

Chapter 6

Conclusion: Hume's Program (and Ours)

I began this exploration by entirely approving of a remark of Barry Stroud's: Hume's attachment to the 'Theory of Ideas' (TOI) is 'unshakable'. But I think that it was quite wrong of Stroud to say that Hume "never gives any argument in support of it." To the contrary, one might well read the whole first book of the *Treatise* as an elaborate argument to the best explanation, the conclusion of which is that TOI is indispensable to any foreseeable naturalistic cognitive psychology. I find this argument, as Hume presented it, enormously persuasive; and I think it's gotten even better over the succeeding couple of centuries. For a number of interlocking reasons, it is fully plausible that cognitive processes are constituted by causal interactions among semantically evaluable mental particulars. Either that, or we really are entirely in the dark.

That line of thought was extensively present, though with varying degrees of explicitness, in the preceding chapters. I now propose, by way of conclusion, briefly to survey a number of problems about the mind for which TOI offers what appear to be plausible solutions; some that Hume suggested and some that he didn't but (save for the anachronisms) perfectly well could have, consonant with what I take to be his sense of the enterprise. I can't prove that TOI is the right approach to this galaxy of worries, but I think there's a striking number of straws in the wind, all of them blowing in much the same direction.

1. Compositionality

No doubt you've heard this one before, so I won't go on about it here. But I think one ought to keep in mind that TOI is an essential element in what is arguably the greatest success that cognitive psychology has had so far. I mean bringing together within a single theoretical framework three aspects of propositional attitudes that are on the one hand entirely characteristic,

but on the other *prima facie* unrelated: their systematicity, productivity and compositionality. In a nutshell, you need the contents of the attitudes to be compositional in order to explain how beliefs, desires and the like can be systematic and productive. And you need Ideas cum concepts to explain how the contents of the attitudes can be compositional. That's because, by definition, compositionality is a property of *complex representations*. In particular, it's a property of the relation between complex representations and the simpler representations that constitute their parts. But Ideas, as TOI understands the term, *just are* (mental) representations that have, or can have, other representations as their parts. As I remarked in an earlier chapter, it seems to me that Hume is onto just about all of this.

2. Mental Causation

The theory *in situ* always has squatter's rights, so we should hold onto Intentional Realism if we can. The test par excellence of whether a philosophy of mind achieves Intentional Realism is what it says about mental causation. (That's unsurprising since only Real causes can have Real effects and vice versa.) Whatever the right story about abstracta may be, the proof of ontological commitment to a kind of concrete particulars is that they are acknowledged as links in causal chains. Conversely, philosophers who think deep down that the mental must be somehow ontologically second rate, invariably tip their hands by refusing to take mental causation quite at face value. 'No doubt, the claim that there is some is true if rightly construed' (so their story goes) 'but it's in want of considerable interpretation.' Thus Charles Travis recently: "Attitudes may cause things... the point does not admit of doubt. One might react to this result in either of two ways. First, one might say: 'Now we know one thing a cause may be. The way in which attitudes are related to their effects is one thing causation may consist in.' The other is: 'We know the sort of thing causation is. (We have spent our time in pool halls.) So *that* must be the way attitudes related to their effects. (190)." Travis thinks the first reaction is much the better of the two.

This is in the style of paradigm case arguments familiar from the mid-century logical behaviorism of Wittgenstein and Ryle: 'Of course there are mental causes; here's one and there's another. But don't get your hopes up; it appears, on analysis, that mental causation isn't quite what you'd supposed. In fact, it turns out to be a kind of *dispositional* causation.' Compare

Berkeley on chairs: 'Of course there are chairs; here's one there's another. Only, it appears on analysis that chairs aren't quite what you'd supposed; In fact, they turn out to be a sort of after image'.

Here's Travis running his version of the dispositional line on propositional attitude attribution: '... the fact that [her] thinking X explains Y tells us that Zoe has a certain sort of attitude --- one which consists in her maintaining a certain sort of system in her way of treating things. There is much that her maintaining that system would, or reasonably might, lead to... That is one thing [mental] causation may look like. (Ibid, 102).' And here's Ryle doing much the same in *THE CONCEPT OF MIND*: "...he is now recognizing or following the tune if, knowing how it goes, he is now using that knowledge; and he uses that knowledge not just by hearing the tune, but by hearing it in a special frame of mind, the frame of mind of being ready to hear both what he is now hearing and what he will hear, or would be about to hear, if the pianist continues playing it and is playing it correctly. (227)."¹

The geography around here is familiar; certain of your beliefs explain how you act in the sort of circumstances you're in (and/or how you would act if the circumstances were thus and so). That's because to acquire those beliefs *just is* to become disposed, in consequence, to act in that sort of way in circumstances of that kind. Likewise, certain of your experiences explain certain of your perceptual beliefs. That's because to have that kind of perceptual belief *just is* to have a certain kind of disposition caused by a certain kind of experience. Contrast the paradigm case according to the naïve view: You desist in attempting to sit on the mat when you notice that the cat is in possession. The noticing is one event, the desisting is another, and the first causes the second. To be sure, on this view, we're not so far from billiard balls; but by precisely what argument is that a reductio?

