

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI[®]

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Bodies, Minds and Materialism

Noah Stewart Beggs

A Thesis

in

The Department

of

Philosophy

**Presented in Partial Fulfilment of the Requirements
For the Degree of Master of Arts at
Concordia University
Montreal, Quebec, Canada**

July 1998

©Noah Stewart Beggs, 1998



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-39433-6

ABSTRACT

Bodies, Minds and Materialism

Noah Beggs

It is common to describe subjective experience as a mental, and thus metaphysical, phenomenon. This core assumption is a central tenet of our intellectual history. How this bears on our ability to take seriously new theories of mind in light of progressive discovery is what is at issue in this paper. Herein, it is argued that because the idiom we use to speak about the mind is predisposed to countenance ontological dualism, it begs the question against alternative conceptions of mind. I will thus argue that subjectivity need not entail non-physical phenomena, pace the assumptions of the traditional dualist idiom.

ACKNOWLEDGEMENTS

First, and Foremost, I would like to thank my parents: Thank-You!: Anna, Danica, Hagan, Terry, for the unconditional support which has been abundant, and manifest in every way I can imagine. My Grandparents, Sadie and Stewart Beggs, also, deserve appreciation for their kind philanthropy.

In other relations, I was kept alive by Stephanie Jones and her sweet lovin' goodness. And, I would like to say to John Lee, the `amazing canoe-walker'; C. Greenwood; J. Katz; B. Bass; the espresso merchants of the Concordia quartier; my Muse; M.J.B.; and whomever brings glad tidings to the Philosophy Department of Concordia University to: "Bring it on!"

Finally, I would like to thank my academic adviser and friend, Dr. Murray Cameron Clarke--salutations!--whose extensive knowledge and guiding wisdom have enabled me to see things clearly, or, at least from a perspective that makes the armchair of Archimedes appear to be the back-seat to any Truly Good Philosophy.

FOR MY FOLKS

Contents

Chapter One:

SHAKING OFF THE IVY

1. Introduction	1
2. Duelling with Dualism	3
3. A La Descartes	6
4. ...Au Jour	7
5. Problems with Materialism	8
6. Revolving the Canon	10
7. So Why Be Objective?	11
8. The Glad Scientist	12
9. Revision in Stride	15
10. En Passant	15

Chapter Two:

REDUCTION AND THE PHILOSOPHY OF MIND

1. Intro-re-duction	17
2. What is Reduction?	18
3. Intertheoretic Applications	22
3.1 Successful Reduction	22
3.2 Reductive Confusion	24
3.3 Elimination	27
4. Reduction in The Philosophy of Mind	29
5. A Recapitulation	32

Chapter Three:

FOLK PSYCHOLOGY

1. Folk History	34
2. Meeting the Folks	37
3. Is Folk Psychology Really a Theory?	40
4. The Fray of the Folk	41
5. Functionalism	42
6. Fodor's Functionalist Fortune: On Paper	46

6.1 The Jerry-Go-Round	48
6.2 ...And Around	50
7. An Intentional Interlude	53
8. Eliminative Materialism	54
8.1 The Genesis of the Eliminative Project	55
8.2 Eliminating the Obvious	56
8.3 Gathering Our Wits	58
9. Rocket Surgeons and Brain Scientists	60
10. The <i>Reductio ad Solutio</i>	64
12. Onward	65

Chapter Four:

THE PERSISTENCE OF THE MANTRA

1. Tuning	66
2. Jackson: Some Epistemic Concerns	67
3. Levine: On Equivocation	68
4. Qualia Schmalia: Dennett's Response	70
4.1 Epiphenomenalism	72
5. P.M. Churchland: Some Concerns For Jackson	75
5.1 An Aside About Prerogative	78
6. Searle: Rediscovery of the Idiom	79
6.1 Emergentism	79
6.2 The Reification of Theory	81
7. The "Might" of Materialism	84

Chapter Five:

THE FINAL CUT

1. The Stimulation of Memory	87
2. The End of The Line	89

Notes 91

Bibliography 98

SHAKING OFF THE IVY

1. Introduction

The mind-body problem is one of the central philosophical enigmas. Roughly stated, the problem is how to reconcile mental phenomena such as subjectivity, beliefs, and intentionality, with physical phenomena such as brains, central nervous systems and synaptic transmission. Traditionally, the problem has been conceived of as being how to explain the causal interaction between two distinct classes of entities: mental substances and material substances. This position, dubbed *substance dualism*, constitutes one end of the broad continuum within which we shall search for the most plausible characterization of the human "mind-brain." At the other end of the continuum is the position of (rigid¹) *materialism*. This view argues that *all* phenomena in our universe, from concrete blocks to abstract thoughts, can ultimately be explained by appeal to a finished, and thus causally closed, physical theory of the universe.

A very popular way to approach this problem is to simply explain that the brain executes all bodily processes from performing a scratch right down to mitosis (cell regeneration); whereas, the mind is the locus of all subjective experience from the sensation

of an itch right on up to the construction and appreciation of a complex theory. And further, that these two disparate classes of relata are nomologically conjoined by a principle which—enter what you think it is here—articulates the unity of one's *entire* self. However, this solution, as appealing to common sense as it may be, puts us no further ahead than were the contemporaries of Descartes. This is because we are yet to invent or discover an unequivocal principle which is able to so specify the union of the mind and the brain which does not elicit legitimate criticism.

In contemporary philosophy of mind most theorists have wisely rejected this view because the hypothesis that there are nonphysical substances outside the realm of spacetime, yet still in causal interaction with physical objects and processes, appears fundamentally misguided. Nevertheless, many philosophers still advocate certain forms of dualism (implicitly or otherwise) for the reason that materialism, as glossed over above, seems to "leave out" certain important aspects of human experience which, it is supposed, cannot be explained by the appeal to a straight physical theory alone. I deny this claim. Moreover, I think that this "subjective" stoicism in the philosophy of mind is largely the result of a common tradition which has been unquestionably passed down the generations from one ivy-towered academy to another.

And thus, in the following essay I will argue that there might just be legitimate reasons, if only in nascent form, for accepting a materialist thesis from within the conception of rationality I take to be representative of the contemporary philosophical environment. Nevertheless, I will try to suspend my present opinions on this issue, as much as that is possible, until the end of this essay as this will certainly afford a more objective presentation of the issues which follow. This being said, however, I think it appropriate to

make my core statement now so that it is clear why this essay endeavours to shake off the stultifying vines of convention, even if it is these very threads which hold together the rationale of discussing the mind in the first place. Thus, I opine: subjective ("mental") experience does not entail non-physical phenomena.

Before we truly get started, though, I think it is important to note that the mind-body problem has been approached in myriad ways, too numerous to mention here, and that the resulting literary corpus comprises an enormous and perplexing field of study. Perforce, and because of this sheer magnitude, I have selected the literature which strikes me as the most pertinent to focus on. My plan, then, at the end of this paper, will be to offer a suggestion as to how I think the philosophy of mind can benefit from the acceptance of the natural sciences, and particularly neuroscience, into her realm of salient discourse. With these initial words of introduction to my thesis communicated, let us now turn to the body of *Bodies, Minds and Materialism*.

2. Duelling with dualism

Perhaps it was Plato,² but I am blaming Descartes for the inculcation of the Dualist idiom into our language. For, ever since *Meditations on First Philosophy* was published and modern philosophers embraced its categories as their own, the paradox of unifying incommensurate entities has proven to be a formidable theoretical challenge. Despite this, some contemporary thinkers believe themselves to have met this challenge and to have exorcised the old Cartesian ghost from their current philosophizing by denying any ties with Cartesian dualism. Yet, they have continued merrily on their way with dualist-incognito talk

of, for example, the causal efficacy of the propositional attitudes³ (Fodor), epiphenomenal knowledge arguments (Jackson), or the irreducibility of the subjective (Searle).

Thus, and if this characterization is correct, these philosophers (and many others besides) are, by definition, vainly attempting the impossible. This is the reason I have characterized the above philosophical approaches as "dualist-incognito"; the philosophers who advocate these approaches all claim themselves to be physicalists (of sorts), yet nevertheless maintain, in one way or another, that there is something besides material particles and physical processes which are constitutive of mental phenomena. Usually, the "something besides" article is referred to as a *property*, yet this doesn't really help to clarify what is being discussed here because it only passes the buck along to the following question: "how can we gauge the difference between a non-physical property and a non-physical substance?" Therefore, I claim, certain philosophers are mistaken to think they have surpassed the original ontological problem. And my reason? They are still trying to describe, "nomologically," how distinctly mental states like 'beliefs' or 'intentionality' or 'whatever' can causally affect our observable behaviour. That is, they are still trying to explain how mental states interact with physical states as if these lexical items nominated two distinct classes of existence. Such, I conclude, is the confusion which often results from the overlooked similarity of substance- and property dualism.

This being said, though, I will grant that it is possible and even occurs frequently that we, in our idiom, speak of things which do not exist, things like: flying pigs, evil daemons, holy spirits and round squares. However, I do not accept that these "things" have any real ontological status beyond their ponderous linguistic meaning. Thus in short, the supposed "explanatory gap," or theoretical lacuna between 'the mental' and 'the physical',

is, according to my thesis, mere metaphorical trifling. We do not need to bridge this speculative gap because this gap is only "in" our minds, and yes--our minds are not *in* our brains--our minds *are* our brains. To further clarify this point allow me to provide a truncated explanation: the brain is the cause, (what we call) mentality is the effect, and this putative "nomological relationship" is supported by an equivocation that we have, generally, failed to mark. What I mean by this is that there is no *causal relationship* between these two phenomena, because in each particular instance though their intensions may be different, the extension is the same.

