Harper O. W. Wallace
PHIL 103 / Buehler
27 Nov 2020
Paper 2

I. 981 words
II. 749 words

I. RTM vindicates commonsense explanations of the attitudes (reconstructed)

Fodor argues that the Representational Theory of Mind (RTM) is a plausible scientific-psychological account of how people instantiate attitudes (e.g., beliefs, desires). He uses RTM's consistency with commonsense psychological belief/desire explanations (CPEs) to defend commonsense psychology as a valuable window into mental goings-on.

CPEs describe the kind of reasoning everybody does every day in order to understand and predict other people's behavior. In order to reason this way, we rely on estimates for others' (unobserved) attitudes and assume that their behavior is based on these attitudes. Moreover, when we reason this way, we observe directly how our own attitudes give rise to yet new ones, in meaningful, causal sequence.

RTM explains an organism's attitude A toward proposition P as that organism bearing a particular relation R to the mental representation MP that means P. The nature of R can be seen as depending on the kind of A (e.g., belief, desire, hope). For clarity, Fodor formulates this alternately, where R is the "way" in which the MP is instantiated or "tokened" (e.g., as a belief-kind of MP). As an example, my belief in the proposition, "It will rain today" means my representing that proposition mentally in a belief-kind of way. Note that I'd be representing the same proposition, by this account, were my attitude instead a hope, but rather in a hope-kind of way. Fodor uses "mental symbols" to make explicit the idea that these mental representations, like symbols, have both semantical and syntactic properties. Finally, RTM explains mental processes as sequences of causally-related mental states.

Fodor's defense of RTM is as follows:

(P1) A plausible scientific-psychological account of attitudes must be compatible with CPEs. This is because:
- i commonsense belief/desire psychology—like "the most powerful etiological generalizations" (Fodor 7)—explains all the way down to unobservable causes;
- ii we rely on commonsense belief/desire psychology every day to infer others' intentions and predict their behavior; and
- iii these predictions are nearly always right.

(P2) CPEs are committed to:
- i (at least some) mental states carry content; the content of a mental state is semantically evaluable; i.e., they can be 'checked' against a true world-state;
- ii there exist causal relations between mental states;
- iii there is a plausible explanation for how mental states' causal relations can respect the semantic relations between their content; and
- iv "generalizations" (implicit logical schemata) emerge from mental processes and inform people's behavior.

(P3) RTM is committed to:

- i mental states carry the content of the propositions that correspond to their mental symbols; these propositions are semantically evaluable;
- ii mental states are causally related in mental processes;
- iii sequences of causally related mental states are allowed only when those mental states bear certain syntactic relations; allowed syntactic relations parallel allowed semantic relations, such that both kinds of relation are respected; and
- iv "generalizations" (cf. P2.iv) can be explained as (a) being shaped/structured by the syntactic relations between mental symbols and (b) taking as arguments those states' contents.

[P′] [Two theories are compatible iff their commitments are compatible.][1]

[C′] [The commitments of CPEs and RTM are compatible; cf. (P2), (P3).]

[C″] [CPEs and RTM are compatible.]

(C1) RTM is a plausible scientific-psychological account of attitudes.

Another central part of Fodor's argument is to use (C1) to make a broader claim in support of the commonsense account:

(P4) If there exists a plausible scientific-psychological account of attitudes that is compatible with CPEs, then it lends credence to ("vindicates") the CPEs.

(P5) C1

[P″] [C″ ~ RTM is compatible with CPEs]

(C2) RTM vindicates CPEs.

Fodor defends (P1) by rejecting a key rebuttal against the explanatory power of commonsense psychology. The objection is that commonsense predictions aren't actually any use because they're so hedged. That is, they're taken to be true, unless they're false. However, exceptions to the predictions made by commonsense psychology (like any special science) do not signal an outright failure of the theory. Instead, exceptions to commonsense predictions merely indicate the existence of cases outside the theory's scope, given the assumptions it makes. Since commonsense psychology has such predictive power in so many situations, this fact seems hardly reason enough to discard it.

It seems reasonable to believe that (at least some) mental states can carry semantically-evaluable content (P2.i, P3.i). Indeed, the mental states that correspond to our attitudes must somehow encode what those attitudes are *about*. Further, actual world states have bearing on semantically-relevant properties of attitudes (i.e., whether a belief is true, or a hope realized). So mental states' contents are semantically-evaluable. Causal relations between mental states (P2.ii, P3.ii) seem to be necessary, too. They allow us to engage in the kind of reasoning that Fodor identifies in "trains of thought." However, the relations we actually observe phenomenologically, e.g., in trains of thought, are only semantic ones. E.g., we are aware of how two successive thoughts are related in content, but not how they are related *qua* mental states. And so, the (syntactic) rules governing states' causal succession must also respect those

[1] Forgive the clumsy notation. I'm using brackets and primes here to distinguish intermediate premises/conclusions—for clarity, per a comment on a previous draft—from the main ones.

states' content (P2.iii). Fodor points out that this might seem a heavy burden for the theory. But despite this, he argues, there is reason to believe that the solution in (P3.iii) is actually quite good. In particular, because *it has been done*: computers manipulate symbols based on purely syntactic rules and yet simultaneously obey semantic rules. Finally, (P2.iv) identifies that certain, recurring logical structures seem to be common between different mental processes or trains of thought. E.g., we rely time and again on the same deductive principles to reach logical conclusions. And (P3.iv) offers possible explanations for how these generalizations might be realized: either explicitly, as representations themselves, or implicitly, "in the hardware."

