

## FOLK TERMS AND AGENCY

*Jane Duran*

### *Abstract*

I employ three major lines of argument to support the conclusion that there is a use for folk terminology in applications of the "science of the mind". The first argument attempts to be precise about what the attribution of intentional states amounts to, and how such attribution could be related to assignment of agency, particularly in moral contexts. The second argument asks us to examine why it is that this particular debate - the debate about the utilization of folk psychological terminology - reminds us of other current debates in philosophy (such as relativism vs. absolutism, or Complete Accounts vs. incomplete accounts), and what it is that we can obtain, philosophically, from being aware of the fact that we are so reminded. The third line cites work by Laudan and helps explicate the notion that there is a range of positions available here (with regard to the employment of the folk psychological), and not merely two. It is concluded that folk terminology is being unfairly discarded from contexts that would allege or purport to be scientific.

Those who have attempted to espouse the non-nomological utility of folk psychological terms have frequently tried to steer a dangerous middle course between a non-theoretical liking for commonsensical terms, on the one hand, and the development of some sort of scientific basis for mentalistic terminology, on the other. The literature on this topic is extraordinarily hedged: one attempts to salvage the unsalvageable, as it were, without forgetting what motivated one's attempts to come to grips with the "mental" in the first place.

The core area of concern here is really agency. In other words, one is inclined to want to try to save the folk terms, on some level, because, as Dennett has said,

... the 'validity of our conceptual scheme of moral agents having dignity, freedom and responsibility stands or falls on the question: can men ever be truly said to have beliefs, desires, intentions?'

It isn't merely that it is important to be able to ascribe beliefs and intentions to humans because this gives us some sort of predictive handle on their day-to-day behavior. As Millikan has noted, "The intentional characterization of John 'He wants to meet the ...' where the blank space is filled in and read transparently, does give us a handle on what John might well do...". More substantively, we would like to know what John or Mary or George Bush might do - or at least be able to give ourselves some statistically accurate summary of what they might do - at least partially so that we can decide what we ought to do. But clearly it is more than mere attribution of future predication which is at stake here. We want to be able to characterize Mary or George Bush. That is, we want to be able to make them accountable for their behavior in a manner which takes into consideration both the future and the past. And the difficulty for the folk psychological on this score, as we know, is that it seems to run smack up against its non-nomological character and its inability to do duty in a science of the mind.

In this paper I intend to fill in an account of the folk terms which reimpresses upon us their predictive utility, and which reestablishes the sense in which they satisfy constraints on the notion of moral agency. Then I want to tie together this line of argument to other recent developments in philosophy which I think are not unrelated. But I will begin at the beginning, and this requires our looking at folk psychological terms in view of their predictive utility for behavior.

## I

When we discuss moral agency, we frequently have examples ready-to-hand of failed attributions of agency. All of the hoary examples from ethics come to mind - in some cases, we cannot attribute agency meaningfully to a subject because the subject was not in his or her right mind; in some cases we can make no such attribution because of the subject's being developmentally disabled; and in some cases where a certain sort of agency is attributed, we find that mitigating circumstances somehow diminish the agency sufficiently as to render the question of responsibility a difficult one.

Questions regarding moral agency all have at least this much

in common, however: in the case where we can meaningfully speak of agency, what we are really talking about, it turns out, is intentionality. If the agent is mentally healthy, not developmentally disabled, and under no "undue duress", then the agent can be held responsible for something morally only if we can show that the agent intended a certain course of action which resulted in moral harm.

When we say that the agent "intended" a certain course of action, we are, naturally, also making a somewhat philosophically more sophisticated statement. We are saying that certain persons, places, objects, and/or situations were taken as intentional objects by the agent, and that the taking of these persons, etc., as intentional objects resulted at least partially in the set of ensuing events for which we are attempting to hold the agent responsible.

If I hold Mary morally responsible for John's death, normally the framework for attribution is much like what I have just described above. Mary is in a state of comparative mental health, of a normal level of intelligence and intellectual development, and was not under such unusual duress at the time of John's death as to render her incapable of ordinary judgment. So Mary must have intended John's death in some way; that is, one or more of the intentional states in which Mary found herself during the time of John's demise had John's death (or grievous harm) as its object.