Thus my thesis is simple. It agrees with the general premises of Evolution⁴ and claims that once human beings became aware of themselves as thinkers, and developed sufficient linguistic capabilities to describe this phenomenon, a new order of ontological being did not "pop" into existence. It follows that talk of 'the mental', 'the propositional attitudes', 'epiphenomenalism' and even 'the social sciences', can certainly be discussed, and discussed rigorously as theory. However, to think that we are uncovering truths about the world by reifying the lexical objects these theories posit--i.e., by describing them as causally potent entities--is as mundane as it is false. What is required for clarity, I then suggest, is a more coherent and systematic theory for explaining why the "mind" seemingly steers "the ship," without lapsing into this type of antique metaphor. Despite these strong claims, however, I have no idea what such a "completed" and unambiguous theory might look like--only, that the idiomatic framework which is currently in use in the philosophy of mind has some very speculative commitments to phenomena which might not really be what it is supposed they are. For this reason, it is also my contention that although fallible, the synthesis of science and philosophy can assist us greatly in literally coming to terms

with what we have heretofore believed to be real: because science provides us with the best, as yet, evidential support for both observable and unobservable phenomena, and philosophy assists us with casting this information into a meaningful and sturdy conceptual framework.

With my position clearly stated, let us now turn back the pages of history, while turning forward the pages of this essay, and be reminded of how all this schizophrenia started in the first place. We will thus begin section four with the alleged perpetrator and his dualist crime.

3. *A la Descartes*

If we remember from our history, the "mind-body" problem was given its first⁵ comprehensive articulation by Rene Descartes. Descartes was a substance dualist with respect to this problem in so far as he pondered how a "mental substance", i.e., the mind, could causally interact with a "physical substance", i.e., the body. This was a perplexing problem for Descartes. He conceived of mind as a spatially non-extended phenomenon whose primary quality, or "essence," was that it was a thinking thing--a *res cogitans*; whereas by contrast, he conceived of the body as being extended in space, yet essentially inanimate. Further, due to the properties he ascribed to these two types of things, viz., for the mind: thinking, willing and believing--what some people call today 'the propositional attitudes'--and for the body: size, shape and mass, Descartes found himself in a quandary. He was certain that one's thoughts were able to motivate action, and likewise, that certain physical events were able to precipitate transformations in his mind, but he was unable to

explain how this union of incommensurate substances was formed. The best analogy he could come up with to describe this mysterious relation was to posit a "*ghost* in the machine."

4. ...Au Jour

Thankfully most of us have come along way since Descartes: we no longer speak of there being two types of "stuff." There is only room in our scientific picture of the world for one type of stuff, and this stuff is physical matter. Curiously, however, the mind-body problem, with its attendant conceptual difficulty, continues to offer a seemingly endless supply of philosophical connundra. These connundra go by a variety of names, and under the rubric of each new name there has developed a different branch of study. Hence, in contemporary philosophy of mind people question, among other things, the "nature of mind"; or they speak about the credibility of "reducing mental states to physical states"; or, and this one is more germane to chapter three, they argue about whether the "propositional attitudes" (beliefs, desires, intentions, et al.) actually have *causal* force in motivating behaviour. Now on first pass these questions might appear to be quite relevant to our present exploration of the mind. However, is this new form of Cartesianism, this "contemporary academia," really that much more conceptually advanced than good old ghosts in machines? And if it is, then why?

I would have to say yes—with reservations, of course. Because of the sheer complexity and ingenuity of these new dualist theories (despite what their proponents call them), we now have a richer conception of ourselves as cognitive beings. Such is the

persuasive magnetism of finely tuned rhetoric which many authors who support dualism have at their eloquent disposal. However, there are also many authors who, with equally tuned rhetoric, argue against the perspicacity of maintaining theoretical ties with *any kind* of dualism. These latter authors frequently base their arguments on the grounds that the dualist hypothesis is a fundamentally flawed conceptual base from which to conduct our studies of mental phenomena; because "mental" phenomena, despite what has been taken for granted in the past, have emerged from our scientific framework as being--if anything at all--ultimately neurophysiological in their origin. Therefore, and on this new view, metaphysical "bridge laws" (which, roughly, endeavour to connect minds and brains) begin to look like extraneous suppositions because fundamentally, the only thing which needs to be fully understood here is the exceedingly complex, indeed labyrinthine, mechanism inside our craniums.

5. Problems with Materialism

Nevertheless it is perhaps too easy to say, for example, that "mental phenomena are caused by neurophysiological processes in the brain and are themselves features of the brain" [Searle, 1992, p. 1.] and leave it at that. For inevitably, someone will question how concepts such as 'qualia' and 'intentionality' (to name only two dualist bugbears) can be identified with the excitation of neural populations or the myriad synapses occurring between axons and dendrites inside our heads. It is obvious, our hypothetical interlocutor might continue, that these "raw feels" or "subjective pushes," respectively, are precisely the phenomena which are in question; and which, further, find no relevant description in

physical terms. So let us get clear how these terms are used as they will figure in the following pages.

Qualia: One of the greatest stumbling blocks for the advancement of materialist theories of mind have been 'qualia'. Qualia are the raw feels of subjective experience: the "what it is like for me" part about enduring pain or feeling pleasure, seeing the colour azure, smelling a favourite scent or hearing one's given name called. These types of experiences are notoriously personal. Thus it seems rather unlikely, some will say, that these same experiences could be analyzed from an objective or scientific point of view while retaining their essential "subjective" character. Materialist theories of mind have foundered on the finer points of qualia time and again for the apparently obvious reason that there is simply something that it is like to *experience* an experience from the privileged, subjective perspective, and that this state of awareness is irreducible to any kind of physical explanation. (This last point is worth emphasizing because it is one of the central background assumptions which is often taken for granted in the philosophy of mind; i.e., that there *are* qualia.)

Intentionality: Intentionality is a philosopher's term of art. Handily resurrected from medieval scholasticism by Franz Brentano in the late nineteenth century, the term 'intentionality' refers to the "directedness", or "aboutness," of a person's thoughts: insofar as a thought, metaphorically speaking, can "reach out" to the object under consideration. Thus for example, the thought 'New York City' seems to pick out, or refer to, the object it is about in a way that goes beyond one's own spatial location (that is, the location of the thought 'New York City'). A good way to get clear about what people are talking about when they use this term is to think about the act of *concentrating*; for when one

concentrates there appears to be an active focusing going on, and it is precisely this deliberate act which is referred to as the "intentional force," or "impetus" behind mental behaviour. Furthermore, intentionality is often cited as being the "mark of the mental" because, by definition, it names a non-physical phenomenon.

6. Revolving The Canon

It is perhaps curious that Qualia and Intentionality have one thing in common: namely, the canonical presumption that the first person (subjective) perspective is the necessary starting point for making any sense of disparate worldly phenomena including, of course, mental states. And indeed, the subjective perspective has been characterized as providing a "privileged access" to the immediacy of one's own mind for the simple reason that there is something that it is like to be someone and that this is distinct and better known from other observable phenomena. What is often concluded from this is that what is made manifest in introspection is something entirely different from what is arrived at empirically. However, because of the character of this "privileged access," i.e., because it is described as a private and better known realm, if one accepts that we, all of us, enjoy this epistemic assuredness then we must also accept that the world beyond our senses and imagination can be known only derivatively--if it can be *known* at all. This strikes me as a peculiar theoretical feature to retain in the light of so much worldly evidence. This is again why it is one of the central contentions of this thesis that this vestige of Idealism be removed from our contemporary theories of the mind. For, if we are ever going to arrive at a tractable explanation of subjective experience which does not begin to wobble at the

mention of 'I', 'me', or 'mine'—currently estimated at six billion different prerogatives (one for everybody)—then we will have to seek a more emotionally neutral theory.

My suggestion is that we first try to abstract our personal opinions from the theories which endeavour to explain mental phenomena; that is, that we try to survey the strengths and weaknesses of these ideas as *theories* about objective worldly phenomena—like other people—and not, as customized models whose sole endeavour is to explain why the world appears as it does to "ME." Yet before I am misunderstood allow me to say that I am well aware that it is precisely this phenomenon, this "self-awareness" or "introspective consciousness," which is exactly what we are trying to get a conceptual grip on—and we will soon return to this concern—but in the meantime, let us stylishly dispense with our vanity and recognize that we too are part of the external world. And, that if we can first develop an objective theory of the mind-brain for our friends and foes alike, then we might just be able to fit our own experiences into this single paradigm and come to have a better understanding of the fleeting and contingent experience we call our selves: without having to pander to the pretensions of each new charlatan who claims himself the center of the world.

