

From Francisco Garzon and John Symons (eds.) The Routledge Companion to the Philosophy of Psychology, London: Routledge, 2009: pp. 251-279.

This is the penultimate draft; please quote only from the published version.

Problems of representation II: naturalizing content

Dan Ryder

University of British Columbia

<3>Introduction</3>

<4>The project</4>

John is currently thinking that the sun is bright. Consider his occurrent belief or judgement that the sun is bright. Its content is that *the sun is bright*. This is a *truth-evaluable content* (which shall be our main concern) because it is capable of being true or false.¹ In virtue of what natural, scientifically accessible facts does John's judgement have this content? To give the correct answer to that question, and to explain *why* John's judgement and other contentful mental states have the contents they do in virtue of such facts, would be to naturalize mental content.

A related project is to specify, in a naturalistically acceptable manner, exactly what contents *are*. Truth-evaluable contents are typically identified with abstract objects called "propositions," e.g. the proposition that *the sun is bright*. According to one standard story, this proposition is constituted by further abstract objects called "concepts": a concept that denotes the sun and a concept that denotes brightness. These concepts are "combined" to form the proposition that *the sun is bright*. This proposition is the content of John's belief, of John's hope when he hopes that the sun is bright, of the sentence "The sun is bright," of the sentence, "Le soleil est brilliant," and possibly one of the contents of John's perception that the sun is bright, or of a painting that depicts the sun's brightness.² This illustrates the primary theoretical role of propositions (and concepts). Saying of various mental states and/or representations that they express

a particular proposition *P* is to pick out a very important feature that they have in common. But what exactly is this feature? What are propositions and concepts, naturalistically speaking? Having raised this important issue, I will now push it into the background, and focus on the question of how mental states can have contents, rather than on what contents are, metaphysically speaking. (That said, the most thoroughly naturalistic theories of content will include an account of propositions and concepts – compare the thoroughly naturalistic Millikan [1984], for instance, with McGinn [1989].)

Whatever the ultimate nature of contents, the standard view among naturalists is that content is at least *partly* constituted by truth conditions (following e.g. Davidson [1967] and Lewis [1970] on the constitution of linguistic meaning). This review, then, will focus on naturalistic accounts of how mental states' truth conditions are determined. That said, "content" is clearly a philosophical term of art, so there is a large degree of flexibility as to what aspects of a mental state count as its content, and therefore what a theory of content ought to explain. For example, is it possible for me, you, a blind person, a robot, a chimpanzee, and a dog to share the belief "that the stop sign is red," concerning a particular stop sign? Clearly, there are differences among the mental states that might be candidates for being such a belief, but it is not immediately obvious which of those differences, if any, are differences in *content*.

It seems that contents pertain both to certain mental states (like John's judgement) and to representations (like a sentence).³ It would simplify matters a lot if contentful mental states turned out to be representations also. This is a plausible hypothesis (see Chapters 7, 10, 17, and 23 of this volume), and almost universally adopted by naturalistic theories of content. On this hypothesis, the content of a particular propositional attitude is inherited from the content of the truth-evaluable mental representation that features in it. What we are in search of, then, is a naturalistic theory of content (including, at least, truth conditions) for these mental representations, or in Fodor's (1987) terms, a "psychosemantic theory," analogous to a semantic theory for a language. Soon we will embark on a survey of such theories, but first, a couple of relatively uncontroversial attributive (ATT) desiderata.

<4>The attributive desiderata</4>

On one view, we begin philosophical study with an a priori grasp of our subject matter; in this case, mental content. Our task is then to elucidate exactly what it is that we have in mind by coming up with a list of *a priori* accessible necessary and sufficient conditions for something to, say, have a particular mental content. But the naturalistic philosopher need not accept this *modus operandi* in order to come up with a list of conditions that a theory of content ought to meet (if it is possible to meet them jointly). We have no *a priori* accessible definition of water, but we would be rightly suspicious of a theory that claimed none of the things we thought were water were actually water. Absent a convincing explanation for our massive error, we would rightly accuse such a theory of changing the subject. A theory of content that said we were massively mistaken about some central aspects of ordinary content attributions should receive the same treatment. This gives rise to two relatively uncontroversial desiderata for a psychosemantic theory:

ATT_{Self} – A theory of mental content ought not have the consequence that we are usually radically mistaken about the contents of our own mental states.

ATT_{Others} – A theory of mental content ought not have the consequence that we are usually radically mistaken about the contents of the mental states of others.

Any theory of content can rescue itself by rejecting the two ATT desiderata, including the theory that the content of all our mental states is *that Bozo is a clown*, or the radical eliminativist option that none of our mental states have content. Absent a convincing explanation for our radical error, any theory of content that fails to satisfy one or both of the ATT desiderata should stand accused of either being false or changing the subject. As it turns out, all extant theories have been so accused based on strong reasons. Psychosemantics isn't easy.

<3>Informational theories</3>

<4>Information</4>

At least at first glance, perceptual systems resemble measuring devices or indicators. An alcohol thermometer is a very simple measuring device, but there are of course measuring devices of much greater complexity. A gel electrophoresis unit, for instance, can measure the relative charges and sizes of DNA molecules in a sample, and a cell phone can measure the level of the local cell's carrier signal ("signal strength," typically shown on the phone's LCD display). It is natural to suppose that perceptual systems – and perhaps even the mechanisms that give rise to judgements and beliefs – are best thought of as highly sophisticated measuring devices. For instance, perhaps the auditory system measures (among other things) the loudness, pitch, and timbre of emitted sounds, while the visual system measures (among other things) the spatial location and extent of objects in the observer's view.

Measurement yields *information* (in a sense to be explained below), and information can be further *processed*. For instance, visual information about the spatial layout and auditory information about timbre could be combined to yield information about the identity of the musical instrument in front of one: a tuba. In a way, the identification of the tuba may be thought of as just another form of measurement, with graduations (regions in multivariate spaces) marked "tuba," "trumpet," "violin" etc., just as the thermometer's tube is marked "5°," "10°," and "15°." Or perhaps the identification of a tuba may be thought of as a light flashing on a "tuba indicator," preceded by a lot of complicated measurement and information processing, like an airport security door beeps to indicate metal. Concepts could then be thought of as banks of indicators preceded by specialized information processing. (Henceforth nothing important will hang on whether I talk about indication or measurement.)

No doubt you are inclined to say that the thermometer doesn't represent anything, only the person *reading* the thermometer does. Perhaps; but the thermometer can be read only because the state of its alcohol column is so intimately related to the

temperature, i.e. to the content of the representation in question. This intimate relationship between an indicator's state and its content is suggestive. The hope of the information-based theorist is that this intimate relation, assuming it can be characterized naturalistically, may ultimately serve as a reductive base for content generally – including the content of your belief when you read the thermometer.

The naturalistic relation available is that of *carrying information about* or *carrying information that*. This relation has been elucidated in many different ways. (Related views date back at least to Reichenbach and possibly to Locke and even Ockham, but modern versions are due to Stampe [1977], Dretske [1981, 1986, 1988], Barwise and Perry [1983], Fodor [1987, 1990], Stalnaker [1984], Matthen [1988], Jacob [1997], and Neander [1995]). It is beyond the scope of the present chapter to examine these in detail, but the general idea can be conveyed quite easily. It is best expressed as a relation between facts or states of affairs: for instance, the fact that the thermometer's alcohol column is 3.5 centimetres carries the information that the ambient temperature is 10° Celsius (C). (Note that the formulation in terms of that-clauses makes information *truth evaluable*.) More generally, the fact that *r* is *G* carries the information that *s* is *F* if and only if the fact that *r* is *G* *guarantees* or *makes probable* the fact that *s* is *F*. If you *knew* that the thermometer's height was 3.5 centimetres, you would be in a position to know or predict with reasonable confidence that the ambient temperature is 10°C because there exists a certain dependence between those two facts. As it is often put, *r*'s being *G* “indicates,” or “is a sign of,” *s*'s being *F*.

As Paul Grice noted in 1957, there is a sense of “means” that is correctly applied to this sort of “natural sign.” When we say that “smoke means fire,” we are using the term in this sense. However, there are a number of reasons why the contents of our mental representations cannot just be what they mean as natural signs, at least if we are to take the ATT desiderata seriously.

<4>Violating the ATT desiderata: let me count the ways</4>

<5>The specificity problem</5>

First, suppose a thermometer's alcohol column currently carries the information that the ambient temperature is 10°C. Then it *also* carries the information that the ambient temperature is between 5 and 15°C, since the former entails the latter. Or suppose a flashing indicator carries the information that a tuba is present; it will also carry the information that a brass instrument is present. Similarly, if I perceive a tuba as such, my perception will also carry the information that a brass instrument is present. Yet I need not recognize that a brass instrument is present; I may not even know what a brass instrument is. This "specificity problem" (or "qua problem" – Devitt [1981]) will obviously be quite general; if a signal carries some piece of information, it will usually carry many pieces of more general information. In such a case, the informational content of my perception does not match what I shall call its "intuitive content," i.e. the content that one would normally apply to it introspectively, or attribute to someone else in the same perceptual state. This mismatch violates the ATT desiderata.⁴

<5>Disjunction problems</5>

How strong a guarantee must r 's being G provide of its representational content, i.e. of s 's being F ? On Dretske's formulation, the conditional probability of s 's being F , given that r is G , must be equal to 1: in this case, call s 's being F part of the "strict informational content" of r 's being G . Then for the thermometer's alcohol column to carry the strict information that the ambient temperature is 10°C, the column's being at that height must *absolutely guarantee* that the temperature is 10°C. Which it doesn't: if the glass has a small hole in it, some of the alcohol will leak out as it moves, and the column's height will be at that level even if the temperature is 12°C. So it seems that the column's being at that height *doesn't* carry the (strict) information that the temperature is 10°C. Rather it carries the strict information that either there is no hole in the glass and the temperature is 10°C, or there is a hole of size x in the glass and the temperature is 10.5°C, or there is a hole of size y in the glass and the temperature is 11°C, etc. This strict informational content is a long disjunction (*very* long, considering there are many other things that might interfere with the column's height other than a hole in the glass).

