The Components of Content[*]

David J. Chalmers

Department of Philosophy
University of Arizona
Tucson, AZ 85721.

chalmers@arizona.edu

*[[This is a heavily revised version of a paper first written in 1994 and revised in 1995. Sections 1, 6, 7, and 9 are similar to the old version, but the other sections are quite different. Because the old version has been widely cited, I have made it available (in its 1995 version) at http://consc.net/papers/content95.html.]]

1 Introduction

Here are six puzzles about the contents of thought.[*]

*[[For background material on the six puzzles, see: (1) Putnam 1975, Burge 1979; (2) Frege 1892; (3) Kripke 1979; (4) Perry 1979; (5) Schiffer 1990; (6) Kripke 1980.]]

(1) *Is content in the head?* Oscar believes that water is wet. His twin on Twin Earth, which is just like Earth except that H2O is replaced by the superficially identical XYZ, does not. His thoughts concern not water but twin water: Oscar believe that water is wet, but Twin Oscar believes that twin water is wet. This suggests that what a subject believes is not wholly determined by the internal state of the believer. Nevertheless, the cognitive similarities between me and my twin are striking. Is there some wholly internal aspect of content that we share?

(2) *Frege's puzzle*. In thinking that Hesperus is Hesperus, I refer to the same objects as in thinking that Hesperus is Phosphorus. But the first thought is trivial and the second is not. How can this difference in cognitive significance be reflected in a theory of content?

(3) *Kripke's puzzle*. In France, Pierre is told that "Londres est jolie", and he believes it. Later, he arrives in London and thinks it is ugly, never suspecting that "London" and "Londres" name the same city. It seems that Pierre simultaneously believes that London is pretty and that London is not pretty. But Pierre is highly rational, and would never believe a contradiction. What is going on?

(4) *The problem of the essential indexical*. When I believe that I am in danger, I will take evasive action. This belief state seems to be essentially indexical, or self-directed; if I merely believe that *x* is in danger, where *x* happens to be me, I might do something else entirely. How can we square this indexical aspect with an account of the contents of thought?

(5) *The mode-of-presentation problem*. If Jimmy says "Lois believes that Superman can fly", he speaks truly. If he says "Lois believes that Clark Kent can fly", he speaks falsely. But on many accounts, the proposition that Clark Kent can fly is the same as the proposition that Superman can fly. If so, it seems that to believe that Clark Kent can fly, it is not enough to believe in the corresponding proposition; one must believe it under an appropriate mode of presentation. But what is a mode of presentation, and how can these be integrated into an account of belief ascription?

(6) *The contingent a priori*. Say it is stipulated that one meter is the length of a certain stick in Paris. Then it seems that one knows *a priori* that the stick is one meter long, if it exists. But it seems the corresponding proposition is contingent, as it might have been that the stick was longer than one meter, or shorter. How can one have *a priori* knowledge of the truth of a contingent proposition?

These puzzles are not unrelated. All of them suggest incompleteness in a familiar view of thought content, on which the content of a belief that *P* may depend directly on the external objects of the belief. In particular, most of them raise questions about how well such an account of thought content reflects *rational* or *cognitive* aspects of thought. Because this sort of content is strongly tied to external factors, these accounts seem to have trouble capturing the rational relationships between thoughts (as witnessed by puzzles 2, 3, and 6), and their role in guiding cognition and action (as witnessed by puzzles 1 and 4).

To resolve these and other puzzles, many have postulated a separate dimension of content - so-called "narrow content" - that depends only on the internal state of a thinker, and is more closely tied to cognition and action.[*] But the road from intuition to theory has been a difficult one, and no account of narrow content has yet gained widespread acceptance. It is widely held that because narrow content is internal, it lacks the sort of relation to the external world that is required to qualify as *content*. Many have thought that narrow content is not the sort of thing that can be true or false, for example, as the Twin Earth cases show us that truth-conditions are not determined internally.

*[[Arguments for narrow content can be found in Dennett 1981, Fodor 1987, Lewis 1994, Loar 1988, Segal 2000, and White 1982.]]

I think that these problems are illusory, and that there is a robust and natural notion of narrow content such that narrow content has truth-conditions of its own. This can be seen by developing the idea that content has two dimensions. On the account I will given here the content of a thought can be decomposed into two components: its *epistemic* and *subjunctive* content. Subjunctive content is a familiar external variety of content. Epistemic content is a less familiar component, with the following properties: (1) it is determined by the internal state of a cognitive system; (2) it is itself a sort of truth-conditional content; (3) it governs the rational relations between thoughts. The first property ensures that epistemic content is a variety of narrow content. The second ensures that it is a truly semantic variety of content. The third ensures that it is central to the dynamics of cognition and action. These three properties together help to resolve many problems in the philosophy of mind and language.[*]

*[[What follows is an application to these issues of a framework I have developed in other papers (Chalmers 2002e, 2002f, 2002s). The discussion often passes over details that are explored in more depth in those papers. The framework presented here has much in common with existing ideas in the philosophy of mind and language, especially Kaplan's (1989) and Stalnaker's (1978) two-dimensional analyses of language and Lewis's (1979) analysis of the contents of thought, and various proposals that have been made about the nature of narrow content. For some connections between these ideas, see section 8 and Chalmers (2002f).]]

2 Intensions

In what follows, a *thought* is a token propositional attitude that aims to represent the world: for example, a belief, an expectation, or a hypothesis. A *concept* is a mental token that aims to pick out something (such as an individual, a class, or a property) in the world. I will take it that thoughts have *truth-values* (truth, falsity, and possibly others), and that concepts have *extensions* (for example, individuals, classes, or properties). Thoughts can usually be expressed in language by sentences, and concepts can usually be expressed in language with terms. For ease of exposition I will restrict attention to concepts and thoughts that can be expressed in language; if there are concepts that cannot be so expressed, the account can be extended to these.

It is a familiar idea that concepts and thoughts can be associated with an *intension*: a function from possible worlds to extensions or truth-values. The intension of a concept maps a world to the concept's extension in that world. The intension of a thought maps a world to the thought's truth-value in that world. In effect, a concept's intension captures the way that its extension depends on the nature of the world, and a thought's intension captures the way that it's truth-value depends on the nature of the world.

It is a somewhat less familiar idea that a concept or thought can be associated with *two* intensions. First, there is an *epistemic* intension, picking out a thought or concept's extension in a world *considered as actual*.[*] This intension captures the *epistemic* dependence of extension or truth-value on the way the actual world turns out. Second, there is a *subjunctive* intension, picking out a thought or concept's extension in a world *considered as counterfactual*. This intension captures the *subjunctive* dependence of extension or truth-value on counterfactual states of the world, given that the character of

the actual world is already fixed. On the picture I will develop, a thought's epistemic intension is narrow content, while a thought's subjunctive intension is often wide content.

*[[The phrase "consider a world as actual" is due to Davies and Humberstone (1981), developing ideas presented by Evans (1979). The explication given here differs from that given by Davies and Humberstone, who do not talk explicitly about epistemic possibility, but it is in much the same spirit.]]

3 Epistemic intensions

When we consider a world as actual, we consider it as an *epistemic possibility*: a way our world might actually be, for all we can know a priori.[*] For all I know a priori, there might be H2O in the oceans and lakes, or there might be XYZ. Let the H2O-world be a specific world with H2O in the oceans and lakes, and let the XYZ-world be a specific "Twin Earth" world with the superficially identical XYZ in the oceans and lakes. Then for all I know a priori, my world might be like the H2O-world, or it might be like the XYZ-world.

Let us say that a thought is *epistemically possible* (in a broad sense) when the thought cannot be ruled out by a priori reasoning. Then my hypothesis *water is H2O* is epistemically possible, as is my hypothesis *water is XYZ*: no amount of a priori reasoning can falsify either of these thoughts. These hypotheses can be associated with all sorts of *specific* epistemic possibilities, which we can represent using possible worlds. For example, lying behind the hypothesis that water is H2O lies the epistemic possibility that my world is like the H2O-world. Behind the hypothesis that water is XYZ lies the epistemic possibility that my world is like the XYZ-world.

For any world W, I can consider the hypothesis that W is actual: that is, I can consider the hypothesis that my world is qualitatively just like W. Such a hypothesis cannot be ruled out a priori: for all I know a priori, my own world could be like the H2O-world, or the XYZ-world, or any of a huge number of very different worlds. In effect, these worlds can be seen as constituting my *epistemic space*: the space of specific epistemic possibilities that are open to me a priori. If I had no empirical beliefs, all of epistemic space would be open to me. As I acquire empirical beliefs, epistemic space is narrowed down. Any given belief will typically *divide* epistemic space into those epistemic possibilities that it endorses and those that it excludes. The basic idea I will pursue is that the narrow content of a thought is given by the way that the thought divides epistemic space.

When one considers the hypothesis that a given world W is actual, this hypothesis *verifies* some of our thoughts, and it *falsifies* others. Here, we can say that one thought verifies another when it is rationally inconsistent to accept the first but deny the second; and one thought falsifies another when it is rationally inconsistent to accept both. For our purposes here, is is natural to say that two thoughts are rationally inconsistent when their conjunction is epistemically impossible: that is, when their conjunction can be ruled out by a priori reasoning. On this interpretation, a hypothesis verifies a thought when acceptance of the hypothesis can lead by a priori reasoning to acceptance of the thought.

Take my thought *water is H2O*. When I consider the hypothesis that the H2O-world is actual, this verifies my thought: if I accept that the H2O-world is actual, I must rationally conclude that water is H2O. That is, it would be rationally inconsistent to accept that the H2O-world is actual (i.e. that the liquid surrounding me with a certain appearance and distribution is and always has been H2O, and so on) but deny that water is H2O. When I consider the hypothesis that the XYZ-world is actual, this falsifies my thought: if I accept that the XYZ-world is actual, I must rationally conclude that water is not H2O. That is, it would be rationally inconsistent to accept that the XYZ-world is actual (i.e. that the liquid surrounding me with a certain appearance and distribution is and always has been XYZ, and so on), and at the same time to accept that water is H2O. So the H2O-world verifies my thought *water is H2O*, but the XYZ-world does not. Rather, the XYZ-world verifies a thought such as *water is not H2O*, or *water is XYZ*.

There is nothing here that contradicts the claim by Kripke and Putnam that water is necessarily H2O. Kripke and Putnam are dealing with what is often called "metaphysical" possibility and necessity, which is usually sharply distinguished from epistemic possibility and necessity. Even if it is not metaphysically possible that water is XYZ, it is epistemically possible that water is XYZ: we could *discover* that water is XYZ, for example. If Kripke and Putnam are right, then when the XYZ-world is considered as a metaphysical possibility (in effect, considered as a counterfactual world different from ours), it is best described as a world where water is XYZ. But it is clear that when considered as an *epistemic* possibility (i.e. considered as a way our own world may be), and when verification is defined as above, it verifies the hypothesis that water is XYZ.

There is a small complication. If we consider an objective world W as actual, this does not yield a fully determinate epistemic possibility. Take a world W containing both H2O and XYZ, in the oceans and lakes of separate planets. Then if I consider W as actual, I am not in a position to determine whether water is H2O or XYZ, since I do not know which planet I am on. In effect, a fully determinate hypothesis must include information about my *location* within a world. To handle this, we can represent epistemic possibilities by *centered* worlds: worlds marked with an individual and a time at their "center".[*] A centered world corresponds to a world *from a perspective*, marked with a viewpoint at its center. In the case above, there will be many centered worlds corresponding to W, some of which are centered on individuals on the H2O planet, and some centered on individuals on the XYZ planet. Now, when I consider the hypothesis that a centered world W' is actual, I consider the hypothesis that my world is qualitatively just like W', that I am the individual marked at the center of W', and that now is the time marked at the center of W'. Given that sort of information in the case above, I will be in a position to determine which planet I am on, and I will be in a position to determine whether water is H2O or XYZ.

*[[This notion is introduced by Quine (1968), who defines a centered world as a world with a marked space-time point. The definition above is due to Lewis (1979).]]

