WHAT WAS I THINKING?
AN ESSAY ON THE NATURE OF PROPOSITIONAL ATTITUDES

Felipe De Brigard

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Masters of Arts in the Department of Philosophy

Chapel Hill
2007

Approved by:

Jesse J. Prinz

Thomas Hofweber

Joshua Knobe

ABSTRACT

FELIPE DE BRIGARD: What Was I Thinking? An Essay on the Nature of Propositional Attitudes
(Under the direction of Jesse J. Prinz)

The thesis I defend in this paper is that the truth—or lack thereof—of our ascriptions of propositional attitudes need not carry ontological weight onto our theories about the nature of mental states. This claim would not be surprising if it weren't for the fact that both Fodorian realists and eliminative materialists about propositional attitudes take it as a premise in their arguments. They do so, I argue, because both assume a realist stance regarding scientific theories. I claim that we would be better off if we reject this underlying assumption. At the end I suggest an alternative strategy for interpreting our ascriptions of propositional attitudes inspired by an anti-realist view on scientific theories. This view, I hope, may relieve the philosopher of mind from awkward ontological concerns regarding the nature of propositional attitudes.

To Anne Harris

TABLE OF CONTENTS

Section

1. INTRODUCTION

Briefly stated, the thesis I want to defend in this paper is that the truth—or lack thereof—of our ascriptions of propositional attitudes need not carry ontological weight onto our theories about the nature of mental states. This claim would not be surprising if it weren't for the fact that both intentional realists and eliminative materialists about propositional attitudes take it as a premise in their arguments. They do so, I will argue, because both assume a debatable realist stance regarding scientific theories. I claim that this underlying assumption underwrites their acceptance of truth as a matter of correspondence between words and things in the world, and it powers the idea that the things named by (or referred to by) our true theories must exist, never mind if they refer to numbers, neutrinos or—as with our case at hand—propositional attitudes.

Such a scientific realist stance is not ungrounded, of course. It is motivated by considerations regarding the success and failure of folk psychology. On the one hand, propositional attitude realists like Jerry Fodor take the *success* of our folk psychology as good evidence for the theory's truth, and then go on to suggest that our best theory of the mind should take the syntactic objects of our propositional attitudes as real entities—specifically, mental representations realized in the brain. On the other hand, eliminative materialists like Paul Churchland take the relative *failure* of folk psychology as sufficient ground for its falsehood, and then go on to suggest that folk psychology is false because it wrongly assumes the existence of unreal entities like beliefs, desires and so forth. The upshot of eliminative materialism is that, being a false theory, folk psychology is doomed to extinction, just as other obsolete theories we used to have—like that of phlogiston or the caloric fluid. Except for some dissident voices (e.g. Donald Davidson, Daniel Dennett, Andy Clark), this dichotomy has largely set the agenda in philosophy of mind (Fodor, 1985): either you are a realist about propositional attitudes, embracing thus one or another version of a representational theory of mind, or you become an eliminative materialist, claiming that the scientific categories of neurophysiology are going to replace those of folk psychology.

Against that background, I'd like to suggest an alternative view in which both eliminative materialists and intentional realists about propositional attitudes turn out to be partially wrong. Briefly stated, the idea is that these views represent two cardinally opposed ways of reading off ontological implications from the same underlying scientific realist assumption. I suggest that we would be better off if we reject such assumption and ground our folk psychology on a (moderate) anti-realist perspective on scientific theories. In order to make my case, I begin by explaining the origins of the dispute between intentional realists and eliminative materialists. This will happen in part 2. I claim that it spawns from disagreements about a single argument—an argument I dub (inspired by Phillip Kitcher, 2001) *the success-to-truth argument*. In part 3, I talk about eliminative materialism. Here I argue that Churchland's arguments in favor of the claim that folk psychology is false are unsound. Then I will claim that since there is no good reason to believe that folk psychology is false, the thesis of eliminative materialism cannot really get off the ground. In part 4, I move onto the critical discussion about intentional realism. My criticism here is two-folded. On the one hand, on the basis of some evidence coming from linguistics and philosophy, I argue that we do not have enough a priori reasons to believe in the reality of ‘that’-clauses’ referents. On the other hand, I suggest that Fodor’s inference to the best explanation vis-à-vis the reality of language-like mental representations can be challenged as well, casting thus more doubts about its ontological implications. Finally, in part 5, and as a way of getting out of the problem, I suggest an alternative strategy of interpreting *the success-to-truth argument*, inspired by a (moderate) anti-realist view of scientific theories. This view, I hope, may relieve the philosopher of mind from awkward ontological concerns regarding the nature of propositional attitudes.

2. THE SUCCESS-TO-TRUTH ARGUMENT

Whenever we find two philosophers who line up exactly opposite on a series of half a dozen points, we know that in fact they agree about almost everything.
Ian Hacking (1983)

Before I move onto the origin of the debate between eliminative materialism and intentional realism, let me make a prefatory terminological clarification. The term 'realism' is, as Hacking (1983, 33) notes, substantive-hungry: it needs to qualify a noun to be properly understood. For instance, one can be a realist about propositions while at the same time anti-realist about numbers. Similarly, one can be a scientific anti-realist and a possible-world realist. I will always qualify my use of 'realism' to avoid confusions. Consequently, by 'intentional realism' I mean a view according to which propositional attitudes are real entities; in particular, I mean Fodor's version, according to which they are sentence-like mental representations in the language of thought (this will become clearer in part 4). Conversely, 'intentional anti-realism' will be the view according to which there are not propositional attitudes. Eliminative materialism is one version of intentional anti-realism, but the terms aren't synonymous. Similarly, whenever I talk about 'scientific realism' and 'scientific anti-realism' I will clarify the terms. As we will see, they do not overlap—Churchland, for starters, is a scientific realist and an anti-realist about propositional attitudes. End of digression.

Let us move on, then, to the formulation of what I want to call *the success-to-truth argument*:

(Assumption) Folk psychology is a theory

(P1) Folk psychology is a successful theory

(P2) If a theory is successful, then it is true. Therefore,

(C1) Folk psychology is true.

Each statement needs some explaining. The Assumption holds that the so-called *'theory'-theory* is true. Barring some idiosyncratic differences in is formulation, the 'theory'-theory can be seen as the conjunction of two claims—the first of which, it appears, is contained by the second (Lycan, 2004). The first claim says that mental terms are explanatory: they got inserted into our language to help us out at the business of predicting and explaining other people's behaviors.[1] The second claim is that these mental terms perform their explanatory and predictive role in virtue of being part of a theory, a *folk* theory, commonly known as *folk psychology*.

Folk psychology can be first approached analogously. Folk psychology is to scientific (organized, systematic) psychology as folk physics is to scientific (organized, systematic) physics. As we grow up and get to know how to go around in the world, we begin to develop an understanding of the structure of everyday objects, about the way in which they behave, how they react with each other or under different conditions, and so forth. In general folk physics works pretty well. Parental teachings instruct us to estimate with accuracy the trajectory of a baseball traveling with increased acceleration, and to

[1] Some might say that without mental terms we could not even make sense of our own behavior, and some may go as far as to claim that without them we could not make sense of the behavior of some large non-human animals. I do not have a view about these last two interpretations. For the success-to-truth argument to work, I just need the first, weaker interpretation of the 'theory'-theory.

then catch or flee accordingly. Less friendly classrooms have taught us to pick out tree branches apt to resist the stress produced by the gravitational force acting upon our well-fed seven-year-old bodies. Thanks to experience we accrue piles of physical folklore that help us in the business of explaining and predicting the behavior of good old middle-sized objects. Likewise, mutatis mutandis, when it comes to folk psychology. Repeated encounters with energetically voiced instructions teach us when it may be wise to cut it out and do as our mother wishes. And our occasional interactions with persons whose behaviors we deemed as questionable, rightly suggest that they follow some beliefs we do not share. Just as we live in a world packed with middle-sized objects we also live in a world populated with people. Folk psychology is the understanding we develop to make sense of people's complex behaviors.