Identifying our mental states' representational content with their strict informational content is not an attractive option. There are even more ways that our perceptual systems can go wrong than the thermometer can, and it would be a significant breach of the ATT desiderata to admit that our conscious visual states have massively disjunctive content, with part of that content being that our retina is not currently being interfered with (by analogy with a hole in the thermometer) and that we are not dreaming. This is an example of a “disjunction problem.”

The most famous disjunction problem is the problem of misrepresentation, where some of the states of affairs included in the long disjunction of strict informational content are misrepresentations – and so should not be included – according to the intuitive content of that representation. For example, the holed thermometer intuitively misrepresents the temperature, and (to use a famous example of Fodor's) a perceptual representation that a horse is present intuitively misrepresents things when it is caused by a cow on a dark night. These misrepresented states of affairs will be included in strict informational content, however, with the ATT-violating consequence that misrepresentation is impossible.

Any type of state of affairs that is disposed to cause the tokening of a representation will be found among the disjuncts of that representation's strict informational content. Some of those states of affairs will intuitively be cases of misrepresentation. Others do not fit that description very well, e.g. when bone thoughts are disposed to cause dog thoughts (so one strict informational disjunct for dog thoughts will be that there is a bone thought present). So the disjunction problem is broader than just the problem of misrepresentation (Fodor 1990).

<5>The distality problem</5>

A related problem faced by information-based theories is the “chain problem” or “distality problem” or “transitivity problem” (Jacob 1997: Ch. 2; Sterelny 1990: 120–21, 2002[[Please include in the ref. list]]). Consider a gas gauge: it measures and represents the level of gasoline in a car's tank. However, the level of gas in the tank is not the only

state of affairs type included among the strict informational disjuncts of a state of the gas gauge: also included are other states in the causal chain leading to the response of the gauge, for example the amount of electrical current in its lead wire. (These other states of affairs are not *alternative* causes for the representation; they will all be present simultaneously with its intuitive content – so Jacob [1997] calls this a “conjunction problem,” as opposed to a disjunction problem.) The analogs in the mental case are the various proximal stimuli responsible for a perceptual judgement, e.g. the state of the intervening light or sound waves, the state of the sensory receptors, etc. So this is yet another way in which strict informational content includes states of affairs that are not part of the intuitive content of measuring instruments, indicators, or mental states.

<4>Proposed fixes for ATT violations</4>

One legitimate way to get from the thermometer’s strict informational content to its intuitive representational content would be *via* the *user’s* intentional states. We, as the thermometer’s user or interpreter, read the thermometer as saying the temperature is 10°C, thus our own mental representations reduce the disjunction. However, an infinite regress threatens if we apply the same move to our perceptual states and other mental representations; it seems that *some* representations must not depend on use or interpretation in order to have content. These representations have *original* rather than *derived* intentionality (Searle 1983). (The interpretivist [e.g. Dennett, see below] may be seen as rejecting that apparent implication.) There have been a number of proposals for how an informational psychosemantics could “get the contents right.”

<5>Nomic dependence</5>

The theoretical tweak normally introduced to deal with the specificity problem appeals to nomic dependence. Fodor, for instance, requires that the tokening of a representation be nomically dependent on its content, e.g. the instrument’s property of being a tuba (Fodor 1990: 102). This means the representation is insensitive to French horns and trombones, but is tokened in the presence (but not the absence) of a tuba. It exhibits a

causal or nomic dependence on tubas, not brass instruments. Dretske (1981: Ch. 7) achieves much the same result by constraining the content of a mental representation to being among the most *specific* information it carries, which happens also to be the information to which it is causally sensitive (180).⁵

Fodor makes use of the nomic dependency condition in response to the distality problem as well (Fodor 1990: 108–10). In general, there is no particular pattern of sensory receptor activations necessary for identifying an object. For belief-level representations, in fact, almost *any* pattern of receptor activations at all can cause a “there is a tuba” representation, since having the thought “There is a tuba” can be highly theory mediated. To borrow his figure of speech, all one would need is a ripple in tuba-infested waters – or, for that matter, a report that there’s a tuba in any language whatsoever (even an alien one), as long as one understands it. This means that the disjunction of receptor activations that can cause “there is a tuba” representations is open-ended. Fodor reasonably asserts that open-ended disjunctions cannot participate in laws, so the open-ended disjunction of receptor activations fails the nomic dependency requirement and is not a legitimate content. Tubas (the distal item), by contrast, are. Of course, this response could only work for belief-level (i.e. theory-mediated) representation, it does not work for hardwired, modular perceptual representation. (But that is perhaps a bullet one could bite – perhaps perceptual representations really *do* mean the disjunction of receptor activations.)

<5>“Lax” information</5>

In order to (partially) avoid the disjunction problems, many informational psychosemanticists reject Dretske’s (claimed) exclusive reliance on strict information (Jacob 1997; Fodor 1998; Usher 2001[[Please include in the ref. list]]; Prinz 2002). On the “lax information” view, representation does not require that *r*’s being *G* guarantee *s*’s being *F*. For instance, one might say that the height of the alcohol column in the thermometer represents that single state of affairs type with which it exhibits the highest correlation, or the one that is most probable given that height (see e.g. Usher

2001). Usually, however, the exact nature of this relaxation from a probability of one is not fully explained.) If it is not very probable that there is a hole in the thermometer, then the disjunct “there is a hole of size x in the glass and the temperature is 10.5°C ” will not be included in the representation’s content.

A major problem for this kind of move is in identifying what counts as “the” state of affairs type with which the representation is most highly correlated. Think of the set of all states of affairs that are nomically/causally related to “there is a tuba,” and which exhibit some correlation with that representation: tuba sideways at distance d_1 in bright light, tuba vertical at distance d_2 in dim light, French horn at d_3 in dim light next to a piccolo ... an extremely large set. Isolating “the” state of affairs type that is the representation’s content involves picking out just the right subset of this large set, and in a non-question-begging manner. This is a tall order. Those states of affairs that exhibit the highest correlation will include optimal epistemic conditions (e.g. “there’s a tuba in close range in good light”), but these conditions are not part of the intuitive content (“the problem of ideal epistemic conditions”). As we move to lower correlations, we include more states of affairs that are misrepresentations (some of which may be highly probable, e.g. if a small person holds a euphonium). On top of all this, the probability of a judgement being true *depends on what is being measured or judged*. (Compare judging temperature with judging anger.) Therefore it is impossible to define content-determining levels of correlation piecemeal, representation by representation, without cheating and taking a peek at each representation’s content in advance – a circularity that violates the naturalism constraint. (See also Godfrey-Smith [1989].)

<5>Channel conditions/optimal conditions</5>

Dretske’s early response (1981) to disjunction problems was to maintain informational strictness but to make the information carried relative to certain *channel conditions*, for instance the channel condition that the thermometer’s glass tube be intact. The channel conditions for the perceptual judgement “There is a tuba” to carry information about the presence of a tuba would perhaps include good light, the subject being awake and

attentive, the absence of trick mirrors, the absence of pesky neurophysiologists injecting neurotransmitters into one's retina, etc. Different channel conditions would of course determine different representational contents for the indicator; the trick is to find some non-question-begging way of assigning the channel conditions that make "There is a tuba" carry non-disjunctive information about the presence of tubas, and thus match intuitive content.⁶ This presents a challenge because, as for the probability of a representation being true, the relevant channel conditions seem to depend on the content of the representation: recognizing tubas may require good light, but recognizing stars requires the opposite. (The problem is only exacerbated when we consider theory-mediated judgements – see McLaughlin [1987].) It seems we need to know the content of a representation in order to know which channel conditions to specify, but this violates the naturalism constraint.⁷

<5>Incipient causes</5>

Dretske (1981) pursues the "incipient-cause" strategy in order to try to isolate the intuitive content from informational content, and solve the problem of misrepresentation; a more recent proponent of this strategy is Prinz (2002) (from whom I take the term).⁸ On this view, the content of a mental representation is limited to the thing or kind of thing that caused (or, on Dretske's view, *could* have caused) the representation to be acquired. For example, although the strict informational content of a judgement might be disjunctive between "there is a Monarch butterfly" and "there is a Viceroy butterfly," if the concept figuring in the judgement was acquired through exposure to Monarchs (and not Viceroy), this rules out the Viceroy disjunct (at least on Prinz's view, if not Dretske's). While this move can help with the problem of misrepresentation (since it is plausible that misrepresented items rarely play a role in representation acquisition), it cannot rule out items that normally do play a role in acquisition: for instance proximal stimuli (the distality problem) and epistemic conditions like "in good light" (the problem of ideal epistemic conditions). (Prinz handles the distality problem by appeal to nomic dependence, but it seems he still faces

the problem of ideal epistemic conditions.)

One final point: the incipient-cause approach promises to handle the tricky matter of reference to individuals, something we have not yet considered. Two individuals can be exact duplicates, yet it is possible to have a concept that is determinately about a particular individual rather than any of its duplicates (parents will be familiar with a child who wants *that very toy that was lost*, not another one exactly the same). It seems, however, that informational and nomic relations are ill-suited for distinguishing between duplicates. Supplementing the informational theory with an historical factor like incipient causation might be a sensible way to link a representation to an individual.