7. So Why Be Objective?

When I tell you that I believe in flying pigs and that the only reason you do not is because *you* have never seen one, there are many legitimate ways you can prove me wrong. One method that is not advisable, however, is for you to say that you have seen a lot of pigs, and that none of them have been flyers: this subjective gambit will obfuscate the facts

and lead you nowhere. If this type of argument is all you could come up with in the attempt to rebuke my claims, then, to be fair, all I would have to counter with is "I HAVE SEEN FLYING PIGS" and that you, simply, have been looking in the wrong places. Now admittedly this is a childish argument; I will not insult your intelligence by carrying on in this manner. I am well aware that there are legitimate standards which, whether "I" like them or not, are embedded in the mature understanding of our cultural history and communal standards of rationality. Hence the arguments of the tyrant, the *enfant terrible* and the solipsist, can be revealed for what they are: unrefined, opportunistic and often egocentric versions of the world which are, despite their emotive bleatings, patently false.

The above being said, however, it is important to note that this does not mean that one must always be suspicious of arguments which have as their evidence only subjective reports. Indeed, often it is the case that reliable witness leads to legitimate testimony, which in turn, does much to expand our understanding of phenomena which we ourselves have not been in a position to observe. Nevertheless, my point here is to clarify that subjective conviction alone is often a weak strategy for proving one's claims about the world to others when these claims are seemingly fantastic or run counter to common opinion, as in the above case of the flying pigs. A better method for proving one's claims about the world is to offer not only subjective conviction, but also, legitimate reasons for why such a conviction is to be regarded as true. But where do these reasons come from, and what makes them legitimate.

8. The Glad Scientist

As you might suspect, given the section heading, good reasons for explaining worldly phenomena (or for objugating wild claims about flying pigs) come from science, and what makes science legitimate is that it is methodologically objective. But what makes "objectivity" such a good thing--besides being able to undermine claims about fabulous swine? Well, on one reading, that which is arrived at objectively is external to the mind, actually existing, and uncoloured by feelings or opinions. These requirements are felicitous because they limit which propositions can be counted as scientific by imposing on them a sober and systematic rigour. Thus, if one accepts the fundamental tenets of the scientific enterprise as legitimate, then fanciful notions, crazy beliefs, and far fetched ideas about flying pigs no longer have currency in our discourse about the nature of reality, because the credibility of such claims runs counter to our cognitive economy.

Nevertheless, it would be rash at this point to just assume that the methods of science are incorrigible. In fact, it would not only be rash, but inconsistent with the received wisdom regarding the *revisability* of science. Indeed, one of the defining features of scientific hypotheses is that they can be proven wrong. As Philip Kitcher notes with regards to this condition, "falsifiability is the hallmark of science" [Kitcher, 1982, p. 33]. Therefore, let us now appraise this condition and some others under which science proceeds, so that we are not merely casting empty aspersions on an approach which may be flawed, with the methods of one we have not yet explained. Hopefully this survey will not lead us away from the issues at hand. Instead, I am hoping, it will reacquaint us with the idea that philosophy and science ought not to be construed as disparate fields of enquiry, but as continuous with one another as they are both continuous with our all too human pursuit of truth.

The first condition for science we have already encountered: that it be objective. Thus, what is further required here is that we come to an agreement about how science originates, what provides evidential support for science, whence justification, and if these criteria are agreed upon, *globally*, then we will have the requirements sufficient for deeming this pursuit not only communally verified and widely accepted, but, as an outcome of this, objective as well.

Some Quine: An excellent articulation of these initial requirements can be found in Quine's *Pursuit of Truth*, and it is from this volume that I now take my cue. For Quine, the entry wedge into a natural language, and thus an idiom, is gained by employing *observation sentences*. These sentences, or sometimes just single words, when uttered with appropriate gesticulation allow us to gain our original command of nouns such as: 'house', 'daddy', "'Jasper" the stuffed jaguar', and for Quine puritans, 'rabbit'. These terms name the relation of stimuli to lexical theory and, for the given language speaker, enable him to make the link⁶ between what is heard and what is meant. From this initial cache of nouns compound sentences can be formed so that an infant can learn the meaning of, for example, 'daddy plays with Jasper'. By looking at 'daddy', looking at the "stuffed toy", and then observing the relationship between these phenomena while 'mommy' points at the event and utters the lexical conjunction of 'daddy *plays with* Jasper' an infant, call him "Holden," begins to learn the semantics of complex speech patterns.

It follows that the observation sentence can be conceived of as a semantic anchor which keeps words and compounds of words from drifting around aimlessly and ambiguously. Hence, when we are doing science and uttering sentences like 'this *x* weighs ten pounds', we have an empirical method at our command for checking the truth of this

sentence--for checking out whether the meaning is correct-- simply, we put x on a 'scale' and weigh it. Quine expresses this well when he says "Observation sentences are the link between language, scientific or not, and the real world that language is all about" [Quine, 1990, p. 5]. Much more could be said about the realistic virtues of objectivity. However, I wish neither to flog Quine nor my explanation; the reader will either agree with the facts, as *facts*, or s/he will not.

9. Revision In Stride

A curious, yet desirable, feature of the scientific enterprise it is falsifiable. This means that, in principle, there could be some new piece of evidence that would force us to either emend any of the theories that we currently accept as good ones, or, to abandon any one of them altogether in favour of something *better*--even if what must be abandoned is central to an entire culture's *weltanschauung*. The reason that this is a "curious" outcome of our collective scientific pursuits is that it arises out of its counterfactual; which is, the need to have a method for deeming a scientific theory potentially correct. Indeed, just as there needs to be evidence which could count *against* a theory, so too does there need to be evidence which counts in *favour*⁷ of a theory. Otherwise, as we have previously noted, we would no longer be doing progressive science as much as we would be stagnating in received dogma.

10. En Passant

At this point in our discussion we have outlined the problem we will further discuss below and we have taken note that methods both scientific and philosophical will be employed in pursuit of a solution to it. We shall thus continue on in the next chapter exploring the relationship between theory and evidence which, if clarity is in fact our goal, will assist us with understanding why our ideas about the world must often undergo radical change if they are to remain consistent with the evolving nature of our collective scientific enterprise.

REDUCTION AND THE PHILOSOPHY OF MIND

1. Intro-re-duction

Let's begin with the obvious: the term 'reduction' has a bewildering range of applications. One can speak of reducing stress, reducing the odds, reducing personal expenditure, reducing the national deficit, reducing a bloated ontology, reducing one's waistline or reducing theoretical posits, to name just a few of the possibilities at large. And because of this semantic flexibility, the term 'reduction' is often used vaguely as a diffuse synonym for 'making smaller'. Hence, of the relata which can be said to undergo reduction, it is often understood that the resulting objects (theoretical or otherwise) which emerge from this process will be nothing more than diminished representations, or impoverished instances, of their former selves. In many cases this is just as it should be. However, in the philosophy of mind--where 'reduction' refers primarily to a relation between theories (see below)--this common understanding is, I suggest, superficial and banal. Certainly not everyone would agree with me on this last point. And indeed, some philosophers⁸ claim that reduction has no relevant application to the central problems in the philosophy of mind; because when employed in this locale of explanation this process seeks to re-explain, or abandon, that which is considered to be the significant phenomena: a most intolerable

consequence for many.

Thus my own opinion on the matter, an opinion I have formed largely from reading the work of "pro-reductionist" philosophers, might itself be charged as being banal and superficial. Whatever the outcome of this hypothetical skirmish may prove to be I'll let the reader decide. At present, my concern is to provide a lucid and accurate account of *reduction*--as this theoretical curio is an important element of our discussion and a loadstone concept in the philosophy of mind.

2. What is Reduction?

I will herein illustrate by stressing two points. First, that 'reduction' is generally considered to be a *relation between theories*. This claim is supported explicitly by P.S. Churchland who writes "...reduction is first and foremost a relation between theories" [1986, p. 278]. (Further agreement can be found in: Quine, 1960; Kuhn, 1962; Feyerabend, 1963; P.M. Churchland, 1979, 1984; and Fodor, 1987, to name but a few. For the classic treatment, see: E. Nagel, 1961.) Fine! But what is reduction? Simply put, reduction--or more accurately for present explanation: *intertheoretic* reduction--occurs when one theory (the "reducing" theory), because of its greater explanatory power, can fully explain the posited phenomena of a "higher level"⁹ theory (the "reduced" theory), and in so doing, clarify the ontological status of the theoretical entities which the reduced theory postulates. Another way to explain this has been offered by P.M. Churchland who writes:

A new and more comprehensive theory *reduces* an older theory just in case

the new theory, when conjoined with appropriate correspondence rules, logically entails the principles of the older theory. (The point of the correspondence rules or "bridge laws" is to connect the disparate ontologies of the two theories; often these are expressed as identity statements, such as *Temperature = mv²/3k.*) [1989, p. 47.]

The outcome of all of this has also been described as "an ontological housecleaning; the number of things in the universe is reduced, and the universe is simpler and more understandable because a smaller number of basic entities explains more of the natural world than before the reduction" [Christensen & Turner, 1993, p. xix]. Thus reduction emerges from the literature as a concept we all know and love, and one which we all employ: when we explain why something is happening by offering the best explanation of the underlying causes of this event, e.g., apples fall down when dropped, *because of gravity*, we are engaged in the reductionist program whether we like it or not.