So a thought's epistemic intension can be seen more precisely as a function from centered worlds to truth-values. We can say that the epistemic intension of a thought T is true at a

centered world W when W verifies T, and is false at a centered world W when W falsifies T. As before, W verifies T when it is rationally inconsistent to accept that W is actual and deny T, and W falsifies T when it is rationally inconsistent to accept that W is actual and accept T.

For a more precise definition of epistemic intensions, we would need to be more precise about what it is to consider a centered world as actual. We can say that to consider W as actual is to consider the hypothesis that D is the case, where D is a canonical description of W: a sort of neutral qualitative description of the character of W (including its physical and mental character, for example), and of the center's place within it. A formal account will restrict canonical descriptions to semantically neutral terms (roughly, terms that are not themselves vulnerable to Twin Earth thought experiments, thus excluding most names, natural kind terms, and terms used with semantic deference), augmented by indexical terms such as 'I' and 'now'. We can then say that W verifies a thought T when the conjunction of T with the hypothesis that D is the case is epistemically impossible. Or equivalently, W verifies T when a material conditional 'if D, then S' is a priori, where S is a linguistic expression of T. These matters are explored in much more depth in Chalmers (2002f); for our purposes here, the informal understanding will suffice.[*]

*[[On a full account, a canonical description involves an epistemically complete statement in an idealized language that uses only semantically neutral terms and indexicals. A statement D is epistemically complete when D is epistemically possible and there is no S such that both D&S and D&~S are epistemically possible. A semantically neutral term is roughly one that is not vulnerable to Twin Earth thought experiments, or one that behaves in the same way under epistemic and subjunctive evaluation. The indexicals allowed include 'I', 'now', and any others required to characterize the center of the world (see footnote XX below). This treatment requires that for every centered world, there exists an epistemically complete description using only semantically neutral terms and indexicals. This can be supported by noting (i) that there will be an epistemically complete description for every world (a consequence of the idealization of the language), and (ii) that semantic non-neutrality does not in itself add expressive power in characterizing epistemic possibilities (at most, it affects the description of metaphysical possibilities). See Chalmers 2002f for more here.]]

In the case of *water is H2O*, it seems that the thought's epistemic intension will be true at the H2O-world (a world centered on Oscar surrounded by H2O, say), and will be false at the XYZ-world (a world centered on Twin Oscar, surrounded by XYZ). To a first approximation, one might suggest that the thought's epistemic intension will be true in a centered world when the clear, drinkable liquid around the center of that world has the molecular structure of H2O. This seems to capture *roughly* what it takes for us to judge that water is H2O in the actual world, depending on how that world turns out. But this sort of approximation is no replacement for the real intension. The intension itself is best evaluated by considering specific worlds as epistemic possibilities, and determining the consequences for the truth-values of our thoughts.

The existence of epistemic intensions is grounded in the fact that given sufficient information about the actual world, we are in a position to know whether our thoughts are true. For example, given sufficient information about the appearance, behavior, composition, and distribution of objects and substances in my environment, I am in a position to determine whether water is H2O. And if the information had turned out

differently, I would still have been in a position to determine whether water is H2O. So given enough relevant information about a centered world, I am in a position to determine whether, *if* that information is correct in my own world, water is H2O. The same goes for all sorts of other thoughts. It may be that in some cases, a complete specification of a centered world does not settle the truth-value of a thought one way or another. In that case, we can say that the thought's epistemic intension is *indeterminate* at that world. But otherwise, the thought's epistemic intension will be true or false at the world.

To help evaluate an epistemic intension at a world, one can use various heuristics. One useful heuristic for evaluating the epistemic intensions of one's own thought T, expressible by a sentence S, is to evaluate an indicative conditional: 'if W is actual, is S the case?' Here, as with other indicative conditionals, one evaluates this epistemically: one hypothetically accepts that W is actual, and uses this to reach a rational conclusion about whether S is or is not the case. If yes, then W verifies T; if not, then W falsifies T. To stress the epistemic character of the conditional, one can also appeal to "turns-out" conditionals such as the following: 'if W turns out to be actual, will it turn out that S'? For example, it seems reasonable to say that if the XYZ-world *turns out* to be actual, then it will *turn out* that water is XYZ.[*]

*[[Unlike the official definitions, these heuristics appeal only to rationality, and not to apriority. For this reason, they can provide at least an approximation to the notion of an epistemic intension for one who rejects the notion of apriority. Still, on my own view, the definition in terms of apriority is more fundamental; if the heuristics and the official definition give different results in special cases (see e.g. Yablo 2002, Chalmers 2002c), one should use this official definition.]]

Some thoughts have a very straightforward epistemic intension. For example, it is plausible that the epistemic intension of my thought *I am a philosopher* will be true at precisely those centered worlds where the individual at the center is a philosopher. The identity of the individual at the center does not matter: it might be David Chalmers, and it might be Immanuel Kant. After all, my knowledge that I am not Immanuel Kant is a posteriori, so the Kant centered world represents an epistemic possibility for me in the broad sense: and it seems clear that *if* I accept that the Kant world is my actual world (i.e. that I am Kant philosophizing at the center of that world), then I should conclude that I am a philosopher.

As for a thought such as *Hesperus is Phosphorus*: it is plausible that this thought will be verified by roughly those worlds where the bright object visible in a certain position in the evening sky around the individual at the center is identical to the bright object visible in the morning sky around the individual at the center. Again, this captures roughly what it takes for us to judge that Hesperus is Phosphorus in the actual world, depending on how the world turns out.

With a mathematical thought such as *2+2=4*, or *pi is irrational*, the thought's epistemic intension will be true in all worlds. This reflects the fact that these thoughts can be justified a priori, so that the negations of these thoughts will not be rationally consistent with any a posteriori hypothesis (the conjunction will itself be epistemically impossible).

The same goes even for complex mathematical thoughts whose truth we are not in a position to know ourselves. The notion of epistemic possibility and necessity involves a rational idealization away from our contingent cognitive limitations: if it is even possible for a thought to be known a priori, then the thought is epistemically necessary. If so, it will have a necessary epistemic intension.

It is tempting to say that the reverse is also the case: that when a thought has a necessary epistemic intension, it is knowable a priori. Or equivalently, when a thought is epistemically possible, it is verified by some centered world. I think this claim is correct, and have argued for it elsewhere, but it is nontrivial. Some philosophical views entail that there are counterexamples to this claim. For example, on some theist views a god exists necessarily, but the existence of a god cannot be known a priori. If so, then *a god exists* is not a priori, but its epistemic intension will be true in all worlds. In effect, there are not enough possible worlds on this view to represent all epistemic possibilities. A similar result follows from some views on which the laws of nature in our world are the laws of all worlds: there will be no worlds with different laws to represent the epistemic possibility of different laws. And a similar result follows from some materialist views on which zombies are epistemically possible but not metaphysically possible: on some such views, no possible world will correspond to the zombie epistemic possibility.

All of these views are controversial, and I have argued elsewhere that they rest on an incorrect conception of necessity. Sometimes these views are presented as drawing support from Kripkean a posteriori necessities such as 'Hesperus is Phosphorus' and 'water is H2O', but the Kripkean examples are all compatible with the thesis that every epistemic possibility is verified by a centered possible world. So these views require a much stronger sort of necessity, one that there is reason to doubt exists. Still, one who accepts these views will deny the thesis that every epistemically possible thought is verified by a centered world. One can preserve a version of this thesis on such a view by defining epistemic intensions using something other than centered worlds: for example, one can define a space of "maximal epistemic possibilities" (or scenarios) using epistemically consistent hypotheses alone, and then make the case that every epistemically possible thought is verified by some maximal epistemic possibility. I have taken that approach elsewhere for full generality (see Chalmers 2002e), but for simplicity I will stay with the centered-world approach here.

One important note. It is tempting to suppose that the epistemic intension of a thought T can be evaluated in a centered world W by asking: what is the truth-value of T, as thought in W? But this is not so. On the present proposal, T's epistemic intension can be evaluated in worlds containing no copy of T; and even when a copy of T is present, it usually plays no special role. For example, my thought *I am a philosopher* is true of a centered world regardless of whether I think I am a philosopher there. To take a more extreme example, the epistemic intension of my thought *someone is thinking* is false in a centered world involving no thought. In these cases, all that matters is the epistemic relation between the hypothesis that W is actual and the thought T, and nothing here requires that T be present in W. One might define a *different* intension (a "contextual intension"; see section X and Chalmers 2002f) using the heuristic above, but such an

intension behaves in a quite different way, and will not have the same sort of epistemic properties as an epistemic intension. This will be important later.

One can define epistemic intensions for *concepts* as well as thoughts. A concept's epistemic intension picks out its extension in a world considered as actual. A precise definition involves some tricky details (see Chalmers 2002f), so here I will simply illustrate the idea intuitively. Let us take a singular concept C expressible by a term B. To evaluate C's epistemic intension in a centered world W, one considers the hypothesis that W is actual, and asks: 'what is B?' One can appeal to the indicative conditional 'if W is actual, what is B?', or one can appeal to the rational consistency of judgments of the form *C is such-and-such* with the hypothesis that W is actual.

For example, in the XYZ-world, the epistemic intension of my concept *water* picks out XYZ. As before, I can say: *if* the XYZ-world is actual, then water is XYZ. In the H2O-world, on the other hand, the epistemic intension of my concept *water* picks out H2O. More generally, one might say as a first approximation that in a given centered world W, the epistemic intension of my concept *water* picks out the dominant clear, drinkable liquid found in the oceans and lakes around the individual at the center. As before, however, this is just an approximation, and the true intension corresponds to the results of considering and evaluating arbitrary centered worlds as epistemic possibilities.

One can do something like this for an arbitrary concept. Even for a seemingly nondescriptive concept, such as *Gödel*, it will still be the case that *given* full information about a centered world and given the hypothesis that this information obtains in the actual world, I will be in a position to make a rational judgment about the identity of Gödel under that hypothesis. This mirrors the fact that given relevant information about the *actual* world, I am in a position to determine the identity of Gödel, and more generally am in a position to determine the extension of arbitrary concepts. For an concept, this rational dependence of judgments about extension on information about the character of the actual world can be encapsulated in an epistemic intension.[*]

*[[For this reason, the current framework can be seen as neutral between "causal" theories of reference (on which reference is determined by a causal chain) and "descriptive" theories of reference (on which reference is determined by a description). Even a causal theorist should allow that relevant information about the actual world dictates rational judgments about our concept's extension. It is this very methodology that underlies Kripke's arguments *for* the causal theory: in effect, he considers epistemic possibilities that we could discover to be the case (e.g., that a man called 'Gödel' stole the proof of the incompleteness of arithmetic from a man called 'Schmidt'), and reaches judgments about a term's extension on that basis (here, we judge that 'Gödel' will pick out the stealer, not the prover). So even on the causal theory, a term will plausibly have an epistemic intension: it it is just that this epistemic intension may have a causal element. For example, for the epistemic intension of my concept *Gödel* to pick out a given individual in a centered world, it may be required that that individual stand in the right sort of causal relation to the subject at the center of the world. See Chalmers 2002s for more here.]]

These epistemic intensions are often difficult to characterize in independent terms, but for some concepts this is straightforward. If we take a quasi-descriptive concept such as *Hesperus* (where we assume this functions to rigidly pick out the evening star in the actual world), we can say that the epistemic intension of *Hesperus* picks out the evening

star around the center of an arbitrary centered world. Or if *Julius* functions to rigidly pick out the inventor of the zip, the epistemic intension of *Julius* will pick out the inventor of the zip in a given centered world.

The epistemic intension for an indexical concept is also very simple. The epistemic intension of my concept *I* picks out the individual at the center of a centered world. The epistemic intension of *now* picks out the time at the center. The epistemic intension of "here" picks out the location of the individual at the center at the time of the center. The epistemic intension of *today* picks out (roughly) the day that includes the time at the center. And so on.