<5>Asymmetric dependence</5>

Fodor's attempts to wrestle with the disjunction problems centre on his asymmetric dependence approach (original version in Fodor [1987]; later version in Fodor [1990: Ch. 4]). Fodor focuses on the sub-propositional components of truth-evaluable representations (e.g. concepts). According to the asymmetric dependence theory, the content-determining informational relation between a representation and its object is fundamental, in the sense that any other causal or nomic relations between the representation and the world depend on the fundamental nomic relation, but not the other way around (thus the dependence is asymmetric).

This approach can be made intuitive in the case of perceptual error, for example when a carrot looks misleadingly like a pencil. Since "pencil"s (the mental representations that denote pencils) can be caused either by pencils or by carrots, there must be laws connecting "pencil"s to both pencils and carrots. There's some pencil → "pencil" law that obtains because of the way pencils look, and there's also a carrot → "pencil" law that obtains because of the way carrots sometimes look – they sometimes look like pencils. That carrots can sometimes cause "pencil"s depends on some shared appearance that carrots and pencils have. Thus the carrot → "pencil" law rides piggyback on the pencil → "pencil" law, via a shared appearance. If there were no pencil

→ “pencil” law there would not be the carrot → “pencil” law. So the existence of the carrot → “pencil” law depends on the existence of a pencil → “pencil” law. However, the reverse does not hold. There could perfectly well be a pencil → “pencil” law even if, for instance, carrots and pencils did not share an appearance, so carrots did not cause “pencil”s. So although the carrot → “pencil” law depends on the pencil → “pencil” law, the pencil → “pencil” law does not depend upon the carrot → “pencil” law. That is, dispositions to commit perceptual errors are dependent upon dispositions to correctly apply a representation, but not the other way around. If you extend this to epistemic routes that go beyond shared appearances (e.g. theory-mediated routes), then you get Fodor’s theory.

Fodor’s theory spawned a small industry producing counterexamples to it (Adams 2000; Mendola 2003). Whether any of these counterexamples succeed is a controversial matter, and beyond the scope of this chapter. But one concessive response that Fodor makes should be mentioned: he points out that his theory presents merely *sufficient* conditions for content, not *necessary* ones. So if a blow-to-the-head → “pencil” law applies to someone (why not?), and this law does not depend on the pencil → “pencil” law, violating asymmetric dependence, Fodor can just say “That’s just not a case to which my theory applies – I didn’t say it applied to *all* representations.” This reduces the interest of the theory significantly, perhaps even rendering it vacuous (if, for example, all of our representations can be caused in non-standard ways, like blows to the head, or specifically designed electromagnetic apparatus – Adams [2000].) See also Mendola (2003) for a general critique of all asymmetric dependence approaches.

<5>Basic representations plus composition</5>

It might be thought that there are some basic, perceptual representations (e.g. colours, tastes, sounds) that fit informational semantics rather well, i.e. they are less susceptible to problems of misrepresentation, distality, etc. Perhaps the informational psychosemantic story applies only to these primitives, and other mental representations are simply constructed out of these basic ones; in this way, the various violations of ATT

might be avoided (Dretske 1986; Sterelny 1990). This compositional strategy is also the preferred response to an as-yet-unmentioned difficulty for informational theories, the problem of empty representations, like the empty concept of unicorns. It is unclear how information can be carried about the nonexistent, but if these concepts can be decomposed into more basic ones that are susceptible to an informational treatment (e.g. of horses and horns), there's no problem (Dretske 1981).

To succeed with this compositional project, we would need plausible analyses of complex concepts, and given the poor track record of conceptual analysis (Fodor 1998), this appears unlikely. Consider also the evidence that we need not know the individuating conditions for kinds or individuals (including mythical or fictional kinds and individuals) in order successfully to refer to them (Kripke 1972; Putnam 1975; Burge 1979; Millikan 1984). This suggests that conceptual analysis cannot individuate such concepts, perhaps explaining its poor track record. Overall, the compositional strategy does not seem very promising. The right balance of a range of primitive contents and a plausible individuating analysis of complex concepts would have to be found, and we are certainly nowhere near that.

<4>Informational teleosemantics</4>

One possible panacea for all of these problems is teleology. Returning to the thermometer, we could say the teleological function or job of the thermometer is to make the height of its mercury column co-vary with temperature such that the mercury column's being at "12" is *supposed* to carry the information that the temperature is 12°C. While its informational content is disjunctive, its semantic content is just the disjunct singled out by the teleology. If the thermometer has a leak, it is failing to do what it is supposed to do, and therefore misrepresenting. Also, the mercury column is not representing any proximal cause of its height (e.g. the pressure and volume of the glass tube), because that is not what it is supposed to carry information about. This teleological version of "informational thermostosemantics" seems to match intuitive content quite easily – thus the pursuit of a plausible teleological version of informational

psychosemantics. The most familiar version of such a psychosemantics comes from Dretske's more recent work (1986, 1988, 1995; see also Neander 1995; Jacob 1997; Shea 2007).

The most difficult task is to justify the teleology. In the case of the thermometer, it is the designer's and/or user's intentions that endow it with its function, but the psychosemanticist must look elsewhere, usually to some selective process, like evolution by natural selection or some variety of trial-and-error learning.⁹ A human designer might select a variety of materials in a particular arrangement to pump water out of a cistern – so that the artefact produced has the function of pumping water, and not of making thumping noises (although it does both). Similarly, Darwinian evolution selects a variety of materials in a particular arrangement in order to pump blood through the body – so that the heart has the function of pumping blood, and not of making thumping noises, although it does both (Wright 1973; Millikan 1984). In both cases, the object (pump or heart) is there *because* it does the thing that is its function (pumping). In both cases, there are limitations on what materials are available, and what arrangements are possible. In both cases, there are random elements involved: what ideas the designer happens across (discarding most of them), and what mutations occur. The contention is that the analogy is close enough to justify applying teleological terminology in the biological world quite generally, even if it rests most naturally upon artefacts (Millikan [1984]; for refined analyses, see Neander [1991], Allen et al. [1998]; for discussion, see Ariew et al. [2002]).

Assuming we are willing to accept natural teleology, how do we get natural *informational* teleology? We need an indicator to be naturally selected for indicating its intuitive content, just as a gas gauge is selected (by a car's designer) for indicating the level of gas in the tank (Dretske 1988). The gas gauge represents the level of gas in the tank (and not the current in its lead wire, or the incline of the slope the car is on) because that is what it is supposed to indicate, that's its job or function. For mental representations that are a product of development (as opposed to learning), the relevant selection may be accomplished by evolution (Dretske 1995). "Hardwired" visual circuits for detecting edges might be an example. A mutation results in a particular type of visual

neuron being sensitive to edges in a single animal. This results in improved vision in that animal, which navigates its environment very successfully and consequently leaves many offspring, more than its average competitor. These offspring also compete well, and the presence of the modified gene gradually predominates in the population. So that neuron type is there (in the population) because it carries information about edges; that is its function. If one of these neurons responds instead to a bright flash, it is falsely representing the presence of an edge – this counts as an error because the neuron is not doing what it is supposed to do. *That* detection disposition of the neuron type did not confer any advantage onto its host organisms, and so was not causally responsible for the spread of that neuron type through the population.

That is the informational teleosemantic story for innate indicators, indicators that are products of development. Learned representations may be accommodated in one of two ways: either (1) analogously to innate representations, but where the selective process is trial-and-error learning (Dretske 1988; Papineau 1987); or (2) by attributing to the learning mechanism an evolutionarily derived generic function of creating mechanisms with more specific indicator functions (i.e. an informational version of Millikan 1984 – it should be emphasized that Millikan’s theory is *not* informational).¹⁰

<5>Objections to teleosemantics</5>

There are many objections that have been raised against teleosemantic accounts of content generally, and some against informational teleosemantic accounts. However, the two principle problems are Swampman and the disjunction/distality problems.

Swampman is a molecular duplicate of (usually) Davidson, who arises by a massively improbable chance in the Florida Everglades, after a lightning strike perhaps.¹¹ Since Swampman lacks any evolutionary or learning history, none of his “mental representations” have any content, on the standard teleosemantic views. Yet he behaves exactly like Davidson, successfully navigating his environment and (apparently, at least) engaging in philosophical conversation. Ignorant of his past, we would unhesitatingly attribute mental states to Swampman, so the example is meant to show

that teleosemantics violates the ATT_{Others} desideratum (albeit for a rather unusual “other”).

Swampman is a problem for teleosemantics only if teleology depends upon *historical* facts, e.g. evolutionary or learning facts. Therefore some teleosemanticists have responded to the Swampman objection by attempting to formulate a non-historical teleology, usually dependent upon cybernetic ideas of feedback and homeostasis (see Chapter 21 of this volume; Schroeder 2004a, b). Another response is a thoroughgoing externalism (like Millikan’s, for instance). It is relatively uncontroversial that Swampman lacks a concept of Davidson’s mother, since he has never had any causal contact with her. If it could be made intuitive that prior causal contact (or other relation) is necessary for *any* concept (Burge 1979; Millikan 2000; Ryder 2004), then it should be intuitive that Swampman’s concepts don’t refer and so his mental states lack truth-conditional content (though they could have *narrow* content – see below). In addition, Millikan (1996), and Papineau (1993: 93) insist that their teleological theory (among others) is a “real-nature” theory, rather like the chemical theory that water is H₂O. While it might seem to us that water could be composed in some other way – it is imaginable to us – it is not really possible. Similarly, while a content-possessing Swampman is imaginable, he is not really possible. (See Braddon-Mitchell and Jackson [1997], Papineau [2001], and Jackson [2006] for further discussion.)