The second point I wish to stress is that the *theoretical* efficacy of this very basic principle is often confused. Consider Searle's treatment in his 1992:

Most discussions of reductionism are very confusing...At the very outset it is important to be clear about what the relata of the relation are. What is its domain supposed to be: objects, properties, theories, or what? I find at least five different senses of "reduction"--or perhaps I should say five different kinds of reduction --in the theoretical literature, and I want to mention each of them so that we can see which are relevant to our discussion of the mind-body problem. [*Ibid.*, pp. 112-113.]

I have provided such an extended quote to exhibit the very confusion of which Searle himself mentions, and obviously thought he was avoiding.

Searle goes on to list these five types of reduction: "Ontological, Property

Ontological, Theoretical, Logical or Definitional, and Causal" [*Ibid.*], but these curious inventions shall not concern us. The reason they shall not concern us is because, on my account, they are all the same--that is, they themselves can all be reduced to the "theoretical" kind because this is the only place where reduction does any work! This is most assuredly contra Searle, who thinks that all of these different types aim at being reduced to *Ontological* reduction; that is, to the type of reduction where "...objects of certain types can be shown to be nothing but objects of other types" [*Ibid.*]. Indeed, he states within this section the claim that "it seems to me this form of reduction [Ontological] is what the other forms are aiming at" [*Ibid.*]. Thus, by borrowing his ontological example, we can see how a chair may be reduced to "nothing but collections of molecules." Yet on this reading, one could get the idea that the chair itself (and others of its kind) vanishes, only to be replaced by a *comfortable* swirling eddy of charged particles--for how else could one see an "Ontological" reduction? And indeed, this is how Searle would have it, for he says: "In each case, the causal reduction leads naturally to an ontological reduction by way of a redefinition of the expression that names the reduced phenomenon" [Searle, p. 115]. This is a tricky argument, yet I think we can begin to see the problem when we analyze Searle's claims with regards to the "causal reduction" which then leads to a "reduced phenomenon."

The reason I have spent so much time with this above distinction is to draw attention to a very severe mistake that can be made with regard to understanding the reductionist program. The mistake occurs when (as above) it is implicitly assumed that when one offers a reduced explanation of certain phenomena, phenomena like chairs for example, that this type explanation is causally connected to the actual discounting of

ontological constitution. This is perhaps how Searle would have it. But no; for clarity's sake, we should see that any given conceptual apparatus does not change reality--A's do not become B's *because* of reduction. Rather, this theory is simply deployed so that we can better understand what the causal elements are of the phenomena at issue. And this applies in scientific procedures as well as with the input from our natural senses. Therefore, and for present consideration, when we speak about the reduction of: observable phenomena, unobservable phenomena, abstract relata, tables, chairs, H₂O, *bouillabaisse*, or anything else for that matter (even, say, the propositional attitudes) it will be understood that this pertains to articulating the conceptual relations between our theories of these phenomena--and *ex hypothesi*, between our theories of these theories of relations between phenomena.

Contra Searle, the only type of reduction one does in philosophy is theoretical reduction (or, "intertheoretic reduction"), because the only thing that is being here reduced is one *theoretical framework* by another. I think this move helps to break the hold on our imaginations that would have us think, erroneously, that any kind of "reductionist" program forces us to understand the world as if seen through a very powerful microscope: devoid, as it were, of all its macro integrity. Here, my claims about reduction are supported by P.S. Churchland who, with reference to theoretical priority in reduction writes: "What gets reduced are theories, and the stuff in the universe keeps doing whatever it is doing while we theorize and theories come and go" [1986, p. 288].

This last point about theoretical contingency is important enough to emphasize because it helps to further explain the (and I hazard) "fact" that theories in philosophy, science, and everywhere else, are revisable. That is, even theories that we tacitly take to be indubitable proofs are eligible for reduction (or "elimination") if better, more concrete,

reduction is by emulating E. Nagel's account of the reduction of Thermodynamics to Statistical Mechanics--and I will not depart from this venerable tradition. I will, however, offer only a truncated version of this account, yet in recompense, we will cut quickly to the chase.

The reduction here being described is one which takes certain (higher level) distinctive concepts found in the taxonomy of Thermodynamics and reduces them all to a blanket theory found in Statistical Mechanics. What this means is that notions found in the former theory such as 'temperature', 'heat' and 'entropy'--notions which prior to intertheoretic reduction each enjoyed an autonomous explanatory status--can all be reduced to one notion found in the latter theory, namely, 'mean kinetic energy.' The reason the above reduction occurs is because it can be shown that the temperature (or heat) of any given object just is the kinetic energy of its constitutive molecules. Therefore, according to this model of reduction an *identity* is wrought, because the reduction shows us that the two postulated notions above explain one and the same phenomenon:¹¹ the former, a higher level explanation, while the latter, a lower level explanation of the same object.

Ideally, intertheoretic reduction is considered to be a good thing because it provides us with an *explanatory unification*: the older theory gets reduced to the newer theory, and as a result, our understanding of the relevant phenomenon (say, temperature) is greatly enhanced because we come to understand why the laws of thermodynamics are as they are and why temperature affects things as it does, and thus, ultimately, what it really is. Further, reduction is considered a good thing because as a consequence of it we may achieve *ontological simplification* as well. Patricia Churchland explains this latter possibility deftly:

Ontology pertains to what entities and properties exist, and in the event of intertheoretic reduction it may turn out that where we had thought there existed *two* different kinds of phenomena [e.g., 'temperature' and 'kinetic energy'] characterized by the laws of two different theories, there is in fact but *one* kind of phenomenon that is described by both theories. [1986, p. 280.]

This "identity relation" is important to bear in mind because it is characteristic of the classical version of reduction which views this relation between theories as being, as we have alluded to, a *smooth* process. It is considered "smooth" because the entities and processes of one theory can be mapped onto the entities and processes of another theory with neither the entities in question nor the general conceptual framework in which both theories figure being significantly altered.

3.2 Reductive Confusion

The above gloss of Nagelian reduction has hopefully provided a clear enough account of the classic view so that we may now continue. The related point of interest I wish to discuss presently is how people in general often continue *conceiving* of various phenomena in terms of an outdated conceptual model, even after a new and more comprehensive model has been developed, demonstrated to have superior explanatory clout, and embraced in the common idiom. This failure to be rational, or at least logically consistent with the popular conception of rationality, can well be demonstrated by considering popular opinion about the causes of motion, in light of differences between Aristotelian and Newtonian mechanics. (For other contemporary versions of human

[ir]rationality see: Nisbett & Ross, 1980; Perner & Howes, 1995; and Gopnik & Wellman, 1995.)

On the Aristotelian account, an object in motion will continue in motion only if a continuous force is applied. This, of course, is consistent with the ancient claim that bodies have, as an internal principle, the desire to maintain a state of rest. Thus, as Aristotle reasoned, if one stops pushing a boulder it will stop moving; and when sufficient force is reapplied to the boulder it will begin to move again, and it will continue to move as long as sufficient force is maintained. This much was intuitively¹² obvious. Despite this empirical observation of the behaviour of rocks, however, Aristotle's theory could not explain the behaviour of projectiles. For, when a stone was hurled from a catapult it continued to move after the force of the catapult was no longer being applied, and this ran counter to the theory. Something else was needed here to explain why moving bodies *continued* to move. And so, the theory of "impetus" was developed to explain how an "internal force" could be imparted from one object (like a catapult) to another object (like a stone) and that this derivative force acquired would then dissipate as the weight of the moving object slowly taxed its acquired cache of animation. [P.S. Churchland, 1986.]

On the Newtonian account, and by contrast, a body in motion will continue in motion so long as there is no external force to halt or impede its progress--the standard impediments cited here are gravity and friction. More important, however, is that in the Newtonian theory there is *no such thing as* 'impetus'; this category, and the specious phenomenon it denotes, gets eliminated from accurate discourse on moving bodies and force. Despite this elimination, though, and relevant to note, many continue to use this term in the folk idiom (even at the dawn of the twenty-first century) as not just a metaphor, but

as a useful explanatory concept as well: one which describes the *driving force* behind a moving body (see 'impetus' in any good dictionary). Hence, in the conception of rationality in which "impetus" figures, a conception we might call "folk physics," a term which has been demonstrated to be conceptually vacuous is still embraced as a tangible theoretical entity.

This is not to say that "impetus" should be construed as a meaningless and dispensable word in our idiom (it does have metaphorical import), but only that in proper discourse about the causes of motion this term is, indeed, an idle one. A different way to explain this is to say that the two terms have a different intension, yet share the same extension. Thus "lightning bolt" and "electromagnetic discharge" may have different meanings, but they denote the *same* phenomenon. This linguistic point is often ignored in the idiom which can result in the inconsistent conflation of popular semantics (or "folk theory") with what has been shown to be a more precise and accurate¹³ conception of ontology.

This ethos is illuminated well in a study conducted by McCloskey (1983) and brought to my attention in P.S. Churchland (1986). She explains how:

In one test college subjects were asked to jog along and try to "bomb" a stationary target by dropping a golf ball on it. Typically, they thought the ball would fall straight down when released, and some even thought it would move backward. Therefore, they did not drop the ball until the hand holding it was directly over the target or even past it. [p. 290.]

The above example should further exhibit that even when a given instance of intertheoretical reduction is a well known, verified, and established success, this success

alone does not guarantee that normal people will incorporate this new theoretical development into their understanding of "the way the world turns"—even when (as above), the *new* theory is centuries old, and the one it displaces dates back to classical antiquity.