In many cases, when a thought is composed from concepts, the thought's truth-value will be determined by the concepts' extensions. For example, a thought of the form *A is B* will be true when the extension of *A* coincides with the extension of *B*. In these cases, the thought's epistemic intension will equally be determined by the concepts' epistemic intensions. For example, the epistemic intension of *A is B* will be true at a world when the epistemic intensions of *A* and *B* pick out the same individual there. One will find a similar compositionality of epistemic intensions wherever one finds compositionality of extensions.

4 Subjunctive intensions

In contemporary philosophy, epistemic intensions are much less familiar than another sort of intension: what we can call a *subjunctive intension*. To evaluate a thought's subjunctive intension, one evaluates it in a world *considered as counterfactual*. To consider a world as counterfactual, one considers it as an *subjunctive possibility*: as a way our own world might have been, but probably is not. In our world as it actually is, the liquid in the oceans and lakes is H2O, but the liquid in the oceans and lakes *might have been* XYZ. So we can say that the XYZ-world *might have* obtained, and that the XYZ-world represents a subjunctive possibility.[*]

*[["Subjunctive" because this sort of possibility is grounded in the semantically subjunctive notion of what might have been the case (Kripke is explicit about this), and because the evaluation of such possibilities reflects the use of subjunctive conditionals. See Chalmres 2002s here.]]

The subjunctive intension of a thought T in a world W picks out the thought's truth-value in W when W is considered as counterfactual. Here, we grant that the character of the actual world is already fixed and ask what *would have been* the case if W had obtained. If T is expressible by a sentence S, we can evaluate T's subjunctive intension by asking: if W had obtained, would S have been the case? If yes, then T's subjunctive intension is true in W; if no, then T's subjunctive intension is false in W. When T's subjunctive intension is true in W, we can say that W *satisfies* T.

For example, if Kripke and Putnam are correct, then if the XYZ-world had obtained - that is, if the liquid in the oceans and lakes *had been* XYZ - then nevertheless, XYZ would not have been water.[*] XYZ would merely have been watery stuff, and water would still

have been H2O. If so, then the XYZ-world satisfies my thought *water is H2O*, and the subjunctive intension of my thought is true at the XYZ-world. More generally, if Kripke and Putnam are correct, then the subjunctive intension of my thought is true at all possible worlds.

*[[I think that it is not obvious that Kripke and Putnam are correct about this, and a case can be made that it might have been that water was XYZ. But for the purposes of this discussion, I will go along with the common view that Kripke's and Putnam's intuitions are correct here. I also think that even if Kripke and Putnam are right about language, it is not obvious that this extends to thought. But again, for the purposes of this discussion, I will go along with the common view that the modal properties of a term such as 'water' mirror modal properties of the underlying concept *water* that the term expresses.]]

It is clear that subjunctive intensions can behave quite differently from epistemic intensions. We have seen that the XYZ-world verifies *water is not H2O*, but it satisfies *water is H2O*. This difference is rooted in the difference between epistemic and subjunctive possibility, and the corresponding difference between considering a world as actual and as counterfactual. This is mirrored in the different behavior of indicative and subjunctive conditionals: it seems reasonable to say indicatively that if the liquid in the oceans and lakes is XYZ, then water is XYZ; but if Kripke and Putnam are right, it is not reasonable to say that if the liquid in the oceans and lakes *had been* XYZ, then water would have been XYZ. In considering a world as counterfactual, empirical facts about the actual world make a difference to how we describe it; in considering a world as actual, they do not.

Something similar goes for an indexical thought such as *I am David Chalmers*. If Kripke is right, it could not have been that I was not David Chalmers. If so, then *I am David Chalmers* is true in any world considered as counterfactual (or at least in any world where I exist). Note that there is no special need for a center in the world here: once we know all the objective facts about a counterfactual state of affairs, we know all that we need to know, even to settle indexical claims. So subjunctive possibilities can be represented by ordinary uncentered worlds, and subjunctive intensions are defined over uncentered worlds.

We can associate subjunctive intensions with concepts in a similar way. A concept's subjunctive intension picks out its extension in a world considered as counterfactual. For a concept C expressible by a term B, we can use B to ask: 'if W had been actual, what would B have been?' For example, in the case of *water*, we can say that if the XYZ-world had been actual, then water would still have been H2O. So the subjunctive intension of *water* picks out H2O at the XYZ-world, and plausibly picks out H2O in all possible worlds.

For many concepts, the concept's subjunctive intension picks out its actual extension in all possible worlds. This applies in particular to *rigid* concepts: those expressible by rigid designators, such as names or indexicals, picking out the same object in all worlds. For example, Kripke argues that 'Hesperus' is a rigid designator: if Hesperus is actually Venus, then Hesperus could not have been other than Venus. If so, then the subjunctive intension of *Hesperus* picks out Venus in all possible worlds. Similarly, it is plausible

that 'I' is a rigid designator: if so, then the subjunctive intension of my concept *I* picks out David Chalmers in all possible worlds.

For a purely descriptive concept such as *circular* or *the inventor of the zip*, by contrast, the subjunctive intension is plausibly very similar to the epistemic intension. Both the epistemic and subjunctive intensions of *the inventor of the zip*, for example, plausibly pick out whoever invented the zip in a given world. Note the difference with *Julius*, which has the same epistemic intension but whose subjunctive intension picks out the actual inventor in all worlds. The difference reflects the intuition that if (for example) Ned Kelly had invented the zip, he would have been the inventor of the zip, but he would not have been Julius. Some concepts behave in an intermediate manner: for example, the subjunctive intension of *the discoverer of water* does not pick out the actual extension in all worlds, but it is nevertheless quite different from the epistemic intension, due to the presence of the rigid concept *water* as a constituent.

The subjunctive intension of a concept or thought usually depends in some way on the concept's epistemic intension and the actual world. For a purely descriptive concept, the subjunctive intension may simply be a copy of the epistemic intension, across uncentered worlds. For a rigid concept, the subjunctive intension will be correspond to the value of the epistemic intension at the actual world, projected across all possible worlds. In other cases, the dependence may be somewhat more complex, but it will still exist.

We can encapsulate this dependence by associating concepts and thoughts with a *two-dimensional intension*. This intension maps an ordered pair (V,W) consisting of a centered and an uncentered world to an extension or a truth-value in W. When a thought T is evaluated at (V, W), it returns the truth-value of T in the counterfactual world W, under the assumption that V is actual. (Heuristic, where S expresses T: if V is actual, then if W had obtained, would S have been the case?). Like an epistemic intension, a two-dimensional intension can plausibly be evaluated without relying on empirical knowledge, since all the empirical knowledge one needs is given in the first parameter V. Then to evaluate a thought's subjunctive intension at W, one evaluates its two-dimensional intension at (A, W), where A is the actual centered world. To evaluate a thought's epistemic intension at a centered world W, we can evaluate its two-dimensional intension at (W, W'), where W' is an uncentered version of W.[*] This two-dimensional intension is useful for certain purposes, but most of the time we need only appeal to a thought's epistemic and subjunctive intensions.

Within this framework, we can easily analyze the Kripkean "necessary *a posteriori*". Let us say that a sentence S is *subjunctively necessary* when it is necessary in the familiar Kripkean sense: that is, when 'it might have been the case that S' is true. A thought is subjunctively necessary when it is expressible by a subjunctively necessary sentence. Then it is easy to see that when a thought is subjunctively necessary, its subjunctive intension is true in all worlds, and vice versa. Cases of the Kripkean "necessary *a posteriori*" (e.g. *water is H2O*) arise when a thought has a necessary subjunctive intension (the thought is true in all worlds considered as counterfactual) but a contingent epistemic intension (the thought is false in some world considered as actual). Cases of the

Kripkean "contingent *a priori*" (e.g. *Julius invented the zip*) arise when a thought has a contingent subjunctive intension but a necessary epistemic intension.

There should be no question of whether the epistemic or the subjunctive intension is *the* intension associated with a given concept. The full story can only be given two-dimensionally. One or the other may be more useful for various specific purposes. In matters of linguistic content, the subjunctive intension often plays a central role, as different users of a name or natural kind term may have quite different associated epistemic intensions, whereas the subjunctive intension will generally be constant. For questions about the rational properties of thought and its role in governing action, however, we will see that the epistemic intension is central.

5 Wide and narrow content

Let us call a thought or concept's epistemic intension its *epistemic content*, and a thought or concept's subjunctive intension its *subjunctive content*. Let us say that when a thought or concept's content depends on only on the intrinsic state of the thinker (that is, when every possible intrinsic duplicate of the thinker has a corresponding thought or concept with the same content), the content is *narrow*. And let us say that when content does not depend only on a thinker's intrinsic state (that is, when an intrinsic duplicate could have a corresponding thought or concept with different content), the content is *wide*. Then one can make the case that epistemic content is narrow, while subjunctive content is often wide.

It is clear that subjunctive content is often wide. For example, Oscar (on Earth) and Twin Oscar (on Twin Earth) are more or less intrinsic duplicates (abstracting away from differences due to the presence of H2O and XYZ in their bodies), and have corresponding concepts that they express by saying 'water'. But the subjunctive intension of Oscar's concept *water* picks out H2O in all worlds, while the subjunctive intension of Twin Oscar's concept *water* picks out XYZ in all worlds. Something similar applies to most rigid concepts, including *Hesperus* and even *I*. Here, a subjunctive intension depends on a concept's extension, which usually depends on a subject's environment, so two intrinsic duplicates can have different subjunctive intensions. In other cases, subjunctive content will not depend on the environment: for example, purely descriptive concepts such as *circular* and *the inventor of the zip*, will plausibly have subjunctive intensions that are shared between duplicates. But in cases where a concept or thought's subjunctive intension depends not just on its epistemic intension but on the way the actual world turns out, we can expect that subjunctive content will be wide content.

This environment-dependence does not extend to epistemic content. A concept's epistemic content is usually quite independent of its actual extension, and of the way the actual world turns out more generally. An epistemic intension encapsulates the way in which our rational judgments about extension and truth-value depend on how the actual world turns out, so can be evaluated without any knowledge of the actual environment, and plausibly does not depend on that environment.

This can be illustrated by looking at familiar cases. Take Oscar's and Twin Oscar's respective thoughts T1 and T2, expressed by saying 'there is water in my pool'. Let W1 be the Earth world centered on Oscar, with H2O in the oceans and lakes and H2O in Oscar's pool. Let W2 be the Twin Earth world centered on Twin Oscar, with XYZ in the oceans and lakes and XYZ in Twin Oscar's pool. Then clearly, W1 verifies T1 and W2 verifies T2. But also, W2 verifies T1: if Oscar hypothetically accepts that W2 is actual, he must rationally accept T1. Equally, W1 verifies T2: if Twin Oscar hypothetically accepts that W1 is actual, he must rationally accept T2. So the epistemic intensions of T1 and T2 are on a par with respect to these worlds.

Something similar applies to other worlds. Let W3 be a Twin Earth world centered on Twin Oscar with XYZ in the oceans and lakes, but an isolated amount of H2O in Twin Oscar's pool. Then W3 falsifies both T1 and T2. If Oscar accepts that W3 is actual, he should reject T1; if Twin Oscar accepts that W3 is actual; he should reject T2. The same goes for any other world: if W verifies T1, it will also verify T2, and vice versa. The same goes equally for any intrinsic duplicate of Oscar. We can even imagine Vat Oscar, who is a brain in a vat receiving artificial stimulation. Even so, Vat Oscar can entertain the hypothesis that W1 (or W2 or W3) is his actual world, and can reach rational conclusions on that basis, and the conclusions that he reaches will mirror those of Oscar and Twin Oscar. So Vat Oscar has a thought with the same epistemic intension as Oscar's, and the same holds for intrinsic duplicates in general.[*] So the epistemic content of Oscar's thought is narrow.

