

CONSUELO PRETI

Belief and Desire Under The Elms

1. Introduction

This paper is broadly concerned with the implications of semantic externalism for mental causation and for the nature of folk psychology. I have argued elsewhere that concepts subject to *weak* externalism pose a threat to folk psychology.¹ Here I will show that the tension between concepts subject to *strong* externalism and folk psychology remains, notwithstanding recent arguments marshaled by Fodor (1994). The threat of externalism to folk psychology may therefore appear to be intractable, but I will close by suggesting a strategy for a defense of the mutual tenability of externalism and folk psychology, one that deploys a familiar line of argument in a novel and promising way.

2. Content and Causation, or, "If Psychological Processes are Computational, How Can Psychological Laws be Intentional?" (2)

Call the following the mutual tenability thesis (MTT):

Mental content has causal powers, and mental content is externalistically individuated.

Familiarly, there is pressure on MTT for two reasons. First: a naturalistic account of the causal powers of intentional properties is elusive. Second – to add to the difficulty – the fundamental intuitions about causation and the basic tenets of externalism appear to be in direct contradiction.

Causation is commonly understood as a relation that holds between entities in virtue of their intrinsic properties, so a causal relation between mental states and action will have to be a relation between *intrinsic* properties of mental states and the *intrinsic* properties of action. It is reasonable to suppose that these properties will be the intrinsic physical properties of the central nervous system, which are what they are independently of the world extrinsic

¹ See Preti (1998). The difference between so-called *strong* and *weak* externalism can be understood by way of a supervenience claim. Strong externalism says that the supervenience base consists of the environment and the *causal relations* the environment bears to a psychological subject. *Weak* externalism says that the supervenience base is restricted to the environment impinging on a psychological subject and does not include causal relations.

to the psychological subject. And the first puzzle is how propositional attitudes such as belief could have causal powers that affect (and effect) action.

Any property will need to meet certain conditions of efficacy (instantiation, for example), in order to play a causally relevant role in a system. Our best theories about causation purport to be accounts of how casual connections obtain between naturalistic properties of the entities in a casual relation, and there is still widespread controversy in the philosophy of mind over how it is that intentional properties are materially implemented. If a mental state is conceived of wholly physically, then how can it have intentionality? But if it isn't, then how can it have casual powers?

Now apart from the problem of explaining how *any* wholly material entity or process can be understood to implement semantic properties like content, there is the further additional difficulty generated by the externalistic individuation of content. According to externalism, the content of a mental state is a relation to the particulars and properties in the world external to the psychological subject. The thesis is provocative precisely because it claims that the mind itself is (partly) constituted by the mind-independent world of particulars and properties. But if the existence-and-identity conditions of a propositional attitude are those of its objects, then it difficult to see how a mental state satisfies the plausible conditions on casual efficacy. Externalism in effect pulls content apart from causation, by showing that sameness of intrinsic properties is not sufficient for the sameness of intentional ones¹.

The puzzle of accounting for the causal properties of the intentional is only one aspect of the general issue. The other concerns the character of psychological explanation. The claims of folk psychology, for example, refer to what a person believes to *explain* why she did what she did. But the attraction of the claims of folk psychology is that the etiology of action devolves on these properties, and that to cite them in explanation is to cite *casually relevant* explanatory properties of action, a claim that is threatened by the externalistic individuation of content.

3. Robust Folk Psychology

One point of formulation: I think there is a way to spare folk psychology from the threat of externalism, and, most important, to do this in such a way that is consistent with a stalwart commitment to folk psychology as genuinely *intentional*, genuinely *casual*, and genuinely *explanatory*. I emphasize this to make the point that some formulations of folk psychology could (and do)

¹ At least for concepts subject to *strong* externalism, which is our concern here. The situation is more complex for concepts subject to weak externalism (cf. Preti, 1998).

forswear one or more of those elements. What I will defend as *robust folk psychology* (hereafter RFP), however, is to be understood as follows: RFP is set of naturalistic/scientific casual generalizations whose explanatory features are modeled on the covering-law or deductive-nomological model of scientific explanation. The claims of RFP are (held to be) causally explanatory and predictive in a naturalistic sense. Most importantly, the explanatory claims of RFP advert to the content of mental states – content bears the explanatory burden.