Thus, even though we may agree that new developments in science help us to get clear on what the causes of certain phenomena are, we may still be unwilling to draw the conclusion, as others have, that "when truth becomes a paramount virtue, folk physics simply becomes antiquated superstition" [Christensen & Turner, p.xvii]. Further parallels could be drawn here by examining folk theories about Optics (which has been reduced to Electromagnetic Radiation), Intrinsic Weight (which has been reduced to mass and gravity), and Genetic Heredity (which has been reduced to the material structure of the double helix, deoxyribonucleic acid, or DNA). However, the point I think has been made.

3.3 Elimination

We saw above how ideas like 'impetus' get eliminated from the scientifically accepted ontology when intertheoretic reduction shows that such ideas do not, really, correspond to anything outside of rhetorical flourish. In this section I want to further develop the notion of *elimination* as it is an important concept and one which will play a key role in the next chapter.

Elimination occurs when an old idea, or term, is dropped from the vernacular because it no longer communicates the relevant information it was designed to. Often this happens because the phenomenon it was selected to nominate has been proven vacuous. A good example of this displacement can be provided by looking at how the "caloric" theory

of heat was eliminated in favour of the corpuscular/kinetic theory of matter and heat (which was then reduced to Thermodynamics). In the old theory, "caloric" (heat) was likened to an imperceptible *fluid* which could be held by a body like water in a sponge and, like a sponge, distribute this fluid over objects it came into contact with. A fairly decent theoretical framework was built around this notion which did much to explain how caloric produced observable effects such as melting or boiling. However, as science progressed and new discoveries were made, it became apparent that this conception of the aetiology of thermal phenomena was flawed. Paul Churchland describes this realization well:

By the end of the last century it had become abundantly clear that heat was not a substance at all, but just the energy of the motion of the trillions of jostling molecules that make up the heated body itself. [1984, p. 43.]

What is important to recognize here is that the ontological posits of an older theory can be shown to refer to nothing at all by a newer, more comprehensive theory. This happens when the newer theory can explain everything the former could--and more--and even why the older theory worked as well as it did. Ideally the new theory has this effect, when successful, because it organizes all the relevant phenomena into a superior conceptual scheme. For example, where heat was formerly understood to be a *material substance* on the old theory, it is now understood, and indeed "proven," that heat just is mean kinetic energy, i.e., it is a *form of motion* (analogously, this is akin to the reduction of substance-to-"property" dualism). This last point is important because it helps to clarify, further, that what gets eliminated (or reduced) in these types of relations is nothing more than one theoretical framework by another. With this in mind, and the background of intertheoretic

reduction set—including the often daft conception of rationality it which it figures (as with the college students above)—let us now examine more specifically how all of this relates to the philosophy of mind.

4. Reduction in the Philosophy of Mind

There are perhaps many ways of explaining how 'reduction' figures in the philosophy of mind. The one we shall use will be to explain how this conceptual method can be used to formulate what is referred to as the (type) *Identity Theory*. This will provide the paradigm method some philosophers have used (see esp: J.J.C. Smart, 1962; and, D. Armstrong, 1968) in their endeavours to reduce mental states to brain states. But more importantly, it will provide the context in which "Functionalism" and "Eliminative Materialism" (see next chapter) have been developed as a reply to the this model.

The (type¹⁴) Identity Theory (hereafter the I.D. theory) holds, as its central premise, the claim that minds simply are identical to brains. In the vocabulary of intertheoretic reduction this can be articulated thus: "The corrected theory of mind will be entailed by a mature neuroscientific theory of the brain" [Christensen and Turner, 1993]. The I.D. theory is, then, perhaps the obvious outcome of the work conducted by thinkers in the philosophy of science (Nagel specifically, and the Logical Empiricists more generally). This can be seen by taking the above example of Thermodynamics and Statistical Mechanics and applying it to mental states. With this analogous relation in effect, it was thought that perhaps mental states could be "smoothly" reduced to states of the brain.

Thus for any type of mental state, 'say stubbed-toe-pain', it was hypothesized that this corresponded to a type of brain state; or more properly, that the "experience" of the pain corresponded to a type of brain state. According to this model, then, if it were possible to duplicate the state of one's brain as it was when it had formerly undergone the pain, then once again, the subject of the initial pain (and presumably owner of the brain) would have the exact same "experience" of 'stubbed-toe-pain' that was previously endured. J.J.C. Smart explains this reduction of mental states to brain states well in his article "Sensations and Brain Processes," from which I quote:

When I say that a sensation is a brain process or that lightning is an electric discharge, I am using "is" in the sense of strict identity...When I say that a sensation is a brain process or that lightning is an electric discharge I do not mean just that the sensation is somehow spatially or temporally continuous with the brain process or that the lightning is just spatially or temporally continuous with the discharge...I distinguish these two senses of "is identical with" because I wish to make it clear that the brain-process doctrine asserts identity in the *strict* sense. [1962, p. 57.]

If the I.D. theory were true there would be at least two benefits. First, the old problem of how two distinct ontological types of existence (i.e., minds and bodies) can interact with one another would be solved: it could be shown how talk of minds can be reduced to talk of bodies and, as a result of this conflation, that all this "mind-brain" talk is about the same stuff. And second, it would enrich our understanding of how the world works because we would be able to rationally explain why "mental states" affect "physical states" (and vice versa) without having to eliminate mentalistic talk from our scientific vocabulary: mental states could be described as "higher order" phenomena which (like

'temperature' above) could be reduced to a "lower order" of explanation. However, the I.D. theory is not usually assented to. I provide two reasons below.

First, if the identity postulated by the I.D. theory were true (i.e., the strict identity of Leibniz's Law¹⁵), then mental terms and physical terms should be logically interchangeable *salva veritate*. However, this is plainly not the case. Consider again the case of pain. If the above mentioned identity held then there would be a nomological correlation between the *experience of pain* and the relevant *set of neurons* which are supposed to have caused this pain; but this law-like correlation has not yet been conclusively demonstrated. Or, and again if the I.D. theory held, there should be a correlation between typical mental states like 'beliefs' and 'desires' and their instantiating neural substrate; but, of course, this identity has not been discovered either.

Second, it is often said that we have "direct access" to our mental states (*I know what I am thinking as I think it*), but only mediated access to our brain states if any at all (one usually cannot observe the functioning of one's own neuroanatomy). Therefore, it is objected, mental states and brain states cannot be identical. For even if one could observe his own neuroanatomy (say with a glass plate in the skull and a mirror) it would certainly be nothing like the *feeling* one is apprised of in conscious thought. Therefore, it is concluded, mental states and brain states correspond to two different sets of phenomena.

There are of course other arguments against the I.D. theory specifically, and of reduction in general. And moreover, there is some positive criticism as well. However, these are the issues which will occupy us in the next two chapters. I will thus conclude this chapter in the following section by making two points I consider to be relevant to our discussion in its entirety.

5. A Recapitulation

Two final words in favour of the reduction program before we move on. First, if inter-theoretic reduction had not been embraced as a useful methodology, we could still be claiming that water was the primordial element, testing the buoyancy of witches, and left with only the vapid, middle sized and sundry objects of the world with which to argue about. And this, certainly, would have long since bored us into dogmatic submission. Thus let us endeavour to agree: scientific theories are not bad things--*au contraire*--they enrich our cognitive horizons by explaining what all those beautiful colours really are.

And second, despite the virtues of the reductionist program, and notwithstanding the conceptual clarity we have heretofore gained by employing this methodology (a clarity which promises better things to come), there are many philosophers who still argue that--no matter what is discovered--mental categories are irreducible to any other vernacular, abstract or otherwise, because these categories are autonomous in their own domain. These thinkers, some of whom we will visit, believe that a reduction of mental states (most notably 'the propositional attitudes') will never, and could never, occur because the science (viz., "folk psychology") which bespeaks of their causal potency is an autonomous framework: one which has been constructed by making rigorous our common sense articulation of mental behaviour. Further, many of these thinkers suggest that the very suggestion of reducing mental states to brain states *simpliciter* is absurd, because the very nature of a "mental state" is that regardless of its aetiological connection with the brain, its defining characteristic is that it is simply something *more* than the exceedingly complex relation between the 100 billion or so neurons which make up our cognitive hardware.

Ironically, however, many of these thinkers maintain that they are still committed to a physicalist program despite their sympathetic alignment with the mental categories of folk psychology.

If, and we shall see whether, this is what folk psychology is all about, then folk psychology is (on my account) an approach to explaining mental states which wants to maintain its yoke with *Descartes*, yet power this ancient wagon with the steam and efficacy of the brand new mechanical engines offered by the (cognitive) neurosciences. I offer a more detailed explanation of this slighted folk theory, and perhaps something better, in the following chapter.

FOLK PSYCHOLOGY

1. Folk History

We will begin this chapter with a brief, albeit putative, look at how folk psychology arose in the first place. I say "putative," here, only because this look we are about to take is through lenses coloured by the general premises of *Evolution*. Therefore--reader be warned--the following chapter, as elsewhere, agrees with the underlying sentiment behind George Graham's articulation of what he calls "Adaptationalism." Says Graham: "An important theme of nineteenth-and twentieth-century science is that men and animals are natural creatures and products of biological evolution. Also men's minds and animal's minds are products of biological evolution" [Graham, 1993]. Further agreement with this theme can be found in (P.M. Churchland, 1981, 1989; Cosmides & Tooby, 1995; and Gopnik, 1995). Thus, with any tinting removed, let us continue with what I am assuming will be clearly obvious.