Note that the teleosemanticist can agree with her opponent that Swampman has *conscious* states, at least phenomenally conscious states (see Chapter 29 of this volume), as long as those are not essentially intentional. (So this move is not available to Dretske, for instance; see his 1995.) Perhaps that is enough shared mentality to account for the problematic intuition that Swampman is mentally like us. Nevertheless, the Swampman issue remains a contentious one – it is probably the most common reason for rejecting a teleological approach to content.

The second main problem for teleosemantics is that it isn’t clear that it can fully overcome the disjunction and distality problems. Fodor (1990: Ch. 4) is particularly hostile to teleosemantic theories for this reason. Much of the discussion has focused on

the case of the frog's "fly detector." The detection of flies, nutritious blobs, small dark moving spots, or a disjunction of receptor activations could all be implicated in the selectional explanation for why frogs have such detectors, so it seems that a teleosemantic theory cannot decide among those content assignments. The nomic dependency condition may help a little here, since the frog's representation arguably exhibits a nomic dependence on small dark moving dots, but not on nutritious blobs. However, proximal stimuli exhibit both the requisite nomic dependence (albeit a disjunctive one) and selectional relevance, so the distality problem appears to be particularly problematic for informational teleosemantics (Neander 2004; see also Godfrey-Smith 1994). Millikan claims a teleosemantic solution is available only if we abandon the informational approach (see below).

<4>The grain problem<4>

In this, the final section on informational semantics, I turn to a rather different problem. It will take us in new directions entirely, towards a variety of psychosemantics that is not information based: conceptual or causal role psychosemantics.

Frege noted how fine grained the contents of linguistic utterances and propositional attitudes are, in that it is possible to believe that P and disbelieve that Q even though P and Q are equivalent in one of multiple ways: extensional, nomological, or logical. The nomic-dependence condition allows informational semantics to distinguish judgements that are extensionally equivalent. For example, creatures with hearts and creatures with kidneys might *actually* cause all the same representations to be tokened, but nomicity requires that they exhibit the same *counterfactual* tendencies as well, which they do not. However, for equivalence stronger than extensional, information-based theories run into trouble. Consider the following pairs of predicates: "is an electron" and "has charge e "; "is fool's gold" and "is iron pyrite"; "is equilateral" and "is equiangular"; and the following pairs of concepts: "the morning star" and "the evening star," "Clark Kent" and "Superman," and "silicone" and "polysiloxane." None of these can be distinguished counterfactually.

One possible response is to say that these representations are not really distinct, at least not in terms of their contents (roughly equivalent to the modern “Russellian” position in philosophy of language [Richard 1983; Salmon 1986; Soames 1989; Braun 2000]). (In a few cases they may be distinguished syntactically: this is particularly plausible for the “is an electron” – “has charge e ” pair.) The usual response, though, is to maintain that the distinct representations that are strongly equivalent play different causal roles in the cognitive economy (e.g. Prinz 2002; Neander 2004). (This corresponds roughly to the Fregean position in philosophy of language [Church 1951; Evans 1982; Peacocke 1992], where the causal roles are to be identified with Fregean senses.) Although the concepts linked to the words “silicone” and “polysiloxane” denote the same thing, they might be differentiated by their internal cognitive roles.

Controversially, informational content and cognitive role are sometimes divided into explanations of two kinds of content, external or *broad content* and internal or *narrow content*. Broad content is linked to the phenomena of reference, truth, and (more generally) satisfaction conditions, while narrow content is linked to the phenomenon of cognitive significance (e.g. the different cognitive significance of “morning star” and “evening star,” despite common reference). Narrow content, by definition, supervenes upon the intrinsic state of the representing mind, while broad content does not so supervene. On hybrid informational-cognitive role theories, narrow content is only related to truth (and satisfaction) by attaching to representations that also have broad contents *via* external informational relations (Field 1977; Loar 1982). Because of its distant relation to truth conditions, and (on the standard view) the intimate link between truth conditions and content, it is questionable whether narrow content deserves the name of “content” on such “hybrid” or “two factor” theories. (Recall, however, that “content” is a philosophical term of art.) By contrast, there are theories on which internal cognitive role is supposed to be much more directly related to truth conditions – we now turn to those.

<3>Conceptual role semantics</3>

Naturalistic versions of conceptual role semantics (CRS) descend directly from functionalism in philosophy of mind (Chapter 10 of this volume) and “use” theories of meaning in philosophy of language (Wittgenstein 1953; Sellars 1963). Use theories of meaning say that the linguistic meaning of an expression is determined by its use or role in a language (in inference and other aspects of the “language game”). Functionalism says that the identity of a mental state is determined by its causal role in the perceptual, cognitive, and behavioural system. Since the content of a mental state is essential to its identity, and since linguistic meaning and mental content share deep analogies (e.g. similar belief contents and linguistic meanings are typically expressed using the same “that” clauses), the two theories naturally come together to say: mental content is determined by causal role (especially inferential role) in the perceptual, cognitive, and behavioural system.

For example, take the thought “it is raining.” The CRS theory will characterize the content of this thought (at least in part) by the inferences that it is disposed to participate in (either as premise or conclusion). For instance, if one has the thought “it is raining,” perhaps one is disposed to infer the thought “there are clouds outside.” This thought will also have its own content-characterizing inferences, to other thoughts which will, in turn, have their own content-characterizing inferences, etc. Depending on the type of theory, any of these inferential-cum-causal patterns might be relevant to the content of the thought “it is raining.” Again depending on the type of theory, causal relations to items in the external environment may also be relevant.

There are several ways to divide CRS theorists. First, there are those who accept the representational theory of mind (RTM), and those who do not (Armstrong 1973; Lewis 1994; see previous chapter). Naturalists have generally found the arguments in favour of RTM persuasive, so most naturalistic CRS advocates apply the theory to mental representations; however, most of the discussion below will apply to similar non-RTM theories as well. Another division is into the teleological and non-teleological camps, a distinction we shall consider later when looking at objections to CRS. Finally, there are different ways of characterizing the content-determining causal roles, in terms of their density and whether they extend beyond the mind into the environment.

<4>Characterizing causal roles</4>

<5>Short vs. long armed</5>

CRS theorists divide on whether content-determining causal roles extend into the environment. A long-armed theory (externalist, e.g. Harman 1982, 1987) allows external objects to enter into the functionalist analysis, while a short-armed theory (internalist, e.g. internalist computational functionalism [see Ch. 10 of this volume]) analyses contents only in terms of perceptual states, motor commands, and the complex systemic causation that occurs in between.

On a short-armed theory, causal roles are initially characterized as relations among mental states characterized by their contents (“it is raining” is disposed to cause “it is cloudy,” etc.) These contents are then abstracted away and one is left with particular causal patterns. These content-characterizing causal patterns are entirely abstract, or purely relational (*a*, *b*, and *c* jointly cause *d*; *d* and *e* jointly cause *f*, etc.). Such a theory is particularly vulnerable to complaints of being too liberal – perhaps the molecules in my wall exhibit the relevant causal pattern, or a set of water pipes could be set up to exhibit the pattern, but these things do not represent that it is raining (or so goes the intuition; see Searle 1980, 1992). A possible response is to require that the variables denote *representations*, with some stringent additional requirements for what counts as a representation (e.g. that they enter into *computations* [Chapter 10 of this volume]; see the previous chapter for some approaches).

While CRS is designed neatly to solve the grain problem, a short-armed theory runs into trouble with the flipside: twin cases. Kripke (1972) (on individuals) and Putnam (1975) (on kinds) persuaded most philosophers of mind (and language) that at least some mental contents are determined, in part, by conditions external to the mind. Oscar, who lives on Earth (where the rivers, lakes, and oceans are filled with H₂O) is a functional duplicate of twin Oscar, who lives on twin Earth (where the rivers, lakes, and oceans are filled with XYZ). Yet (Putnam persuades most of us) Oscar’s water-role thoughts are of H₂O, while twin Oscar’s water-role thoughts are of XYZ. A short-armed

CRS, being an internalist theory, does not appear to have the resources to account for this difference in content.

A long-armed theory is somewhat resistant to the charge of liberalism and twin-case worries. The charge of liberalism is not as problematic because the causal pattern characterizing a particular content is less abstract. Included in it are causal links to types of items in the environment, which need not be abstracted away (Cummins 1989: 122). Furthermore, as long as those causal links to external items are not construed purely dispositionally (e.g. a disposition to respond perceptually to clear, potable liquids), a long-armed theory may not be vulnerable to the twin problem either. The causal role of Oscar's thoughts may link him only to H₂O, while the causal role of twin Oscar's thoughts may link him only to XYZ (Block 1998).

<5>Causal role density</5>

There are several choices as to what sorts of causal relations to include in the content-determining causal patterns. One possibility is to include, as content-determining, *all* the causal relations that a token contentful mental state enters into. This is not very plausible – a mental state token's disposition to reflect light is clearly not relevant to its content. How to narrow down the relevant dispositions, though?

One very restrictive possibility is to include only relations that are definitional, so that the representation “*x* is a bachelor” is characterized by the disposition to infer, and be inferred by, “*x* is an unmarried adult male person.” On a naturalistic theory, these contents would then be abstracted away, leaving a causal pattern. There are three serious problems with this idea. First, after decades of effort in the twentieth century alone, very few concepts seem to be definable (Fodor 1998). Some of the few exceptions include logical concepts, for which a CRS theory is particularly plausible.¹² The second major problem with the definitional route is that the possibility of isolating definitional relations depends upon there being a determinate difference between claims that are analytic, and those that are synthetic (Fodor and Lepore 1991) – and it is far from clear that the analytic-synthetic distinction can be maintained in the face of the objections

raised by Quine (1953, 1976). Third, the meagre causal roles that are supposed to be content determining are highly vulnerable to the charge of excessive liberalism. This problem will also apply to possible alternative “sparse-role” versions of CRS, for example a naturalized version of Peacocke’s theory, where the content-determining roles are only those that are “primitively compelling.”