It is customary to trace back the historical origins of folk psychology to Sellars' celebrated myth of Jones (Sellars, 1956/1963). Details aside, Sellars' fable purports to convey the idea that mental terms are theoretical terms inserted in our folk psychology to refer to inner, unobservable episodes of others' mental lives—episodes which, *allegedly*, are causally responsible for their overt and observable behavior. That way, whereas our Rylean ancestors' theoretical repertoire was limited to mere observational/dispositional expressions, Sellars tells us that "Jones develops a *theory* according to which overt utterances are but the culmination of a process which begins with certain inner episodes" (Sellars, 1956/1963, 186). These unobservable 'inner episodes' are to be taken as the referents of the theoretical mental terms Jones uses to explain the rich mental life unreachable to the behaviorist. To sum up: the Assumption says that folk psychology is a theory; that just like any other scientific theory it works in part by introducing theoretical terms; that our mental terms are those theoretical terms; and that, hypothetically, mental terms refer to inner episodes.

The first premise (P1) insists that folk psychology is a successful theory. This premise, in fact, is the Rubicon dividing eliminative materialists and intentional realists. On the one hand, intentional realists suspect that, as far as it goes, folk psychology works just fine. In general, predictions and explanations couched in mental terms seem to work, their generalizations seem to apply to novel cases, and their exceptions seem to be somewhat easily explainable away, either by theory itself, or by pointing at some violation of a ceteris paribus clause. On the other hand, eliminative materialists take folk psychology to be a complete failure, a stagnant science at most, with all sorts of predictive and explanatory shortcomings. Arguments in favor and against (P1) are, therefore, the main topic of the next section.

Finally, the second premise (P2) corresponds to what Philip Kitcher (2001, 177) calls "the success to truth inference". The motivation behind (P2) is the belief that if scientific success is systematic, nothing miraculous must be going on: scientific accomplishments must not to be cashed out in terms of repeated coincidences but—at least intuitively—in terms of truth. As we will see in section 5, most scientific realists take (P2) as one of the main arguments in favor of scientific realism. The idea, in brief, is that a scientific realist view is the only view that does not make the success of science look like a sheer collection of systematic miracles. Needless to say, if this was the only option, one would seem to face an unfortunate dilemma: either to embrace scientific realism or to accept the preposterous thesis that the success of science is pure luck (Votsis, 2004). In section 5, I will discuss an alternative view upon which to build my rejection of (P2) and my solution to the realism/anti-realism debate about propositional attitudes.

3. THE PERSISTENCE OF FOLK PSYCHOLOGY

If it isn't literally true that my wanting is causally responsible for my reaching, and my itching is causally responsible for my scratching, and my believing is causally responsible for my saying. ..if none of that is literally true, then practically everything I believe about anything is false and it's the end of the world.
Jerry Fodor (1990)

Eliminative materialism, according to Paul Churchland, "is the thesis that our common-sense conception of the psychological phenomena constitutes a radically false theory, a theory so fundamentally defective that both the principles and the ontology of that theory will eventually be displaced, rather than smoothly reduced, by completed neuroscience" (1981, 67). The force of this view, I contend, stems from the rejection of (P1). Notice, however, that Churchland needs (P2) to be stronger than the version I provided. He needs the implication in (P2) to be a bi-conditional. As it stands, it may very well be possible for folk psychology to be an unsuccessful theory and yet still be true. Success is a practical concept, not an ontological one. Scientists may (or may not) agree on the accuracy of a certain theory, and the theory could turn out to be true, but due to human practical limitations we may not be able to do anything with it[2]. To be sure, then, Churchland needs (P2) to read:

[2] It is not impossible to find an example of a theory having produced no successful predictions, not because of the falsity of its premises, but because scientists don't even know how to apply it in experimental or practical situations—i.e. how to test its observational consequences. Consider Schrodinger equation. Although it is sufficiently clear which mathematical outcomes could be expected from calculations involving it, the empirical interpretations to be correlated with such calculations are still unclear or impracticable. The American physicist John Cramer (e.g 1988), for instance, has suggested a novel interpretation of the nature of wave equations, such as Schrodinger's, according to which a mixture of real and imaginary numbers is required. The problem is that these complex variables—as the mixed numbers

(P2*) A theory is successful if and only if it is true

This way, if he can prove that folk psychology is actually an unsuccessful theory, its falsehood will be warranted—that is, C1 would be false. To that effect he cites "three major empirical failings of folk psychology" (Churchland and Churchland, 1998, 8 [but see also Churchland, 1981; 1988]):

(a) Folk psychology cannot explain a considerable variety of psychological phenomena, including mental illness, dreams, and concept acquisition by pre-linguistic children, amongst many others.

(b) Folk psychology has remained unaltered for the past 2500 years, showing no signs of development and many of stagnation.

(c) Folk psychology does not seem to be easily integrable with the other disciplines in its theoretical vicinity, like physics, chemistry, biology, and physiology.

The upshot, then, is that folk psychology is unsuccessful and as a consequence it should be deemed as false.

---

are often called—are written as +/- numbers, by virtue of which there are always two possible solutions. Alas, when used in equations involving the behavior of a system in time, the change in sign is supposed to be understood as "reversing" the direction of time, and that—as far as I understand—is still not quite easily interpretable in terms of empirical success. This impossibility, however, purports no harm to the acceptance of the equation as being true, and I suspect there may be similar examples in other areas of physics, perhaps even beyond quantum mechanics.

Despite the apparent appeal of these alleged empirical reasons, I think all of them can be contested. Let us begin with (a). I thought that the *main* moral we were supposed to draw from Sellars' myth of Jones was that mental terms were introduced in our folk psychology in order to help us explain the observable complex behavior of other people. More specifically, mental terms were supposed to contribute to the systematization of laws the purpose of which was to explain and predict the observable behavior of other persons. Now, Churchland considers that folk psychological explanations fail on two grounds: firstly (1) because their theoretical terms depict a "radically inadequate account of our internal activities" (Churchland 1981, 570), and secondly (2) because they prove ineffective when applied to a subset of psychological phenomena (e.g. mental illness, sleep, etc.). However, rejecting folk psychology on the grounds of (1) does not seem fair once one realizes that "our internal activities" was not its proprietary domain of evidence and explanation in the first place. When it comes to scientific explanations it is always important to keep the notion of success relative to the kind of objects over which its predictions and explanations are supposed to operate. And it seems clear that in the case of folk psychology these objects are persons. Mental states never got into our folk psychological language in order to stand in place of neural events. (How could this have been possible?—the idea that mental events are brain events is clearly newer than folk psychology itself). It is true that Jones *hypothesized* that theoretical-mental terms—perhaps because they *seem to be* referential terms—were supposed to refer to inner linguistic episodes. However, this consideration, as well as any other further considerations regarding the *nature* of such episodes, is going to be either gratuitous or dependent upon subsidiary hypothesis (e.g. that our inner mental life mirrors our overt linguistic life; that mental states are to be correlated with brain states; that there are not non-linguistic inner episodes causally responsible of overt utterances, etc). If you want to claim that inner episodes are brain events you may provide these subsidiary hypotheses. Nonetheless, for the purpose of the effectiveness of the myth, you need not to. For all Jones knows, dualism could be true (mental states could be carried out by some sort of immaterial soul whose operations may in no way be “inner”), the extended cognition hypothesis could be true, in fact, even people could be zombies, and yet folk psychology would still be vindicated. Why is that? Because the assumption of mental terms—that is, of theoretical terms—serves *primarily* the purpose of systematization: “it provides connections among observables in the form of laws containing theoretical terms” (Hempel, 1958/1965, 186). Theoretical terms in our laws are, as it were, operational shortcuts posited in place of a bunch of observational data, which are further used to infer observational conclusions there-from. They do not serve primarily a referential purpose. Therefore, as long as they serve *their* purpose within the laws, whether they fail to refer to our internal neural activities doesn’t really matter.