Having a mind gives an animal an advantage. When an animal is endowed with this mental privilege it gains awareness of its environment; and with this awareness comes the ability to observe. With the ability to observe, an animal becomes able to recognize certain natural patterns of sensory input; although at this stage of development it is perhaps

debatable whether such an animal is *aware* of said awareness. But no; at this stage an animal is in a position to be affected by what it observes in its environment and to react in accordance with its genetically programmed reflexes coupled with what it has learned through experience.

What makes humans distinct from other animals is that they, or "we" in this case, have a highly sophisticated *conception* of mind. We are able to, seemingly,¹⁶ reflect on some of our mental states, and to describe that we are doing so. This "extra-sensory" faculty (nothing metaphysical implied,) has given human beings the conceptual ability to invent and manipulate language--our proudest achievement to date--and to understand the returned signals of our fellows. Originally I suspect, a long, *long* time ago, the use of language was limited to a few observational grunts which, when accompanied by gesticulative behaviour, meant something like 'food', 'danger', 'it is raining', 'you are attractive to me' and other terms of pre-theoretical ilk.

This above conjecture is supported by observations in the field with vervet monkeys who effectively distinguish and communicate the presence of the three major predators which vervets are hunted by. Three different yet distinct "calls" alert vervets to 'leopards', 'eagles', and 'snakes', and cause "different evasive action: climbing a tree (leopard), looking up in the air or diving straight into the bushes (eagle), or standing on hind legs and looking into the grass (snake)," [Cosmides and Tooby, 1995]. Moreover, it is supported by observing our infant members of society and their globally structured progress in the acquisition of natural languages: from observation sentences to the semantics of complex speech patterns.

From these rudimentary beginnings Humans have learned to call attention to a huge

variety of objects in the world, and also, to call attention to many different *relationships* between these objects. (Remember, if you will, how our model human "Holden" was able to learn the name for the abstract concept '*playing with*'.) And, as this linguistic ability developed into a full blown faculty, human beings became able to call attention to an enormous variety of not only physical objects (and observable relationships), but also, we became able to deploy metaphor. And with metaphor, we became able to describe objects and events in peculiar yet effective new ways and to hypothesize about the ontological status of unobservables.

For example, certain ancestors were able to describe the sea as *angry*, the Moon *seductive*, and the Sun as *kind*. Generations later, and with the reification of metaphor, later ancestors were able to further develop this personification of natural causes. In this new linguistic environment mystical deities often became the causal forces behind observable events in nature. And thus, it became possible for certain other ancestors to delineate natural causes by positing either an Apollonian or Dionysian influence.

Yet later in our collective culture and rich linguistic history, and with the further development of our concept of mind—here is the big leap—our predecessors began to explain how certain mental states were responsible for causing certain spectacular, physical events. In some cases, I claim, these mental productions were called "miracles," in so far as they were conceivable by the mind and expressible in language, yet plainly metaphysical in their origin. There were, of course, other physical events which were less spectacular to observe, yet equally mystifying. Examples of these lesser phenomena might be called 'prophecies', 'prognostications', or simply 'projections'. My point, however, is that it was often explained in the folk idiom that these mere divinations themselves when uttered, or

otherwise made manifest, had a causally effective hand in the realization of their own prediction--much the same way a "spell" is supposed to work in a fairytale. And thus, possibly, we have arrived at the nature of human belief.

2. Meeting the Folks

When we make common sense predictions or offer explanations about the mental life and accompanying behaviour of ourselves or other people we are practising what many contemporary philosophers, and cognitive scientists alike, are wont to call *folk psychology*. There are many different ways to characterize this conventional practice but the following reference introduces it well enough:

Briefly, folk psychology is the tag given to ordinary talk about the mind. It does not refer to talk about the biology of the brain and central nervous system; rather it refers to talk about beliefs and desires, intentions and fears...It is the vocabulary of the mental. [Christensen & Turner, p. xvi.]

In line with this above reading, then, we can interpret folk psychology as being a colloquial "common sense psychology" or as a prescientific method for understanding the mental character of human beings. This "theory," or rough collection of theories rather, has a history dating back at least twenty-five centuries; and has, because of its global applicability, managed to stay optimistically afloat on the common tongue as well as on the philosophical tides of new cognitive discovery since ancient times. It is most certainly because of this ubiquitous success that folk psychology has proved itself to be a good

explanatory catalyst for gaining useful ideas about the mind—it gets the point across well enough when a scientific, or strictly causal, understanding of the world is not at issue.

However, when these two demands clash, i.e., when the folk and the scientific community disagree about the causes of our mental lives, it becomes a philosophical problem. Our goal in what follows, then, will be to distinguish which approach has greater explanatory power with respect to the entities in question, or, to devise ways in which the posits of each theory can remain explanatorily autonomous in each of their respective domains.

With the incredible, and for some incredulous, new empirical findings of the neurosciences there has arisen such a disagreement. Hence the main motivation for this chapter: to discuss whether the chief categories of folk psychology (i.e., the propositional attitudes) should be "reduced" to those of a more sophisticated science like neurology; whether they should be regarded as functionally autonomous in the realm of psychology; or, whether these folk categories are so intrinsically flawed that they ought to be "eliminated" outright in favour of the new, and purportedly more powerful, conceptual framework being constructed by the neurosciences. These are the questions we shall pay close attention to in what follows.

Neuroscience is a very young discipline,¹⁷ however, and so it is only recently that the underlying principles of folk psychology have even met with serious conceptual challenge; and here, by only a minority of philosophers interested in offering a new paradigm for understanding how the mind works. One such philosopher is P.M. Churchland.

Churchland argues that the categories of folk psychology are simply vestiges of a

more ignorant time. Ones which have, after long and frequent use, simply become embedded in the collective understanding of what defines our cognitive lives. Thus as he suspects:

Folk Psychology has survived for so very long, presumably, not because it is basically correct in its representations, but because the phenomena addressed are so surpassingly difficult that any useful handle on them, no matter how feeble, is unlikely to be displaced in a hurry. [1984, p. 46.]

And therefore, as we make further advances in the natural sciences, and particularly neuroscience, Churchland argues that we will come to agree with his "eliminative materialist" convictions and see that, as he says, "our common-sense psychological framework is a false and radically misleading conception of the causes of human behaviour and the nature of cognitive activity" [*Ibid.*, p. 43]. We will hear much more from Churchland(s) in the pages that follow.

Other thinkers¹⁸, such as Jerry Fodor, think that the categories of folk psychology are just fine. As cognitive science progresses we will be able to hold on to our 'beliefs' and 'desires', 'hopes' and 'fears', and explain the causal nature of these mental states under the rubric of "functionalism." To be sure, the central premises of functionalism argue in favour of the ontological realism of mental states. These premises state that the propositional attitudes can be defined in terms of (1) the causal relations between properties of the brain, (2) input from the environment, and (3) the further intra-relations between other mental states.

Fodor's view¹⁹ can thus be generalized under the banner of functionalism, and his

commitments to the categories of folk psychology can be well observed in his writing. Consider: "[w]e have no reason to doubt--indeed, we have substantial reason to believe--that it is possible to have a scientific psychology that vindicates common sense belief/desire explanation" [Fodor, 1987, p. 135]. Functionalism in general, and Fodor's position specifically, will be presented with greater detail below.

3. Is Folk Psychology really a Theory?

I will digress for a moment here to observe that there are some philosophers (see esp: Gordon, 1995a,²⁰ 1995b; see also: Wilkes, 1991; Goldman, 1995; and Millikan, 1993) who argue that folk psychology may not be a theory at all, and thus, that it is not susceptible to either reduction or elimination. This is an interesting move made by those who wish to fortify the "common sense" approach of folk psychology and its attendant vernacular; and one which will be prudent to bear in mind. However, our immediate concerns require us to treat folk psychology as if it were in fact a theory, as this assumption enables us to judge its categories, and its explanatory power, with the set of criteria with which we judge other theories. And moreover, we shall discuss folk psychology as if it were a theory for the reason that this is how it is represented in most of the literature we shall examine (and from what I have observed, inclusive of most of the "folk psychological" literature, period).

I here offer two such examples which support this theoretical alignment from philosophers who are, perhaps, as divided on the issue of folk psychology as to be the definitive poles; and correspondingly, the two we shall pay the closest attention to:

Seeing our common-sense conceptual framework for mental phenomena as a theory brings a simple and unifying organization to most of the major topics in the philosophy of mind, including the explanation and prediction of behaviour...the problem of other minds, the intentionality of mental states, the nature of introspection, and the mind-body problem. [Churchland, 1989, p. 3.]

And:

The point--to repeat--is that the theory from which we get this extraordinary predictive power is just good old common sense belief/desire psychology. [Fodor, 1987, p. 3.]

Therefore, if we can agree (even if only temporarily) that 'folk psychology' names a theoretical framework, we can continue on with our discussion of its relevant virtues and vices, and examine the way in which it characterizes mental states. This marks the end of my "anti-theoretical" digression.