My point, in any case, is that it's because he is committed to the Theory of Ideas that Hume can tell the naïve story if he's so inclined. Token Ideas are concrete particulars, so tokenings of Ideas are events in good standing; a fortiori, they have every right to relate to one

¹See also Baker (19xx).

another as causes to effects. Conversely, however, if TOI goes, the naïve story about mental causation goes with it. Since, as far as anybody know, the choice between TOI and some dispositional theory of mental states exhaust the options available to a PA Realist, if the dispositional story doesn't work, that would vindicate TOI.

And, of course, the dispositional story doesn't work. That it doesn't is a point that's been pretty well explored for the last fifty years and is notorious by now. But I do want to call your attention to an aspect of its failure that hasn't been widely emphasized.

Once upon a time, theologians worried a lot about why God made the world on a Monday. I think they were right to worry. Consider, in the first place, that God is free to do, or to refrain from doing, whatever He chooses; so He wouldn't have made the world if he hadn't been disposed to. But, in the second place, God is not frivolous. He doesn't change His mind. So, if He was ever disposed to make the world, He must always have been so disposed. But, finally, God is omnipotent, so once He was disposed to make the world, there was nothing that could have stopped Him. Why, then, did He make the world when he did rather than at some other time? Why did He wait till Monday? Surely God doesn't act without a sufficient reason? The problem, in a nutshell, is that *something has to happen* to make a disposition manifest itself. But nothing can happen to God.

This theological puzzle doesn't, in fact, keep me up at night. But I do think it's revealing about the relations between dispositions and causes. In particular, it raises hard problems for any account of mental causation that proposes to dispense with mental events, a fortiori for any purely dispositional account of mental causation. Suppose, for example, that a certain belief is a standing disposition to perform a certain action: Maybe believing that it's raining is having a standing disposition to say that it's raining. Well, if you do believe it's raining, why aren't you saying that it is *right now*. Likewise, if you believe that it's not, why aren't you right now saying that it's not? What are you waiting for, next Monday? Come to think of it, why aren't we all talking about the weather all the time?

Logical behaviorists did, of course, have an answer to this sort of question. Ascriptions of dispositions are really assertions of *hypotheticals*. For the glass to be fragile is for it to break *if*

struck (or dropped; or stepped on; or whatever. And then only *ceteris paribus*). Likewise, for Jones to believe it's raining is for him to say that it's raining *if he's asked* (or whatever. And again, only *ceteris paribus*). The point, is just the one I made above: something has to happen to make a disposition manifest itself. Notice that 'what happens' to make a disposition manifest is always *an event*. This is a truism; an events *just is* something that happens and vice versa.

The moral is that if mental causes are to be dispositional causes, there have to be the kinds of occasions on which the dispositions are (or would become) manifest; and these occasions must involve some or other kind of event taking place. One can see (sort of) how that kind of story might work this in cases like perception, where something nonmental causes a belief; and, likewise, at the other end, in cases where a belief causes an action. Perhaps it's *proximal stimulations* (construed as events) that cause perceptual beliefs to be acquired. And perhaps your believing that P (that being a disposition) can cause your saying that P (that being an action) *if someone asks you whether P* (that being an event). All right so far. But how does the dispositional causation story go in the case where mental things cause other mental things? Surely there are such cases? For example (I suppose; don't you?) that *thinking* is a mental process in which some thoughts cause other thoughts to follow them. I take it there would be something deeply wrong with a theory of the mind that made thinking, so construed, seem problematic.

If propositional attitudes and the like are dispositions, then I suppose what happens when one belief causes another consists in a creature's disposition to X being manifested in the creature's becoming disposed to Y. (If it's the kind of creature that doesn't like to get its hair wet, then maybe its believing that it's raining manifests itself in its becoming disposed to carry an umbrella.) Well, the manifestation of a disposition requires an event to cause it (see above), and that's true even if the way that the disposition manifests itself is by causing another disposition. So now, if beliefs are dispositions, what sort of event could cause a belief to manifest itself by causing some other belief? Search me; but not, anyhow, a *mental* event. By hypothesis, there aren't any mental events; all there are is mental *dispositions*.