*[[Thus even a brain in a vat might have thoughts with epistemic content. This can be used to address Putnam's (1981) anti-skeptical argument that if he were a brain in a vat, he could not think "I am a brain in a vat". A brain in a vat could think a thought with the appropriate epistemic content, if not the appropriate subjunctive content; it could also think thoughts such as *I am in a skeptical scenario*, which have more or less identical epistemic and subjunctive content. The epistemic contents of these thought seem sufficient to express a significant skeptical possibility, true only in worlds in which the individual at the center lacks the usual sort of epistemic contact with the surrounding world.]]

The same goes for other thoughts and concepts. Even though I may have a twin whose concept expressed by 'Hesperus' has a different extension and subjunctive intension, this concept will nevertheless have the same epistemic intension as mine, picking out roughly the evening star near the center of any world. Although the *I* concepts of my twins will have an extension and subjunctive intension that differs from mine, they will have the same epistemic intension, picking out the individual at the center of any world. And so on.

One can even apply this analysis to the cases used by Burge (1979) to argue for the social nature of content. Bert has a belief that he expresses by saying 'arthritis sometimes occurs in the thighs'. In fact, arthritis is a disease of the joints and cannot occur in the thigh, so it seems that Bert has a false belief about arthritis. Twin Bert, an intrinsic duplicate of Bert, also has a belief that he expresses by saying 'arthritis sometimes occurs in the thighs'. But Twin Bert lives in a community in which the word 'arthritis' is used for a different disease, one that affects the muscles as well as the joints: we might call it 'twarthritis'. It seems that Twin Bert has a true belief about twarthritis. Where Bert believes (falsely) that he has arthritis in his thigh, Twin Bert does not: Twin Bert believes (truly) that he has

twarthritis in his thigh. Burge concludes that in this sort of case, belief content is not in the head.

Here, the crucial factor is that Bert uses the term 'arthritis' with *semantic deference*, intending (at least tacitly) to use the word for the same phenomenon for which others in the community use it. We might say that this term expresses a *deferential concept* for Bert: one whose extension depends on the way the corresponding term is used in a subject's linguistic community. It is clear that for deferential concepts, extension can depend on a subject's environment, as can subjunctive intension: the subjunctive intension of Bert's concept *arthritis* picks out arthritis in all worlds, while the subjunctive intension of Twin Bert's concept picks out twarthritis in all worlds.

Let T1 and T2 be the thoughts that Bert and Twin Bert express by saying 'arthritis sometimes occurs in the thighs'. Let W1 be Bert's own centered world, with a surrounding community that uses the term `arthritis' to refer to a disease of the joints. Let W2 be Twin Bert's centered world, with a surrounding community that uses `arthritis' to refer to a disease that can occur in the thigh. Then clearly W1 falsifies T1 and W2 verifies T2. At the same time, W2 verifies T1: if Bert accepts that W2 is actual - that is, if he accepts that his linguistic community uses 'arthritis' for a disease that can occur in the thighs - then (since his concept is deferential) he should rationally accept that arthritis can occur in the thighs, and so should accept T1. Similarly, W1 falsifies T2: if Twin Bert accepts that W1 is actual - that is, if he accepts that his community uses 'arthritis' only for a disease of the joints - then he should reject his thought T2. So the epistemic intension of T1 is false at W1 and true at W2, and exactly the same is true for T2.

Something similar applies to any other centered world that Bert and Twin Bert evaluate. In general, the epistemic intension of their *arthritis* concepts in those worlds will pick out the extension of the term 'arthritis' as used in the linguistic community around the center of those worlds. (In worlds where the term is not used, the epistemic intension will arguably be empty or indeterminate.) And the same goes for any intrinsic duplicate of Bert. Any such duplicate can entertain the hypothesis that a given world W is actual, and will rationally reach conclusions similar to Bert's.

One can apply the same reasoning to Putnam's case of 'elm' and 'beech', in which a subject can use the terms with different referents despite users having no substantive knowledge to differentiate the two. In this case, the terms are being used deferentially: the epistemic intension of the subject's concept *elm* picks out roughly whatever is called 'elm' around the center of a world, and the epistemic intension of her concept *beech* picks out roughly whatever is called 'beech' around the center of a world. Here again, the epistemic intension is independent of the environment. So we can see that semantic deference and "the division of linguistic labor" is quite compatible with thoughts and concepts having internally determined epistemic content.

Putnam suggests that such terms such as 'water' and 'elm' show that the intension of a concept cannot determine the extension, if an intension is internally determined. The current analysis shows that this is only half-true. The epistemic intension of a concept

determines its extension, and the epistemic intension is internally determined. Of course, the epistemic intension is a *centered intension*, taking a centered world as argument, so Putnam's claim still holds for uncentered intensions. But any intension requires facts about the actual world to determine extension, and it is most natural to regard the actual world of a thinker as centered, so an internally determined centered intension is very useful here.

Why is epistemic content narrow? Intuitively, this is because a thought's epistemic content is rationally prior to any knowledge of a subject's environment: it captures the way a thought's truth-value *depends* on the character of the environment, and so is independent of the environment itself. More deeply, it may be because epistemic content is defined in terms of the rational properties of thoughts, and these rational properties are internally determined. For example, if one subject has a thought that is justifiable a priori, a corresponding thought in any intrinsic duplicate of that subject will also be justifiable a priori; if so, a thought's epistemic necessity is determined by the internal state of the thinker. This observation can be combined with the observation that when one subject entertains the hypothesis that a world W is actual, any duplicate of that subject is also entertaining the hypothesis that W is actual. This second observation is grounded in the fact that these hypotheses involve semantically neutral descriptions of worlds, so there is no possibility of a "Twin Earth" difference between thinkers here. Putting these two observations together, it follows that if the hypothesis that W is actual epistemically necessitates a thought in one subject, it will also epistemically necessitate the corresponding thought in any duplicate subject. So epistemic content is narrow.

(Of course, the epistemic content of a thought will almost always depend *causally* on the external world, but it will not depend *constitutively* on the external environment. Whenever the external environment affects the epistemic contents of our thoughts, it will do so by affecting the internal state of the thinker.)

As promised, this sort of narrow content is truth-conditional. The epistemic content of a thought delivers conditions that one's actual centered world must satisfy in order for one's thought to be true. We might think of these as a thought's *epistemic* truth-conditions, as opposed to a thought's *subjunctive* truth-conditions, which govern truth across counterfactual worlds. Of course these truth-conditions can come apart at a given world: at the XYZ-world, the epistemic truth-conditions of *water is XYZ* are satisfied, but the subjunctive truth-conditions are not. This is to be expected, given the different functions of epistemic and subjunctive evaluation. One might worry that because of this, a thought could turn out to be both true and false, in the actual world, but this is impossible: when evaluated at the actual world, epistemic intensions and subjunctive intensions always give the same results.

6 The Advantages of Epistemic Content.

In recent times, the "content" of a thought has usually been identified with something like its subjunctive content;[*] but the epistemic content seems to be an equally good candidate. As before, there is no need to decide which is *the* content; but that being said,

there many ways in which the epistemic content of a thought is responsible for most of the explanatory work that we would expect a notion of content to do.

*[[Alternatively, content is often identified with a structured proposition composed from either subjunctive intensions of the concepts involved, or from the extensions of the concepts involved (when the concepts are rigid). This sort of structured content is more fine-grained than a subjunctive intension, but it has the same truth-conditions, and depends on the environment in a similar way. What I say below about subjunctive intensions applies equally to structured propositions. Likewise, what I say about epistemic intensions can easily be adapted to a view on which the contents of thoughts are structures composed from epistemic intensions of concepts.]]

First, epistemic content determines the rational relations between thoughts. If one thought implies another thought *a priori*, the epistemic intension associated with the first entails the epistemic intension associated with the second. (One intension entails another whenever the second is true at all worlds where the first is true.) If I know that it is hot where I am now, I know that it is hot here, and vice versa; this is reflected in the fact that the epistemic contents of the two thoughts are the same. The subjunctive contents of the thoughts are very different, however: there is no obvious relation between the intension true when it is hot where DJC is at time *t* and the intension that is true when it is hot at place *P* then.

It is straightforward to see why this is so. If one thought entails another *a priori*, then any centered world that verifies the first will verify the second. Conversely, it is plausible that if the epistemic intension of one thought entails that of another, a thinker should in principle be able to infer the second from the first by (idealized) *a priori* reasoning. (As before, this converse claim will be false on views of possibility on which not all epistemic possibilities can be represented by centered worlds: for example, on such a view the epistemic intension of *a god exists* might be entailed by any intension without the thought being a priori. Again, the claim can be preserved on such a view by moving to an expanded space of epistemic possibilities; but I will set this issue aside here.) This is not so for subjunctive intensions: entailments between these may turn on facts about the external world that are not accessible to the thinker.

This can be applied straightforwardly to explain the informativeness of a thought such as *Hesperus is Phosphorus*. Although its subjunctive intension is equivalent to that of the trivial *Hesperus is Hesperus*, its epistemic intension is quite distinct, so it is not cognitively trivial. In effect, epistemic intension here plays the role of Fregean sense. Again, it is epistemic intensions that reflect the rational properties of thoughts.

We can also invoke epistemic content in the case of Kripke's Pierre, who paradoxically seems to believe that London is pretty and that London is not pretty, without any breakdown in rationality. Pierre's concepts *Londres* and *London* have quite different epistemic intensions: in a given centered world, the first picks out (roughly) the famous city called 'Londres' that the individual at the center has heard about, whereas the second picks out (very roughly) the grimy city in which that individual has been living. The subjunctive intensions are identical, picking out London in every world. So Pierre's two beliefs *Londres is pretty* and *London is not pretty* have contradictory subjunctive

intensions, but their epistemic intensions are quite compatible. Rational relations are determined by epistemic content, so contradictory subjunctive intensions support no charge of irrationality.

Intuitively, Pierre's two beliefs are rationally compatible because there are specific ways the actual world could be that are consistent with both: that is, there are centered worlds that verify both. There is a centered world in which 'Londres' names a faraway, beautiful city (maybe it is in India), and in which the individual at the center inhabits an entirely distinct ugly city called 'London'; and for all Pierre knows and believes, such a world could be actual. As long as there is such a world, satisfying the epistemic content of all Pierre's thoughts - that is, as long as the *epistemic* contents of his thoughts are compatible - his rationality is not in danger.

This brings out the relation between this account and Dennett's (1981) suggestion that the narrow content of a thought is reflected in the *notional world* of the thinker. The notional world we can take to be a centered world (really a class of centered worlds) that verifies all of a subject's beliefs, or at least as many as possible. Pierre's notional world is a world in which there is a beautiful faraway city called 'Londres', and a grimy city close at hand called 'London'. If Pierre really lived in his notional world, he would be right about everything and rarely surprised.

*[[Dennett suggests that the relevant worlds are "the environment (or class of environments) to which the organism as currently constituted is best fitted." This class may be quite different from the class of worlds that verify all of a subject's beliefs: subjects are sometimes better fitted to worlds that falsify their beliefs (when they are pessimistic or altruistic, for example); they often have beliefs about distant matters that are irrelevant to fitness; and their fitness often turns on matters about which they have no beliefs. See also the criticisms in Stalnaker 1989 and White 1991, and White's more refined account. Dennett's and White's suggestions might be seen as a first attempt at giving a naturalistic reduction of something in the vicinity of epistemic content. Such a reduction is likely to be a major project in its own right.]]

On similar grounds, one can make the case that epistemic content reflects the *cognitive* relations between thoughts. Here there is an important qualification, as epistemic content as I have defined it does not distinguish the various cognitive relations that might hold between thoughts that are deductively equivalent. From the point of view of epistemic content, a complex mathematical proof is as trivial as modus ponens; so the fine-grained cognitive dynamics of deduction lies beyond the reach of epistemic content as I have defined it here. I think a more fine-grained variety of epistemic content can handle these cases better (see Chalmers 2002e), but I will set these issues aside here, as subjunctive content does not handle them any better, and they are largely independent of the issues at play in this paper.