Now the difficulty for current psychological models has been simply this. Most of the computationally-oriented models eschew lengthy or even brief) discussion of intentionality, since the focus in these models has always been at a level beneath that of intentionality, no matter how intentionality is characterized. Although some theorists have steadfastly maintained that part and parcel of the motivation for the computational model in the first place was a desire to get from the inner contents of one's mind, as it were, via causal chains, to some sequence of events, most of the literature on the topic has focused on the logical/semantical relations nature of the model itself. Thus Pylyshyn had remarked:

Plainly, what is going on is, my behavior is being caused by certain states of my brain. Yet - and this is the crux of the problem - the only way to explain why those states caused me to type the specific sentences about walking, writing, the mountains and so on is to say that these states are in some way related to the things referred to (writing, walking, mountains)... My brain states are not, as we have

noted, causally connected in appropriate ways to walking and to mountains. The relationship must be one of content: a semantic, not a causal, relation.<sup>3</sup>

But the fact that this desire to account for the causal may have originally constituted some of the motivation for portions of the computational model does not detract from the lacuna in the literature on the topic of moving from the realm of the mental to the outer realm. If intentionality can be glossed in terms of the computational model as representations, then the model has tended to focus on relations of representations to syntactic encoding, rather than possible relations of representations to material outside the mind.

More specifically, and to recapitulate existing arguments, if one attempts to tie a folk psychological term to a representation (and thence to the encoding), one can aim in either the direction of type-type or type-token identity. But there are convincing counterarguments to both these sorts of identity. Stich has argued forcefully against type-type identity, and Baker has an elaborate (and apparently irrefutable) argument against type-token identity in *Saving Belief*. The latter argument involves the clever twist that one can hypothesize strings of output which are of indistinguishable phoneme and morpheme type, but which would have to be of distinguishable type at some level, since they involve different languages and hence different meanings.<sup>4</sup>

Now if neither type-type nor type-token identity goes through for the relation between folk psychological terms and brain states, then (the conclusion usually goes) the folk psychological can have, at best, an instrumental value, while failing to exhibit the nomological regularity which is necessary for the formulation of a cognitive science. So we are back to square one with the folk psychological terms: if they do in fact help us with notions of agency, they do so in a way which is instrumentally applicable only at the grossest level, and which is completely inapplicable at the microlevel.

Now the gloss on Mary's intending John's death begins to look something like this. As we said earlier, Mary could have intended John's death only if Mary was in intentional state  $x$  with regard to John at some time  $t$  either preceding John's death or during the sequence of events leading up to it. Now Mary's intentional state would ordinarily be expressed in English with a "that" clause, and the object of this state, thought of as a proposition or sentence, involves a semantic representation which itself is correlated with a syntactic encoding. The fact that no hook-up for the encoding can be given which would hold

on the basis of regularity (either intra- or interpersonally ) does not mean that we cannot attribute agency to Mary, and it certainly does not mean that we cannot make large scale predictions about the behavior of those who would exhibit some intentionality. It does mean, apparently, that we would not be able to invent a brain-probe or scanning device which would correlate brain state A, B or C with some set of intentional attitudes, but, come to think of it, we have been doing without these devices these many years.

If I know that Mary wished for John's death, and was especially hopeful that the chandelier would fall on him - which is why she seated him under it - I know a good deal about Mary's causative role in John's death. If I know that George Bush has the destruction of the Medellin cartel as an intentional object, (and the best evidence for this, of course, is verbal/behavioral, although other evidence will suffice, perhaps with auxiliaries), then I can make some predictions about Bush's behavior, and, more importantly, when and if the destruction of the cartel is accomplished, I can attribute it (at least partially) to the agency of George Bush. Our desire to establish agency is, then, in almost all circumstances, related to intentionality and folk psychological constructs. Merely behavioristic criteria are not, of course, enough. In some cases these criteria would lack the appropriate motivation, and even if they did not lack the appropriate motivation, we would be inclined to think that we need more direct evidence of intentionality (again, largely verbal) before we could properly assess motivation and then make an attribution of agency. The fact that we take agency for granted merely underscores the import of the notion of agency for almost all of our moral assessments, and the difficulty in getting along without the ability to make such attributions. Without a notion of agency, we cannot properly predict the future or retrodict the past. And our notions of agency are reliant on intentionality.