Perhaps a better option is to be less restrictive about what causal relations are content-determining. Armchair philosophy might be given the task of elucidating all manner of conceptual relations and platitudes that individuate contents (a tall order!) (Lewis 1994; Jackson 1998), and these conceptual roles could then be naturalized by mapping them onto causal ones. Alternatively, it could be left to psychology to determine what causal roles characterize states with particular contents. These roles may have a probabilistic structure, as in prototype theory, for instance (Rosch 1978). On prototype theory, no particular perceptual or other sort of representation, or set of such representations, need be necessary and sufficient to make a thinker token a mental state with a particular content. Rather, such an inference will occur only with a certain probability. Alternatively, the causal roles may be characterized somewhat as one characterizes the role of a concept in a scientific theory (Gopnik and Meltzoff 1997), with a many-layered and revisable inferential structure.

<4>CRS theories and truth conditions</4>

One fundamental problem for naturalistic CRS theorists is in relating causal roles to truth conditions. Prima facie, a psychosemantic theory ought to make plain the general principles by which internal, physical states are mapped onto truth conditions. This mapping is made relatively transparent by informational theories. On a CRS theory, by contrast, the mapping is difficult to make explicit. On a short-armed theory, there are no content-determining relations to external items that could be used to identify truth conditions, and on long-armed theories, there are *too many* such (potentially) content-determining relations – as Fodor observed (see the section, above, “Disjunction problems”), a contentful mental state can be tokened in response to multifarious

external conditions.

One strategy is to make use of informational relations as determinative of truth conditions, producing the hybrid theory mentioned at the end of the section, “The grain problem,” above. This theory will inherit most of the virtues and problems that come with informational theories, and will not be further discussed here. Another strategy is to take the mass of related inferences to be similar or *isomorphic* to related states of affairs in the environment. This isomorphism-based strategy will be considered below. The third strategy, which comes in two forms, is to reject the demand for a general naturalistic formula or “recipe” for determining truth conditions. That is the strategy we shall examine first.

<5>“No-recipe” CRS theories</5>

The most radical rejection of the demand for a general formula for determining truth conditions is *deflationism* (e.g. Field 1994; Horwich 1998). The CRS deflationist denies that “the problem of content” relates to any genuine mind-world connection that needs explaining. The deflationist accepts the standard schemata,

“*P*” is true if and only if *P*

If *b* exists then “*b*” refers to *b* and nothing else; if *b* doesn’t exist then “*b*” doesn’t refer to anything

and claims that that is all there is to say about truth and reference (whether for language or for mental representations). Meaning and content are determined by use, and there is no need for a theory of content to provide a substantive account of how truth conditions are determined. The debate between deflationists and “inflationists” (primarily in the arena of language) is voluminous and complex, defying any quick summary here (see Blackburn and Simmons [1999] and Lynch [2001] for discussion). If deflationism works, however, it would in one fell swoop remove all the difficulties we have seen in

matching intuitive content. The major cost is that our success in dealing with the environment could not be explained by our mental states “corresponding to the facts” in any robust sense (e.g. carrying information about or being isomorphic to the facts).

The second version of the “no recipe” strategy effectively takes the relation between a causal role and a truth condition to be real, but primitive and unexplained. Certain causal roles or patterns determine certain truth-conditional contents, and that is all there is to say. However, this sort of primitivist theory is severely hampered in trying to characterize content-determining causal roles. A recipe for determining truth conditions would provide a test for whether one has got the right causal role for a particular content or set of contents. If it is the right causal role, it must determine the right (i.e. intuitive) truth conditions. This introduces a welcome constraint. Without such a recipe, one has only raw intuition to go by in figuring out what aspects of a mental representation’s causal role are content-determining. Given the clash of intuitions one finds in the literature, one’s hopes should not be high that philosophers of mind will ever settle on a particular “no-recipe” CRS theory. That, by itself, is reason to hope for a truth-condition determining recipe. Luckily, there is a very ancient one available for the CRS theorist to make use of: *resemblance*.

<5>Relational resemblance as a recipe for determining truth conditions</5>

While first-order resemblance between mental states (e.g. brain states) and represented states of affairs is clearly a non-starter, the *relational* resemblance that fits with a CRS theory is much more plausibly a content-determining one. On this view, the complex, internal web of causal roles resembles (is isomorphic or homomorphic to) the complex structure of that which is represented. That which is represented could be the causal or regularity structure of the environment (Cummins 1989; McGinn 1989; P. S. Churchland 2002; Ryder 2004), or (less naturalistically) computational functions or entailment structures among propositions (Cummins 1989; McGinn 1989). The key point is that CRS, on this version, turns out to be a species of “structural representation,” where a pattern of relations among a group of representations mirrors

a pattern of relations in what is represented (Swoyer 1991).

A nice example is a solar system model, or orrery. The gears in the orrery ensure that the pattern of regularities in the motions of the model planets mirrors the motions of the real planets. As Swoyer points out, this allows for the process of “surrogate reasoning”: if you don’t know the position of Venus in two months, but you do know where the Earth will be, simply rotate the Earth into that position and read off the future position of Venus. Similarly, if the pattern of inferences we make about the world mirrors real regularities in the world, we can understand how we are able to make predictions (this is the key insight of the internal-model theory of thinking [Craik 1943]).

If the relational resemblance is meant to serve as a recipe for determining referential content and truth conditions, however, it runs into serious trouble. The problem is that relational resemblance (whether isomorphism or homomorphism) is too cheap. Given a set of representations exhibiting a number of causal relations, this structure will be mirrored by an indefinitely large number of environmental (or propositional entailment) structures. Thus the relational resemblance strategy faces a serious problem of indeterminacy. Some other element would need to be added – perhaps causal, informational, or teleological. Thus, for example, McGinn (1989) and Ryder (2004) propose teleological isomorphism theories, and must contend with the standard problems for teleosemantics (especially Swampman and indeterminacy).

<4>Further problems with CRS</4>

<5>Holism</5>

Perhaps the central problems for non-definitional versions of CRS that appeal to fairly dense causal roles are those that arise as a result of their *content holism* (to be contrasted with the *representational holism* described in the previous chapter). With the rich conceptual cum causal relations that are content-determining on such theories, each mental representation gets constitutively linked to many others, which are linked to many others, and pretty soon the content of a single mental state gets determined by

the entire web of conceptual relations. Since no two actual thinkers believe all the same things, the theory dictates that no two thinkers have *any* beliefs in common. This clearly violates ATT_{Others}; we regularly say that two people believe the same thing.

CRS theorists typically respond that our attributive practice requires only *similarity* of content (e.g. P. M. Churchland 1996). For practical purposes, we need not pay attention to the small differences in belief contents when attributing the judgement “the puppy is black” to two people, only one of whom also believes that dog hair is made of protein and that Jung was a charlatan. The conceptual-cum-causal webs of these two people are similar enough to justify treating them as the same, for ordinary psychological explanation and other purposes. The problem, then, is to define a measure of similarity, and then to determine how similar is “similar enough.” Critics, most notably Fodor, deny this can be done. Part of the problem is that complex relational structures are extremely difficult to compare objectively, especially if they differ in a large number of respects (e.g. number of nodes as well as pattern of connectivity). And webs that *do* differ in a large number of respects nevertheless seem to be able to support the same mental contents: both you and Helen Keller can believe that the sun is shining, for instance. This debate is unresolved, with CRS advocates turning to the details of neural network theory for help (P. M. Churchland 1998; Goldstone and Rogosky 2002), and critics responding that this doesn’t change a thing (Fodor and Lepore 1999).

<5>Compositionality</5>

Fodor and Lepore have pressed another issue faced by non-definitional CRS, which is compositionality. We are able to think an indefinite variety of thoughts that are systematically related by their contents; the most obvious explanation for this capacity is that our thoughts have components (e.g. concepts), and the propositional contents of our thoughts are a function of the sub-propositional contents of those components (see the section “Syntactic structure,” in the previous chapter, and Chapter 17 of this volume). The problem for non-definitional versions of CRS is that conceptual roles do not seem to compose in the right way – for example, the conceptual role of a complex

concept need not be a function of the conceptual roles of its component, simpler concepts. One of Fodor's favourite examples is the complex "pet fish." Take, for example, an aspect of conceptual role focused on by prototype theorists: typicality judgements. The typical pet is a cat or dog, and the typical fish is a trout or salmon. But the typical pet fish is not the average between cat/dog and trout/salmon; rather it is a goldfish or guppy. More generally, it is hard to see how the non-definitional conceptual role of "pet fish" could be *any* function of the conceptual roles of "pet" and "fish" – it is certainly not a simple additive function. Perhaps some more complex function could be determined (see Prinz [2002] for an attempt), but the holistic nature of the content-determining roles makes it a difficult task. It is fair to say that the compositionality problem is an unresolved issue, with CRS (at least the "no-recipe" sort) on the defensive. (The relational resemblance strategy might provide a solution, since truth conditional/referential content *is* compositional.)

<5>Normativity</5>

CRS also runs into an analogue of one of the principal problems that plagues informational semantics, namely the problem of misrepresentation, or more generally the problem of error. It rears its ugly head in even broader form: people make inferential errors not only when the conclusion involves *applying* some representation to an item or state of affairs (giving rise to the classic problem of misrepresentation), but when making other sorts of inferences as well (e.g. inferences in practical reasoning). A naturalistic version of CRS needs to make sense of these inferences as being erroneous, and as Kripke has famously argued (Kripke 1982), it is far from obvious how this is to be done with only causal notions at one's disposal. Kripke asks, what distinguishes between two people who are both disposed to infer "5" from "57 + 65," one of whom does so incorrectly (meaning by "+" the plus function), the other correctly (since she means something else by "+," call it "quus"). What aspect of their functional profile could distinguish them? Perhaps the "plusser" is disposed to correct himself, but that seriously risks begging the question – why suppose that this is disposition to "correct" as opposed

to, say, second-guess incorrectly?