A qualified thesis would be the following: insofar as epistemic content or subjunctive content reflect the cognitive relations between thoughts, the contribution of epistemic content *screens off* the contribution of subjunctive content. That is, in cases where two thoughts are cognitively related, then (1) in related cases where the epistemic content of the thoughts is held constant but the subjunctive content is varied, the cognitive relations are preserved (except insofar as cognitive relations can be affected by varying factors

independent of *both* epistemic and subjunctive content, as in the deductive case); and (2) in related cases in which the subjunctive content is preserved but epistemic content is not, the cognitive relations are damaged. One can make this case straightforwardly by examining cases; the details parallel those of the discussion of the explanation of behavior, below, so I will not duplicate them here.

A third advantage of epistemic content is its suitability for a role in the explanation of behavior. It is often noted that subjunctive content seems slightly out of synchrony with what one would expect of an explanatory psychological state. To use an example of Kaplan's (1989), if you are watching me and my pants catch fire, our respective beliefs that my pants are on fire now will have the same subjunctive content (true in all worlds in which DJC's pants are on fire at time *t*), but will lead to very different actions (I might jump into a river, while you just sit there). The difference between our actions does not seem to be something that a characterization in terms of subjunctive content alone can explain. In a similar way, belief states can produce very similar behavior for apparently systematic reasons, even when the beliefs have very different subjunctive content: witness the behavior that my twin and I produce when we think about twin water and water respectively, or the similarity between the actions of two people who think "I am hungry". A whole dimension of the explanation of behavior hard for subjunctive content to explain.

These explanations can be easily handled in terms of epistemic content. If you and I think *I am hungry*, the epistemic contents of our thoughts are the same, and that similarity is reflected in the similarity of our actions. When you and I both believe that my pants are on fire, on the other hand, our epistemic contents are very different, and our actions differ correspondingly. Note that this provides a straightforward solution to Perry's problem of the essential indexical: it is epistemic content, not subjunctive content, that governs action, and epistemic content, consisting in a centered intension, is a sort of *indexical* content.[*]

*[[Perry (1979) considers the possibility that centered ("relativized") propositions might provide a solution, but dismisses it on the grounds that believing that such a proposition *P* is "true for me" does not distinguish me from third parties who also believe that *P* is true for me, but act differently. The trouble is that Perry's locution "true for me" introduces an unnecessary extensional element. What distinguishes me from the third parties is rather that I believe *P simpliciter*, or better, that my belief has *P* as its epistemic content.]]

Epistemic content also accounts for the similarity of action between twin cases; this similarity reflects the fact that my beliefs about water and my twin's beliefs about twin water have the same epistemic content. But we need not move to the realm of science fiction to see the point. Two thoughts can share epistemic content even when two thinkers are quite different, as our thoughts *I am hungry* show, and even in these cases, similarities in epistemic content will lead to similarities in action, other things being equal. Say I think that Superman is across the road, and I want to have Superman's autograph: then other things being equal, I will cross the road.[*] If you have thoughts with similar epistemic content to mine, then you will do the same. But if your thoughts only share subjunctive content with mine, but have different epistemic content - say you

think that Clark Kent is across the road, but want Superman's autograph - then your behavior may be correspondingly quite different.

*[[To simplify the discussion, I make the happy assumption that Superman is actual and is identical to Clark Kent.]]

In general, whenever the content of a thought is causally relevant to behavior, its contribution is screened off by that of epistemic content in the following sense: if a thought had had the same epistemic content but different subjunctive content, the behavior would have indiscernible (except insofar it might be affected by changing factors independent of both sorts of content), whereas if it had had the same subjunctive content but different epistemic content, the behavior might have been correspondingly different.

To see the latter point, we need only examine cases like those above. The thoughts *I am hungry* and *The guy over there is hungry* (unknowingly looking in a mirror) will lead to very different behavior, even though their subjunctive content is the same. When Lois Lane is trying to cut Clark Kent's hair, her observation "Clark's hair breaks the scissors" will prompt a reaction very different from that provoked by a corresponding thought concerning Superman. If I hear that Cary Grant is starring in a movie, I might be more likely to watch than if I hear that the movie stars Archie Leach. In all these cases, different reactions are provoked by a difference in the epistemic content of a thought. In general, whenever the epistemic content of a thought is varied, different consequences can be expected, even if subjunctive content is preserved throughout.[*] Given that epistemic content governs cognitive relations and that cognition governs action, this is just what we would expect.

*[[Of course, thoughts like *Cary Grant is in the movie* and *Archie Leach is in the movie* might lead to the same actions despite their different epistemic content, if I know that Cary Grant is Archie Leach. But even here, there exist circumstances under which the thoughts might play a different role - if someone tells me that Cary Grant is not Archie Leach after all, for instance. In general, whenever two thoughts have different epistemic content, there are at least hypothetical circumstances under which the action-governing roles of the thoughts will differ.]]

By contrast, if the subjunctive content of a thought is varied but epistemic content is kept constant, behavior stays indistinguishable throughout. Perhaps, unbeknownst to me, Cary Grant is an elaborate hoax, a co-operative construction by avant-garde animators and the mass media. In such a case, my thought about Cary Grant will have no nontrivial subjunctive content, but as long as it has the same epistemic content, my behavior will be indistinguishable as in the case in which he is real. Or perhaps Cary Grant is really Ludwig Wittgenstein in disguise: if so, the thought has a very different subjunctive content, but the same behavior results. Similarly, when my twin and I think *I need some more water for this pot*, the subjunctive contents of our thoughts differ, but we both go to the sink.

We can make a similar point within a single system. Take Evans; example of 'Julius', which functions to rigidly designate whoever invented the zip. Then the epistemic

intensions of my concepts "Julius" and "the inventor of the zip" will be the same, but the subjunctive intensions will be very different. Despite the difference in subjunctive intensions, however, it is clear that any thoughts that *Julius is such-and-such* will play very much the same role in directing cognition and action as thoughts that *The inventor of the zip is such-and-such*. The rigidification and consequent difference in subjunctive intension is largely irrelevant. (One exception: the two concepts may behave differently in subjunctive thought, as when one judges that Julius might not have been the inventor of the zip, but not that Julius might not have been Julius. But even here the difference is accounted for by a difference in the internally determined two-dimensional intension, rather than by a difference in subjunctive content *per se*.)

Some might object that there are cases in which we individuate behavior extrinsically - for example, Oscar drinks water while Twin Oscar drinks twin water - so there is a dimension of behavior that escapes epistemic content. But even in this sort of case, subjunctive content does not usually help. Even Twin Oscar, with his different subjunctive content, would drink water if he was in Oscar's present environment. What is relevant to behavior here is not subjunctive content but current environment, as we can see by an extension of the varying-factors strategy; and I certainly do not wish to deny that current environment is relevant in the explanation of behavior.

The only cases in which there is a direct tie between subjunctive content and behavior are cases in which behavior is individuated by an intentional object, such as that in which we say that Oscar *searches for* a glass of water whereas Twin Oscar searches for a glass of twin water. This connection holds across all environments, as behavior only counts as water-searching if it is caused by water-thoughts. But for the same reason, this is a very weak sort of relevance for subjunctive content: as Fodor (1990) notes, the subjunctive contents of thoughts are not *causally* relevant to action here, but instead are *conceptually* relevant, in effect determining the category the action falls under.[*] And subjunctive content gives us very little purchase in the *explanation* of action here, as we will only know that some behavior is water-searching if we already know that water-thoughts lie behind it. In a causal (as opposed to a conceptual) explanation of the action, epistemic content will still play the central role.

*[[See Fodor 1990 for a detailed argument along these lines. I note also that one can individuate this sort behavior intentionally but still narrowly if one individuates by epistemic content.]]

Why is epistemic content primary? To answer this question, it is useful to think of my belief contents as constituting a model of my world, a kind of map by which I steer. This is a model of the world as I find it, a centered world with me at the center, and my beliefs are constraints on that world. Beliefs constitute a model by constraining *epistemic space*: the space of epistemic possibilities that were open to me a priori. One belief might rule out *these* epistemic possibilities as a candidate for the world where I am, another might rule out *these*, until only a limited class of worlds is left. I operate under the assumption that my world is one of those worlds, and if I am lucky I will not be too surprised.

My world-model is ultimately a *notional* world: a set of epistemic possibilities, such that none of these would overly surprise me if they turned out to be actual. The constraints on these possibilities are those of epistemic content. Any further constraints imposed by subjunctive content are not useful to me. The subjunctive content of my belief that the liquid in thermometers is mercury endorses only those worlds in which thermometers contain the element with atomic number *X*, but this constraint is so distant that if it turned out that the liquid has atomic number *Y*, I would not be in the least surprised. In an important sense, this constraint is not reflected in my world-model at all. Insofar as my world-model is useful to me in guiding cognition and action, the constraints on it are entirely those of epistemic content.

It is worth noting that in making a case for the primacy of epistemic content, I have not appealed to any *a priori* methodological principles such as the dictum that what governs behavior is in the head. The case for epistemic content has been made directly, independently of questions about physical realization. Indeed, it should be stressed that nothing I have said implies that facts about a thinker's environment are irrelevant to the explanation of behavior. Facts about the proximal environment will clearly play an important role insofar as they affect the thinker;[*] facts about the current environment are crucial to explaining the success or failure of various actions; and facts about environmental history will at least be central to a causal explanation of a thinker's current cognitive state. All that follows from the present framework is that the environment is not relevant to the explanation of behavior *in virtue of its role in constituting subjunctive content*. The kind of *content* that governs behavior is purely epistemic.

*[[It may even be that in certain cases, epistemic content can itself be constituted by an organism's proximal environment, in cases where the proximal environment is regarded as part of the cognitive system: if a subject's notebook is taken to be part of a subject's memory, for example (see Clark and Chalmers 1998). Here, epistemic content remains internal to a cognitive system; it is just that the skin is not a God-given boundary of a cognitive system. This is another way in which the issue between epistemic and subjunctive content runs deeper than the issue between internalism and externalism.]]

7 Belief Ascription and Psychological Explanation

All this raises a puzzle about the role of belief ascriptions in psychological explanation. If what has gone before is correct, the kind of content that governs cognition and action is epistemic content, which is narrow. But at the same time, there is strong evidence that the kind of content attributed by belief ascriptions is often wide. Does this mean that the common-sense framework of explanation of behavior in terms of belief ascription should be discarded? Alternatively, is the success of the common-sense framework evidence that something in these arguments has gone badly wrong?

Neither conclusion is justified. The present framework shows how it can at once be true that (1) belief ascriptions ascribe wide content, (2) narrow content governs action, and (3) belief ascriptions explain action. In short: Belief ascriptions ascribe a combination of epistemic and subjunctive content. It is in virtue of the subjunctive component that the ascribed content is wide, and it is in virtue of the epistemic component that the ascribed content is explanatory.

A full justification of this answer requires two things: first, an analysis of what is attributed in belief ascriptions, so that we can see precisely what sorts of epistemic and subjunctive content are attributed; second, an analysis of the role of belief ascriptions in psychological explanation, so that we can see that even in ordinary practice, it is the epistemic content attributed that carries the explanatory burden. I cannot provide anything like a complete treatment of these matters - the analysis of belief ascriptions deserves entire volumes of its own - but I can provide a preliminary sketch.

It is easy to see that ordinary belief ascriptions ascribe both epistemic and subjunctive content. If I say 'Ralph believes that Clark Kent is muscular', in order for my utterance to be true Ralph must have a belief that satisfies two sorts of constraints. First, the belief must have the subjunctive content of the proposition that Clark Kent is muscular (perhaps we can allow a certain amount of variation in the subjunctive content, if for example his concept of muscularity is slightly different from the norm). But that alone is not enough: a belief that Superman is muscular would have the same subjunctive content, but would not make my ascription true. As is often noted (e.g. Schiffer 1990), for the ascription to be true, the belief must involve a concept that refers to its object (Clark Kent) under an appropriate mode of presentation.[*]

*[[Some views (e.g. Salmon 1986) take ascriptions such as 'Lois believes that Clark can fly' to be strictly speaking true, so that modes of presentation are irrelevant to truth. Even if these highly counterintuitive views are accepted, the current account can be viewed an account of the (pragmatic) intuitive correctness conditions of belief ascriptions. Either way, we need an account of these intuitive correctness conditions to explain the function of belief ascriptions in psychological explanation.]]