All of the foregoing merely reestablishes the usefulness of folk terms in some contexts. But the heat generated by this topic currently is, I believe, related to other areas of debate in philosophy, and it is to those areas I now turn.

## II

The normative nature of philosophy and the pretensions it has always had to provide a Complete Account have recently come under withering attack.<sup>5</sup> We tend to associate these attacks with areas of philosophy which we can easily pinpoint as having

origins in antiquity. Metaphysics, classically construed, or epistemology - in its most rigorous and technically counterexamplifying guise - are two of the areas we associate with the desire to save the possibility of a Complete Account and the desire to abolish the belief in such a possibility.

But I claim that it is no less true that philosophy of mind itself has fallen into the same sort of difficulties as the areas mentioned above, and for much the same sort of reasons. The science - mimicking aspects of philosophy in this century - the growth of work in formal logic, the paring down of metaphysics, the formulation of questions in theory of knowledge or even ethics in terms of ever more logically rigorous necessary and sufficient conditions - all of these, I argue, are attempts to continue philosophy in the tradition which dates back to Plato.

I need not argue strenuously for this rather uncontroversial thesis, for it seems to be accepted as a commonplace. But what may be more difficult for us to swallow is that philosophy of mind, qua subdiscipline of philosophy, has developed along the same lines. The fact that the classic mind/body split, which was the locus for almost all of the original work in philosophy of mind, is probably best characterized as a metaphysical split may be one factor which helps explain why it is that the desire to obtain the one true story in philosophy of mind reminds us of some of the same sorts of desires in other areas of philosophy, even down to the more minor corollaries which might be thought to adhere to them. In a way, philosophy of mind was better off when dualists still abounded, and when those who would be content to think of mind as something completely apart from the body were given to frequent and vociferous expressions of their views. On a dualistic view, the difficulty of obtaining a completely accurate account is understood, a characterization of mind which, however vague, has the virtue that one cannot really begin to think of it in lawlike terms. But the demise of dualism and the rise of identity theories at an earlier point in this century virtually guaranteed that work in philosophy of mind would begin to suffer from the same sorts of conundra which infected, for example, philosophy of science. The computational model, work in cognitive science, physiological evidence about brain states - all of this is designed, again, to lead in the direction of the Complete Account, even if it has not been immediately obvious. Although some have associated naturalization in epistemology with a move away from the Account (and epistemology may be a slightly different story, partly because of the manner in which the questions constitutive of it were originally formulated, and partially because of the age of the ques-

tions), naturalization in philosophy of mind, as it were, seems to be in accord with the notion that an Account could be given which adequately details Mary's brain state in such a way that I could know once and for all whether, on an externalist view, Mary was truly in possession of evidence  $y$ , and so forth.

The appeal of the possibility of reduction, then, is that, if one were able to reduce at the most fundamental level, one would then have the sort of account that is consonant with our notions of a Complete Account, and one would have moved theory of mind as far from its dualistic origins as possible. We might be, theoretically, somewhat ashamed of the dualistic beginnings of philosophy of mind, especially from our contemporary vantage point which forces us to view all such beginnings as conceptually murky. But the difficulty with all of the foregoing is that, reduction notwithstanding (and the arguments for it are powerful), we are not currently in a position to give the definite answer about brain states. More importantly, as we have just seen in the previous section, there are powerful counters to the argument that information about brain states could ever be decisive on an internalist or intentionalistic account. And, unfortunately for the tradition, the tradition has always been couched in intentionalistic and internalistic terms, so that further external evidence, of whatever sort, does not really seem to do the trick.

Alas for the future of a science of the mind, allusion to brain states clearly will not solve the sorts of difficulties - difficulties with identity of brain states and intentional states - which I have referred to in the earlier parts of this paper. But this need be a serious problem for philosophy of mind only, I claim, if philosophy of mind insists on allying itself with the Complete Account outlook. If one can think in terms of abandoning the Complete Account view, then one can begin to come to grips with the fact that a good deal of what we usually use on the folk psychological level will never, under any circumstances, help us with an Account-with-a-capital-"A". But the obsession with such Accounts is really, historical, and there are powerful arguments to the effect that it is high time it is dropped.<sup>6</sup> Might not philosophy of mind benefit from another sort of turn? The utility of folk terms for accounts of agency is, I claim, directly related to one's lack of pretension about philosophy of mind.