Now, given that the intuition the RFP has explanatory virtues worth defending, notwithstanding the tension generated by externalism, the so-called casual role problem is a considerable obstacle. But I believe it *appears* more intractable than it really is, and I will argue below that the problem is generated by the tendency to formulate the necessary conditions of RFP too inclusively. How much do we really need to stay true to the robust conception folk psychology as genuinely intentional science? We may need less than we think. I shall argue in closing that a new look at the old debate concerning reasons and causes may help to resolve the problem of content in explanation¹. But first we must make clear the severity of the threat to RFP from externalism.

4. Nomological Supervenience and Content-Individuation

Fodor's claims (1994) are a useful perspective from which to focus on the threat to RFP posed from strongly externalist content individuation². I will argue here that his arguments do not blunt the threat to folk psychology coming from Twin Earth cases, by showing that his attempt to defuse Twin Earth cases by emphasizing the *nomological* conditions for their obtaining does not succeed. This is because (as I will defend) Twin Earth cases are ways of highlighting that an appearance/reality distinction is what is at work in the content-individuation of strongly externalist concepts. The threat to RFP is a metaphysical one, which Fodor's reformulation does not address.

In (Fodor 1994), Fodor claims that there *are no* Twin Earth cases of the sort that could genuinely threaten the claims of folk psychology. For Fodor, twin cases are cases that, among other things, show that content can come apart from causation, since they show that the same implementing mechanism (some internal computational state, say) is *metaphysically* insufficient to gua-

¹ This is not the place to argue for the independent plausibility of folk psychology. This is a subject treated extensively in literature; see (Guttenplan, 1994) for an overview.

² It goes without saying that Fodor is the most ardent defender of RFP. Stich (1991), among others, is a useful start for the case *against* RFP.

rantee the same intentional state. As Fodor rightly sees it, Twin cases separate the key properties that underpin the virtues of the explanatory claims of RFP. So for RFP to be true is for these cases to fail to obtain (Fodor 1994, 24–25). This way of putting it, of course, makes the argument straightforwardly vulnerable, by contraposition, to the rebuttal that should Twin Earth cases be shown to proliferate after all, then RFP remains vulnerable. I argue in what follows that in fact they do proliferate quite naturally.

What Fodor does is to reformulate the Twin Earth thought-experiments with an emphasis on the nomological, rather than metaphysical, conditions that govern their obtaining. His position is that there are no laws that require or necessitate Twins, because there are no laws that require concept-distinguishability even under conditions where kinds may be distinct from one another. So, he claims, if there are no laws necessitating Twin cases then Twin cases are accidents. No respectable folk psychology is under any obligation to cover my mental states and any *accidentally* divergent mental state of any of my molecular duplicates. And as a result, folk psychology can withstand the purported threat of externalism¹.

The upshot of what Fodor calls Putnam's story (Fodor 1994, 28) is that

[it] shows something about our *concept* of content: it shows that the supervenience of the broadly intentional upon the computational isn't conceptually necessary. But it doesn't argue against the *nomological* supervenience of broad content on computation...

But Fodor's claims with respect to Twin Earth cases, as formulated in terms of nomological supervenience of the intentional on the computational are too weak². Recasting the issue in terms of nomological supervenience of the intentional on the computational cannot meet the genuine problem, because, as I show below, it is precisely the *point* of Twin Earth cases to show that your concept *C* and that of your twin can be distinct *irrespective* of the nomologi-

¹ It could be that Fodor is conflating:

(1) $\sim (\exists L)$ (There are twins)

with

(2) $(\exists L) \sim$ (There are twins)

The stronger claim is (2), but doesn't follow from (1), nor is it obvious that it could be defended.