In the current framework, modes of presentation are naturally seen as epistemic intensions. If Ralph refers to Clark Kent under an epistemic intension that picks out whoever is called 'Clark Kent', or one that picks out whoever is that reporter with glasses at the Daily Planet, or some more complex intension in the vicinity, my belief ascription will be true. If Ralph refers to Clark Kent under an epistemic intension that picks out the guy in the cape, or one that picks out the strongest man in the world, my belief ascription will be false. One might say that for the ascription to be true, Ralph must refer to Clark Kent under a 'Clark Kent'-appropriate epistemic intension. Here, the conditions on a 'Clark Kent'-appropriate epistemic intension are somewhat vague and unclear, and they may well be context-dependent, but it is clear from an examination of cases that they are substantive.

To take another case, if I am right in saying 'Tom believes that he is hungry', then Tom must have a belief with more or less the appropriate subjunctive content, true of all those worlds in which Tom is hungry at time *t*, but there is a strong constraint on epistemic content too. In particular, Tom must refer to himself via the epistemic intension that picks out the individual at the center in every centered world. If he sees someone in the distance clutching their belly, without realizing that he is in fact looking into mirrors, then a thought that that person is hungry has the right subjunctive content, but on the most natural reading it does not make my ascription true. The ascription will only be true if Tom's belief refers to himself under a *self*-concept, which requires a very specific sort of epistemic content. One might say that here, Tom must refer to himself under a 'he'-

appropriate epistemic intension, where in context the only 'he'-appropriate epistemic intension is the purely indexical intension.

The general principle here is something like the following. A belief ascription 'x believes that S' is true when the ascribee has a belief with the subjunctive intension of S (in the mouth of the ascriber), and with an 'S'-appropriate epistemic intension. Here, the epistemic intension is usually much less strongly constrained than the subjunctive intension. The conditions on 'S'-appropriateness may well be complex and context-dependent; their precise nature is one of the hardest questions in the theory of belief ascriptions. One can make a few generalizations: much of the time, an epistemic intension that is not too different from the ascriber's will be 'S'-appropriate, and much of the time, an epistemic intension that involves the terms in 'S' itself will be 'S'-appropriate. But this does not yield any sort of general condition. Rather, the appropriateness-conditions are best revealed by careful investigation of judgments of the ascription's truth in specific cases involving various different epistemic intensions.

In effect, this yields truth-conditions on belief ascriptions that parallel those of what Schiffer (1992) calls a "hidden-indexical" theory of belief ascription (although I have remained neutral on the ascriptions' logical form), with epistemic intensions playing the role of modes of presentation.[*] If something like this is correct, then epistemic intensions yield a solution to Schiffer's "mode of presentation" problem.[*] Epistemic intensions are perfectly suited to satisfy what Schiffer (1990, p. 252) calls "Frege's constraint" on modes of presentation: roughly, that a rational person may both believe and disbelieve that *y* is such-and-such only if the two beliefs involve different modes of presentation of *y*. If "rationality" is interpreted to involve idealized a priori reasoning, then the satisfaction of this constraint follows from the fact that epistemic intensions reflect a priori connections between thoughts.

*[[See also Crimmins 1991 and Richard 1990. Many of the insights of these and other philosophers on the semantics of belief ascription should be straightforwardly adaptable to the present framework.]]

*[[This sort of possibility is not mentioned in Schiffer's (1990) otherwise thorough survey of potential modes of presentation.]]

We can apply this to the case of Pierre, and the ascriptions 'Pierre believes that London is pretty' and 'Pierre believes that London is not pretty'. To satisfy these ascriptions, Pierre must have beliefs with the specified subjunctive intension, referring to London under a 'London'-appropriate epistemic intension. Pierre's *London* and *Londres* concepts have different epistemic intensions, but both satisfy the conditions on 'London'-appropriateness. So by virtue of his belief *Londres is pretty*, Pierre satisfies the first ascription, and by virtue of his belief *London is not pretty* he satisfies the second. Before, we explained Pierre's *beliefs* by noting that his two beliefs involve have contradictory subjunctive intensions but compatible epistemic intensions, and only the latter is relevant to rationality. Now, we can explain the apparent contradiction in the belief *ascriptions* by noting that two different epistemic intensions can both be 'London'-appropriate, so the two ascriptions do not in fact ascribe a rational contradiction to Pierre.[*]

*[[So Kripke's "Principle of Non-Contradiction" is false: someone can rationally believe that S and believe that not-S, as long as the beliefs involve different epistemic intensions both of which satisfy the appropriate constraints.]]

We have seen that content decomposes naturally into epistemic and subjunctive content; we now see that belief ascription puts strong constraints on both. Ideally, a full theory of belief ascription will specify the nature of these constraints for any given ascription, telling us the conditions that beliefs' epistemic and subjunctive contents must satisfy in order to make the ascription true. We can think of a belief ascription as marking out a subspace in the space of (epistemic content, subjunctive content) ordered pairs.

Given that epistemic content governs action, it follows that if belief ascriptions are to causally explain action, it must be in virtue of the epistemic content ascribed; the subjunctive content ascribed is redundant to the explanation. To make this case properly requires examining many specific cases, but the general point can be straightforwardly illustrated. One way to see the primacy of epistemic content is to consider belief ascriptions involving empty names, such as 'Santa Claus'. These ascribe no nontrivial subjunctive content, but ascription of beliefs about Santa Claus seem to function in precisely the same way in the explanation of action as do ascriptions of beliefs about real people. We might explain Karen's agitation on Christmas Eve in terms of her belief that Santa Claus is coming, that he will not fit down the chimney, and so on. Santa's non-existence and the corresponding absence of subjunctive content make little difference to the success of such an explanation. What governs Karen's actions are her *notions* of Santa Claus; and what governs the success of the explanation is the epistemic content that these ascriptions ascribe to her. And this is a *typical* case of the role of belief ascriptions in explanation: even when non-trivial subjunctive content is ascribed (as when the referent of the name exists), it makes little difference to the patterns of explanation.

In a very wide variety of cases in which content explains action, we can see that the explanation succeeds even if the subjunctive content attributed is ignored. For instance, if we explain my opening the refrigerator in terms of my belief that there is water in the refrigerator and my desire for a glass of water, we never need to invoke the H2O-involving subjunctive content. The explanation gains sufficient purchase from the epistemic content ascribed alone - roughly, the content that there is some of the liquid with the appropriate properties in the refrigerator, and that I want some of that liquid, and so on.[*]

*[[I leave aside here the important question of the epistemic content of desires, and the semantics of desire attributions. On my view, the epistemic content of a desire cannot in general be represented by a simple intension: rather, it is a sort of two-dimensional intension that can endorse a different set of worlds depending on which centered world is actual. This is clearest in cases such as "I wish I was two inches taller" or "I want to be over there". The moral is that the content of desires is perhaps more deeply two-dimensional than the content of beliefs.]]

It might be objected that there are cases in which the constraints on the epistemic content ascribed by a belief ascription are weak, so that subjunctive content must be doing any explanatory work. I think that ascriptions putting weak constraints on epistemic content

are rare, but assume they occur - perhaps an attribution of a belief about Smith constrains the relevant epistemic intension very little.[*] Even so, if we look at explanatory practice, we see that epistemic content is still doing the real work. For example, perhaps we explain why Bev goes to the pub by saying that she wants to see Smith and believes that Smith is at the pub. Leaving aside constraints in the concepts of seeing, the pub, and so on, there is a constraint on epistemic content implicit in the 'Smith' attributions. It is implicit here that the two 'Smith' concepts in Bev's thoughts have the same epistemic intensions. If her belief and her desire had very different epistemic content associated with 'Smith' is concerned - perhaps she wants to see Batman and believes that Smith is at the pub, not knowing that Smith is Batman - the inference from those states to her action would fail. So there is a strong joint constraint on epistemic content: despite a lack of constraint on the individual beliefs, Bev is implicitly ascribed the belief that a person she wants to see is at the pub. It is *this* ascribed belief that is doing the real explanatory work, and this ascription clearly puts a heavy constraint on epistemic content. To make the case that all such examples can be similarly analyzed requires a detailed treatment, but this illustrates the general pattern.

*[[How should one analyze so-called *de re* belief attributions, of the form 'S believes of x that it is y'? In the current framework, one might adapt the proposals of Kaplan 1967 and Lewis 1979 by holding that such an ascription is true when S has a belief with the appropriate subjunctive intension, true in worlds where A has property P (where A is the referent of 'x' and P of 'y'), picking out A under a *de-re*-appropriate epistemic intension. Here, a *de-re* appropriate intension is one that entails acquaintance: that is, one such that in any centered world the subject at the center is acquainted (in the contemporary non-Russellian sense) with the extension of the intension at that world, if that extension exists.]]

It follows that the centrality of narrow factors in the causation of action need not overthrow the role of belief ascriptions in explaining behavior, as some (e.g. Stich 1984) have suggested it should. At most we have shown that belief ascriptions are a somewhat rough-edged tool, due to the way they wrap both components of content into a single parcel, bringing the idle subjunctive content into play alongside the epistemic content that does all the work. But this should not surprise us; we cannot expect a folk theory to be maximally efficient.

(Why is subjunctive content ascribed at all? I think the reasons are tied to language. First, we ascribe beliefs in the same language we use to describe the world, and when we use world-involving language to ascribe epistemic content, world-involving constraints come along naturally in the package. Second, subjunctive content is important to understanding the success of communication and of collective action. When I tell you that I have a cold, you acquire a thought whose epistemic content is different from mine, but whose subjunctive content is the same. Communication very frequently involves transmission of subjunctive content, and our collective cohesion (if not our individual actions) can often be understood in terms of shared subjunctive content. But both of these points deserve a much more extensive development.)

In moving from common-sense psychology toward a developed cognitive science, we might expect that the kind of content that is invoked will become more purely epistemic, and that subjunctive content will be relegated to a secondary role or dropped entirely.[*]

We might also expect that better tools will be developed to specify the epistemic contents of thoughts than the current rough-and-ready language of belief ascription. This might qualify as a revision of our folk notion of belief, emphasizing and refining the elements of epistemic content that are already present within it. But precisely because those elements are already present and playing a central role in our practices, such a development would fall well short of elimination.

*[[It is arguable that cognitive psychology is already mostly concerned with epistemic content rather than subjunctive content, insofar as it is concerned with content at all. For example, the psychological literature on concepts seems to be largely concerned with how concepts are applied to the actual world, concentrating on something like the epistemic intensions of the concepts involved. See Smith and Medin 1981 and Patterson 1991.]]

8 Related Suggestions and Some Objections

The framework outlined here is related to a number of existing proposals. There is a clear structural resemblance to other broadly two-dimensional frameworks, such as proposals by Kaplan (1989) and Stalnaker (1978) for analyzing the content of language, and proposals by White (1982) and Fodor (1987) for analyzing the contents of thought. The idea that this sort of proposal can be used to yield a sort of narrow content has been criticized by Block (1991), Stalnaker (1990), and others, and extended to an earlier version of the present proposal by Block and Stalnaker (1999). So we need to examine the relationship between these proposals, to see whether the criticisms apply. I think that on examination, the current framework differs in fundamental respects from the others, and their problems do not arise here.

The relationship can be brought out by contrasting epistemic intensions with *contextual intensions*. A thought's contextual intension is defined by the heuristic discussed earlier: T is true in a centered world W (with T present at the center) if T is true as thought at the center of W. Likewise, the contextual intension of a concept C will return C's extension in worlds with C at the center. (One can define contextual intensions for sentences and other linguistic expressions similarly.) There are various possible variations here: one might have different requirements for what counts as a token of T in a world, or one might require only a token of T's type (for some relevant type) rather than T itself. But however one does things, the centered worlds here are functioning as *contexts* in which a thought (or concept) occurs, and a contextual intension encapsulates the *context-dependence* of a thought's truth-value or a concept's extension.