On the same token, to reject folk psychology on the grounds of (2) does not seem reasonable either. Suppose we agree that we have always used mental terms to make sense of people’s behaviors. Now: insofar as we have used mental terms in *this* way, psychological explanations and predictions are actually quite successful. In general we are pretty good at interpreting someone else’s needs, for instance, or her hopes, what to expect from her given what we know about her, or even what we don’t know, as when we hire lawyers to write down our contracts. Indeed, the success of folk psychology in everyday life is so ubiquitous that it is “practically invisible” (Fodor, 1985, 3). It is true that, at times, our explanations at the folk psychological level seem to fail. But there are failures and there are failures. Suppose I ask you to meet me tomorrow at school at three in the afternoon. Suppose further that you say “yes, I’ll be there”. From that piece of information I infer that you have formed the desire to meet me at school tomorrow and that you have formed the belief that I will be there at three in the afternoon. Then I put belief and desire together and I predict the following action: that you will go to school tomorrow at three in the afternoon for our meeting. The prediction fails, alas: you forgot the date. What did go wrong? Here one has (at least) two options: one can either blame the entire predictive apparatus (i.e. folk psychology) or one can simply argue that your obliviousness constitutes a violation to a tacit ceteris paribus clause. Blaming the entire apparatus of folk psychology on the basis of just one failure seems a bit exaggerated. For one, I can provide an explanation of the failure in terms of the very same theory: if you hadn’t *forgotten* the date, my prediction would have worked just fine. Secondly, it is true that similar extrapolations have proved successful in the past (last Wednesday—remember?—you did actually make it to our appointment). Finally, I can also be confident that the new inference I make—right after I talk to you, you apologized, you swore me that this time you’d be there on time, etc.—is actually going to work, ceteris paribus of course. Then again, maybe the problem is that you may not like ceteris paribus clauses at all. Fair enough. However, if that is so, your concerns can be generalized across the board, for they may actually affect most of our scientific theories (including neuroscience!), not only folk psychology (see, for instance, Lange, 2002).

Surely Churchland does not have *those* cases of failure in mind when he claims that folk psychology cannot accommodate certain phenomena. He has in mind *big* failures—like epilepsy (everyone's favorite example) and witchery. I have something to say about both. On the one hand, it seems to me that epilepsy is rather an exceptional disturbance whose behavioral characteristics are "less psychological" than the prototypical folk psychological phenomena. It is not only that epilepsy was not easily explainable by reference to folk psychology's ceteris paribus clauses; it is rather that it was a very odd behavior, like hiccups or somnambulism, and it just did not seem to be the product of typically hypostasized psychological states. Perhaps that was *precisely* the reason why people introduced demonic possessions in order to explain it: for not being part of the domain of characteristic behaviors folk psychology usually explained, a different discipline was required to do the job. It is true that theology failed to explain the phenomena and that now neuroscience can explain epilepsy alright. But it is not clear to me how this achievement of neuroscience is supposed to harm the success of a folk theory for which epilepsy was not clearly a proprietary explananda. For not being able to explain epilepsy in terms of demonic possessions psychology should not be blamed, but theology!

On the other hand, the idea of "modern theories of mental dysfunction" explaining away the phenomenon of witchery seems problematic. According to Churchland, our current "theories of mental dysfunction led to the elimination of witches from our serious ontology" (1984, 44): they show that instead of being possessed by demons these women were psychotic. I'm dubious as to how accurate this explanation is. To begin with, it is very unclear what he means by "modern theories of mental dysfunction". Does he mean psychoanalytic theories? Does he mean theories according to which a mental dysfunction is a behavioral dysfunction? Or maybe theories that conform to the standards of the bio-psycho-social model (like those underwriting the DSM III)? Perhaps he has in mind a biologically oriented theory, like the one underwriting the nosology of the DSM IV. The truth is that no modern theory of mental dysfunction knows for sure what psychosis *really* is, in part, because no modern theory of mental dysfunction knows what a mental dysfunction *really* is. Second, I'm tempted to think that the elimination of witches from our ontology had, actually, very little to do with advances in psychiatry and much more to do with advances in folk psychology. It is true that old witches had some features in common, and that it was largely on the basis of *these* features that they were ruled out of our current "serious ontology". However, these features were not so much neurological as they were social: so-called witches happened to be largely old rich widows living in terrains suitable to be expropriated by the state or by the church. We had to undergo serious changes in our *beliefs* about the powers of the church and the state, for instance, before we were able to rule witches out of our "serious ontology"—or, rather, before we came to understand how the word "witch" was being used. But if this is going to be the right kind of explanation (or at least part of the right kind of explanation), then I feel that the changes responsible for this alleged "elimination" are better explained from the point of view of folk psychology than from the point of view of neuroscience.

Similar points can be made regarding other cases of *big* failures Churchland mentions. Take dreams for instance. Dreams do not elicit typical overt behaviors. Very rarely people behave when they are dreaming. And when they do, very rarely their behavior is elicited by any inner episode they are aware of—or, at least, that they could causally respond to in virtue of their content. In that regard, then, dreams do not seem to be proprietary explananda of folk psychology. Therefore, insofar as they do not belong to the domain upon which folk psychological explanations were supposed to operate, it is unfounded to use them as counterexamples. A similar conclusion can be found in Horgan and Woodward (1985, 402) for whom “There is no good reason, a priori, to expect that a theory like [Folk Psychology], designed primarily to explain common human actions in terms of beliefs, desires, and the like, should also account for phenomena having to do with visual perception, sleep, or complicated muscular coordination” (Horgan and Woodward, 1985, 402).

What about (b)? There is a longstanding line of argumentation against the stagnation objection trying to show that, in reality, folk psychology has actually progressed in the past 2000 years. To that effect, philosophers and psychologists have shown that psychology, at the social and personal levels, make constant use of belief/desire talk in the process of pushing forward their research programs: “for instance, temperament seems to be more useful in predicting behavior than other sorts of personality traits, according to social psychology; short-term memory holds about seven “chunks” of information, whether these are numbers or names or grocery items, according to cognitive psychology; and so on” (Schroeder, 2006, 69). I think this line of argument is basically right, and it should be taken much more seriously by eliminative materialists. I’d just like to add one more ingredient to the mix: folk psychology not only proves necessary to the process of concocting research programs but, *more importantly*, to the process of carrying out those researches. It seems undeniable that true ascriptions of mental states are necessary when interpreting and producing neuroscientific data in situ, both inside and outside of the laboratories (from hospitals and asylums to urban places, as when neuropsychologists examine their patients at their work places or homes). Neuroscientists ought to believe that their subject's introspective reports are veridical no less than they should trust in the word of their co-workers. However, all these intersubjective data would be useless unless we had the network of folk psychology up and running. I take this to be an obvious point, and I won't argue further in its favor.