² It looks a bit as if Fodor is just smuggling in narrow content in disguise, in spite of his claims in (1994) that narrow content is now *unnecessary* – by defending a kind of narrow *nomological* content. Fodor claims that there is no obstacle to claiming that nomological similarity between me and my Twin fixes our intentional similarity. But this seems to turn depend on her presence in my environment or one where (as Fodor sees it) the same nomological regularity of concept discrimination and deployment rules. Note the familiar counterfactual formulation that he used in his defense of narrow content in (1988). It's still an argument to causal powers, but this time the causal powers attach to the nomological regularity with which we make concept discriminations.

cal/empirical conditions that govern your dispositions to apply C (Fodor 1994, 30). It may be true that Twin Earth cases may show something about our *concept* of content, or about the conditions that apply to our abilities for concept-discrimination. But the relevant issue with respect to the role of content in explanation is what Twin Earth cases entail about the nature of mental content (in Russellian terms, about propositional constituency). The heart of the issue remains a metaphysical one, Fodor's redirection of emphasis notwithstanding¹.

5. Twin Earth: Appearance and Reality

It bears repeating that Twin Earth cases are a device in an argument that more generally concerns what the determining conditions might be for meaning and, ultimately, the nature of mind – the externalist thesis that it is the *world outside the subject's head* necessarily determines mental content. I wish to suggest that the true essence of a Twin Earth case – its relevance to the issues arise concerning content-individuation – is what it illustrates about the relation between mind and world. And this is that the way the world appears is not always the way the world is – but that, significantly, it is the way the world *is* that governs the existence-and-identity conditions of mental content. Twin Earth cases, in short, should be thought of as a vivid way of showing how the head-independent environment has a life of its own with respect to content-individuation. The head-independent environment amenable to Twin Earth cases is so amenable because it is subject to an appearance/reality distinction, a pervasive and systematic characteristic of all worlds that contain psychological subjects.

¹ I should say here the erroneous picture may be due in part to Fodor's view about informational semantics, about which he is inconsistent. On the other hand, he claims that (Fodor 1994, 4-5):

The basic idea is this: The content of a thought depends on its external relations; On the way that the thought is related to the world, not the way it is related to other thoughts...

But on the other (Fodor 1994, 37):

Semantics, according to the informational view, is mostly about counterfactuals; what counts for the identity of my concepts is not what I do distinguish but what I could distinguish if I cared to (interalia, what I could distinguish by exploiting instruments and experts).

I think Fodor tends to slide between nomological/epistemological considerations and semantic ones in his zeal to counter holism in semantics (and maybe for other reasons, too). And while I agree with him about holism, I think this preoccupation steers him wrong on externalism and mental causation.

(i) Natural kind concepts

The natural kind concept Twin Earth cases – water/xyz; jade/jadeite; aluminum/molybdenum; elm/beech – are the simplest illustrations of the point that part of the head-independent environment is subject to an appearance/reality distinction which can affect the individuation of mental content. Recall that Putnam formulates the thought experiment so as to undermine that Fregean semantic theory, which he achieves by showing that, to take one example, even *identical* descriptions associated with the term “water” on both Earth and Twin Earth, generated by the appearance of the stuff in each environment, is insufficient to guarantee that both Earthlings and Twin Earthlings *mean* the same thing by their term “water.”

And this is because *what looks like water is not necessarily water*. XYZ presents the same phenomenological *appearance* to me, and to my twin, as H₂O does, but does not have the same chemical composition, so *it isn't the same stuff as the stuff I call “water.”* My twin and I use the same term (“water”) to apply to what surrounds us in our respective environments, but it doesn't *mean the same thing* because it doesn't apply to the same stuff¹.

This point holds likewise for the jade/jadeite case (Putnam 1975, 241). Putnam describes this case as involving two very different microstructures that produce the same surface properties; here the exploitation of an appearance/reality distinction is clear. Neither the aluminum/molybdenum case nor the elm/beech case is fundamentally different, although these introduce the notion of deference to experts that we will discuss separately below. For now, note that, as stipulated by Putnam (1975, 225-6), the only people who can tell that ‘aluminum’ refers to *molybdenum* on Twin Earth are (both sets of) metallurgists, who are in possession of additional facts about metallurgical reality; the rest of us, by stipulation, *cannot distinguish the two elements*.

And similarly with the elm/beech case, as described by Putnam (1975, 226): “Suppose you are like me and cannot tell an elm from a beech tree...we still say the extension of ‘elm’ in my idiolect is the set of all elm trees.”