This is the classic problem of the normativity of content, a challenge for any naturalistic psychosemantics. Similar responses are available to the CRS theorist as to the informational theorist, accompanied by much the same problems (see the sections, “Proposed fixes for ATT violations,” and “Informational teleosemantics,” above). Appeal may be made to “ideal conditions” (in an individual or population) of some sort, though it is unclear how this is to be done in a non-question-begging manner. Alternatively, the correct inference could be determined in some way by its incipient cause, although isolating “the” incipient cause would be difficult. (Explanations for our acquisition of inferential habits are typically rather complex, not a matter of simple conditioning.) Perhaps the incorrect inferences are asymmetrically dependent on the correct ones (Mendola 2003). Finally, one might introduce naturalistic teleology (also further dulling the charge of liberalism) (Sober 1985), and face Swampman and disjunction/distality problems.¹³

<3>Success psychosemantics</3>

Mental representation is involved in perception, inference, and action. We have seen psychosemantic theories founded upon the first two (informational and conceptual role theories); what about the last one? We now turn to this action-oriented strategy, which I call, following Peter Godfrey-Smith (1994), “success (psycho)semantics.”

<4>Papineau and success as desire satisfaction</4>

The pragmatist view is that beliefs are true when they are *successful* in some sense. Given a naturalistic definition of success, we can get a psychosemantics out of the pragmatist view of truth. One plausible account of success is desire satisfaction; but desires are intentional mental states themselves, and so must have their contents accounted for in turn. David Papineau (1987, 1993) turns to teleology to solve this problem. He claims that a desire’s satisfaction condition is the effect it is selected (by evolution or learning) to produce. So, for example, the desire for food was designed

(presumably by natural selection) to produce the effect of the animal ingesting food. The function of the desire is to bring about that state of affairs, which is also the desire's satisfaction condition.

Of course, desires can only bring about their satisfaction conditions with the cooperation of beliefs (following a standard belief-desire explanation for action). Beliefs, then, also must have a general function of bringing about desires' satisfaction conditions – the satisfaction conditions of the desires with which they are cooperating. Each belief does this in its own particular way: it has the function of causing actions that will satisfy desires *if a particular condition* obtains. This condition is that belief's truth condition.

When Papineau says that beliefs will lead to desire satisfaction if their truth conditions obtain, he means something quite strong: that desire satisfaction is *guaranteed* if the belief is true. (So a belief's truth condition is “that condition which guarantees that actions generated by that belief will fulfil its biological purpose of satisfying desires” [1993: 80].) This might seem obviously false: if I believe that the sun is 93 million miles away (true), but I also believe that I can jump that far, a desire to jump to the sun will fail to be satisfied (in magnificent fashion!) by my ensuing action. His response to this problem is to go holistic (1993: 72–3): a belief is guaranteed to satisfy desires if it is true *and* if the other beliefs it is acting in concert with are also true. This approach makes the truth conditions of whole sets of potentially related beliefs interdependent, to be determined simultaneously. This helps with the counterexample above, where I have a false belief that I can jump 93 million miles high. But it seems that there are other cases where the fix will fail. Perhaps all my beliefs are true of the situation I'm involved in, but some outside circumstance conspires to interfere with satisfaction of my desires, e.g. a meteorite knocks my ice cream to the ground. The truth of my operative beliefs have failed to guarantee satisfaction of my desires, unless we maintain, implausibly, that I had a belief operating to the effect that no meteorite would knock my ice cream to the ground. (Perhaps I had a belief to the effect that nothing would interfere with the satisfaction of my desire to eat ice cream – but this threatens to trivialize the identification of a belief's truth condition. For instance, we could say that my belief that I can jump 93 million miles high is true, except for the interference of my

weakness.)

<4>Millikan and representation consumers</4>

Ruth Millikan (Millikan 1984, 1989, 2004) offers a version of success semantics that forges a weaker link with success than the pragmatists or Papineau. Hers is probably the most ambitious theory in our survey, attempting, as it does, to account for the contents of representations in simple animals all the way to high-level cognitive contents, and linguistic meaning as well. Unlike Papineau, she does not define success in terms of desire satisfaction. Rather, she divides a system up into representation *producers* and representation *consumers*. The notions of “producer” and “consumer” are very broad: they include two honeybees communicating via a dance, a speaker and a listener, two parts of a cognitive system, or a cognitive system and its attached motor system. (We will be focusing on psychosemantic cases like the latter two, of course.) For Millikan, success is defined in terms of the proper performance of the representation consumer’s function(s).

For instance, if the representation consumer is the motor system, and one of its functions is to bring food to the mouth, success may be defined as actually performing that function. She also does not require that a representation’s truth condition *guarantee* proper performance of the consumer’s function. It just needs to be an “historically Normal condition” for the performance of the consumer’s function. This means that it forms part of an explanation for how it was that the consumer managed to perform its function in the past, which performance explains why this producer-consumer pairing was selected. For example, the function of eating was accomplished in the past by the belief that I’m holding an ice cream being compresent with the condition of my actually holding an ice cream, and this accomplishment selectionally explains why my belief and the motor system have the interaction that they do. If Papineau were to adapt his account to be similar to Millikan’s in this respect, he would say, not that a belief’s truth condition guarantees satisfaction of desires (which causes trouble), but that a belief’s truth condition is an essential part of an explanation for why that belief

was selected for desire satisfaction.

Millikan complicates things further. On Papineau's model, beliefs have the function of aiding in desire satisfaction, and they do so by covarying with a particular state of affairs type. On Millikan's model, representation producers have the function of aiding in consumer function satisfaction, and they do so by producing representations that *map* onto world affairs according to a particular *rule*. The representations that are produced are articulated or complex (see previous chapter, the section, "The degree-of-systematicity axis: from detectors to speakers"), such that aspects of the representation that may vary can be mapped onto ways in which the world can vary. For example, activity in visual areas of the cerebral cortex may vary in intensity and location, where those aspects of the representation map onto variations in contrast and location of discontinuities (lines) in the environment. So it is not compresence with a truth condition that selectionally explains the satisfaction of consumer function, but mapping according to a particular rule. This allows for novel representations, that never before occurred in history, e.g. a novel combination of a line in a particular location at a particular contrast.

Selected mapping rules, according to Millikan, may be extremely broad and general. For instance, her theory can apply to a general-purpose learning system, such as that presented in Ryder (2004). On that theory, the cerebral cortex was selected for structuring itself isomorphically with particular regularities (of a certain type) upon encountering them in the environment, just as a camera's film was selected for structuring itself isomorphically with reflected light patterns that enter the camera's lens. Just as a camera's film can represent many different patterns of light, so the cerebral cortex can represent many regularities (and the items that participate in them).

While Papineau's link between representation and success seems too strong to yield correct truth conditions, it may be that Millikan's is too weak. Pietroski (1992) has objected that her theory violates the ATT desiderata because the selected mapping rule may identify contents that, intuitively, a creature is incapable of conceiving. In his example, the simple kimus evolve a producer-consumer mechanism that causes them to

pursue red light, allowing them to avoid their predators (the snorfs) at sunrise and sunset. Intuitively, the kimus know nothing of snorfs (they wouldn't even recognize one), they are just following red light. But the mapping rule that applies to their representations maps onto the presence/absence of snorfs, so their representations (according to Millikan's theory) are about the presence/absence of snorfs. Millikan accepts this consequence with equanimity (Millikan 2000); others claim that the problem does not infect complex belief-desire psychologies (Rountree 1997).

<4>Problems for consumer semantics</4>

Both Papineau's and Millikan's theories are subject to the standard worries about teleosemantics, namely Swampman and disjunction/distality. I commented earlier (the section, "Objections to teleosemantics") on their responses to the Swampman problem. Millikan believes that her particular focus on consumer functions solves the disjunction and distality problems. Turning to the fly/frog example, it must be a Normal condition for the representation consumers to perform their function that the frog's active detector map onto the presence-of-a-fly-now, while the silent detector map onto the absence-of-a-fly-now. However it is not clear how this is supposed to resolve the disjunction (or specificity) problem, for it seems that the fly-now mapping and the nutritious-blob-now mapping are equally good candidates for being Normal conditions for the representation consumer (the frog's motor and digestive system?) performing its function. This is because the representation consumer's function is equally indeterminate – is it supposed to ingest and digest flies, or is it supposed to ingest and digest nutritious blobs? As Fodor puts it, "Millikan's strategy is to solve the disjunction problem for the signals by describing the 'consumers' of the signals in intentional terms (in terms of what they 'care about') and then to beg the disjunction problem for the consumers" (Fodor 1991[[Fodor and Lepore? Or please include in the ref. list]]).

Papineau concedes that indeterminacy attends to simple systems like the frog, but believes there is no such indeterminacy in the case of complex belief-desire systems like ours, because of the specificity of desire functions (1998, 2003). However, it isn't clear

that he solves the disjunction problem either, because he doesn't consider the possibility of a desire's content being a disjunction of all the various means historically used to satisfy it, rather than the desire's intuitive end.

<3>Further issues</3>

A set of problems that tend to cause trouble for *any* naturalistic psychosemantics is accounting for seemingly abstract concepts, like numbers, causation, logical concepts, democracy, etc. These things are either very difficult to fit into the causal order (e.g. numbers), or, in one case (causation), they *constitute* the causal order. Either way, it is far from obvious how they could participate in informational relations, be selectionally relevant, or participate in long-armed causal roles. It is only slightly more promising to suppose that they could be isomorphic to mental representations, or somehow contribute to success. Here, we hearken back to age-old disputes between rationalism and empiricism, with most naturalized psychosemantics tending towards the empiricist camp. In general, psychosemanticists have focused on solving the “easy problems” of representing everyday physical states of affairs, saving these more difficult cases for later. For some sample attempts, however, see Millikan (1984) and Prinz (2002).