Oh well, events are cheap; there are always lots of them around. Maybe it's *neural* events that cause some mental disposition to cause other mental dispositions? Or maybe it's

meteorological events? Or geological events? Whatever. So be it; but there's a price to pay. We no longer have a robust notion of mental causation. For, a robust notion of mental causation would require that some mental things are *causally sufficient* for others. And, though there is indeed such a thing as dispositional causation (the vase broke because it was fragile) , dispositional causes aren't sufficient to bring about their effects. Dispositions manifest themselves only when something that's not a disposition causes them to do so. It's not sufficient for the vase to break that its fragile; *something has to happen that causes its fragility to become manifest.*

Likewise in the present case, *mutatis mutandis*. If mental causes are dispositional causes, then one belief's causing another has to look something like this: *e causes (M1 causes M2)* where *e* is an event (hence not mental according to the present assumptions) and M1 and M2 are mental (hence not events according to the present assumptions). Notice the critical distinction between this case, where an event causes one mental disposition to cause another, and the sort of case where an event causes some mental disposition, which in turn causes another one ($e \rightarrow M1 \rightarrow M2$). In the second, but not the first, the causation is robust; i.e. M1 is causally sufficient for M2. But, precisely because the causation is robust, the cause can't be a disposition. Dispositional causation is ipso facto *not* robust; dispositional causes are ipso facto *not* sufficient for their effects. (See above).

So the moral is that we can have mental things be dispositions, or we can have the robust causation of some mental things by others; but we can't have both, and we will have to choose. I think it's pretty clear, if the choice is indeed forced, it's the robust mind/mind causation that we must hold onto. That's because it's very so very plausible that mental processes *just are* causal chains in which each link is causally sufficient for its successor. The generalization is, say, 'thinking of cats causes you to think of dogs.' In particular, it's not 'thinking of cats *together with something else* causes you to think of dogs.' Thinking, associating, and the like are thus paradigms of cases of mental causation; or, at least, everybody in the trade supposes that they are. That's why they're the bread and butter of intentional explanation as psychologists

understand it.² But if it's a priori that that mental causes are dispositional causes, then it's likewise a priori that nothing that's mental is causally sufficient for anything else that is. But that *can't* be a priori (true, maybe. a priori? Surely not.) So, simply, the dispositional story about mental causes must be wrong. No wonder Ryle never talks about mental processes; he can't, in principle, allow there to be any.

By contrast, none of this needs to cost Hume a moment's sleep. Hume takes for granted a Theory of Ideas, according to which a thought is a bona fide mental particular and having a thought is a bona fide mental event. So Hume can likewise take for granted that thinking is, par excellence, a mental process in which some mental events are causally sufficient for others.³ Indeed, according to Hume, the laws of psychology are precisely ones that govern mind/mind causation; they're the laws of Association in virtue of which Ideas succeed one another in thought. That it can (and, in point of historical fact invariably did) allow psychologists (and layfolk too, come to think of it) to take thinking at face value strikes me as quite a good argument for TOI. Truth to tell, it strikes me as pretty near knock down.

It's perhaps worth mentioning, to conclude this section, some connections between the present line of thought and the issues about 'traces' that were raised in Chapter 5. I think *water*, and that makes me think *wet*. By assumption, this is because the Ideas WATER and WET have been paired in my previous experience. But how does my mind know *now* that the ideas WATER and WET have been paired *in the past*? We say that my experiences *then* have left a 'memory trace' that's *still there now*, where a memory trace is a mental particular stored, I suppose, somewhere in the head. That is to say, we appeal to TOI to explain how one's

² In passing: It's important to distinguish between, on one hand, the question whether the antecedent of a generalization articulates causally sufficient conditions for the satisfaction of its consequent; and, on the other hand, the question whether the generalization is 'strict' in the sense of being *exceptionless*. I assume that the typical special science law (a fortiori, the typical psychological law) claims (approximately) that the satisfaction of its antecedent is causally sufficient for the satisfaction of its consequent, *all else equal*. And I assume that that's *not* equivalent to claiming, as it might be, that the satisfaction of the antecedent is a *contributing cause* of the satisfaction of the consequent. For discussion, see Fodor 19xx)

³ Conversely, when TOI went out of fashion, psychology and philosophy ceased to offer theories of mental processes. This was a historical watershed. Dewey, Quine and Ryle (for example) are in various ways modern heirs of Hume's empiricism, but not of his psychology. Hume was interested in thinking, but Dewey, Quine and Ryle weren't.

experiential history can effect the causal powers of one's current Ideas. Well, since Hume took TOI for granted, why didn't he say that too?

I think it's just an historical accident that Hume didn't have a trace theory of memory. (In fact, as I remarked in Chapter 5, Hume really has no theory of memory at all; what he has instead is an (unconvincing) epistemological story about how you tell veridical memories from false ones). Presumably it was because memory traces are, practically by definition, mental particulars of which one isn't conscious, and Hume is wedded to the doctrine that there are no unconscious Ideas. For Hume, what's *in* the mind is ipso facto *present to* the mind. This is, I think, one of the many places where his epistemological agenda costs him a treatment that he would have jumped at if his only concern had been to construct a naturalistic psychology of cognition. If he had allowed traces to be unconscious Ideas, as indeed he ought to have, Hume would have lost a standard Empiricist argument against Nativism, hence against the Cartesian account of apriority. The methodological moral is: don't expect your theory of mind to do your epistemology for you. Be grateful if you can get it to do your psychology.