(ii) Artifact concepts

Artifact concepts like “pencil” (Putnam 1975, 242-43), also fall neatly in line. Earth-pencils can't be organisms, because earth-pencils are necessarily artifacts. Another world, however, might contain *pencil-counterparts* whose microstructure is that of an organism, but whose macrostructure is phenomenologically indistinguishable from the artifacts that we use in this world. In

¹ See Putnam (1975).

that world, the term "pencil" refers to *pencil-counterparts* which are organisms; in our world, "pencil" refers to an artifact.

(iii) *Indexical concepts*

Putnam (1975, 234) explicitly considers the distinction between the phenomenological properties that arise in employing the concept "I" to refer to oneself, and the particular to whom one refers. Indexicals such as "I" have context-dependent reference; each use of "I" by a different person *refers* to a different person and so yields a different meaning, according to externalism. Yet each time one uses the concept to refer to oneself, one presents oneself with the same appearance as does one's twin; that of *oneself*. Same appearance; different reality.

(iv) *Deferential concepts (expert cases)*

Fodor's (1994) position on deferential concepts is somewhat murky, in my view because he may be implicitly grappling with the consequences of externalism for the intuition that we have privileged access to our mental states, and the upshot of this issue for the retention of RFP. We can assess Fodor's claims here without going too far afield, however, by noting that there is indeed an issue here as to what commitments folk psychology makes (or ought to make) with respect to whether I can introspectively distinguish my own concepts, and how externalism bears on this.

Remember that according to the usual story, I myself am in the same computational/intrinsic state whether I am thinking about elms or beeches, but *elm-thoughts* are about *elms*, not *beeches*, and thinking things about elms is not thinking things about beeches. So there appears to be a problem: I am not able to distinguish by introspection my elm-thoughts and my beech-thoughts, because what distinguishes them is a question of what botanical environment my thoughts are directed onto, something to which I do not have introspective access.

Fodor claims instead that my failure to discriminate concepts in cases like these is an epistemological failure, and claims that such a failure is no threat to folk psychology. His reasoning looks to be as follows: semantically speaking, elms individuate elm-thoughts according to the principles of externalistic content-individuation. If I can't tell elms from beeches, however, the problem of concept-distinguishability arises; I can't distinguish my own concepts of elm and *beech*, but externalism claims they are distinct. Generalizations fail; and RFP is under siege.

But Fodor claims I *can* tell the difference between elms and beeches, and so I do have the ability to distinguish my elm-thoughts from my beech-thoughts: what I do is, I consult the experts. It is just false, on his view, to claim that I can't distinguish concepts that *are* distinct. My own particular abilities may be limited, but in the wider community there are experts whose distinguishing abilities I can help myself to, or, as he puts it, co-opt (Fodor 1994, 36). Deferential concept cases, for Fodor, are cases where I *can* make the requisite conceptual distinctions, in the sense that I get the experts to do the work for me. So externalist principles of content-individuation do not bear so threateningly on the claims of folk psychology when it comes to deferential concepts, precisely because they are *deferential*¹.

There are a number of objections to this. Arguably, one of the things that is subject to pressures from externalism is my ability to discriminate my own thoughts with no reference to the external environment. As many have noted² the issue of privileged access is thrown into relief in the debates on externalism and its consequences. A further element of this issue, as Fodor is aware, bears on the intuitive case for the robust conception of folk psychology. Part of that conception – possibly implicit – seems to be we do have privileged access to our own mental states. We (traditionally) know when we are in pain, we know what kinds of intentional *perceptual* states we are in, and, further, that we know that we have beliefs and desires and we know what they are³.

Fodor counters the putative threat with an emphasis on the deferential concept cases. My own (externalistically individuated) thought might be introspectively closed to me, he claims, but I can thwart this difficulty by consulting experts. My distinguishing abilities are more than just enhanced by *theirs*; my mental states expand to include those of the experts to which I defer⁴.

1 Fodor could be following Putnam's line of argument here (Putnam, 1975), which introduces the so-called socio-linguistic hypothesis about the division of linguistic labor with respect to *recognizing* the reference of such terms as "elm," and "aluminum." My objection to this is that to the shift the emphasis to *method of recognizing* reference is ultimately of no use against the problem posed by externalistic individuation of content.