Some philosophers maintain that there is a necessary connection between intentionality and consciousness (Searle 1983). If this is true, and it turns out that consciousness cannot be naturalized, then the prospects of naturalizing representational content would be dim indeed. However, the arguments for this necessary connection are based upon intuitions that naturalists have not found compelling.

This chapter has been organized around the ATT desiderata, but there are other important desiderata for a theory of content. Some we have come across along the way, e.g. that there be a general formula for determining a representation's truth (or satisfaction) conditions, and that content be shareable. However, there is one important desideratum that has not made an explicit appearance, although it is related to the ATT desiderata:

A theory of content ought to cohere with our everyday content-dependent explanations of behaviour.

We explain behaviour by appeal to what people think, believe, and want, not by what is actually the case extra-mentally. For example, we might explain Fred's behaviour in opening the fridge by saying "he wanted a beer, and he thought there was a beer in the fridge." Actual beer in the fridge need not figure in this explanation at all; rather it is the content of Fred's belief that figures in it. After all, there need not have been any beer in the fridge, perhaps Fred was mistaken. This is why our explanations of behaviour are typically "content-involving" (and why the problem of misrepresentation is so crucial – misrepresentations play an essential role in the explanation of behaviour). It seems that a theory of content which made this sort of explanation unfathomable could not be about our target phenomenon: the content of the propositional attitudes and other truth-evaluable mental states. The problem is made particularly pressing in view of the fact that it is unclear how broad, truth conditional content – which involves states of affairs external to the mind – can causally explain what a mind does (Braun 1991; Yablo 1997).

Almost all of the theories we have examined or mentioned have something to say about how they can accommodate the explanation of behaviour (Fodor 1987; Dretske 1988; Cummins 1989; Block 1990; Millikan 1993). There are also serious problems that attend to the proposed solutions (Stich 1983; Godfrey-Smith 1994; Melnyk 1996). Clearly this issue will be of fundamental importance in evaluating the various psychosemantic proposals, but unfortunately it would require its own chapter for a proper treatment (see Heil and Mele [1993] for a good introduction).

Finally, I should explain why I have not examined interpretivism (e.g. Dennett's "intentional stance" theory; see the section in the previous chapter, "The thin notion"): it is because such theories are not reductive. For the interpretivist, "all there is to really and truly believing that *P* (for any proposition *P*) is being an intentional system for which *P* occurs as a belief in the best (most predictive) interpretation" (Dennett 1987, p.

29). As Dretske (1986) and Byrne (1998) point out, interpretations involve beliefs with various contents, so there is no analysis here of belief content in terms of something *else*. That said, such theories are quite friendly to naturalism since they maintain a strong supervenience of mental representational content upon the physical, plus they clearly have no problem meeting the ATT desiderata. Thus interpretivism might be an attractive backup option if no reductive account can be made to work. However, I hope the reader has not concluded from this review that it is time to give up on the reductive project.

<3>Notes</3>

- 1 I shall paper over the distinction between conceptual and non-conceptual truth-evaluable content because it has not played a prominent role in debates about naturalizing content. For a nice discussion of the distinction, see Heck (2007).
- 2 The latter two are often taken to be non-conceptual, so not involving concepts (see note 1).
- 3 It is controversial to what extent sensations and emotions, for example, are contentful mental states. Some maintain that the sensation of pain, for instance, represents that there is damage to tissue in a certain location, and so that it is contentful (Tye 1995), while most deny that pain represents anything.
- 4 The specificity problem is what makes “the causal theory of reference” a nonstarter for our purposes. This “theory,” whose recent history traces back to the work of Kripke (Kripke 1972), is really just a programmatic suggestion and does not provide reference-determining conditions (see Devitt 1981; Kim 1977).
- 5 In Dretske’s terminology, this is the information it carries “digitally.”
- 6 Some identify representational content with the information a representation carries “under ideal epistemic circumstances” (or “fidelity conditions” [Stampe 1977] or “relevant normal conditions” [Stalnaker 1984]). This is essentially the same strategy, and suffers from the same problem.

- 7 See Dretske (1981) and Jacob (1997) for attempts to specify the channel conditions as those which lack relevant alternative possibilities, i.e. whose reliability is such that the *fact* that they are reliable isn't news to the receiver of the information.
- 8 Fodor suggested something similar in his 1990 book, but later repudiated the idea.
- 9 For the rest of this section, "function" will mean "teleological function" unless specified otherwise.
- 10 On Dretske's view, the indicator is selected to cause some movement, via operant conditioning. This is, in effect, a selectional explanation for its presence in the cognitive system, since Dretske agrees with Ramsey (Ramsey 1931) that it isn't a belief unless it "steers."
- 11 The original idea is due to Boorse (in the form of "swamprabbits") (1976) and made famous by Davidson in his 1987.
- 12 See Peacocke (1992) for a detailed account, though he stops short of making the move to naturalistic causal roles. Even Fodor takes CRS to be the correct theory for logical concepts (Fodor 1994: Ch. 3).
- 13 Another source of teleology besides the biological and the cybernetic should be mentioned in this connection: linguistic norms. On Sellars' account (1963), norms have their primary place in a public language using community, and the norms that apply to mental representations are derived from the public ones. See Gillett (1997) for an attempt to naturalize this sort of strategy.

<3>References</3>

- Adams, F. (2000) "Asymmetrical Dependence"; available:
<http://host.uniroma3.it/progetti/kant/field/asd.htm>
- Allen, C., Bekoff, M., and Lauder, G. (eds) (1998) *Nature's Purposes: Analyses of Function and Design in Biology*, Cambridge, MA: MIT Press.
- Ariew, A., Cummins, R., and Perlman, M. (eds) (2002) *Functions: New Essays in the*

- Philosophy of Psychology and Biology*, Oxford: Oxford University Press.
- Armstrong, D. M. (1973) *Belief, Truth and Knowledge*, Cambridge: Cambridge University Press.
- Barwise, J., and Perry, J. (1983) *Situations and Attitudes*, Cambridge, MA: MIT Press.
- Blackburn, S., and Simmons, K. (eds) (1999) *Truth*, Oxford: Oxford University Press.
- Block, N. (1990) "Can the Mind Change the World?" in G. Boolos (ed.), *Meaning and Method: Essays in Honor of Hilary Putnam*, Cambridge: Cambridge University Press, pp. 137–70.
- (1998) "Conceptual Role Semantics," in E. Craig (ed.), *The Routledge Encyclopedia of Philosophy*, London: Routledge [[Page range?]].
- Boorse, C. (1976) "Wright on Functions," *Philosophical Review* 85: 70–86.
- Braddon-Mitchell, D., and Jackson, F. (1997) "The Teleological Theory of Content," *Australasian Journal of Philosophy* 75: 474–89.
- Braun, D. (1991) "Content, Causation, and Cognitive Science," *Australasian Journal of Philosophy* 69: 375–89.
- (2000) "Russellianism and Psychological Generalizations," *Nous* 34: 203–36.
- Burge, T. (1979) "Individualism and the Mental," *Midwest Studies in Philosophy* 4: 73–122.
- Byrne, A. (1998) "Interpretivism," *European Review of Philosophy* 3: 199–223.
- Church, A. (1951) "A Formulation of the Logic of Sense and Denotation," in P. Henle, M. Kallen, and S. K. Langer (eds), *Structure, Method, and Meaning: Essays in Honor of Henry M. Scheffer*, New York: Liberal Arts Press, pp. 3–24.
- Churchland, P. M. (1996) "Fodor and Lepore: State-Space Semantics and Meaning Holism," in R. McCauley (ed.), *The Churchlands and their Critics*, Cambridge, MA: Blackwell [[Page range?]].
- (1998) "Conceptual Similarity and Sensory/Neural Diversity: The Fodor/Lepore Challenge Answered," *Journal of Philosophy* 95: 5–32.
- Churchland, P. S. (2002) *Brain-Wise: Studies in Neurophilosophy*, Cambridge, MA: MIT Press.
- Craik, K. J. (1943) *The Nature of Explanation*, Cambridge: Cambridge University Press.

