Some Notes on “Conscious Meets Lewisian Interpretation Theory” for M&L seminar students

I am an old-fashioned analytic metaphysician, in pursuit of hypotheses about what things are the elements of being and about how all else may be reduced to patterns of these elements. – David Lewis

Since my paper makes lots of claims, I thought it might be helpful if I told you what I consider to be the important guiding idea (the one I care most about), and what I consider to be the less central and more sketchy ideas about how to work out the details (ideas I’m less sure about). I’ll also say where I could most use some help!

(Incidentally, although my paper is more than enough for my visit, here are some additional optional things of mine: a paper on Fodor, my Williams review, my Mendelovici review, my comments on Morrison on representation in cognitive science.)

Main idea: a consciousness-first reductive thesis

In my paper, following Lewis in “Radical Interpretation”, I focus on an imaginary person, Karl. Even though I don’t really say so until the end of the paper (§3.5), the main underlying idea is:

Consciousness-first reductivism: all of Karl’s intentional states "reduce to" his actual and potential conscious experiences with “thin” contents together with the functional and physical facts about him.

Here is how I see things –

In the 80s and 90s, many (Lewis, Fodor, Neander) hoped all the intentional facts about Karl reduce to the austere physical-functional facts. I don’t think that works. For instance, when Karl is conscious of the ostensible state of something being red and round, this “conscious-of” relation can’t be reduced to a tracking relation or anything else. I think it is irreducible – similar to what Bertrand Russell thought about “acquaintance” (§1, §3.1). And maybe there are additional facts about what Karl ought to believe and desire given his experiences.

But I favor the alternative reductive vision above. The idea is that, while the austere physical-functional facts aren’t enough, once you throw in Karl’s basic conscious experiences with relatively “thin” contents, you do have a sufficient basis for reducing everything else. For example, Karl’s cognitive intentional states with “thicker” contents (his believing that all emeralds are green, his thinking that democracy is in trouble) reduce to this richer set of facts. (This is
close to Dave's PQTI scrutability thesis for intentionality in *Constructing the World* p. 275, but it adds that Q only needs to include thin content experiences.)

This consciousness-first reductivism is what I care about the most. The reason I care about it is that I am really bothered by what I call “further fact” theories of cognitive intentionality. On such theories, even once you have this augmented “base”, you must throw in *still more* facts to explain intentionality. Ugh! On *one* natural way of developing it, the recently popular cognitive experience theory is an example of such a further fact view (p.10ff). The idea is that Karl can perform a primitive mental act of judging or grasping that 68 + 57 = 125, for instance. It's a little cognitive experience. It's not reducible to anything else but somehow depends on Karl’s underlying internal physical states. That is supposed to solve Kripkenstein’s plus-quus problem. I argue that a further fact view is *especially* weird and complicated in the case of such cognitive intentional states.

So another way of putting my main idea is:

**No further fact views.** Such further fact theories of cognitive intentional states are wrong.

As long as there are no such “further facts”, I'll be pretty happy!

**Working out the details: a four-stage reductive proposal?**

I give some general arguments for my brand of consciousness-first reductivism (§3.5). I also give arguments against specific “further fact” views of cognitive intentionality (§2). So I feel pretty confident about the main idea of the paper.

In §3, I also stick my neck out and say something about how the details go. That’s difficult. But to me the details are less important than the main idea. The reason is that, even if it turns out that we cannot supply the details, I don’t think that’s a strong enough reason to reject the main idea – to reject my kind of consciousness-first reductivism and go in for “further facts”. We often have good reason to accept some general reductive hypothesis if we cannot supply a reduction (p. 34).

(In addition, in my view, the difficulty in supplying the details is also not a strong enough reason to go in for some kind of some “radical indeterminacy” view for cognitive intentionality, or some kind of “deflationism” about reference – views I have a tough time understanding.)

Also, since I like a kind of “content pluralism”, I think that, in *some* cases at least, there may not be a very substantive disagreement between different theories of content-determination.

Still, I think that the details are very interesting. I develop a four-stage reductive view. One guiding idea here is that, even though Karl’s basic conscious experiences only have relatively “thin contents”, those conscious
experiences are very important to intentionality, in two ways. First, they are important to fixing the contents of his other mental states. For instance, I think they are at least important to fixing the contents of his beliefs and desires within the “perceptual circle” (§3.1). However, they are less helpful when it comes to fixing the contents of his more sophisticated mental states (about plus vs quus, democracy, etc.). Second, Karl’s conscious experiences play a role in anchoring his “understanding” all contents. I think most people would say that an insentient robot doesn’t really “understand” anything (p.30).

I could use a lot of help with the details! For one thing, in my third stage (§3.3), I vaguely gesture at a really unsystematic “anchored use theory” of content-determination. I am open to alternative, more systematic accounts here.

For instance, in my third stage, I am open to co-opting Dave’s more uniform anchored inferentialism set out in his draft “Inferentialism and Analyticity” (which was already an influence), especially if it can be combined with his further suggestion that . . .

. . . one might try to ground the normative facts about inference in certain non-normative facts about inference—most obviously, the fact that we are disposed to make certain inferences—along with further factors in order avoid problems concerning error and incompleteness here. [This will] then yield as a reductive dispositional inferentialism, not just a nonreductive normative inferentialism. (p.10)