2 See (Burge 1988) for a start; and additional discussion by, among others, Boghossian (1989); Brueckner (1992); McKinsey (1991); Warfield (1992).

3 The question as to whether a robust formulation of folk psychology *need* be based on Cartesian intuitions about the mind is one to pursue elsewhere.

4 This suggestion is also open to a reductio in that it appears to trespass illegitimately on the notion of ability. For instance, it seems absurd on the face of it to claim that my partner has the ability to have a baby because he can defer the baby-having to a baby-having expert (i.e. a mother, even the mother of his child). For Fodor to be right here, my ability to make conceptual discriminations has to be different from this sort of ability, and it has to be different enough to make my deference to experts plausible in the right kind of way.

But now we face a dilemma (horns (H1) and (H2)). Externalism threatens privileged access and folk psychology. To save it, I build into my mental states deference to the mental states of experts (among other things). But then this suggests that the mental states of *others* are somehow individually part and parcel of my own. So (H1): either I don't have privileged access to my own mental states or, remarkably (H2): I have, in effect, privileged access to the mental states of *others*. If the mental states of experts are "co-opted" by mine in the strong way that Fodor claims – their ability to make distinctions is now mine – then my mental states include the mental states of others. In short, either we have first-person acquaintance with the mental states of *others* as well as our own or we don't have first-person acquaintance with our own mental states¹.

The concession to make to this, it seems to me, is to admit the "co-opting" the mental states of experts is not so strong a claim as to lead to H1 and H2. But to weaken the claim puts it in line with the claims we made above on the nature of the individuation of natural kind, indexical, and artifact concepts. If the mental states of experts impinge on the existence-and-identity conditions of my mental state *whether or not I can tell*, then the situation is precisely that of canonical externalist content-individuation. It is the world outside my head that independently settles the question of what mental states I possess².

There are other objections to Fodor's treatment of deferential concept cases. For instance: there is no reason to believe that experts are less subject to the pressures of externalist content-individuation than the rest of us. For one thing, they are subject like anyone else to appearance/reality distinctions. Suppose there are entities whose macrostructure is identical but whose microstructure is opaque in principle – then there would be entities for which there *could be* no expert discriminators³.

So by what means could experts mitigate the problem of distinct entities that *can't* be distinguished? And what of the mechanics of deference? Is there any reason to suppose that we can in principle rule out a *chain* of deference between me and my expert? Suppose the chain were to telescope into an infinite regression of deference? These points seem to undermine the kind of deferential relation that is bearing the weight here.

¹ This point, incidentally, would apply to Burge's claims that the social environment also plays a role in the individuation of my mental states; not surprisingly, since Burge's point is an extension of Putnam's socio-linguistic hypothesis (Burge 1975). But Burge (wisely, it turns out) makes the weaker point, and so evades the dilemma.

² The nature of the world outside of my head should be understood to encompass properties, particulars, and the mental states of others.

³ See (McGinn 1991).

6. Summary

If I am right, then Fodor's attempt to refashion Twin Earth cases by way of nomological supervenience is insufficient to blunt the threat of externalism to folk psychology. That threat, as I spelled out, instead comes from the underlying metaphysical state-of-affairs that Twin cases illustrate. The conditions that bear on the *application* of our concepts, or our discriminatory capacities, may well be subject to all sorts of interference, and some of the interference may well be "accidental, fortuitous, or wildly exceptional"¹. But the disruptive conditions that may in principle negatively affect cases of concept-application and concept-distinguishability are ultimately irrelevant to the problem of content in explanation. Those latter are *epistemological* concerns; the problem of content in explanation is a *metaphysical* one.

We can support this claim by way of a broader question – relevant to all sciences – as to what kinds are to be relevantly referred to in an explanatory theory. RFP, as we know, requires its kinds to be *intentional*. But then it is reasonable to expect that the taxonomical principles that will be at work to deliver intentional kinds in the right way will be more than a local nomological matter. When content-individuation is externalist, we must ask: can content-individuation be simply a matter of the nomological properties (or lack thereof) at work in local/contingent examples of head-world relations? Or does content individuation need more genuinely *constitutive* conditions to be satisfied?