- Cummins, R. (1989) *Meaning and Mental Representation*, Cambridge, MA: MIT Press.
- Davidson, D. (1967) "Truth and Meaning," *Synthese* 17: 304–23.
- (1987) "Knowing One's Own Mind," *Proceedings and Addresses of the APA* 60: 441–58.
- Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, Mass.: MIT Press.
- Devitt, M. (1981) *Designation*, New York: Columbia University Press.
- Dretske, F. (1981) *Knowledge and the Flow of Information*, Stanford: Center for the Study of Language and Information Publications.
- (1986) "Misrepresentation," in R. J. Bogdan (ed.), *Belief: Form, Content and Function*, Oxford: Oxford University Press, pp. 17–36.
- (1988) *Explaining Behavior*, Cambridge, MA: MIT Press.
- (1995) *Naturalizing the Mind*, Cambridge, MA: MIT Press.
- Evans, G. (1982) *The Varieties of Reference*, Oxford: Oxford University Press.
- Field, H. (1977) "Logic, Meaning and Conceptual Role," *Journal of Philosophy* 69: 379–408.
- (1994) "Deflationist Views of Meaning and Content," *Mind* 103: 249–85.
- Fodor, J. (1987) *Psychosemantics*, Cambridge, MA: MIT Press.
- (1990) *A Theory of Content and Other Essays*, Cambridge, MA: MIT Press.
- (1994) *The Elm and the Expert*, Cambridge, MA: MIT Press.
- (1998) *Concepts: Where Cognitive Science Went Wrong*, Oxford: Oxford University Press.
- Fodor, J. A., and Lepore, E. (1991) "Why Meaning Probably Isn't Conceptual Role," *Mind and Language* 6, no. 4: 329–43.
- (1999) "All at Sea in Semantic Space: Paul Churchland on Meaning Similarity," *Journal of Philosophy* 96: 381–403.
- Gillett, G. (1997) "Husserl, Wittgenstein and the Snark: Intentionality and Social Naturalism," *Philosophy and Phenomenological Research* 57, no. 2: 331–49.
- Godfrey-Smith, P. (1989) "Misinformation," *Canadian Journal of Philosophy* 19, no. 4: 535–50.
- (1994) "A Continuum of Semantic Optimism," in S. Stich and T. Warfield (eds),

- Mental Representation*, Oxford: Blackwell, pp. 259–77.
- Goldstone, R. L., and Rogosky, B. J. (2002) “Using Relations within Conceptual Systems to Translate across Conceptual Systems,” *Cognition* 84: 295–320.
- Gopnik, A., and Meltzoff, A. (1997) *Words, Thoughts, and Theories*, Cambridge, MA: MIT Press.
- Grice, P. (1957) “Meaning,” *Philosophical Review* 66: 377–88.
- Harman, G. (1982) “Conceptual Role Semantics,” *Notre Dame Journal of Formal Logic* 23: 242–56.
- (1987) “(Nonsolipsistic) Conceptual Role Semantics,” in E. Lepore (ed.), *Semantics of Natural Language*, New York: Academic Press, pp. 55–81.
- Heck, R. (2007) “Are There Different Kinds of Content?” in J. Cohen and B. McLaughlin (eds), *Contemporary Debates in Philosophy of Mind*, Oxford: Blackwell[[Page range?]].
- Heil, J., and Mele, A. (eds) (1993) *Mental Causation*. Oxford: Oxford University Press.
- Horwich, P. (1998) *Meaning*, Oxford: Clarendon Press.
- Jackson, F. (1998) *From Metaphysics to Ethics: A Defense of Conceptual Analysis*, Oxford: Oxford University Press.
- (2006) “The Epistemological Objection to Opaque Teleological Theories of Content,” in G. Macdonald and D. Papineau (eds), *Teleosemantics*, Oxford: Oxford University Press, pp. 85–99.
- Jacob, P. (1997) *What Minds Can Do*, Cambridge: Cambridge University Press.
- Kim, J. (1977) “Perception and Reference without Causality” *Journal of Philosophy* 74: 606–20.
- Kripke, S. (1972) *Naming and Necessity*, London: Blackwell.
- (1982) *Wittgenstein on Rules and Private Language*, Cambridge, MA: Harvard University Press.
- Lewis, D. (1970) “General Semantics,” *Synthese* 22: 18–67.
- (1994) “Reduction of Mind,” in S. Guttenplan (ed.), *A Companion to the Philosophy of Mind*, Oxford: Blackwell, pp. 412–31.
- Loar, B. (1982) “Conceptual Role and Truth-Conditions,” *Notre Dame Journal of*

- Formal Logic* 23, no. 3: 272–83.
- Lynch, M. (ed.) (2001) *The Nature of Truth*, Cambridge, MA: MIT Press.
- Matthen, M. (1988) “Biological Functions and Perceptual Content,” *Journal of Philosophy* 85: 5–27.
- McGinn, C. (1989) *Mental Content*, Oxford: Blackwell.
- McLaughlin, B. (1987) “What Is Wrong with Correlational Psychosemantics,” *Synthese* 70: 271–86.
- Melnyk, A. (1996) “The Prospects for Dretske’s Account of the Explanatory Role of Belief,” *Mind and Language* 11, no. 2: 203–15.
- Mendola, J. (2003) “A Dilemma for Asymmetric Dependence,” *Nous* 37: 232–57.
- Millikan, R. (1984) *Language, Thought, and Other Biological Categories*, Cambridge, MA: MIT Press.
- (1989) “Biosemantics,” *Journal of Philosophy* 86, no. 6: 281–97.
- (1993) “Explanaton in Biopsychology,” in J. Heil and A. Mele (eds), *Mental Causation*, Oxford: Oxford University Press, pp. 211–32.
- (1996) “On Swampkinds,” *Mind and Language* 11, no. 1: 103–17.
- (2000) *On Clear and Confused Ideas*, Cambridge: Cambridge University Press.
- (2004) *Varieties of Meaning: The 2002 Jean Nicod Lectures*, Cambridge, MA: MIT Press.
- Neander, K. (1991) “Functions as Selected Effects,” *Philosophy of Science* 58: 168–84.
- (1995) “Malfunctioning and Misrepresenting,” *Philosophical Studies*, 79, 109–41.
- (2004) “Teleological Theories of Mental Content”; available: <http://plato.stanford.edu/archives/sum2004/entries/content-teleological/> (accessed summer 2004)
- Papineau, D. (1987) *Reality and Representation*, Oxford: Blackwell.
- (1993) *Philosophical Naturalism*, Oxford: Blackwell.
- (1998) “Teleosemantics and Indeterminacy,” *Australasian Journal of Philosophy* 76, no. 1: 1–14.
- (2001) “The Status of Teleosemantics, or How to Stop Worrying about Swampman,” *Australasian Journal of Philosophy* 79, no. 2: 279–89.

- (2003) “Is Representation Rife?” *Ratio* 16, no. 2: 107–23.
- Peacocke, C. (1992) *A Study of Concepts*, Cambridge, MA: MIT Press.
- Pietroski, P. (1992) “Intentionality and Teleological Error,” *Pacific Philosophical Quarterly* 73: 267–82.
- Prinz, J. (2002) *Furnishing the Mind: Concepts and Their Perceptual Basis*, Cambridge, MA: MIT Press.
- Putnam, H. (1975) “The Meaning of Meaning,” in K. Gunderson (ed.), *Language, Mind and Knowledge*, Minneapolis: University of Minnesota Press [[Page range?]].
- Quine, W. V. O. (1953) “Two Dogmas of Empiricism,” in *From a Logical Point of View*, Cambridge, MA: Harvard University Press, pp. 20–46.
- (1976) “Carnap and Logical Truth,” in *Ways of Paradox and Other Essays*. Cambridge, Mass.: Harvard University Press [[Page range?]].
- Ramsey, F. P. (1931) *The Foundations of Mathematics and Other Logical Essays*, London: Routledge & Kegan Paul.
- Richard, M. (1983) “Direct Reference and Ascriptions of Belief,” *Journal of Philosophical Logic* 12: 425–52.
- Rosch, E. (1978) “Principles of Categorization” in E. Rosch and B. Lloyd (eds), *Cognition and Categorization*, Hillsdale, NJ: Erlbaum, pp. 27–48.
- Rountree, J. (1997) “The Plausibility of Teleological Content Ascriptions: A Reply to Pietroski,” *Pacific Philosophical Quarterly* 78, no. 4: 404–20.
- Ryder, D. (2004) “SINBAD Neurosemantics: A theory of Mental Representation,” *Mind and Language* 19(2): 211–40.
- Salmon, N. (1986) *Frege’s Puzzle*, Cambridge, MA: MIT Press.
- Schroeder, T. (2004a) “Functions from Regulation,” *The Monist*, 87, 115–35.
- (2004b) “New Norms for Teleosemantics,” in H. Clapin, P. Staines, and P. Slezak (eds), *Representation in Mind: New Approaches to Mental Representation*, Amsterdam: Elsevier, pp. 91–106.
- Searle, J. (1980) “Minds, Brains, and Programs,” *Behavioral and Brain Sciences* 3: 417–24.
- (1983) *Intentionality: an Essay in the Philosophy of Mind*, Cambridge:

- Cambridge University Press.
- (1992) *The Rediscovery of the Mind*, Cambridge, MA: MIT Press.
- Sellars, W. (1963) *Science, Perception and Reality*, London: Routledge & Kegan Paul.
- Shea, N. (2007) “Consumers Need Information: Supplementing Teleosemantics with an Input Condition,” *Philosophy and Phenomenological Research* 75, no. 2: 404–35.
- Soames, S. (1989) “Direct Reference, Propositional Attitudes and Semantic Content,” *Philosophical Topics*, 15, 44–87.
- Sober, E. (1985) “Panglossian Functionalism and the Philosophy of Mind,” *Synthese* 64, no. 2: 165–93.
- Stalnaker, R. (1984) *Inquiry*, Cambridge, MA: MIT Press.
- Stampe, D. (1977) “Toward a Causal Theory of Linguistic Representation,” in P. French T. Uehling, and H. Wettstein, H. (eds), *Contemporary Perspectives in the Philosophy of Language*, Midwest Studies in Philosophy, vol. 2, Morris, MN: University of Minnesota, pp. 42–63.
- Sterelny, K. (1990) *The Representational Theory of Mind*, Oxford: Blackwell.
- Stich, S. P. (1983) *From Folk Psychology to Cognitive Science*, Cambridge, MA: MIT Press.
- Swoyer, C. (1991) “Structural Representation and Surrogate Reasoning,” *Synthese* 87: 449–508.
- Tye, M. (1995) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*, Cambridge, MA: MIT Press.
- Usher, M. (2001). A statistical referential theory of content: using information theory to account for misrepresentation. *Mind & Language*, 16(3), 311-34.
- Wittgenstein, L. (1953) *Philosophical Investigations*, Oxford: Blackwell.
- Wright, L. (1973) “Functions,” *Philosophical Review*, 82: 139–68.
- Yablo, S. (1997) “Wide Causation,” *Philosophical Perspectives* 11: 251–81.