3. Intentionality and 'which rule are you following?'

Hume knew that there aren't any unicorns and that there isn't any New Jerusalem. But he took it for granted that, whereas thinking about unicorns is one thing, thinking about the New Jerusalem is quite another. So I guess Hume wasn't an extensionalist. So I guess he was an *intentionalist*. To be sure, practically everybody is an *intentionalist* in this relaxed sense of the term. It takes a very determined philosopher indeed to argue that a thought about Superman is ipso facto a thought about Clark Kent. Inconvenient though it may be it would seem that Lois Lane can believe of Superman, but not of Clark Kent, that he is a frequent flier.

But if, in such cases, it's not their extensions that distinguish thoughts, what on earth does? And how could coextensive thoughts being different in that way, whatever that way is, account for the sometimes egregious differences in the consequence of thinking them? I think that Hume had exactly right answer to these sorts of questions. What distinguishes coextensive thoughts is that different mental representations are entertained in the course of having them. What makes thinking about unicorns different from thinking about The New Jerusalem is that

different Ideas mediate the processes. In particular, the Ideas that are entertained in thinking about them are tokens of different mental representation types. Hume gets this agreeable treatment of intentionality free; it's his reward for his fidelity to TOI.

I assume, and Hume did too, that a necessary condition for mental representation types to be distinct is that their tokens differ in certain of their intrinsic properties. (Roughly, the relevant 'intrinsic' properties of such tokens are ones they have in virtue of their relations to their (possibly improper) parts. So, an Idea's being simple or complex is one of its intrinsic properties, but its being tokened when and where it is, is not). Up to this point, my version of TOI is just a quotation of Hume's, in case anybody cares. Since, however, Hume held to the picture theory, he supposed that the salient difference between Ideas that are coextensive but distinct is in their (as it were) geometrical structure. By contrast, I hold to a Language of Thought view of mental representation, so I suppose that Ideas that are coextensive can be distinguished by (inter alia) the way they decompose into their syntactic constituents. Insofar as the project is to understand intentionality, this disagreement is perhaps not very substantial; it's just an in-house argument between ways of running TOI. In either case, what matters to intentionality is that concepts can be distinguished *either* by their extensions *or* intrinsically or both.

Ideas, as TOI understands them, are semantically evaluable, causally active, mental particulars; in effect, they're 'modes of presentation,' only psychologized. The thought that concepts might be distinguished by their modes of presentation (in effect, that the same extension may be presented to the mind in lots of different ways) has, of course, been around at least since Frege. The big divide is not, I think, between LOT and the picture theory; it's between philosophers like Hume and me (and, if I read them right, such 'neo-Fregians' as Christopher Peacocke), who think that MOPs are mental and particular, and philosophers like Frege who thinks that they are neither. According to Hume, considerations of explanatory adequacy settle this argument insofar as anything can: TOI makes it immediately clear why mental states that are coextensive but intentionally distinct can differ in their causal roles in mental processes. To say that it's hard to understand this on an account of MOPs as pure abstracta would be an extravagance of understatement.

Here's an example of how TOI's sort of story about the nexus between mental content and mental representation might be supposed to go in an otherwise puzzling case. There has been a long running debate in the philosophy of mind about what (if anything) distinguishes behavior that *follows* a certain rule from behavior that merely 'accords' with it; and, assuming that what one has is indeed a case of rule following, about what decides *which* rule it is that's being followed. It's often supposed that these are, at heart, issues about consciousness: In the paradigm case, the agent is fully aware of (and is able to report) what rule he follows. Tendentious cases grade off from this paradigm, the question being how much consciousness can be attenuated consonant with ones following a rule at all. Opinions range from 'hardly at all' to 'to whatever extent explanatory adequacy may require.'

In fact, however, what's at issue isn't consciousness but intentionality; what distinguishes the rule one follows from other merely coextensive rules is what distinguishes *any* equivalent thoughts that differ in their causal powers; viz., it's the mode of mental (re)presentation of the thought.

It looks like you can infer from $P \ \& \ (P \rightarrow Q)$ to Q in any of three ways:

-by modes ponens.

-By any of indefinitely many equivalent rules (by contraposition, say).

-By no rule at all, even though your behavior is just as it would be if you were following one. (That is, you might have a kind of mind in which this sort of inference is 'hard wired'.)

The pertinent questions are: 'what's the difference between these four cases?' and 'Why is behaving by rule such that it can come in any of these four kinds?'

But these are (so I suppose) just the kinds of questions that *always* arise when equivalent descriptions correspond to distinct psychological states. So, famously, John's believing that Orcutt is a spy may explain John's behavior, though his believing that the man in the hat is a spy does not; and this can be so even though the man in the hat is Orcutt. That's presumably because of the way that John's (mentally) represents Orcutt (viz as Orcutt, not as the man in the hat.)