Still, there is another reason why one would be skeptical about the force of (b). 'Development' is a tricky word. In which sense does a theory develops? If developing counts as fostering research programs, then folk psychology is all set. On the other hand, if development means something like "refinement" of a theory's axioms and principles, then I agree: folk psychology hasn't showed that much of it. But then again this sort of "immobility" need not be a sign of unsuccessfulness. It may be a sign of proper functioning instead. If a theory constantly proves unsuccessful and does not undergo revisions and changes, I think it is right to accuse it of being a bad theory. But if a theory works just fine when it has to, why would we want it to change at all? Consider basic arithmetic. Nobody would reject basic arithmetic on the grounds that it has not undergone any significant changes in the last 2000 years. Basic arithmetic—the primary school arithmetic that most people operate with—hasn't changed because it works just fine for most everyday tasks. (Ways of *teaching* basic arithmetic may have changed in the past, but not new ways of getting 4 out of 2 + 2.) A similar point can be made about folk physics. People keep making the same rough generalizations and predictions about middle-sized mundane objects on the feeble basis of previous successful experiences; yet, so far as quotidian life goes, folk physics works alright and hasn't showed signs of severe alterations. And the same goes, mutatis mutandis, for folk psychology.[3]

Let me conclude with a comment about (c). To being with, it seems unclear what the objection amounts to. For the objection to be *really* an objection against the success of folk psychology the following claim should be true: that if a theory A is not integrable to a theory (or a set of theories) B, then A is unsuccessful. Call this claim *the integrablility condition*. The key word here, of course, is "integrable". I take it that by "integrable" Churchland means "reducible", and by "reducible" he basically means what he meant in his 1979 book, namely that a theory A is successfully reduced to a theory B so long as two conditions are met: (1) that we can provide a set of rules (so-called "bridge laws") according to which the terms in A are mapped onto terms of a subset of sentences in B, and (2) that the expressions in B which the terms of A were mapped onto are axioms of A. (Churchland, 1979, 81ff). That way, A will be "contained" in B, i.e. B will explain as much as A explains and more. However, sundry arguments in the philosophy of science should have convinced us by now that (1) is not the case for most—if not for all—(special) sciences, and that since (2) presupposes the success of (1), (2) may prove impractical as well[4]. Therefore, given the correct rendering of *the integrability condition*

[3] A different concern is to accuse folk physics of being unable to solve puzzles in the domain of scientific (organized, systematic) physics. This is also an unfair claim. Scientific physics deals with highly idealized objects and situations whereas folk physics has a more mundane domain and a very different purpose. I think it would be a mistake to reject folk physics on the basis that its generalizations don't coincide with the generalizations of scientific (organized, systematic) physics. The same, I think, goes for folk psychology.

[4] I have in mind the arguments as in Oppenheim/Putnam (1958) and Fodor's "Special sciences" (1974). For instance, the latter, very briefly, goes like this: a successful reduction of the psychological law like

(1) $S_1x \rightarrow S_2x$

is achieved as long as we can provide bridge laws of the form

(2a) S_1x *iff* P_1x and

(2b) S_2x *iff* P_2x,

guaranteeing the reduction of the psychological predicates S_1 and S_2 to neurophysiologic predicates P_1 and P_2 in a law of the form

(3) $P_1x \rightarrow P_2x$.

(if a theory A isn't *reducible* to another theory B, then A is unsuccessful), and given the arguments against the tenability of such reductions, the acceptance of *the integrability condition* required for the success of (c) would force us to reject as unsuccessful any theory that proves irreducible. Sadly, that would include basically all special sciences (not only psychology, but also economy, sociology, and so forth) and some lower-level sciences, like ecology, biology and perhaps neurology. To argue that none of these sciences is successful is preposterous. Irreducibility just cannot be the mark of scientific unsuccessfulness.

Churchland scholars may object at this point that I am being unfair to his theory. After all, Churchland soon realized that "the classical account of intertheoretic reduction appeared to be importantly mistaken", so he proceeded to perform some "necessary reparations". (Churchland, 1985/1989). Fair enough. I'm willing to assume, for the argument's sake, that his new account actually circumvents the previous theoretical obstacles alright. Still, there is another reason to be suspicious of the idea that reducibility speaks in favor of the success of a theory. If the success of a science is to be accounted for in terms of its explanatory and predictive achievements, then a successful reduction needs to impact negatively on the explanatory power of the reduced science. That is, it can't be the case that a reduced science can provide a better answer for a certain question

---

Alas, this sort of reduction is impracticable because bridge laws connecting type-psychological predicates with type-neurophysiologic predicates are, if not impossible, highly improbable ("an accident on a cosmic scale"). At most, all we can get are correlations between type-psychological predicates with heterogeneous disjunctions of type-neurophysiologic predicates like

(4) Sx *iff* P_1x or P_2x or … or P_nx

in which case the right side of the bi-conditional won't correspond to a natural-kind of neurophysiology. Ultimately, the reduced law that uses type-neurophysiologic predicates would look like

(5) P_1x or P_2x or … or $P_nx \rightarrow P'_1x$ or P'_2x or… or P'_nx

where P_i and P'_i are nomologically related. The problem, however, is that if the identity relation in the bridge laws (like 4) isn't between natural-kinds, then they aren't laws. But if they aren't laws then (5) isn't a law either. And when no laws, no reduction. QED.

than its reducing science. But this is hardly the case with folk psychology. Often times, the kind of explanations users of folk psychology require look for answers that aren't neurological. Sometimes we demand historical explanations, or accounts in terms of the environment in which the subject is embedded, or even contrastative answers couched in terms of reasons as opposed to causes, as when we wonder why he is doing X as opposed to Y. Reductive accounts may be able to provide us with full-fledged elaborations of the neural underpinnings of those behaviors, but it isn't obvious that an answer couched in neurological terms is going to be always, and for every possible purpose, explanatorily satisfactory. We frequently demand explanations in folk psychological terms, regardless of whether we have or not reductive accounts of the terms being used. I don't think it is clear at all that every why-question we may raise in folk psychological terms is suitable to be satisfactorily answered in neurological terms.

I think that part of the problem has to do with a fact I mentioned earlier: that the explanatory power of a theory is relative to its domain of evidence. Folk psychology's domain of evidence is people's behaviors. Brains may be entirely causally responsible for those behaviors. However, our epistemic limitations make it really hard to successfully explain or predict these phenomena in terms of the operations of an underlying structure we do not have epistemic access to, namely people's brains. And when it comes to evidence our epistemic limitations dictate our possibilities (I will come back to this problem later on in section 5). So, I think it is safe to conclude that issues about irreducibility seem to be basically orthogonal to preoccupations about the theory's success.

4. THERE MAY NOT BE BELIEFS AFTER ALL

Folk psychology may not be playing the same game as scientific psychology,
despite its deliberately provocative and misleading label.
Andy Clark (1989)

If you have been convinced by the considerations in the previous section, then you probably think that the eliminative materialist does not have sound reasons to show that folk psychology is unsuccessful. In addition, if you consider that the success to truth argument introduced in section 2 is a valid argument then you probably think that folk psychology is true. None of the above, however, gives you intentional realism yet. To get to it we still need one further—and a bit more complex—argument. Call it *the truth-to-existence-via-reference argument* and it goes like this:

(P1) Folk psychology is true.

(P2) The statements of folk psychology report propositional attitudes.

(P3) Propositional attitudes are two-place relations between subjects and the referents of 'that'-clauses.

(P4) All things considered, the best candidates we have for referents of 'that'-clauses are mental representations in the language of thought. Therefore,

(C2) There are mental representations in the language of thought.