4. The Fray of the Folk

We are thus poised at the beginning of an argument about the critical status of folk psychology (hereafter FP) and its categories. The *eliminative* side of the debate will argue that the categories of FP mesh with neither the emerging framework of the neurosciences, nor physical theory writ large, and should thus be discarded. The *functionalist* side of the debate will argue that these categories are indispensable components of any comprehensive theory of mental causation and should thus be retained, come what may. A third view, the *I.D. theory* (which, to recall, argues for a theoretical reduction because the categories of FP are thought to be *identical* to the categories of neuroscience--only obscured by their

antique dress), will also be given some reign in the discussion, for the purpose of supporting the broad historical continuum between the two featured positions under consideration.

Common to all of these disparate views, however, and used in innovative ways by each camp as an explanatory aid, is the idea of "reduction." And as we saw in chapter two, this method of ascertaining conceptual clarity, regardless of who employs it and for what end, is crucial to understanding how the posits of any given theory must be *translatable* to a set of sentences--sentences which can be somehow tested for their truth value in accordance with inference rules (as that is the mark of not just an estimable theory, but of *any* theory). Hence, and as a reminder, it will be prudent to keep in mind just how this conceptual device (i.e., reduction) can be deployed to make sense of not only complex theoretical phenomena, but of simple occurrences in nature as well.

5. Functionalism

For the functionalist then, as we have mentioned, mental states are not entirely reducible to the material substrate of the brain. This is because although these "two" phenomena are intrinsically and inter-causally connected, there remains an abstract relation between "purely mental" mental states whose process defies explanation in physical terms alone. This is the main move functionalism makes away from the I.D. theory. For, according to the functionalist, a mental state (such as pain) does not occur without there being a supporting physical state (i.e., a particular²¹ brain state) as these two states are inter-causally connected; but, and contrasted with the I.D theory, the functionalist also

maintains that these psychological tokens are not reducible to this type of physical fact alone because such a material stringency fails to capture the causal, yet nevertheless abstract relationship *between* mental phenomena (such as 'pain', and the belief that 'I want it to stop'). Hence, it is in this way that functionalism aims at establishing the autonomy of a psychological vocabulary for the explanation of human cognition and behaviour. This commitment is pervasive throughout functionalist literature, and no doubt has been imbued by the seminal authors²² of this doctrine. I here provide an example of how a "classical" functionalist views the relation:

I shall, in short, argue that pain is not a brain state, in the sense of a physical-chemical state of the brain (or even the whole nervous system), but another *kind* of state entirely. I propose the hypothesis that pain, or the state of being in pain, is a functional state of a whole organism. [Putnam, 1967, p. 154.]

And suspiciously further:

In particular, the functional-state hypothesis is *not* incompatible with dualism! [*Ibid.*, p. 157.]

Now I am aware that this characterization of functionalism may be construed as offering an *ad hoc* example of the functionalist program. And since I do not wish my own views on this subject to be regarded as opportunistic, I will provide a more recent and perhaps more objective articulation of functionalism:

Functionalism holds that mental states are essentially defined by the set of causal relations between (a) input to the system from the environment, (b) other types of mental states, and (c) output, characterized by behaviour. [Christensen & Turner, 1993, p. xxiii.]

If we accept this gloss of functionalism then we might see how it is an improvement from early versions of the identity theory (which, to again recall, requires a one:one mapping of mental states and physical states to sustain its central premise). Indeed, the functionalist does not have to agree that there is this kind of numerical identity (akin to Leibniz's law²³) between mental states and brain states, because "mental states" are partly instantiated by an abstract relation, and which, according to functional hypothesis, cannot ever be reduced to a strictly material function.

Thus functionalism, as a "physicalist" thesis, avoids the difficulty of *multiple realizability* which has troubled many other materialist versions of the mind. "Multiple Realizability" is the hypothesis that type-identical mental states could be realized in, or "by," a variety of different physical systems²⁴ as long as they were complex enough to support such mentality. There have of course been problems with demonstrating multiple realizability in the flesh, as it were. Nevertheless, multiple realizability has worried many proponents of the I.D. theory because with this notion of potential "multiplicity," the very idea of an "identical relation" is supplanted by the possibility that there need not be such a relation. The important point to note here, however, is that within the FP idiom, functionalism is exempt from explaining how type-identical brain states cause type-identical mental states because according to functionalism, this is not the *sole* process by which mental states get instantiated anyways. This is a crafty move. And one which is made possible because, as an explanatory hypothesis of functionalism, there is also the "internal" connection between mental states which, in the literature, is described as their "functional role" (see Putnam quote above).

Suspiciously, however, this "role" is purportedly causal in at least two interesting

ways. First, any mental state can play a causal role in the production of any other mental state without the "effect" of this interaction being construed as a physical process. And second, the outcome of this interaction, often a further mental state, can be engendered with intentional clout just in case what is formed is a new propositional attitude. Moreover, this "functional" relation (the notorious relation 'R'²⁵), is considered to be a metaphysical relation because the entities it is defined over are not, themselves, purely physical in their description. Despite this, however, the functionalist often maintains that his theory is a physical one (of sorts) because it agrees that a specific brain state (or specific functional state²⁶) is the bi-conditional for any instance of a mental state. And thus by analogy, where there is smoke (or the belief that 'there is smoke') there is also fire (however in this case, the "fire" corresponds to neural activity).

Therefore, to fully articulate the premises of functionalism one needs both the nomological (law-like) relation between the mind and the brain, and the abstract or metaphysical relation (relation 'R') which characterizes the interaction between mental states alone. Thus for a "physicalist" thesis, functionalism relies pretty heavily on what we now have reason to suspect as "the mystical relation 'R'" (see note 26). Perhaps, then, we can now see how the functionalist program buys its explanatory ubiquity with an abstract currency--yet buys it at the expense of not being able to identify, in empirical or verifiable terms, what the "cash value" of this explanation amounts to. This is because the "relation" deferred to cannot ever be succinctly defined. Despite these problems, however, functionalism has emerged as a very popular position to hold in the philosophy of mind.

Let us now fix on just one functionalist in order to see if we can better make sense of this philosophical position. This should provide us with a close study of whether there

are important explanatory features of functionalism which can dispel the confusion, or whether the background assumptions of functionalism, despite their panache, and to continue with the above metaphor, cost more than their worth.

6. Fodor's Functionalist Fortune: On Paper

I have chosen to pay certain homage to Jerry Fodor because this author is a virtual champion of functionalist explanation. Moreover, he is recognized as an ardent defender of the "sentential" property of thought, a proponent of intentional computation, and a realist about the propositional attitudes; and further, he is an exceptionally persuasive writer. Thus, I propose, if we can find one particularly good version of FP to closely examine it will be by looking at Fodor's claims about the functional organization of the human mind.

However, in his 1990 book, *A Theory of Content: and Other Essays*, where Fodor unabashedly lays down the (philosophical) creed of common sense explanation, he seems to be somewhat vague right at the crucial moment. Allow me to explain.

As I understand it, common sense explanations (specifically in this literature) are supposed to clarify the ontological status of mental states, and thus explain: what they are, how they are related, and whence comes their causality. And indeed, this is what Fodor attempts to explain with his "RTM". Unfortunately, this program seems suspiciously circular from the get go. Let us first consider his program, after which we will attempt to justify the objection:

The connection with our topic is this: the full-blown Representational

Theory of Mind (hereinafter RTM, about which a great deal presently) purports to explain how there *could be* states that have the semantical and causal properties that propositional attitudes are commensensically supposed to have. In effect, RTM proposes an account of what the propositional attitudes *are*. [1990, p. 5.]

The reason this program is circular is not because all of the (implicit) categories seem to have a psychological referent, but because all of these categories, including Fodor's unqualified "state" (see quote *supra*) rely on one another for their meaning. They thus derive their semantic content and causal clout from their *metaphysical* relation alone.

Think about it. As we know: Fodor's "state" enjoys semantical and causal properties (the notorious relation R again), and these properties are what defines the propositional attitudes; yet, as we also know, by common sense (and, indeed, more rigorous functionalist theory), the propositional attitudes just *are* states which we can mentally occupy. Now certainly we should not question Fodor's logical inference--here--for he is offering us a perfect logical syllogism:

$$(M.S.) > (P.A.)$$

$$(P.A.) > (C.P.)$$

$$\therefore (M.S.) > (C.P.)$$

(Where: [M.S. = Mental States], [P.A. = Propositional Attitudes], and [C.P. = Causal Properties].)

Q.E.D., Fodor has provided a valid argument. But is it sound? Remember, the argument was designed to give, and I quote, "an account of what the propositional attitudes *are*" in

the first place. Perhaps then, Fodor's theory is beginning to look like it is substantiating the ontological premises of his argument with a tautology (i.e., that mental states exist and have the causal properties he assumes they do). Yet, as logic and even common-sense assures us, tautologies, although true by definition, do not substantiate a non-trivial meaning for their subjects. Therefore, whether Fodor's argument is true or not must depend on some further explanation.