On this view, the puzzle about Orcutt is 'what's going on in John's head when his believing that P explains his behavior and his believing that Q does not, and P iff Q?' Exactly likewise for puzzles about which rule is being followed: Sometimes, and in spite of their equivalence, 'it's because he was following modes ponens' can be true and 'it's because he was following contraposition' is not. TOI says that what distinguishes intentionally distinct but equivalent Ideas is the way they specify their contents: Ceteris paribus, reasoning is guided by modus ponens rather than contraposition when it's a mental representation in the form of modus ponens by which the behavior is caused. Likewise, mutatis mutandis, for reasoning guided by contraposition. The rules are equivalent, but they differ in how they specify their content. And likewise, mutatis mutandis, when rule-according behavior isn't rule-guided at all; that's the 'hard wiring' case where the behavior, though it's consonant with modes ponens, is in fact not caused by any mental representation of a rule.

So, according to TOI, the treatment reduces of the problems about rule following reduce to the treatment of the problems about intentionality; you get the two for the price of one (which, admittedly, is no bargain if you don't like either.)

4. Atomism

It's pretty widely agreed that Hume's version of the Theory of Ideas is basically atomistic; and, as we saw in the Introduction, it's not unheard of to hold this against him. When Hume is berated for not having been Wittgenstein, it's generally his atomism that's the gravamen of the reproach. What's at issue here is a thesis about relations among conditions for concept possession: You're an atomist insofar as you hold that the possession conditions for some concepts are independent of the possession conditions for the others; you're not to the extent that you don't. (So, if you hold that it's not possible to have the concept RED unless you have the concept COLOR, then your theory of concept possession is to that extent an-atomistic.) It matters to philosophers which (if any) concepts atomism is true of. That's because it's very plausible that concepts have their possession conditions essentially; i.e. that possession conditions are concept-constitutive; i.e. that concepts which differ in their possession conditions are ipso facto distinct concepts. Accordingly, if you're wanting to argue that C and C' are different concepts, all you have to do is make a case that they have different possession

conditions. And, of course, philosophers often are in the position of arguing for conclusions of the form: 'C and C' are different concepts. Some think that's all that they do.

Anyhow, I think it's clear that Hume has to be an atomist about primitive concepts. For: on the one hand, he's committed to a resemblance theory of their content and, on the other hand, resemblance is itself plausibly an atomistic relation; plausibly, whether x resembles y depends solely on the intrinsic properties of x and y. (If the world consisted solely of x and y, there would still be a fact of the matter about whether x resembles y (assuming there's a fact of the matter about that as things actually are.)) Conversely, Hume holds that anatomisms do arise as a consequence of relations between complex concepts and their constituents. If C is part of C' (as, ANIMAL is sometimes said to be part of ZEBRA) then you can't have the first unless you have the second. As far as I can tell, Hume departs from this general picture only in his discussion of 'Relations of Ideas'; and when he does so, he departs from the resemblance theory too. Just a word about that for its historical interest.

It's often supposed that you can't have any shape concepts unless you have some color concepts (and vice versa). Let's say that's so. Then, a fortiori, you can't have the concept 'sphere' unless you have some color concepts. But this seems puzzling if you hold, as a matter of atomistic principle, that basic concepts are mutually independent; for it would seem that typical color concepts and typical shape concepts are often both basic. Consider the constituents of the concept WHITE SPHERE. Might its constituent concepts both be basic? On the one hand: 'Well, yes, because it's surely possible to have either of them without having the other.' But on the other hand 'well, no; because if C and C' are independent concepts, then a fortiori, it is possible to *think* either without the other, and you can't think any shape without thinking some color (or vice versa), assuming, as Hume does, that thinking is a kind of imaging. Thus, "*all ideas, which are different, are separable.* [sic]" (I, VII, 72). But "a person who desires us to consider the figure of a globe of white marble without thinking on its color, desires an impossibility... (Ibid) 73" Dilemma.

Here's the solution Hume proposes: "we consider the figure and color together, since they are in effect the same and indistinguishable; but still view them in different aspects. When we would consider only the figure... we form in reality an idea both of the figure and colour, but

tacitly carry our eye to its resemblance with [a] globe of black marble ... [when we would] consider its colour only, we turn our view to its resemblance with [a] cube of white marble. (Ibid 73)." Hume remarks that this way of removing the difficulty has "recourse to [his] explication of abstract ideas (Ibid)". The connection is that, in both cases, one's mentally represents a property by mentally representing (viz by picturing) some individual that has the property.