On Fodor's account, my Twin and I are internally/computationally identical, and what varies, as he sees it, are the nomological conditions that govern

¹ I think there is an interesting analogous point to be made with respect some of the discussions of personal identity familiar in the literature. It is sometimes claimed that it is tiresomely speculative to erect thought-experiments-interruptions of consciousness, split-brain experiments and the like-in order to fix intuitions about personal identity (Cf. For example Wiggins, 1980, who claims that the self is a natural kind). This kind of point could be formulated as follows: To fix the nature of the self we ought to take advantage of the fact that selves supervene on *human beings* with local nomological regularity. On this folk-personal-identity-theory, our taxonomic criteria are commonsense, local and empirical.

This approach is not entirely unattractive. I think, however, that it is unphilosophical. The whole point about *selfhood* is whether such a thing as a self can be attributed to entities in the light of criteria that transcend the contingencies – regular and general though they may be – that obtain in the actual world.

I can't help but suspect something like this may be at work in Fodor's position here. And what I've tried to show, on the contrary, is that there is nothing farfetched, anomalous or unconstrainedly speculative about Twin Earth thought-experiments. What underlies them, and what leads to the problem for content in explanation, is a perfectly straightforward and uncontroversial appearance/reality distinction, an attribute of the very nature of *worldliness* that is at work in the issue.

the application of our respective concepts. These conditions are fortuitous, depending as they do on the accidental local variations in the substance environment impinging on my mental states. So, as Fodor would have it, the important issue is not really that *content* varies between twins; but that what varies are the conditions of content-deployment, so to speak. And folk psychology is under no obligation to generalize over such local contingent matters.

We can go back to Putnam's original example in (1975) to support, in contrast, the view that the significance of externalist content-individuation (and its consequences for the nature of the explanatory concepts of RFP) is in its character as a metaphysically constitutive thesis. Putnam (1975, 224-25) is clear that the natural kind environment (among others) bears a constitutive relation to the mental states of those in relation to it, and it does so whether or not they can tell (Cf. his remarks on what 'water' meant in 1750). When Fodor claims that there are no laws that *require* concept distinguishability (even) where concepts are distinct, I have argued that he is missing the point. It may well be we will be able to tell that our concepts are distinct only under certain contingent conditions, but the point is irrelevant to the problem at hand. We have to face the externalist thesis squarely: our concepts are individuated by factors independent of our capacities to distinguish, discriminate and apply them. Are externalism's consequences for the claims of folk psychology therefore insurmountable?

7. Retaining Folk Psychology

Suppose that we are not convinced that either robust folk psychology or externalism should be abandoned in spite of the problems that so far seem to plague MTT. It may be that to solve the problem we shall need to approach it from another angle. My concluding remarks here will take the form of two brief discussions: one on the nature of the proposed strategy; the other, on how it could work.

(i) *Casual relations and Casual Explanations*

We can start by looking squarely at the source of the impasse. The difficulty of reconciling wide content with RFP appears to arise from taking content-based psychological explanation as a species of casual explanation, *in virtue of its reference to the intentional properties of mental states*. But, as we have seen, content-adverting explanation fails to capture the causal-explanatory properties of content. So what we need is a strategy which will enable us to consistently hold these theses:

Intentional mental states have casual powers
 Intentional mental state are externalist
 Psychological explanation is intentional

without falling afoul of the problem. And the way to do this may be to abandon the condition that the intentional vocabulary of RFP need be casually explanatory.

It is a familiar point that there are (at least) two explanatory roles that an intentional mental state might bear vis-à-vis an action: rationalization and causation. Rationalization, a normative/justificatory enterprise, situates the mental state in question within a logical network of other mental states possessed, and purports to reveal their logical interconnections. But, familiarly, rationalization is not enough, because we may (say) possess all the beliefs and desires that would make a particular action reasonable and still do not do it (*akrasia*); or ultimately, act for different reasons. Rationalization alone can fail to fully explain why an action was performed, so there appears to be need for a further link between mental states and actions. That link is commonly thought to be a casual one.