Once again, each premise needs some clarification. (P1) is the conclusion of *the success-to-truth argument* (i.e. (P1) = (C1)). (P2) is basically a traditional tenet that can be traced back at least to Russell's 1918 lectures on logical atomism (reprinted in Russell, 1985). According to this claim, mental states are to be characterized as ascribing to a subject *S* an intentional verb *Vs* (such as 'believes', 'fears', 'hopes", etc.) and a certain proposition *p*. Propositional attitude reports, thus, conform to the following general form: '*S V*s that *p*', examples of which are "John hopes that it is raining", "Anne believes that having a small wedding is fine" and "Mario cree que el tiempo en Nueva York se siente distinto". Because propositional attitude reports conform to this general form, it is believed that propositional attitudes are better understood as two-place relations between a subject *S* and a proposition *p* which is the referent of the 'that'-clause. Such is the rationale behind (P3). Now, in support of (P3) intentional realists give three reasons[5] (1978/1981; 178-179):

(a) "It is intuitively plausible. 'Believes' looks like a two-place relation, and it would be nice if our theory of belief permitted us to save appearances".

(b) "Existential Generalization applies to the syntactic objects of verbs of propositional attitudes; from 'John believes it's raining' we can infer 'John believes something' and 'there is something that John believes'."

[5] As I said at the beginning, I'm confining my notion of intentional realism to Fodorian sentential realism. Because of that, the arguments in favor of (P3) and (P4) are his. Alternative accounts supporting (P3) and (P4) are not going to be considered. It may be possible that my arguments apply to them as well, but they need not.

(c) “The only known alternative to the view that verbs of propositional attitudes express relations is that they are (semantically) “fused” with their objects, and that view would seem to be hopeless.”

The force of all these reasons comes from linguistic and philosophical analysis of propositional attitude talk. The assumptions that support them will be soon discussed, when I present my arguments against (a), (b) and (c). Finally, (P4) is basically an inference to the best explanation. The suggestion is that once you take into account all the data a theory of propositional attitudes is supposed to account for, the best candidate we end up with is a theory according to which “propositional attitudes are relations between organisms and formulae in an internal language; between organisms an internal sentences, as it were” (Fodor, 1978/1981; 187). I think this inference to the best explanation can be blocked as well. Let us move on, then, to my criticisms.

The first thing one could challenge is the claim, conveyed by (P2)—and, to a certain extent, (a)—that mental states can (and need) be characterized as embedded within ‘that’-clauses. It has been pointed out (e.g. Ben-Yami, 1997) that some bona fide sentences purporting to report mental states cannot be rendered into the canonical form of propositional attitude reports. Consider the following sentences (examples 1 and 3, from Ben-Yami, 1997, 85):

1. I want to sleep
2. Andrew knows how to multiply six digit numbers mentally
3. I trust John

I you like the idea of having your mental states being characterized in the fashion of a propositional attitude report, you may want to offer alternative paraphrases for these sentences. Maybe you would suggest something like:

1*. I desire that I am asleep

2*. Andrew knows that to multiply six digit numbers mentally one needs to ϕ.

3*. I believe that John is trustworthy

But notice that these forced paraphrases introduce several problems. 1*, for instance, sounds odd to my ears. Someone may argue that this is only a problem for English. A quick look at the same proposition in French and Spanish, for instance, dissuades us from that option.[6] On the other hand, it may be argued that in order to get the correct paraphrasing some extra linguistic maneuvering may be required, not at the surface level, but at the level of their deep structure (viz., ‘that’-clause in 1 involves an implicit subject). Perhaps that could solve the problem for these cases, but if so one would like to know why is it the case that we want to force all our mental state reports to fit a certain kind of structure. I know of no argument to that effect (neither does Ben-Yami, 1997, 85). In absence of such an argument it is hard not to conclude that the theory may be forcing the maneuver.

A somewhat related worry can be raised regarding 2*. I take it that all 2 tells us it that within Andrew’s abilities we can count that of multiplying six digit numbers

[6] Contrast 1 with its Spanish translation “Quiero dormir” and its odd rendering into a canonical form: “Quiero que yo esté dormido”. Likewise for French: “Je veux dormir” versus “Je veux que je sois endormi”.

mentally. 2*, however, seems to imply that if one were to ask Andrew how to multiply six digit numbers mentally he would be able to gives us an answer in terms of ϕ. But 2* could be false while 2 be true. After all, Andrew may not know how is it that he manages to multiply six digit numbers in his mind. He just knows that he can do it, but he may not know how or why he can do it. (Notice that this is *not* a problem of expressibility. It isn't that Andrew does not know how to put into words what he does; it is rather that he may have no idea how he does it—he may not even know how to *begin* explaining what he does). And, finally, the same worry goes for 3*. All 3 tells us is that I trust in John. It says nothing as to whether I believe that John is trustworthy or not. I could be a stubborn idiot who still trusts in John despite the fact that I am seriously suspicious about his trustworthiness. Finally, I think that these considerations also speak against the first reason Fodor offers in support of (P3). If not all mental states' attributions are suitable to be translated into statements of the canonical form, those that can may at most constitute a subset of folk psychological statements. That all folk psychological statements are better seen as two-place relations does not seem, therefore, as intuitively plausible as Fodor suggests.

But for the sake of the argument, let's assume for a moment that it is, in fact, intuitively plausible to render our attribution of mental states in the canonical propositional attitude form. That is, suppose we accept that mental states can be paraphrased without semantic loss as expressing a two-place relation between subjects and the referent of 'that'-clauses—whether propositions in abstracta or, as in the case of Fodor, presumably neural concreta. Does that constitute enough reason to believe that the referents of 'that'-clauses are real? The answer is *no*. More assumptions need to get accepted for that conclusion to follow. Fodor gives us two reasons in support of (b): first, that ‘that’-clauses are indeed referential, and second, that existential generalization applies to ‘that’-clauses. Now: why is it the case that these two reasons constitute a good argument in support of there being referents of ‘that’-clauses? It seems to me (although I’m not alone; see e.g. Balaguer, 1998) that what underwrites this claim is basically Quine’s criterion of ontological commitment plus an “intentional” reading of the Quine-Putnam indispensability thesis. Let me elaborate by comparing the case at hand with that of mathematics. Due to the influence of the Quine-Putnam indispensability thesis[7] in mathematics, theoretical irreducibility (and non-eliminability) is often assumed to carry with it a heavy ontological baggage. For it is frequently accepted that if *S* is irreducible to *R* ($=_{df}$ untranslatable to the other via bridge laws [see footnote 4]) and, when regimented, both S_r and R_r turn out to quantify over different variables[8], then one is *eo ipso* committed to the existence of those entities (or kind of entities) picked up by the bounded variables. In the case of mathematics such is the case with numbers (sets). I contend that for (b) to count as ontologically significant, the same should go for propositional attitudes.

I think this argumentative line could be blocked with two moves. The first move is to show that ‘that’-clauses do not behave referentially. The second move is to show that although existential generalization applies to ‘that’-clauses, such a quantificational device can be read as being ontologically innocent, i.e. as conveying no ontological

[7] “The claim, roughly, that if one’s best scientific (physical) theory [after regimentation onto first-order logic] requires existential quantification over certain entities, then one is ontologically committed to such entities” (Azzouni 1998, 1).

[8] “Turn out” is short for: Take *Px* to a formula with a free variable *x*, and take $\exists(x)(Px)$ to be directly deducible from S_r but not from R_r. Given Quine’s criterion for ontological commitment, one is here committed to the existence of *Px*. Now: take $\exists(x)(Qx)$ to be deducible from R_r but not from S_r. I take that if the criterion is correct, then it “turns out” that one is committed also to *Qx*. (All under the assumption that one can have regimented versions of both *S* and *R*, my S_r and R_r).

commitments by itself. Thankfully recent developments in linguistics and metaphysics show us that both moves not only are available but also make sense.