But this story is no good; in fact, it's circular. Hume has to explain how we "turn our view" to the resemblance between a white sphere and a black sphere if we don't already have the concept of a shape as such. Surely 'attending to their resemblance' is just noticing that, although the spheres differ in their color, they don't differ in their shape. But you can't do that unless you have a concept that abstracts from the color of a sphere and applies to it just in virtue of its shape, viz the concept of a sphere as such. So we're back where we started; abstract ideas and distinctions of reason both appear to be hopeless problems for the resemblance theory.

The moral, pretty clearly, is that we can't have a plausible TOI unless we ditch the image theory; a fortiori, we can't have a plausible *atomistic* TOI unless we ditch the image theory. But, on the other hand if we do ditch the image theory, we *can* have an atomistic TOI for all the arguments to the contrary that we've seen so far. That's to say that TOI per se is neutral about conceptual atomism, as far as anybody knows. Not so, however, for the alternatives to TOI insofar as they construe concept possession in terms of dispositions-to-draw inferences; which they practically all of them do. That's because you need more than one concept to draw an inference. You can't infer from RED to COLOR if you've only got RED or if you've only got COLOR. So, if being disposed to draw that inference is a possession condition for RED (or for COLOR), then the possession condition for RED (/COLOR) is ipso facto atomistic.

Now, you may think that not being compatible with atomism is a virtue in a theory of concepts possession (/individuation). But I don't, because I think atomism is quite likely true (at least of nonlogical concepts). At a minimum, since there's some evidence in its favor, we don't want to prejudice the issue by taking for granted that possession conditions are constituted, even inter alia, by inferential dispositions. I won't discuss the evidence at any length; I've done so elsewhere. But, roughly:

First:

Call the thesis that some inferences belong to the possession conditions of (some of) the concepts they deploy 'inferential atomism (=IA)'. Well, prima facie IA is committed to a variety of empirical claims which seem to be, quite simply, false. The situation is egregious when considers the concept *acquisition*. If having C is a condition for having C' (because the inference $C' \rightarrow C$ is constitutive of C'), it looks to be ruled out that the acquisition of C' might be prior to the acquisition of C. Such cases would thus be prima facie counterexamples to IA. But, in fact, they come in plethoras. Suppose that TIGER is the concept of a tigerish animal. Then, by assumption, you can't have TIGER unless you've got ANIMAL and, a fortiori the acquisition of ANIMAL must precede the acquisition of TIGER. But that isn't true according either to commonsense or to any known psychological test. Nor is the analogous conclusion about the priority relations between acquiring PARENT and acquiring MOTHER; between acquiring CAR and acquiring VEHICLE; between acquiring CHAIR and acquiring FURNITURE; between acquiring WINDOW and acquiring APERTURE... and so on. And on. In general, IA wants the acquisition of concepts of kinds to be prior to the acquisition of concepts of their instances. Sometimes this works all right, (DOG is prior to POODLE). But usually it doesn't. In fact, the concepts that get in first are ones that apply to 'middle-sized' objects (see Roch, 19xx). This is *overwhelmingly* the case, but its hard to see how to square it with IA short of a very large number of ad hoc assumptions.

A parallel argument holds for predictions about the order in which concepts are applied in the course of perceptual identifications. If ANIMAL is part of CAT, then prima facie seeing a cat as an animal should be a precondition to seeing it as a cat. But it's not. In fact the same generalization holds here as in ontogeny; midde-sized-object concepts are the first available. It's harder to spot an animal than it is to spot a cat, even if a cat is the animal that you're spotting.

The failure of IA to predict the empirical facts is, in short, as near perfect as any fit between data and theory ever gets in psychology. I wonder why that doesn't bother anti-atomists.

Second:

Once you've started on IA, it's hard to see how to stop short of a really ruinous holism.

For one thing, 'constitutive of' is prima facie transitive. If the inference $C' \rightarrow C$ is C'-constitutive, then if the inference the inference $C \rightarrow D$ is C-constitutive, then it looks like having D is constitutive of C'. This leads to such unhappy conclusions as that if you don't have the concept ORGANISM, you can't have the concept DOG; and, that having PHYSICAL OBJECT (or, who knows, the concept 4-DIMENSIONAL SPACE TIME WORM) is a possession condition for having MOMMY. We could block this slide if we had some notion *immediate* inference, since the notion of *immediate inference from --- to ---*' isn't transitive. But we have no such notion; and the history of attempts to construct one has been unencouraging.

This is the tip of a familiar iceberg. If one is to hold IA but avoid holism, one needs something that plays the role that the 'analytic/synthetic' distinction was traditionally assigned. I don't say there can't be such a thing; but I do say that nobody's got one; and that nobody has a clue how to put one together. The prospects for IA are, in short, at best no better than the prospects for a/s, and the prospects for a/s seem not good.