Now Davidson's claim that rationalization is a species of *casual* explanation (Davidson 1980) was an attempt to render compatible the distinct explanatory properties attributed to reasons. But Davidson's defense of that thesis is obscured by the fact that it appears to proceed by way of the claim that reasons cause action. The former, however, doesn't follow from the latter. There are familiar *logical* gaps in the move from the thesis that reasons *cause* actions to the thesis that reasons *causally* explain actions.

We know that:

Reasons must cause action to explain action; but just because reasons cause actions, it won't follow that citing reasons explains action;
 If a reason causes an action *and* justifies it, it won't follow that citing that reason causally explains the action;
 Reasons might explain actions by justifying them, and reasons can cause actions; but it doesn't follow that reasons *causally explain* actions.
 Causation is a relation between events *however* described, but causal explanation is a relation between events *under a description*. So the way events are described will affect the truth-conditions of the explanation-sentence, quite independently of the causation relation borne by those events as they are in themselves.

Now the distinction between causation and explanation is commonly thought to engender more problems than it solves. I think instead that that very dis-

inction is potentially our best hope to preserve the mutual consistency of externalism and robust folk psychology.

The proposed strategy – call it neo-Davidsonian – exploits the difference between causal relations and casual explanations head on. The claim is that mental states cause actions, but the intentional properties of mental states, while explanatory, are not causally explanatory. Intentional mental states enter into genuine causal relations, but we cannot causally explain their doing so by adverting to their intentional properties.

The significant consideration in favor of this strategy is that it has more promising implications for the threat posed by externalism to the claims of folk psychology. There seems to be no initial prohibition on this view to the claim that all causally relevant properties in cases of mental causation are intrinsic to the agent. Intentional properties, however, can be individuated relationally or externalistically all the same.

My concluding remarks will sketch how this strategy could work without sacrificing RFP. The aim, again, is to make a case for the mutual tenability of externalism and a robust folk psychology, the latter formulated as a set of lawlike claims that are at once intentional, causal, and explanatory, in spite of the fact that those conditions appear to be the source of mutual inconsistency.

But are they? Let's consider how strong a claim the defender of MIT must make to preserve RFP. The pretheoretical intuitions to be preserved – the ones doing the work – appear to be two: first, that propositional attitudes are *causes* with *contents* (Fodor 1986, 24); second, that the intentional vocabulary of folk psychology provides a naturalistic and ineliminable taxonomy for genuinely explanatory psychological statement. What is doubtful is that the preservation of a robust conception of folk psychology really needs the condition that its intentional claims be *causally explanatory*.

To suppose so, I would suggest, might be the result of taking the threat of epiphenomenalism (too) seriously and as a result, conflating two claims: (1) that intentional properties are “causally responsible;” *and* that¹ referring to these properties under intentional descriptions is *causally explanatory*. Epiphenomenalism can't really threaten robust intentional causal realism on the score of intentional description. We cannot take seriously an argument to the effect that intentional properties do not or cannot have causal powers unless referring to them *qua* intentional yields causally explanatory claims. If intentional properties have causal powers they will have them quite independent of choices of vocabulary.

¹ The words are Fodor's (Fodor 1994, 1). For more discussion of this issue, see *Mind and Language* (Vol. 12, Numbers 3 and 4), September/December 1997.

What makes this plausible – I would claim even irresistible – for a defense of RFP, is a commitment to *distinct levels of conceptual explanatory taxonomy*. Now Fodor (for instance) has never quailed from the position that the intentional properties of mental states will need, by the principles of natural causation, implementation in a casually plausible mechanism. But that suggests that we can get what we're looking for. Again, strictly speaking, all RFP needs is: (1) generalizations that advert to content; (2) generalizations that advert to content and are explanatory; and (3) causally responsible intentional properties. The implementation of intentional properties in, say, computational ones grants to them the desired causal role. So to advert to intentional properties is genuinely explanatory. Adverting to them qua intentional may block causally transparent reference to the implementing mechanism, but the prior condition on RFP – the implementation of the intentional in the causal/intrinsic – makes this no sacrifice that hasn't already been made. It may be that this is a small price to pay for a plausible defense of the mutual tenability thesis.

I am grateful to Jerry Fodor and Colin McGinn for discussion of the ideas in this paper.