As we saw before, contextual intensions are quite different from epistemic intensions. An obvious difference: epistemic intensions give no special role to thought tokens within a centered world, and can be evaluated in worlds without any such tokens at the center. Thus the epistemic intension of *I am a philosopher* can be true at a world regardless of what the being at the center is thinking. And a thought such as *someone is thinking* has an epistemic intension that is plausibly false in some worlds (e.g. those without any thoughts), although its contextual intension is true in all worlds. A deeper difference: where contextual intensions represent context-dependence, with centered worlds representing contexts of thought, epistemic intensions represent *epistemic dependence*,

with centered worlds representing epistemic possibilities. This is a very different conception, and yields quite different behavior. This can be illustrated by the frameworks of Kaplan and Stalnaker. Kaplan defines the character of a linguistic expression type as a function from a context of utterance to the expression's content (roughly, subjunctive intension) relative to that context. In some ways this resembles the two-dimensional intension discussed above (in effect a function from centered worlds to subjunctive intensions), but the underlying ideas and resulting behavior are quite difference. For example, on Kaplan's framework names such as 'Hesperus' and 'Phosphorus' have identical characters, picking out the same content in all contexts. This happens because the referent of a name is essential to that name, so any use of the name in any context will have the same referent. For this reason, Kaplan notes that his framework cannot provide a solution to Frege's puzzle in the case of names, natural kind terms, and the like. But as we have seen, the epistemic intension associated with a subject's use of a name behaves very differently, often picking out different individuals in different centered worlds (whether a name has its referent essentially is irrelevant on a non-contextual understanding), and holds out much more hope of dealing with Frege's puzzle.

Stalnaker defines the diagonal proposition of an expression token as a function from a world containing the token to the truth-value of the proposition that the token expresses in that world, as evaluated in that world. This bears a formal resemblance to an epistemic intension, which can be seen as equivalent to the diagonal of a two-dimensional intension. But again, the underlying ideas and resulting behavior are different. On Stalnaker's framework, the diagonal proposition of an expression is defined at worlds where it has a very different meaning: so at a world where 'water is solid' means that snow is white, the diagonal proposition of 'water is solid' will be true is snow is white in that world. This is quite different from an epistemic intension: if my usage is nondeferential, the use of terms such as 'water' in a centered world will be irrelevant to epistemic intensions. Stalnaker (1999) notes that because of this, diagonal propositions are not closely connected to a priori truth. This seems correct, but the problem does not generalize to epistemic intensions, which have a built-in connection to a priori truth.

White (1982) and Fodor (1987) generalize these analyses to the contents of thought. Fodor defines the narrow content of a thought as a function from a context of thought to a thought's (wide) content in that context. White does something similar, although his account is slightly more complex and he requires that a functional duplicate of the original thinker be present in the relevant context. As before, these proposals are based on context-dependence, and give results that differ correspondingly: witness the intensions of *I am a philosopher* and *someone is thinking*.[*]

*[[Related proposals for understanding narrow content in broadly contextual terms are given by Brown 1986 and Loar 1988.]]

Block (1991) gives a number of objections to proposals of this sort. White's proposal is subject to a charge of *holism*: no two different subjects can have thoughts with the same narrow content, unless they are functional duplicates. Further, it seems that the narrow contents of a subject's thoughts all change every time the subject acquires a new belief, or

indeed every time that anything happens in the mind of the subject. This problem does not apply to epistemic intensions. There is no problem with quite different thinkers having thoughts with the same epistemic intension: very different people can have *I am a philosopher* thoughts with the same epistemic intension, for example. And epistemic intensions will not usually change with the acquisition of new beliefs. A change in epistemic intension requires a change in a subject's rational pattern of judgments about centered worlds considered as actual: a change in belief may change the subject's judgments about which centered worlds are actual, but it will not usually change a subject's rational judgment about what will be the case *if* a given centered world is actual.

Block charges Fodor's proposal with underdetermination: it is left unclear how to evaluate the mapping across worlds. The main problem is that of "what is held constant": one need to know just what features of the original thought must be present in the thought token at the center of a world, in order for it to fall in the domain of the intension. If some sort of mental syntax is held constant, the result will be an intension that delivers wildly varying results across worlds: there will be worlds where the mapping for *water* picks out steel, if a token with that mental syntax has a different meaning. If extension is held constant, then the intension will be trivial: the mapping for *water* picks out H2O in all worlds. For an intermediate result, one might suggest that the token's *narrow content* be held constant: but that presupposes what we are trying to explain. So it seems very difficult to set things up so that the mapping yields a notion of narrow content that behaves in an appropriate way.

Again, epistemic intensions do not have this problem. There is no issues concerning what to "hold constant" across worlds here, since there is no need for the original token to be present in different worlds. Rather, we simply appeal to the *original* thought, and to its epistemic relations with the hypothesis that a given world is actual. These epistemic relations are well-defined, being grounded in the idealized rational judgments of the subject. They also do not presuppose any theoretical notion of narrow content; but they can be used to ground such a theoretical notion.

Fodor himself (1987, p. 50) raises the problem that his sort of narrow content is not semantically evaluable (for truth and falsity), and so is not really content; rather, it is just *potential* content, delivering a content in a context. (He later rejects narrow content for this reason.) Again, epistemic content is immune to this problem. An epistemic intension is a sort of first-order content, placing direct constraints on the world, with truth-conditions of its own. Epistemic intensions can also stand in semantic relations such as entailment, and can be analyzed using semantic frameworks involving possible worlds, which allows for significant explanatory power.

Stalnaker (1990) considers the idea that some version of his diagonal proposition (or "realization conditions") might yield an account of narrow content, and raises three criticisms. First, he suggests that we cannot identity a thought independently of its content, so we cannot ask what the content of a belief *would* have been had it been a belief on twin earth. Second, he notes that diagonal propositions are defined only in worlds containing the relevant thought token, and cannot easily be extended to worlds

without the thought token. Third, he notes that on this proposal narrow content is derivative on wide content (since a diagonal proposition is defined using a two-dimensional matrix which is defined using wide content), so it presupposes rather than explains wide content. In response, it is fairly clear that the first two objections apply only to contextually defined narrow content, and not to epistemically defined narrow content. On the epistemic proposal, we never need to ask what the content of a belief would have been if it had been a belief on twin earth, and narrow content is defined in a straightforward way at worlds that do not contain the relevant thought token.

In discussing his second objection, Stalnaker raises a case which is worth addressing. If Bert uses his semantically deferential concept to think *my father has arthritis in his thigh*, how can we evaluate this thought in which there is no word 'arthritis' in Bert's language, and in which Bert has no thoughts about his father's health? On the epistemic framework, it is most natural to say that the epistemic intension of Bert's *arthritis* concept picks out nothing in this world. In effect, the use of a semantically deferential concept *presupposes* that one lives in a community that uses the relevant term, just as a notion such as *The present king of France* presupposes that there is a king of France. If we discover that those assumptions do not hold in the actual world, it is reasonable to judge that thoughts involving these concepts lack truth-value. The same goes for alternative possible worlds, considered as actual: so the epistemic intension of Bert's thought is indeterminate in the relevant worlds. In effect, Bert's thought partitions the space of centered worlds in which the background assumptions are satisfied, and says nothing about those worlds in which the assumptions are false.[*]

*[[Other concepts whose epistemic intensions have a limited domain of determinacy include perceptual demonstratives. When I think something like *That is pretty*, the referent of my demonstrative is often picked out (very roughly) as the cause of such-and-such experience. In a centered world in which there is no appropriate experience at the center, the epistemic intension may lack truth-value. This raises another subtlety: to capture the content of perceptual demonstratives, one may need to build in a "marked" experience to the center of the class of actual-world candidates, as one builds in a marked individual and time. Building this into the center will sometimes be needed to secure reference to otherwise indistinguishable experiences and their perceptual objects, as with (perhaps) a speckle in a large field, or one of the symmetrical red spots in Austin's (1990) "Two Tubes" puzzle (to which the present framework then provides a solution). In certain cases, centers may also require more than one experience, and perhaps a marked thought ("this very thought"). One might suggest that the contents of a center involve objects of "unmediated" reference: oneself, the present moment, the current thought, and perhaps certain experiences and orientations. This matter is closely connected to Russell's suggestions about direct reference; I hope to explore it in more depth elsewhere (see also Chalmers 2002b).]]

As for Stalnaker's third objection: narrow content may be derivative on wide content on the diagonal understanding, but not on the epistemic understanding. Epistemic content can be defined quite independently of subjunctive content, and our definition of epistemic intensions made no appeal to subjunctive evaluation. For this reason, an epistemic intension is not fundamentally a diagonal intension. *After* the fact, one can see an epistemic intension as equivalent to the diagonal of a two-dimensional intension involving both epistemic and subjunctive notions; but this complex construction is quite unnecessary to define epistemic intensions. One can characterize the first dimension of the framework in entirely epistemic terms, independently of the second dimension.

In the current framework, if anything, wide content derives from narrow content. We have already seen that the subjunctive intension of a concept is determined by the epistemic intension in conjunction with the environment. In some cases it is a near-copy of the epistemic intension, as for simple descriptive concepts; in other cases it is determined by rigidifying the actual-world extension of the epistemic intension. By contrast, we can tell the entire story about the epistemic intension without ever involving the subjunctive intension. It therefore seems that if either intension is more fundamental, it is the epistemic intension. Still, there is no need to make too strong a claim here: both epistemic and subjunctive content are important, and both have a role to play in different domains.

Block and Stalnaker (1999) give a number of objections to the version of this framework put forward in Chalmers (1996). Many of these objections echo the objections above, and turn on interpreting the proposal via a contextual rather than an epistemic understanding.[*] Another objection is that the formal two-dimensional apparatus alone does not yield intensions with the relevant properties. This is clearly correct, but on my approach it is a substantive characterization of the intensions that yields the properties in question. Block and Stalnaker also argue that the two-dimensional approach cannot explain or ground a notion of a priori truth. I have not suggested that the framework can do this; rather, I have used the notion of apriority in defining the framework. The notion of apriority, and the specific uses of it in grounding the framework, can be defended on quite independent grounds. The use of apriority in capturing the dependence of judgments about extension and truth-value on sufficient information about the world is defended at length by Chalmers and Jackson (2001).

*[[Block and Stalnaker note, in effect, that Chalmers (1996) does not intend a contextual interpretation, but suggest that a version of the "what is held constant" problem nevertheless arises in using an actual-word thought or concept to evaluate worlds without that thought or concept. I think that when things are understood in the appropriate epistemic terms, this problem clearly disappears. In fairness, it should be noted that the discussion in Chalmers (1996) is not explicit about the difference between contextual and epistemic intensions, and although the discussion tends to suggest an epistemic intension, the precise definition is left unclear. See Chalmers 2002f for discussion.]]

The current proposal also bears a resemblance to "descriptive" accounts of narrow content. It has sometimes been suggested that the narrow content of a concept such as *water* corresponds to the content of an associated description such as 'the dominant clear drinkable liquid in the environment', or some such. In response, a number of philosophers (LePore and Loewer 1986; Taylor 1989; White 1982) have objected that even if such descriptions exist, terms such as 'liquid' are themselves susceptible to Twin Earth scenarios (e.g., where liquids are replaced by superficially identical masses of sand), so that the content of such a description is wide rather than narrow. One might think that this objection will apply to the present proposal, since I have used descriptions of this sort to characterize epistemic intensions. But importantly, the description merely provides a rough handle on the intension for the purposes of illustration. The real narrow content is a function from centered worlds into extensions, and can be characterized fully on by specifying its value at specific worlds. As soon as we move to a summarizing description in language, imperfections are introduced, and the narrowness of the content is impurified.