### III

In the preceding sections I have offered us a somewhat simplified choice between a Complete Account and an account which

utilizes intentionality in a non-nomological way. When I mentioned the pretensions of the Complete Account in the preceding paragraph, I did not mention the other area, currently a subject of much debate, which seems to run smack into the contretemps here under discussion: relativism.

If I insist on a model which allows for the use of the non-nomological, and which affords a place for intentionality as a guide to agency, even if it falls short of the level of a science, I might be accused of relativism. Isn't it relativistic to assume that one can utilize intentionality for certain sorts of endeavors, even if it fails to meet epistemic standards, and then assert that other sorts of endeavors require a nomologically-safe account?

One is tempted to say that there is a middle ground here. It is not only the case that one could insist on a nomologically secure science of the mind. More importantly, one can utilize intentionality without being a full-blown relativist. One can allude to conditions which provide us with predictively accurate information in most situations without claiming that no one would ever be in a position to be more precise about the given situation. A great deal of the literature on this point is concocted as if a sophisticated version of the black/white fallacy were the unvarnished truth here.

This same sort of problem emerges in epistemology *simpliciter* as a version of skepticism vs. the possibility of knowledge. The recent project of attempting to naturalize epistemology runs up against this difficulty almost immediately in some circles, for the normative tradition claims that it is question-begging to try to allude to the manner in which we do acquire knowledge - after all, the very question at issue is supposed to be whether there is indeed a genuine possibility of acquiring knowledge or whether we are working with a set of largely or entirely false beliefs. But the naturalizing epistemologist has more than one way out here, and at least one of the avenues open to her or to him is helpful for the relativist issue with which we are faced.

As Kornblith has phrased it in a recent anthology piece, one version of the problem is:

How is it possible to reject holism without embracing scepticism?

The short answer to this question is that the very factors which insure that inconsistency is kept down - reliability of individual mechanisms, social factors in cognition, and local monitoring of one's beliefs - also insure that our beliefs are kept in reasonable touch with the world. It is by means of these features that our informationally encapsu-

lated mechanisms of belief acquisition may simulate a holistic system to a sufficient degree to ward off scepticism.<sup>7</sup>

In other words, Kornblith is telling us, epistemic success (survival) is a very strong argument against the kind of scepticism with which traditionally epistemology has frequently wanted to concern itself. This argument sounds like, and may well be, simply another form of evolutionary epistemology, as it has come to be known, but it has two important points. Nothing militates against the theoretical possibility (or against the consideration of the theoretical possibility) that there may be more genuine ways to acquire information than those we currently possess. The perfect, global knowledge of the normative tradition is not a logically incoherent notion. It's merely that we are embodied cognizers, and we have to make do with the cognitive apparatuses which we possess. Luckily, those cognitive mechanisms are correct a very high proportion of the time.

All of this, of course, ties back into our original point about relativism. Kornblith's allusion to "reliability of individual mechanisms" and "social factors in cognition" reminds us that many of those mechanisms and factors are functioning (insofar as we are able to construe them with the information we have today) on the level of intentionality. Hence part of the "simulation", as Kornblith has it, involves utilizing those very intentionalistic mechanisms, even if they are not nomologically correct. The desire for the nomologically correct is akin to the desire for the epistemically perfect; neither desire looks likely to be fulfilled in anything like the current state of affairs, and, more importantly, it is mistake to believe that admitting this forces one into a position that is blatantly relativistic, skeptical or anti knowledge.