Let us begin with the first move. In general, objections against the non-referentiality of ‘that’-clauses have been directed toward theories holding that the referent of ‘that’-clauses are propositions. I believe that the force of at least two of these objections carry over to Fodor’s analysis of propositional attitudes as being relational. The first of these objections in known as *the substitution failure*. Briefly stated the substitution failure objection says that if ‘that’-clauses were really referential, and if their referents were really propositions, then they should share their denotations with linguistic constructions of the sort “the proposition that *p*” (Moltmann, 2003, 82ff). However, this sort of substitution often fails. Consider the following substitution case:

4. John fears that Obama will be our next president.
5. John fears the proposition that Obama will be our next president.

Ex hipotesi, “that Obama will be our next president” and “the proposition that Obama will be our next president” share their reference: namely, the proposition that says that Obama will be our next president. But to be afraid of the eventual situation of Obama being the next president is different from fearing a proposition. It seems obvious that 4 and 5 differ in truth value, so we should better conclude that ‘that’-clauses do not refer to propositions (see also Hofweber, forthcoming). Now, does this concern carry over when we aren’t talking about abstracta but neural concreta (i.e. sentences in the language of thought)? What would happen if, instead of 5, we were to have

6. John fears the mental sentence that Obama will be our next president.

Would it change the outcome of the substitution failure objection? I do not think so, at least insofar as the substitution failure objection counts as an argument *against* the relational analysis of propositional attitude reports. In order for (b) to count as a linguistically valid reason in favor of 'that'-clauses being referential, Fodor needs that whatever goes for propositions goes as well for mental formulae. To argue in favor of the latter as opposed to the former on the basis of some property that one but not the other has, is not permitted at this stage. Remember that Fodor wants 'that'-clauses to be referential so he can argue, a priori, that there *must be* referents of 'that'-clauses. Using an alleged property about their nature to justify the argument in favor of their existence is a circular maneuver.

The second objection I have in mind against 'that'-clauses being referential is originally due to Kripke (1979), although more recently has been developed by Bach (1997). The relational analysis of propositional attitudes finds support partly because it seems to reflect the apparent logical form of inferences like:

I1: A believes that *p*

B believes that *p*

→ There is something that A and B both believe.

However, when Kripke introduced his Paderewski-case puzzle he showed us that inferences of the form I1 aren't always valid. Suppose Carl met Paderewski at a business meeting and as a result fixed the belief that Paderewski is a nice guy. Carl is pretty bad with faces, though. Later on he comes across Paderewski at a cocktail party where he strikes him as an annoying guy. As a result he forms the belief that Paderewski is not a nice guy. If the relational account of propositional attitude reports is correct, it seems as though Carl believes contradictory things. Specifically,

> I2: Carl believes that Paderewski is a nice guy.
>
> Carl disbelieves that Paderewski is a nice guy.
>
> → There is something that Carl both believes and disbelieves.

But Carl isn't being irrational; he's just ignorant about the fact that he's taking the name "Paderewski" to refer to two distinct individuals. Notice, however, that this fact is inessential to the problem. As Bach notes, when it comes to the relational analysis of propositional attitude reports, the believer need not have "any familiarity with the name in question or have any name at all for the object of belief" (Bach, 1997, 224). Consequently, it seems that the two premises in I2 have Carl believing and disbelieving different things. If so, then I2 is not a valid inference. But given the fact that there aren't relevant formal differences between I1 and I2, we have no reason to believe that the linguistic appearances in I1 aren't misleading as well. To solve the puzzle Bach suggests that we reject an essential ingredient of the relational analysis of propositional attitude ascriptions: the assumption "that the 'that'-clause in a belief report specifies the thing that the believer must believe if the belief report is to be true" (Bach, 1997, 221). In his account, 'that'-clauses *describe* the content instead. Without this assumption, however, we have very little reason to take 'that'-clauses as referential.

Of course Fodor can reject Bach's solution and stick to a relational analysis under the assumption that 'that'-clauses refer to mental sentences which, unlike propositions, are neither ambiguous nor semantically incomplete. But then again, this would be an unjustified move. Remember that (b)—and for that matter (P3)—was supposed to convey pre-theoretical reasons in favor of 'that'-clauses being referential. Resorting to such alleged properties of hypothesized mental sentences in order to save the linguistic phenomena whose clarity was supposed to motivate the relational analysis in the first place, looks, at least to me, rather circular. (For that matter, if we are to allow the resources of a theory in order to explain this phenomenon, a connectionist approach sensitive to graceful degradation and assignation by omission may turn out to do a better job than the language of thought when it comes to explaining why Carl forgot Paderewski's face to begin with.)

There is, however, a second—and, I think, more powerful—reason to reject (b). Even if one accepts that 'that'-clauses are referential, the only reason Fodor seems to offer to jump from that linguistic fact to the conclusion that their referents exist is a commitment to an ontologically loaded reading of existential generalization. Since belief reports admit of existential generalization ranging over their 'that'-clauses (e.g., the example in I1), and since 'that'-clauses admit no reduction to another language whose ontological commitments we could be more comfortable with ("Behaviorists used to think such translations might be forthcoming, but they were wrong" [Fodor, 1978]; see also footnote 4), then we *should* go ahead, as Quine taught us, and accept the referents of 'that'-clauses as real (Quine, 1948; see also Fodor, 1987, 15).

But why would Fodor wants us to do this? I take it that he *cannot* be suggesting this move on the basis of his acceptance of Quine's theory of reference; after all, Fodor is well know for his rejection of Quine's holism tout court. A more plausible answer is that he is doing so on the basis of a weaker assumption: that the best—if not the only—way to understand existential generalization is by treating it as ranging over domain-independent entities. But this is a contentious claim. One can instead adopt what Hofweber calls "an internalist view" about quantification and deem existential generalization as a logical device to increase expressive power, a logical tool that allows us to talk about infinitary disjunctions of single instances (Hofweber, 2006)—in this case, infinitary disjunctions of instances of attributions of mental states. If so, then, existential generalizations would be ontologically innocent. The internalist view of existential generalization could turn out to be wrong, of course, but it should be noted that is a *good* contender. And without an argument against it—or without an argument in favor of a domain-independent reading of quantification—we should better remain agnostic as to whether we should take existential generalizations as unquestioned carriers of the ontological burden of our regimented theories. As Jody Azzouni pointed out—in a rather different context—without an independent argument of that sort, it seems that the only reason we have to take the ordinary phrase "there is/are" to commit us to the existence of whatever it seems to commits us to, is simply "that the ordinary language 'there is' *already* carries ontological weight" (Azzouni 1998, 4). Does Fodor have an independent argument in favor of his realism about propositional attitudes? He sure does—that's the bulk of the proof for (P4).

Before we gear toward that discussion, however, let me say something very briefly about reason (c) for (P3). In light of the previous considerations, it may be clear that the force of (c) has now diminished. Fodor's original rejection to the "fusion" theory was supposed to mobilize the intuition that *unlike* that theory, a relational account of propositional attitudes faced no problems. But we have seen that relational accounts face severe objections too. Indeed, contemporary attempts to explain away precisely those objections seem to favor instead non-relational accounts of propositional attitude reports (see, e.g., Moltmann, 2003, for a neo-Russellian account, as well as the appendix of that paper for other non-relational alternatives). The fusion theory may not be true, after all. Still, that alone gives us no reason to prefer the also problematic relational account.

So what about (P4)? To tell the truth, Fodor can accept all these objections and reject (P3), and still argue in favor of his intentional realism on the grounds of (P4) alone; he may say that, *all things considered*, intentional realism constitutes the best *empirical* theory we have to "vindicate"—as he says—folk psychology. That is, he may well accept that we do not have either linguistic or a priori metaphysical reasons to accept the reality of sentence-like mental states, and still hold that such a hypothesis needs to be accepted on empirical grounds. This, at the end of the day, has been his preferred strategy. Sheltered by the motto "the only game in town", the hypothesis of the language of thought has been advertised as the best theory we can muster to explain several psychological phenomena. Niceties aside, his argument boils down to an inference to the best explanation for some puzzling phenomena: concept acquisition, the compositional, systematic, and productive character of our thought, the projectability of mental terms in our psychological laws, and some (but not very many!) more. Copious pages have been written in an attempt to provide alternative accounts of these phenomena in terms that do not force us to accept a language of thought (see, for instance, Jackendoff, 1992, Millikan, 1984, Prinz, 2002, just to name a few). I'm afraid I will not contribute to the discussion. Instead, I am going to try a different tack.