II. Objection to the vindication of CPEs

Fodor does not claim that a commitment to the CPEs in turn commits us to RTM.[2] On the contrary, he leaves plenty of room for RTM to fall flat, while CPEs stand (notice the direction of the conditional):

> Commonsense would be vindicated if some good theory of the mind proved to be committed to [certain particulars]. RTM looks like being a good theory of the mind that is so committed; so if RTM is true, common sense is vindicated. It goes without saying that RTM needs to make an [i.e., its own] empirical case. (Fodor 26)

Instead, he defends the commonsense account by offering RTM as one plausible scientific-psychological accounts that is compatible with it. The weight of (C2), then, does not derive from RTM being *beyond doubt*, but instead only from it being *plausible*.

(P1) asserts a theory's compatibility with the commonsense account as a necessary (and *not* sufficient) condition for the theory's plausibility. This is Fodor's own formulation: "An explicit psychology that vindicates commonsense belief/desire explanations must [be compatible with the commitments of the commonsense account, enumerated in (P2)]" (Fodor 14–15). But this criterion does not support the inference in (C1) that RTM is plausible *because* it is compatible with the commonsense account.[3] Such an inference would rely on compatibility with commonsense being a *sufficient* condition for plausibility.[4] And that formulation would easily fold. In particular, there could reasonably be other properties required for an account to be plausible, besides its compatibility with commonsense. This weakness in (P1) leaves open the possibility that RTM be consistent with CPEs and still not be a plausible scientific-psychological account of attitudes.

To be sure, Fodor defends RTM's plausibility on other grounds, unrelated to its consistency with CPEs. He argues that RTM is the best available account of attitudes because, among other reasons, "some version of RTM underlies practically all current psychological research on mentation" (Fodor 17). So we may wish to grant (C1) anyway, not just because RTM is compatible with commonsense, but for other reasons, too.

---

[2] Yes, this is in conflict with the reconstruction presented in TD 5. Specifically, I claim that (P3) in the TD 5 reconstruction is nowhere to be found in Fodor's text and support this reading with a couple decisive quotes from the text.

[3] Yes, I'm claiming that my own reconstruction is not valid. This is, I think, faithful to Fodor's own argument (though, as I discuss in the following paragraph, it doesn't turn out to be a fatal blow).

[4] It would look something like: "Any scientific-psychological account of attitudes that is consistent with commonsense psychological belief/desire explanations is plausible" (or, less conservatively, "Any account at all that is consistent . . . is a plausible scientific-psychological one").

Even granting (C1), (C2) goes too far. Most concerning, (P1) and (P2) conspire to impose functional constraints on how a plausible scientific-psychological account of attitudes must work. (P1), in particular, pre-commits us to CPEs in such a way that a major sub-claim in (C2)—that CPEs are vindicated—becomes trivial:

(P1) A plausible scientific-psychological account of attitudes must be compatible with CPEs:

$$\forall x\ (Px \rightarrow Cx)$$

(P4) If there exists a plausible scientific-psychological account of attitudes that is compatible with CPEs, then it vindicates CPEs:

$$\exists x\ (Px \cap Cx) \rightarrow V(\text{CPEs})$$

(C2′) If there is any plausible scientific-psychological account of attitudes to be had, then it vindicates CPEs:

$$\exists x\ Px \rightarrow V(\text{CPEs})$$

Notice that RTM, in effect, cancels out. It was only ever a dummy theory, albeit a sophisticated and maybe even correct one.

A more appropriate version of (P1) would posit that a plausible scientific-psychological account of the attitudes should have at least as much explanatory power as CPEs. It should have to explain at least as many cases in at least as many contexts. But it is simply not believable that commitment-compatibility with CPEs should be required for a theory to be plausible. Not even (P1), for all its supporting evidence, actually ever makes the claim that CPEs are correct *at the level of explanation*. Fodor is only claiming that they work, and that they're convincing. But the distinction here is the same as that between a heuristic and a rule, or a model and a mechanism. Commonsense can make predictions—accurate ones even—and still be inaccurate at the level of explanation. It becomes an empirical question whether the true mechanisms behind attitudes actually *are* commonsensical, or just make the same predictions as CPEs. In the latter case, RTM and CPEs fail.

Work Cited

Fodor, Jerry A. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. *Vol. 2.* MIT Press, 1987, pp. 1–26.