Laudan has written in a similar vein about the realist controversy in philosophy of science. I quote his material here, since he makes a point pertinent to our topic:

We have overwhelmingly good reasons to suspect that our theories about the world, even our best-tested ones, are not true simpliciter. Yet the realist still wants to cash in on the hunch that the 'truthlikeness' of our theories is responsible for their success ... The core idea here is that an approximately true theory will have consequences most of which are true, or at least which are close to the truth. As I have shown elsewhere in detail, this argument is fundamentally flawed.<sup>8</sup>

Our desire to construct a "science of the mind" is directly related, I claim to the sort of deception going on in our desire to search for scientific theories which are "approximately true".<sup>9</sup> Laudan essays to make the larger point that many of our best-confirmed theories have spots of "truthlikeness", and spots which contain no such relation to truth; furthermore, it is easy to hypothesize that these relationships of intratheoretical material will change over a period of time. One needs to understand, if one is to get past the realist/antirealist dispute (over the dispute about the employment of folk psychological terms), that this flimsy demarcation is far from setting off the entire range of positions here. Philosophy of mind is both richer and more poverty-stricken than we might initially make it out to be - richer in the sense that one need not be caught on the horns of an apparent dilemma, and more poverty-stricken in the sense that we do not have at our fingertips now the material which would enable us to reduce in the manner that an actual science of the mind would require.

Now if we did possess such material, we might indeed have nomologically useful material, and at that point we might be tempted - at least in many circumstances - no longer to employ folk psychological terms. But if one buys some of the distinctions made above (both Kornblith's distinction between naturalized and normative epistemology, and Laudan's between theories which are true *simpliciter* and theories which do a lot of work for us), one sees again how the simple mindedness of the realist/relativist dispute prevents us from seeing that, even given further reduction and more lawlike, neuronally-correct material, we still will not have a Complete Account (nor, I contend, should we have a Complete Account as a desideratum).

A view of science and the scientific enterprise like Laudan's, to which I referred above, helps us understand that we may very well be making a mistake when we take rationality (on whatever construal) as the standard by which scientific work should be judged. Science is a good deal more complicated than that, and only the most rigid realists have tried to leave us with a picture of science wherein rationality and Access to the Real take precedence over all other factors. It might be useful in this vein to cite the beginning of Laudan's *Progress and its Problems*, because he sets out forcefully and concisely what he proposes to do in a manner which puts rationality in something like its appropriate place:

To anticipate some of my conclusions, I propose that the rationality and progressiveness of a theory are most closely

linked - not with its confirmation or its falsification - but rather with its *problem solving effectiveness*. I shall be arguing that there are important *nonempirical*, even 'non-scientific' (in the usual sense), factors which have - and which should have - played a role in the *rational* development of science. I shall suggest, further, that most philosophers of science have mistakenly identified the nature of scientific appraisal and thereby the primary unit of rational analysis, by focussing on the individual theory, rather than on what I call the *research tradition*. This study will show, moreover, that we need to distinguish between the *rationality of acceptance* and the *rationality of pursuit* if we are to make any progress at reconstructing the cognitive dimensions of scientific activity.<sup>10</sup>

Here Laudan is reminding us that there is scientific work to be done, and that we frequently evaluate theories on the basis of what work they actually accomplish within a given paradigm. Such evaluation may mean that features of the theory (or even of its acceptance) which later come to be seen as irrational are deemed sufficiently rational at a given time. But why are they so deemed?

Because they contribute to the resolution of whatever problem is at hand, and it is that resolution which has, so to speak, epistemic priority (even if we would like to believe otherwise).

Finally, as Laudan has argued in still another piece, realist views themselves, when examined carefully, frequently seem unable to pinpoint precisely what the vaunted referent is which is supposed to be the target of our epistemic access and the indicator of the strength of our rationality. Realism frequently seems to want to contend that the successful scientific terms in the mature sciences contain, typically, only terms which refer, but on unpacking, it is not clear what "refer" really boils down to. In some cases it does indeed mean that a term in one of the mature hard sciences like physics does denote or pick out, but realists have been in some disagreement on this. Laudan notes in a short journal piece<sup>11</sup> that individual realists may back off from this position in a manner which does not seem to be compatible with a full-blooded realism:

Hardin and Rosenberg are evidently prepared to jettison this key tenet of scientific realism, at least with respect to theories of the past. Specifically, they are ready to say that there can be theories which have been highly successful, even 'approximately true', even though there is no

reason to think that there is anything like the basic entities which those theories postulate.<sup>12</sup>