So, why does everybody take for granted that IA must be true? Search me. I've been told, however, that conceptual atomism is intuitively implausible. I should only have such convenient intuitions. Anyhow, intuitions come and intuitions go. Atomism didn't seem implausible to Hume, or to hosts of philosophers who preceded and followed him (including, by the way, Wittgenstein circa 19xx). In fact, as far as I can make out, atomism didn't start to be intuitively implausible on a really big scale until around 1950. What they always say about the weather is true in spades of intuitions of philosophical plausibility. If you don't like what you've got now, just wait till next week.

The sum and substance is that if, as would appear, inferential atomism is the only serious alternative to conceptual atomism, then it might be wise to hedge one's bets about how plausible IA is. TOI *does* hedge its bets about how implausible conceptual atomism is. But, of course, IAT doesn't. If dispositions to infer are constitutive of concept possession/individuation, then atomism *can't* be true since, as remarked above, it takes more than one concept to make an inference. All that being so, I hold it to be an advantage of TOI that it's compatible with (but doesn't entail) conceptual atomism. It is a wise philosopher who only burns such bridges as he's sure that he's already crossed.

5. Thought and language:

I promised I wouldn't tell you yet again the story that goes 'it's productive and it's systematic, so it must be compositional;' I'm doing my best to kick the habit. But, at the risk of backsliding, I will permit myself an (Er!) metatextual observation. Early on in telling that story, one always points out that it works equally well for language and for thought, so that it doesn't matter which it's applied to. In effect, one argues that *either* natural language *or* the language of thought is compositional, but one avoids saying which.

Well, but which, in fact, is it? English? Mentalese? Both? This matters for our present line of inquiry. For, suppose it turns out not to be English. Since the usual argument seems to show that *something* is compositional, all that's left is for it to be Mentalese. But to make a case for the compositionality of Mentalese is, a fortiori, to make a case for mental representations, since the former is a species of the latter. The long and short is that, if the argument from productivity and systematicity is good, then evidence that English isn't compositional is ipso facto evidence for TOI.

I think, in fact, the evidence suggests that probably English isn't compositional, hence that such systematicity and productivity as it has, it borrows from Mentalese. On this view, we think in Mentalese and communicate in English. Insofar as we can say what we think and understand what we're told, that's because there is a more-or-less-good procedure for translating from English to Mentalese and back; that's what one learns when one learns English.

It is, as one says, an empirical issue whether it's English or Mentalese that composes, though, as usual in such matters, lots of methodological doctrine comes into play in attempts to resolve it. I won't even try to resolve it here; not least because I don't know how to. But, *prima facie*, viewed naively, English doesn't *look* to be frightfully compositional. I'll stick to familiar and relatively untendentious cases.

Consider a standard account of imperative constructions like (1), according to which

1. Scratch the cat!

it's derived from is something like 'You scratch the cat!'. I take it that this story about the

derivation of imperatives is pretty plausible. For one thing, that's the way they are understood; that someone says (1) to me is a reason for me to scratch the cat (but not for you to scratch the cat; or for Sam to scratch the cat.) For another thing, paradigm English sentences ('Mary runs', 'John scratched the cat' and so forth always have subjects. That is, 'Scratch the cat' seems to be an exception to an otherwise generally reliable syntactic generalization about English. Surely, one ought to prefer analyses that make apparent exceptions to reliable generalizations go away. Third, the 'missing' subject would appear to have what the linguists call distributional reflexes. For example, the subjects of transitive verbs reflexive their objects ('he is scratching himself', but *'he is scratching he'). So, you'd expect that, if the subject of (1) is really 'you', then the sentence that means that you are to scratch you should have a reflexive as its object. Which it does: 'Scratch yourself!' Finally, you'd expect that clauses with overt subjects (other than 'you') can't be read as imperatives ('He'll scratch the cat'); as it were, the schema '___ scratch the cat' is ambiguous between declarative and imperative, but an overt subject resolves the ambiguity. And so, familiarly, forth.

None of this shows, of course, that English imperatives aren't compositional. Rather, according the usual understanding, it shows only that English sentences are compositional at some level of representation more abstract than their surface form. Linguists refer to this (putative) level as 'LF' (approximately 'logical form'), or as the 'semantic level'. The idea is that LF is ipso facto regular and explicit in respect of all sorts of things about which the corresponding surface forms needn't be. It's at LF, but not at the surface, that (1) has an explicit subject 'you'. Likewise, though the scope of the quantifiers is ambiguous in surface forms like (2), it's univocal in each of the two LFs from which such surface forms are said to derive. Or, suppose, Russell was right

(2) Everybody scratches somebody

about what 'the' means. If so, then at LF 'the king of France is bored' is disjunctive and contains two quantifiers. Or suppose Davidson is right about 'He cut the bread in the kitchen'; if so, then there is a quantifier over a variable for events in its LF representation. And so on

That's one way of seeing the situation. If it's right, then English is compositional at the

LFC level and the question whether thought is also compositional is open. (Unless we think in English. In that case, trivially, if English is compositional then so is Mentalese.) Notice, however, that running the story this way takes for granted what is in fact tendentious: that English sentences are what LF representations represent. There's an alternative. Namely that what's represented at LF isn't (eg.) the sentence 'John scratched the cat' but, rather, the thought that that sentence is used to express; viz. the thought that John scratched the cat. On this view, (1) isn't compositional after all, though its Mentalese translation is. Mentalese thus points in two different directions: towards thought, and towards language: The Mentalese sentence that translates to the English 'John scratched the cat' is the same mental representation one uses to make the proposition that John scratched the cat the object of your thought. It's because Mentalese faces both ways that one sometimes manages to say what one thinks.