If Fodor's argument for the truth of intentional realism boils down to an inference to the best explanation, then—in contend—it better be the case that an inference to the best explanation constitutes a *good* reasoning pattern for realism about theoretical or unobservable entities. Folk psychology, after all, is just another scientific theory, less refined if you want, and operational over a slightly different domain than scientific psychology, but a scientific theory none the less. Recall that folk psychology's mental terms are theoretical expressions whose alleged referents are unobservable inner episodes, i.e. mental states. Now, scientific realists usually take inferences to the best explanation as good argumentative patterns in favor of the truth of a certain theoretical hypothesis. In brief, the rationale behind the inference to the best explanation is that if a certain hypothesis H explains a certain phenomena X better than any of its rival hypothesis, then H's explanatory superiority should be taken as a mark of its truth—or, at least, as a mark of its approximate truth. From there, however, scientific realists often jump to the conclusion that the unobservable entities postulated by the theory must be real. Fodor, as we have seen, is no exception here. He takes the hypothesis of the language of thought to be the best hypothesis we have to account for the aforementioned psychological phenomena, and then goes on to claim that such is enough reason to believe that it is true that there are sentence-like representations in our brains.

Notwithstanding the widespread use of inferences to the best explanation by scientific realists, its validity as an argument to support the truth of a scientific hypothesis has been challenged on several grounds. Perhaps the most common attack comes from scientific anti-realism. To begin with, scientific anti-realists—like Bas van Fraassen (1980) and Nancy Cartwright (1983)—have argued that being a good hypothesis is never enough ground for believing that it is true. After all, the set of all rival hypotheses we can choose from may contain only false ones. Moreover, as van Fraassen remarked (1980, 21ff), when a scientist is in the business of accounting for some observational evidence, she does not really choose the best possible explanation *there is*, but rather the best explanation that is available to her. However, it would be a mistake to infer from that fact that such a hypothesis must be true, or closer to the truth than any other hypothesis she may or may not have access to.

Furthermore, van Fraassen also noted that most scientific realists take the thesis of scientific realism *itself* as an inference to the best explanation, insofar as it is the best hypothesis we can muster in order to explain the success of science (see also Fine, 1984). According to them, the success of a theory mustn't be cashed out in terms of sheer luck. Scientific realism is the best hypothesis we can muster to reject that preposterous conclusion. Now: the circularity of the maneuver isn't worrisome, yet it opens the door for a rival hypothesis to scientific realism, namely that "we are always willing to believe that the theory that best explains the evidence, is empirically adequate (that all the observable phenomena are as the theory says they are)" (van Fraassen, 1980, 20). This anti-realist alternative to scientific realism—known as *constructive empiricism*—basically tells us that if a theory is successful then it is empirically adequate, and that a theory is empirically adequate "exactly if what it says about the observable things and events in this world, is true—exactly if it 'saves the phenomena'." (van Fraassen, 1980, 12). Constructive empiricism may be false, of course, and scientific realism may be vindicated through a different path. Still, the point I'd like to highlight here is it that on the basis of the empirical evidence alone, the hypothesis of constructive empiricism cannot be ruled out.

I think my suggested tactic to reject (P4) may be obvious now: if Fodor's argument for intentional realism boils down to no more than an inference to the best explanation, and if inferences to the best explanation aren't conclusive reasons to believe in the reality of postulated entities, then (P4) does not constitute a conclusive reason to infer the existence of mental formulae coded in our brains. Perhaps now an anti-realist perspective (which need not be a constructive empiricist perspective [but see below]) about propositional attitudes must look very appealing to the philosopher of mind who does not want to settle the ontological issue of the existence of mental formulae on the sole basis of an inference to the best explanation. With the previous arguments against (P2) and (P3) I tried to show that the jump from truth to existence via reference was pending on the viability of inferences to the best explanations as valid arguments for the existence of unobservable entities. Van Fraassen's arguments show us that this need not be the case. Even if *all things considered* the language of thought turns out to be the best hypothesis we have to explain some behavioral (i.e. observational) phenomena, to jump from here to the conclusion that there *are* mental formulae imprinted in the brain would be, if anything, a leap of faith. Notice that I'm *not* saying that the hypothesis of the language of thought is false. All I'm saying is that the *truth-to-existence-via-reference* *argument* does not provide us enough reasons to jump from the truth of our ascriptions of propositional attitudes to the reality of mental formulae in our brains. The idea that there are mental formulae *may* be true, but so far we do not have enough *evidence* to embrace that conclusion. And for that reason alone, we should not believe in them. Now: what would constitute enough evidence? That, precisely, is the question I will tackle in the next and last section of this essay.

5. TOWARD AN EMPIRICAL ACCOUNT OF PROPOSITIONAL ATTITUDES

Establishing the truth of a theory—even an empirically adequate theory—is one thing; establishing that what the noun phrases in the theory refer to are existents is quite another.
Jody Azzouni (2004)

Let me recap quickly what happened so far. In section 2 I introduced the *success-to-truth argument* and suggested that both eliminative materialism and intentional realism spawned from different takes on it. In section 3 I argued against Churchland's reasons to consider folk psychology unsuccessful. Finally, in section 4, I presented some objections against the *truth-to-existence-via-reference argument* in order to prove it insufficient to support intentional realism. At the end of last section I suggested a shift from an ontological inquiry to an epistemic one: rather than focusing on the ontological quandary about the reality of mental sentences, we should better get ourselves in the task of finding out what would constitute good *evidence* to justify our beliefs in them. A brief and quite programmatic answer to that question will be suggested in this last section.

Back in the days of the mythical Jones our Rylean ancestors were Positivists as well. They believed in a difference between observational terms and theoretical terms, with the former finding meaning in empirical evidence, and the latter finding meaning only in connection to observational terms. This dichotomy, however, was soon rejected and the idea that our observational terms are theory-laden has prevailed, more or less, ever since. Van Fraassen, however, has taken up the empiricist legacy of the positivists by drawing yet another distinction between theoretical and observational terms, on the one hand, and observable and unobservable entities, on the other. With this distinction at hand he can accept the theory-ladeness of observational terms while keeping a division between entities we can observe and those we cannot. This way, observation becomes once again the warrantor of empirical evidence. Accordingly, his constructive empiricism suggests that we are justified in believing only what our theory says about the observable stuff. Truth is only predicable of those statements pertaining to observational entities. Of the reality of unobservable entities, supposedly named by our theoretical terms, we shall rather remain silently agnostic.

I want to suggest in turn *a similar* strategy for folk psychology: we shall only take as real those parts of the theory we have empirical evidence for. However, unlike van Fraassen, I do not want to take naked-eye observation as our unique criterion of empirical evidence. Ever since its origins, constructive empiricism has been harshly criticized for endorsing a rather chauvinistic standard of empirical evidence dependent on the also chauvinistic and arbitrary notion of 'observable-to-us' (see Churchland and Hooker, 1985). I am sympathetic to these sorts of criticisms. So, instead, I want to take as criterion for empirical evidence what Jody Azzouni calls *thick epistemic access*.