What, then, does realism amount to? Without being unduly facetious, one can suspect that many realisms amount to the hope, rather than the assertion, that key terms refer, or the much weaker assertion (than that they refer) that they point us vaguely in the direction in which the theorists - Galileo, Copernicus, etc. - have been working all the way along. These are very weak realisms indeed, and they are, as Laudan notes, "attenuated".<sup>13</sup>

If we accept a view like Laudan's larger view about science, we no longer have Complete Accountism as a goal. We are acknowledging that some accounts work better than others at given times; and we may even hold the belief that Complete Accountism is a possible worthy goal for some far-off future. But the import of all of this for what we do here is that acceptance of something less than Complete Accountism is, insofar as philosophy of mind is concerned, acceptance of such useful and homely devices as intentional constructs. For the work done by such constructs frequently unravels knots at all sorts of levels, and might very well continue to be helpful in loosening knots even if greater reduction were available to us. It is rigidly realistic views (and their theory of mind counterparts, the Ideal Reduction views) which, if they insist on complete divorce from useful-if-loose constructs, seem wrongheaded.

#### IV

In this paper I have tried to set out what the use of folk terminology in the context of determination of moral agency really requires. I have argued that a coarse-grained gloss on the intentional states of individuals provides us with most of the information we need for attributions of intentionality (in standard cases), without the information's actually constituting nomological regularity. I have also tried to tie the folk psychology dispute into the larger Complete Account/incomplete account debate, and I have reminded us that the terms of this debate may tell us more about the extent to which we have digested the history of philosophy than they do about what is and is not a valuable endeavor, metaphilosophically. Finally, I have cited material from the literature on realism, and commented on the usefulness of the particular slant adopted for the problem under consideration here. Folk psychological terms may not always carry us to the approximate truth which we claim to be one of

the desiderata for the descriptive powers of scientific theories. But the fact that the employment of folk psychological terminology normally occurs in a context without such aspirations is not necessarily a poor reflection on the theoretical utility of folk terms.

University of California  
at Santa Barbara

## REFERENCES

1. Baker, Lynne Rudder, *Saving Belief*, Princeton, NJ: Princeton University Press, 1987, pp. 149-150. Baker is quoting Dennett from "Skinner Skinned", in *Brainstorms*.
2. Millikan, Ruth Garrett, "Thoughts Without Laws; Cognitive Science Without Content", *Philosophical Review*, XCV, No. 1, p. 65.
3. Pylyshyn, Zenon, *Computation and Cognition*. Cambridge, MA: Bradford of MIT Press, 1983, pp. 130; 27.
4. Baker, in *op.cit.*, p. 30.
5. I refer, of course, to Rortian criticisms and other sorts of Continental lines, many of which are set out in the anthology *After Philosophy*, Kenneth Baynes, James Bohman and Thomas Mc Carthy, eds., Cambridge, MA: MIT Press, 1987.
6. Interestingly enough, much of the semi-Realist literature (not the most hard-core variety) points us in this direction. See, for example, the work of Larry Laudan, or realist critics of Richard Boyd.
7. Kornblith, Hilary, "The Unattainability of Coherence", in *The Current State of the Coherence Theory*, J.W. Bender, ed., Hingham, MA: Kluwer Academic Publishers, 1989, pp. 213-214.
8. Kornblith, Hilary, "The Unattainability of Coherence", in *The Current State of the Coherence Theory*, J.W. Bender, ed., Hingham, MA: Kluwer Academic Publishers, 1989, pp. 213-214.
9. Laudan, Larry, "Explaining the Success of Science: Beyond Epistemic Realism and Relativism", in *Science and Reality; Recent Work in the Philosophy of Science*, James T. Cushing, C.F. Delaney, and Gary M. Gutting, eds., Notre Dame, IN: University of Notre Dame Press, 1984, p. 90.
10. This term has been associated with the work of Richard Boyd.
11. Laudan, Larry, *Progress and its Problems*, Berkeley: University of California Press, 1977, p. 5.
12. Laudan, Larry, "Realism Without the Real", in *Philosophy of*

*Science*, Vol. 51, No. 1, March 1984, pp. 156-162.

13. *Ibid*, p. 158.

14. *Ibid*, p. 156.