object*, as I understand it, is a specification of what would explain the representation, i.e., something the actual character of which actually explains the character of the representation – as *the temperature* explains the height of the mercury column – though not necessarily in the way the *content*-specifying hypothesis – *the temperature is 70°* – would explain it. Cf. “Show and Tell,” *op. cit.*
- 6 To put it another way: it is only relative to one thing or to another that the height of the corn can be understood. But to say this is just to say that it is only one relational property of it, or another relational property that can be explained. Both relational properties are perfectly real. There’s no “relativism” in the offing. Needless to say.
- 7 These agricultural phenomena do not fall into any ready taxonomy that gives focus to these considerations. But the representational phenomena that mainly concern us here, utterances and states of mind, certainly do. They are states of *kinds* that have names: perceptions, beliefs, desires, intentions, etc. Or they are events, acts, utterances of certain kinds: decisions, statements, promises, commands. These taxonomies define the presumed explanatory powers of the states, or acts, in question. The notion that such a state is *representational* in character, is the notion that the explanatory power of particular states of that kind, different *beliefs*, say, is further defined, differentiated, by a further property of the state – its having the content it does. It is this property that’s given by the relevant interpretive hypothesis, about the explanation of some particular belief. (Of the belief that it’s raining, or that it’s snowing.)
- 8 There are other things represented by the curvature of the bimetallic strip, e.g., the difference between the co-efficients of expansion of the two metals. The degree of its curvature can be explained accordingly, as well as by the temperature’s having dropped. That is, the fact that strip A bends more acutely than another strip B at a given temperature will be explained by the fact that A is

composed of metals having co-efficients of expansion which differ more greatly than do those of the metals of which B is made. But this fact does not explain the thermometer doing what it does in the thermostat, or the temperature's being kept constant.

- 9 Lucky for us we know about natural selection: otherwise our semantics would have to be theistic instead of naturalistic.
- 10 There is a question, of course, how there happens to be such a mechanism, one such that under certain conditions, only certain facts would cause it to do certain things, e.g. things that have a beneficial effect, under certain other conditions. The answer to that may be: it was so engineered; or it was the product of natural selection. Of course the function of the mechanisms in question are naturally determined, owing to the fact that they have an effect on the organism that is conducive to their own continued existence. But it might be noted that in a certain way this just doesn't matter to the truth of naturalism. Naturalism requires that our semantic descriptions may be true, when they are true, owing to the natural facts of the matter: owing e.g. to the fact that that p *would* under certain conditions cause the state of mind that *would* under certain conditions cause a certain action. But it doesn't matter *why* those happen to be the facts of the matter, whether owing to a natural process or owing to human design, or owing to troops of fairies or angels. That's of historical interest only.
- 11 The property of the peach that would explain the property of the belief that would explain the desire-satisfying property of the act is its being a peach, not (say) its being nearly the shape of a baseball.
- 12 There are explorations in this vein in my essays "Need," *The Australasian Journal of Philosophy*, June, 1988, and "Defining Desire," in *Ways of Desire*, ed. Joel Marks (Chicago, 1987).
- 13 Compare Christopher Peacocke's presentation of this problem, in his contribution to this conference. His text is a formulation from Ruth Millikin: "a belief has p as its content just in case it is the natural function of the belief-manufacturing mechanisms to produce that belief when and only when it's the case that p." The problem is bound up, in his discussion of it, with a thesis about the role of functions in content determination, and a thesis about the analysis of functions. I wish here to separate these matters from the basic problem, which is of course a variation on an old problem for externalist, e.g. empiricist, theories of representation. (Cf. Locke and Berkeley on general "ideas.")
- 14 Such a thought should be familiar from considerations about illocutionary force, or the context determining what speech act is being performed; it is not a novel notion that, e.g., the context of *assertion* is somehow basic to fixing of the meaning of indicative sentences, not just in the context of assertion but in other e.g. story-telling, or command-giving contexts as well, where the function of the act, and thus the conditions of felicity or well-functioning, differ markedly, while the sentence "retains" the meaning it would have in assertion; and likewise when the sentence occurs as a clause in a conditional sentence.
- 15 "Perfect" (or "explicative") induction yields an ". . . assertion concerning all the entities of a collection on the basis of examination of each and every one of them. The conclusion sums up but does not go beyond the facts observed." This is from Curt Ducasse's entry "Induction" in *Runes' Dictionary of Philosophy* (1958, Ames, Iowa).
- 16 But the exact relevance of learning to questions about meaning determination is a complicated one. Fred Dretske places quite a heavy burden on this factor in *Knowledge and the Flow of Information* chapter 9 (M.I.T., 1981).

- 17 *Psychosemantics* (M.I.T., 1987) p. 105.
- 18 The phrase is Fodor's (loc., cit. 108). Fodor (105) represents the present problem as follows: we are trying to define truth conditions in terms of a concept of optimal circumstances, and optimal circumstances in terms of the teleology of the mechanism that produces the representation in question. He notes that there needn't be any such thing as "the" single mechanism of that sort, and some such mechanisms might not be attuned to the facts, as his imagined mechanism of repression is not: ". . . there's no obvious way to weed mechanisms of repression [etc.] out of the definition of optimality unless we can independently identify them as mechanisms of repression; viz., as mechanism tending to the fixation of beliefs that are false." But this is not so. We need only find some unique property of the relevant mechanism, and it is not at all obvious why this cannot be one in terms of dependency relationships, of substantially the same sort, indicidentally, as those Fodor depends on in stating his own positive view.