Well, which way of seeing things is right? I would tell you if I knew; but here's a consideration that may bear on the question: It's usually taken for granted that, given their representations at LF, sentences are ambiguity free. LF is supposed to be, ipso facto, the level at which linguistic ambiguities are resolved. In consequence, there are two sentences with the surface shape (2), one corresponding to each way in which LF orders the quantifiers. But here too there's an alternative; one could say that there is only one sentence with the surface shape (2); viz a sentence that is ambiguous between everybody is such that there is someone... and there is someone such that everybody. If you are of sanguine temperament, perhaps you will say 'it doesn't matter which you choose; talk whichever way you like, so long as you're consistent.' If, however, you're gloomy, it may strike you that there's there's no obvious reason why sentences should be ambiguity free at any level of representation; i.e. there's no obvious reason why there shouldn't be *really* ambiguous sentences. It would seem, prima facie, to be just a matter of fact whether there are. Prima facie, either there are ambiguous sentences or there aren't, and we don't know which.

In short, to say that LF is a level of representation of sentences is to take a stand on whether there are fully ambiguous sentence types; it's to say that there aren't. If LF represents sentences, is in effect, it's to claim that there is a level of description at which all sentences are ambiguity free. The question is: what justifies this claim?

Compare thoughts. You can say (that is, utter) things that are ambiguous; but you can't think things that are ambiguous. Lincoln said 'you can fool all of the people some of the time'; maybe it occurred to him that there are two things he might have said (that is, *asserted*) by saying so, and maybe it didn't. But Lincoln couldn't have had the *thought* that you can fool some of the people all of the time, without, as it were, having it one way or the other. It couldn't be that what he did was: he closed his eyes and took a deep breath and thought, ambiguously, that *you can fool all of the people some of the time*. *Punkt*. It seems to be open whether there are ambiguous sentences, but it's closed whether there are ambiguous thoughts. So far as ambiguity is concerned, thoughts are where the buck stops.

I admit, I don't know why there can't be ambiguous thoughts (if indeed there can't) in the way that there can be ambiguous sentences (if indeed there can). Maybe it's because thoughts have 'underived content, but sentences have only such content as they inherit from the thoughts that they express. On that view, the content of a sentence might be ambiguous between thoughts, but it's unclear what the content of an ambiguous thought could be.

In any case, it's independently plausible that (for some reason or other), thoughts can't be ambiguous; and it's being plausible that they can't doesn't depend on assuming that LF represents them. That will do for our purposes since, by contrast, sentences don't *have to be* ambiguity free; it's just that if they are what LF represents, then they turn out to be. The sum and substance is: we can reduce (if only by one) the number of mysteries outstanding if we suppose that LF represents thoughts rather than sentences. So, all else equal, I guess that's what we'd better suppose.

Maybe that's an OK argument, maybe it's not. I don't begin imagine it settles the matter; all I want is that the question whether LF represents thoughts or sentences is bona fide and empirical. For, as we've seen, it's not in dispute that whatever LF represents is ipso facto explicit and unambiguous; which is in effect to say (what linguists quite generally suppose) that whatever LF represents is ipso facto compositional. So, if it turns out to be thoughts that LF represents, then that would be reasons for thinking that Mentalese is compositional. And if we have reasons to think that Mentalese is compositional, we can take seriously the first blush impression that English doesn't *look* very compositional. We have the option of saying: the reason it doesn't look

very compositional is that it isn't.

Well, enough; this was supposed to have been a book about Hume, more or less; and anyhow, we're now approaching depths at which the cognitive science is much in dispute. Suffice it that there are *all sorts* of interesting and, as far as anybody knows, researchable questions that connect with the issue whether TOI is a viable account of the cognitive mind. Indeed, large tracts of cognitive science depend on assuming that it is and are devoted to figuring out what sorts of things mental representations are. They seem to have made some modest progress. Whereas the alternative kind of PA Realism ---a dispositional account of thought and the attitudes--- appears to be what one calls a 'stagnant' research program. Philosophers would use it to beat skeptics over the head with, except that there aren't any skeptics. Psychology has no use for it at all. To every appearance, and just as Hume predicted, TOI is where the science of the cognitive mind wants to lead us; so why not go there?

What a nice little theory TOI is after all. I do think that Hume was right to cleave to it. I do think that we are too.