Although for Azzouni observation does constitute a good criterion for fixing our beliefs about what is real, he thinks that it is a mistake to take observation as primitive. After all, one could always wonder why observation, in particular, constitutes such a great epistemic warrantor. His answer is straightforward. He thinks that our epistemic confidence in observation has to do with the fact that it meets what he calls *the tracking requirement:* that "the epistemic processes which establish truths that we're committed to, must be sensitive to the objects about which we're establishing those truths" (Azzouni, 2004, 372). And he thinks that observation meets the tracking requirement mainly because it comprises four "neat [epistemic] properties" (Azzouni 2004, 383):

> (1) Robustness: what we do with observation is largely independent of the theory.
>
> (2) Refinement: we have theory-independent ways of adjusting/improving our observations.
>
> (3) Tracking: what is observed can be monitored spatiotemporally.
>
> (4) Explanatory import: the properties of the objects that we see can be used to explain why we see them the way we do.

His suggestion is that if what makes observation so epistemically valuable is the fact that it meets these four requirements, then it may be possible that other epistemic procedures may meet those four requirements as well. Any procedure that meets these requirements is said to provide us with *thick epistemic access* to an object, and thick epistemic access constitutes our best method to form beliefs about what is real. Observation is one of many processes that provide us with thick epistemic access to objects. Some instrumental interventions may do so too. My proposal, consequently, is to take Azzouni's notion of *thick epistemic access* as criterion for empirical evidence.

At the beginning I promised to provide an alternative interpretation to the second premise of the success-to-truth argument. It is now time to pay my debts. I suggest interpreting (P2) as saying only that if a theory is successful then it is empirically accurate. As a result, our original argument should read now:

(P1) Folk psychology is a successful theory

(P2) If a theory is successful, then it is empirically accurate. Therefore,

(C1) Folk psychology is empirically accurate.

And by empirically accurate I mean just what van Fraassen means by empirically adequate—that what our theory tells us about the observable world is true—but with a twist: *that what our theory tells us about the things we have thick epistemic access to is true.* With this interpretation, of course, I'm *not* denying intentional realism. However, I'm shifting the burden of proof to the intentional realists by demanding her to tell us an epistemic story regarding why we need to take sentence-like mental representations as real entities. Are there any epistemic procedures, able to meet the tracking requirement, which can give us thick epistemic access to this kind of mental representations? (Or, if they aren't, can we craft a good account in terms of the mental representation's nature as to why this is not the case?) I know I am leaving here the door open for a realist to come with her own epistemic story so as to undermine my line of attack—that's fine. I just hope I gave enough reasons to demand such a story.

REFERENCES

Azzouni, J. 1998. "On 'On what there is'". *Pacific Philosophical Quarterly.* 79: 1-18.

Azzouni, J. 2004. "Theory, Observation and Scientific Realism". *British Journal for the Philosophy of Science*. 55: 371-392.

Bach, K. 1997. ''Do Belief Reports Report Beliefs?'' *Pacific Philosophical Quarterly*. 78: 215–41.

Ben-Yami, H. 1997. "Against Characterizing Mental States as Propositional Attitudes" *The Philosophical Quarterly.* 47 (186), 84–89.

Balaguer, M. 1998. "Attitudes Without Propositions". *Philosophy and Phenomenological Research*. 58 (4), 805-826.

Cartwright, N. 1983. *How the Laws of Physics Lie*. Oxford: Clarendon Press.

Churchland, P.M. 1979. *Scientific Realism and The Plasticity of Mind*. Cambridge: Cambridge University Press.

Churchland, P.M. 1981. "Eliminative Materialism and the Propositional Attitudes". *Journal of Philosophy*. 78 (2). Reprinted in: Churchland, 1992.

Churchland, P.M. 1985. "Reduction, Qualia, and Direct Introspection of Brain States". *Journal of Philosophy*. 82 (1). Reprinted in: Churchland, 1992.

Churchland, P.M. 1988. Matter and Consciousness. Cambridge, MA: MIT Press.

Churchland, P.M. 1992. *A Neurocomputational Perspective.* Cambridge, MA: MIT Press.

Churchland, P.M. and Hooker, C.A. 1985. *Images of Science: Essays on Realism and Empiricism*. Chicago: University of Chicago Press.

Clark, A. 1989. *Microcognition*. Cambridge, MA: MIT Press.

Cramer, J. G. 1988. "An Overview of the Transactional Interpretation". *International Journal of Theoretical Physics.* 27, 227.

Fine, A. 1984. "The Natural Ontological Attitude". In: *Scientific Realism*. J. Leplin (ed.). Berkeley: University of California Press.

Fodor, J. 1974. "Special Sciences". Reprinted in: Fodor, 1981.

Fodor, J. 1978. "Propositional Attitudes". Reprinted in: Fodor, 1981.

Fodor, J. 1981. *Representations*. Cambridge, MA: MIT Press

Fodor, J. 1985. "Fodor's Guide to Mental Representation". *Mind*. Spring: 66-97.

Fodor, J. 1987. *Psychosemantics*. Cambridge, MA: MIT Press

Fodor, J. 1990. *A Theory of Content and Other Essays.* Cambridge, MA: MIT Press.

Hacking, I. 1983. *Representing and Intervening*. Cambridge: Cambridge University Press.

Hempel, C.G. 1958. “The Theoretician’s Dilemma: A Study in the Logic of Theory Construction”. Reprinted in: *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* (1965) Free Press.

Hofweber, T. (Forthcoming). “Schiffer’s New Theory of Propositions”. *Philosophy and Phenomenological Research.*

Hofweber, T. 2006. “Inexpressible Properties and Propositions”. In: *Oxford Studies in Metaphysics*. V. 2. D. Zimmerman (ed.) Oxford: Oxford University Press.

Horgan, T. and Woodward, J. 1985. “Folk Psychology is Here to Stay”. *The Philosophical Review*. 44 (2), 197-226.

Jackendoff, R.1992. *Languages of the Mind*. Cambridge, MA: MIT Press.

Kitcher, P. 2001. “Real Realism: The Galilean Strategy”. *Philosophical Review.* 110, 2: 151-197.

Kripke, S. 1979. “A Puzzle about Belief.” In: *Meaning and Use.* Margalit, A. (ed.) Dordrecht: D. Riedel.

Lange, M. 2002. “Who's Afraid of *Ceteris-Paribus* Laws? Or: How I Learned to Stop Worrying and Love Them” *Erkenntnis.* 57 (3), 407-423.

Lycan, W. 2004. “Eliminativism.” (Unpublished) Available at: http://www.unc.edu/%7Eujanel/3255H5.htm

Millikan, R.G. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.

Moltmann, F. 2003 “Propositional Attitudes Without Propositions”. *Synthese*. 135 (1), 77-118

Oppenheim. P. and Putnam, H. 1958. “The Unity of Science as a Working Hypothesis”. *Minnesota Studies in the Philosophy of Science*. Vol. II.

Prinz, J. 2002. *Furnishing the Mind.* Cambridge, MA: MIT Press.

Russell, B. 1918. *The Philosophy of Logical Atomism*. Reprinted in Pears, D., 1985, Chicago: Open Court.

Schroeder, T. 2006. “Propositional Attitudes” *Philosophy Compass. 1 (1), 56-73.*

Sellars, W. 1956. “Empiricism and the philosophy of mind”. Reprinted in: *Science, Perception and Reality*. (1963) NY: Routledge & Kegan Paul Ltd.

Quine, W.V.O. 1948. “On what there is”. *Review of Metaphysics.*

Van Fraassen, B. 1980. *The Scientific Image*. Oxford: Oxford University Press.

Votsis, I. 2004. *The Epistemological Status of Scientific Theories: An Investigation of the Structural Realist Account*. PhD Dissertation. UK: London School of Economics.