But the intension itself remains narrow; we should not mistake the linguistic description for the real thing.[*]

*[[This might suggest that epistemic content is "ineffable". But the real problem is simply that it is difficult to capture the *epistemic content* of one expression with the *subjunctive content* of another. Just as one can capture the subjunctive content of a concept such as *water* by appealing to the equivalent subjunctive content of an expression such as 'H2O', one might capture its epistemic content by appealing to the equivalent epistemic content of an expression such as 'the clear, drinkable liquid...'. It is hard to see why the second is any more objectionable than the first, or why it makes epistemic content any more "ineffable". Thanks to Frank Jackson for discussion on this point.]]

One would obtain a more closely related sort of "description" theory of epistemic content if one abstracted away from linguistic characterizations and regarded the relevant "descriptions" simply as properties that a referent must satisfy, or better, as relations to the thinker. If we speak merely of properties and relations, the linguistic contamination is avoided. Schiffer (1978) suggests a description theory of this sort, on which there is irreducibly *de re* reference by a thinker to himself or herself, with reference to everything else is mediated by a property or relation. If we map the irreducible self-reference here to the appeal to centered worlds, and map the properties and relations to epistemic intensions, the resemblance between the accounts is clear, although Schiffer does not appeal to a two-dimensional framework, and addresses his proposal largely to the question of accounting for *de re* thought.

Another closely related idea is Lewis's (1979) proposal that belief involves the self-ascription of a property. The set of individuals satisfying a property corresponds directly to a class of centered worlds, as Lewis notes. Lewis (1994) argues that this sort of content is narrow and is primary in explanation. In effect, Lewis advocates a one-dimensional view of content, where apparent wide content is an artifact of belief ascriptions. While Lewis does not advocate understanding these contents in epistemic terms, and does not give a general characterization of the set of worlds associated with a belief, his examples suggest that these sets of worlds closely resemble those of an epistemic intension. So the present proposal appears to be highly compatible with Lewis's framework.

A residual problem for an account like this is the problem of hyperintensionality. It seems that two beliefs - mathematical beliefs, for example - can have the same epistemic and subjunctive intensions, while nevertheless having intuitively different content, and playing quite different roles in cognition and action. To handle these issues, one needs a more fine-grained sort of epistemic content that goes beyond epistemic intensions as I have defined them. One possibility is that one can appeal to intensions over a more fine-grained space of epistemic possibilities, defined using a more demanding epistemic necessity operator (see Chalmers 2002e for some ideas here). It is also possible that one might appeal to a more basic sort of content that lies behind and determines an epistemic intension. Epistemic and subjunctive intensions are aspects of the contents of thoughts, but I have not suggested that they exhaust these contents. The nature of a complete characterization of thought contents, if such a thing can be given, remains an open question.[*]

Another open question: is it possible to reductively explain the epistemic content of a subject's thoughts in naturalistic terms, in the way that some have attempted to explain wide content in causal or teleological terms? Certainly no such explanation is currently available. A first attempt might exploit the idea that epistemic content is mirrored in the idealized rational dispositions of the subject, although the normative character of the idealization may pose an obstacle to reduction, as will the fact that these dispositions are themselves characterized by appeal to content. My own view is that epistemic content is ultimately determined by a combination of a subject's functional organization and phenomenology,[*] so any attempt at explanation will need to appeal to these factors. In any case, it is arguable that if wide content depends heavily on narrow content, as the current account suggests, any adequate reductive theory of wide content will require a reductive theory of narrow content first.[*]

*[[For an argument that phenomenology is essential to the epistemic content of at least some concepts, see Chalmers 2002b. See also Horgan and Tienson (this volume) for arguments for phenomenally constituted narrow content that can be seen as complementing the current approach.]]

*[[Arguably, contemporary causal theories of content have been unsuccessful precisely because they attempt to account for wide content directly, without taking into account the crucial epistemic dimension involved in its determination.]]

9 Conclusion

What of the six puzzles at the start? To summarize:

(1) A thought's content decomposes into epistemic and subjunctive content, given by its epistemic and subjunctive intensions. Oscar's and Twin Oscar's thoughts differ in their subjunctive contents, and as a result ground different belief ascriptions, but their epistemic contents are the same.

(2) My thoughts that Hesperus is Hesperus and that Hesperus is Phosphorus have the same subjunctive intension but distinct epistemic intensions, as the *Hesperus* and *Phosphorus* concepts have different epistemic intensions. The triviality of the former does not imply the triviality of the latter, as it is epistemic content that governs rational relations.

(3) Pierre's two beliefs have contradictory subjunctive intensions but compatible epistemic intensions. The apparently contradictory belief ascriptions arise because of the contradictory subjunctive intensions and because his two concepts of London have distinct epistemic intensions that can both make 'London'-involving belief ascriptions true. There is no rational contradiction here, since rationality is governed by epistemic intensions.

(4) The essential indexicality of belief reflects the fact that epistemic content, not subjunctive content, governs action, and that epistemic content, unlike subjunctive content, is an indexical centered intension.

(5) The modes of presentation central to a theory of belief ascription are epistemic intensions. Belief ascriptions specify a believer's subjunctive content, and constrain the believer's epistemic content.

(6) Instances of the contingent *a priori* have a necessary epistemic intension but a contingent subjunctive intension. It is epistemic content that constrains one's world-model, so a contingent subjunctive proposition does not indicate a cognitive achievement.

There are many problems about the contents of thought that are not resolved by this framework. These include the problems of hyperintensionality, of a full account of belief ascriptions, and of giving a naturalistic explanation of content. Some of these matters are likely to be much more difficult than the puzzles at issue in this paper, but the two-dimensional approach at least clarifies the lay of the land.

Acknowledgments

Thanks to too many people to mention, but especially Ned Block, Curtis Brown, Frank Jackson, David Lewis, Robert Stalnaker, and two reviewers. Thanks also to audiences at talks between 1994 and 1997 at Arizona, Cornell, Memphis, Princeton, Rice, UCLA, UC Santa Barbara, UC Santa Cruz, Washington University, Yale, and the Australasian Association of Philosophy.

References

Austin, D.F. 1990. *What's the Meaning of "This"?* Ithaca, NY: Cornell University Press.

Block, N. 1991. What narrow content is not. In (B. Loewer & G. Rey, eds.) *Meaning in Mind: Fodor and his Critics*. Oxford: Blackwell.

Block, N. & Stalnaker, S. 1999. Conceptual analysis, dualism, and the explanatory gap.

Brown, C. 1986. What is a belief state? Midwest Studies in Philosophy 10:357-78. Burge, T. 1979. Individualism and the mental. *Midwest Studies in Psychology* 4: 73-122.

Chalmers, D.J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.

Chalmers, D.J. 2002b. The content and epistemology of phenomenal belief. In (Q. Smith & A. Jokic, eds) *Consciousness: New Philosophical Essays*. Oxford University Press.

Chalmers, D.J. 2002c. Does conceivability entail possibility? In (T. Gendler & J. Hawthorne, eds) *Imagination, Conceivability, and Possibility*. Oxford University Press. [consc.net/papers/conceivability.html]

Chalmers, D.J. 2002e. The nature of epistemic space. [consc.net/papers/espace.html]

Chalmers, D.J. 2002s. On sense and intension. [consc.net/papers/intension.html]

Chalmers, D.J. & Jackson, F. 2001. Conceptual analysis and reductive explanation. Philosophical Review. [consc.net/papers/analysis.html]

Clark, A. & Chalmers, D.J. 1998. The extended mind. Analysis 58:7-19. [consc.net/papers/extended.html]

Crimmins, M. 1991. *Talk about Beliefs*. Cambridge, MA: MIT Press.

Davies, M.K. & Humberstone, I.L. 1980. Two notions of necessity. *Philosophical Studies* 38:1-30.

Dennett, D.C. 1981. Beyond belief. In (A. Woodfield, ed.) *Thought and Object*. Oxford: Oxford University Press.

Evans, G. 1979. Reference and contingency. *The Monist* 62:161-89.

Fodor, J. 1987. *Psychosemantics*. Cambridge, MA: MIT Press.

Fodor. J. 1991. A modal argument for narrow content. *Journal of Philosophy* 88:5-26.

Frege, G. 1892. &Uumlber Sinn und Bedeutung. Translated in (P. Geach & M. Black, eds.) *Translations from the Philosophical Writings of Gottlob Frege*. Oxford: Blackwell, 1952.

Horgan, T. & Tienson, J. 2002. The intentionality of phenomenology and the phenomenology of intentionality. In (D. Chalmers, ed) *Philosophy of Mind: Classical and Contemporary Readings*. Oxford University Press.

Kaplan, D. 1967. Quantifying in. Synthese 19:178-214.

Kaplan, D. 1989. Demonstratives. In (J. Almog, J. Perry, and H. Wettstein, eds.) *Themes from Kaplan*. Oxford: Oxford University Press.

Kripke, S.A. 1979. A puzzle about belief. In (A. Margalit, ed.) *Meaning and Use*. Dordrecht: D. Reidel.

Kripke, S.A. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

LePore, E. & Loewer, B. 1986. Solipsistic Semantics. *Midwest Studies in Philosophy* 10:595-614

Lewis, D. 1979. Attitudes *de dicto* and *de se*. *Philosophical Review* 88: 513-43.

Lewis, D. 1994. Reduction of mind. In (S. Guttenplan, ed.) *Companion to the Philosophy of Mind*. Oxford: Blackwell.

Loar, B. 1988. Social content and psychological content. In (R.H. Grimm and D.D. Merrill, eds.) *Contents of Thought*. Tucson: University of Arizona Press.

Patterson, S. 1991. Individualism and semantic development. *Philosophy of Science* 58:15-35.

Perry, J. 1977. Frege on demonstratives. *Philosophical Review* 86: 474-97.

Perry, J. 1979. The problem of the essential indexical. *Nous* 13:3-21

Putnam, H. 1975. The meaning of `meaning'. In (K. Gunderson, ed.) *Language, Mind, and Knowledge*. Minneapolis: University of Minnesota Press.

Putnam, H. 1981. *Reason, Truth, and History*. Cambridge: Cambridge University Press.

Quine, W.V. 1968. Propositional objects. *Critica* 2(5):3-22.

Richard, M. 1990. *Propositional Attitudes: An Essay on Thoughts and How we Ascribe Them*. Cambridge: Cambridge University Press.

Salmon, N. 1986. *Frege's Puzzle*. MIT Press.

Schiffer, S. 1978. The basis of reference. *Erkenntnis* 13: 171-206.

Schiffer, S. 1990. The mode-of-presentation problem. In (C.A. Anderson & J. Owens, eds.) *Propositional Attitudes: The Role of Content in Logic, Language, and Mind*. Stanford: CSLI Press.

Schiffer, S. 1992. Belief ascription. *Journal of Philosophy* 87:602-14.

Segal, G. 2000. *A Slim Book about Narrow Content*. MIT Press.

Smith, E.E. & Medin, D.L. 1981. *Categories and Concepts*. Cambridge, MA: Harvard University Press.

Stalnaker, R. 1978. Assertion. In (P. Cole, ed.) *Syntax and Semantics: Pragmatics, Vol. 9*. New York: Academic Press

Stalnaker, R. 1990. Narrow content. In (C.A. Anderson and J. Owens, eds.) *Propositional Attitudes*. Stanford: Center for the Study of Language and Information.

Stalnaker, R. 1991. How to do semantics for the language of thought. In (B. Loewer & G. Rey, eds.) *Meaning in Mind: Fodor and his Critics*. Oxford: Blackwell.

Stich, S. 1983. *From Folk Psychology to Cognitive Science*. MIT Press.

Taylor, K. 1989. Narrow content functionalism and the mind-body problem. *Nous* 23:355-72

White, S. 1982. Partial character and the language of thought. *Pacific Philosophical Quarterly* 63:347-65.

White, S. 1991. Narrow content and narrow interpretation. In *The Unity of the Self*. Cambridge, MA: MIT Press.

Yablo, S. 2002. Coulda, woulda, shoulda. In (T. Gendler & J. Hawthorne, eds) *Imagination, Conceivability, and Possibility*. Oxford University Press.