Perhaps this is why Fodor suggests in more recent work that "mental states are characteristically computational...[And, that] mental processes are characteristically computational" [Fodor, 1994, p. 2]. For, if we accept these new claims on the strength of the analogy between computers and minds, i.e., that there really is symbol manipulation going on and that these symbols have intentional content, then it behooves us to accept Fodor's further claim that "if there are no intentional laws, then there are no psychological explanations" [*Ibid.*, p. 3]. And since, *ex hypothesi*, it is impossible to *imagine* a state which has no psychological content, it must also be impossible that psychological states do not exist. Once again, Fodor's initial presumption can be flushed out. For consider: he says "The explanatory paradigm of laws and implementing mechanisms is familiar, and I take it to apply in psychology as elsewhere" [*Ibid.*, p. 8].

6.1 The Jerry-Go-Round

We have thus come around again to the brass ring, which, if the following criticism be warranted, is indicative that Fodor's addition of computational processing has done nothing more for his RTM than reverse the direction of his argument and turn it into a

CTM (Computational Theory of the Mind). And this, as we shall see, bodes poorly for Fodor's theory.

As Stephen Horst recognizes, the conceptual "revolution" of CTM's are problematic because they erroneously presuppose their causal engendering by their own theoretical framework. Horst writes:

Thus one important problem with CTM's (Computational Theory of Minds) account of intentionality is that it turns out to be circular and regressive: circular because it explains the meaningfulness of mental states by appealing to the meanings of symbols, while one must also explain the meaningfulness of symbols by appealing to mental states; regressive because the explanation of any particular mental state will ultimately refer back to other mental states [Horst, 1996, p. 113].

If this is true, and applicable to Fodor's program, then Fodor's program is indeed vicious in its circularity. For it must, inevitably, require us to agree that the propositional attitudes (and all they entail) are real because we do really have them. Hardly a *kind* philosophical program would offer this confusion to its (folk) readership. I thus agree with Horst's claims that Fodor's "representational"²⁷ theory of mind is both circular and regressive.

Now with these criticisms of Fodor's general program laid aside for the moment, let us look specifically at Fodor's interpretation of the causal (intentional) nature of the propositional attitudes. For perhaps I am very wrong about his program and there is in theory, if not practice, a way to ground what I have characterized as the metaphysical prime-element of his common sense explanation. Fodor hypothesizes:

When an ordinary chap says he's doing what he is because he has the

beliefs and desires that he does, it is reasonable to read the 'because' as a *causal* 'because'—whatever, exactly, a causal 'because' may be [*Ibid.* p. 4].

Now before we go any further, I ask merely that you re-read the last clause of the above quote. Done? Okay; so the question which is pressing here, and perhaps against functionalist theories in general, is just exactly how it is that the "attitude" embedded in any proposition, which has such an attitude so embedded, can be judged to have "intentional" force—that is: whatever, exactly, a causal 'whatever' may be. For I agree with Fodor's finely tuned grammatical sense of our common language which has it that "beliefs and desires are semantically evaluable; that they have *satisfaction-conditionals*" [*Ibid.*, p. 5].

6.2 ...And Around

Fodor's familiar strategy will again emerge in what follows, (perhaps this is because it is the only strategy). Only here, it will seek to demonstrate how it is that the propositional attitudes are causally efficacious by running a short series of psychological experiments. Yet these, of necessity, have as their apparatus questions about beliefs (desires, et al.) that bear directly on the outcome of the test; which, of course, is to decide about the functional potency of further beliefs (desire's et al.). This move, I now explicitly charge, is a prime example of assuming what you're trying to prove because these sympathetic experiments are already couched in the very idiom they are arguing to substantiate. Consider Fodor's claim that:

We have, in practice, no alternative to the vocabulary of common sense psychological explanation; we have no other way of describing our behaviours and their causes to be subsumed by any counterfactual-supporting generalizations that we know about. [*Ibid.*, p. 7.]

I am at a loss to understand why Fodor does not recognize this argument as begging the question, and further, as erroneous. For certainly there are other ways, besides appealing to the propositional attitudes, to describe the underlying causes of our cognitive behaviour: this present essay (taken as a whole) being one such example, anything written by the Churchlands' being another, and pretty much the entire literary corpus of current neurobiology being a third. Furthermore, Fodor's argument strategy, even if it were accepted, could be shown to work just as well at supporting folk theories we have long since abandoned--one's which Fodor would be a lunatic to endorse. To demonstrate the problem here, consider if you will how a parallel argument could have been offered by an early cosmologist who was seeking to justify his theory:

Ptolemy: we have, in practice, no alternative to the vocabulary of a geocentric cosmos; we have no other way of describing the celestial behaviours of the spheres and their causes with any other generalizations *that we know about*...Therefore, and on final analysis, terra firma must, indeed, be firma.

Admittedly, I have taken some license with my example; but is that not precisely what Fodor has taken with his? If we agree to accept as a premise that "the syntactical character of thought suggests a view of cognitive processes in general as occurring in a languagelike medium, a sort of 'language of thought'" [*Ibid.*, p. 9], then, by the simplest of sentential

logic manoeuvres, i.e., *Reiteration*, (and perhaps some discreet, tendentious grammatical changes) we can easily prove the suspicious conclusion that: cognitive processes in general occur in a languagelike medium—and support our claim by referring to the unquestionable syntactical character of thought. This method could be aptly named "intuitive realism"—if it's not already. However, as all of us well know, and partly because of other people's mistakes (e.g., Fodor 1990, 1994; the college students from chapter two; Aristotle and Ptolemy, as above) "intuition" is hardly an incorrigible epistemic instrument.

Hence, and on our own final analysis, perhaps the underlying problem with Fodor's approach to mental states (and perhaps folk psychology in general) is that the original assumption made—i.e., that mental states can be neatly explained, and thoroughly subsumed, by a comprehensive mental grammar—simply begs the original question. Indeed, this we have seen. Despite this, however, I do agree that if the Functionalist program could find a way out of this dilemma it would be well on its way to solving many problems about the mind; for it is *intuitively* plausible within the idiosyncratic domain of folk psychology to explain that: when 'I want' a glass of water I get one—because I want one—and accept this explanation as a theoretical cliché.

Nevertheless, to posit that the mind-brain operates by shuffling intentional symbols around, i.e., mental states (types/tokens/the propositional attitudes/whatever), in accordance with a theoretically structured grammar, and, that this grammar engenders the meaning and truthfulness of its terms by appealing to its own laws is, in my humble opinion, to be a prime case of assuming what you're trying to prove. I say this because with all that we know about the origins of language acquisition (see: chapter one) and the evolutionary history of our species, it would seem that our cognitive capacity to learn and manipulate

language in the first place (and especially within the "adaptationalist" context) occurs *prior* to our having a fine-grained theoretical understanding of the grammatical demands of language. And anyways, even if this grammar were construed as being some kind of biologically abstract ability or innate structure which we are not immediately conscious of, then it seems—even more—as if this mechanism would not be a "grammar" at all. Rather, it would be smarter to think of it as a biological phenomenon that would, because of its evolutionary selection, respect the mute and illiterate laws of nature first over the later arriving and definitively artificial "grammatical" ones.

7. An Intentional Interlude

I would like to observe the fact that I have made little direct objective reference to *Intentionality*, and virtually none to *Qualia* so far in this chapter. And, that these two categories are often argued by FP proponents, in the first case, and dissidents in the second, to be indispensable aspects of not only this theory, but of *any* theory which purports to offer a philosophical insight into the nature of the mind. These motifs will be discussed in greater detail in the subsequent chapter (and "qualia" will be explicitly mentioned in the next section).

However, I will note here—as a thematic device and a preamble to *Eliminative Materialism*—that for the same reasons the propositional attitudes are thought by some to be defunct categories of an inadequate theory (i.e., FP), so too might the folk (philosophical) notions of *intentionality* and *qualia* be conceived of as no longer tractable components of a mature and comprehensive theory of the mind. And thus, the privileged

and distinct status these concepts have long since enjoyed might possibly be drawing to a close. We will now examine the claims made under the rubric of eliminative materialism and see if this approach can deliver the actual conceptual object which we have, thus far, seen only in the foreshadows.

8. Eliminative Materialism

Eliminative Materialism is primarily a reactionary thesis against the failures of FP. As a physicalist thesis, the most important feature of it is that it denies, outright, the causal efficacy of the propositional attitudes. Paul Churchland makes this point poignantly clear in the opening sentence of his (1989):

Eliminative materialism is the thesis that our commonsense conception of 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. [*Ibid.*, p. 1.]

Our task in what follows, I suggest, will be divided into three parts. First, we will investigate the origins of eliminative materialism as this will demonstrate that, unlike FP, eliminative materialism has evolved since its inception (and even more to EM's credit, it has done so in a fraction of the time--about 1/100th). Second, we will further examine why FP has been characterized as a story of "retreat, infertility, and decadence" (*Ibid.* p. 7). And third, we will examine some positive evidence for eliminative materialism to see if the defiant claims made against FP are warranted. We thus need to proceed systematically,

refrain from jumping to any hasty conclusions, and discover for ourselves whether this unorthodox approach can shine some light into the "black box" of the functional mind.

8.1 The Genesis of the Eliminative Project

In the past, eliminativism was a more difficult position to argue in favour of--think of the bilious responses to Feyerabend's (1963)--because as a materialist thesis, there was relatively little empirical evidence available to substantiate this position. Indeed, neuroscience in the 1960's had really just begun and was, despite its many learned practitioners, a much less sophisticated and commonly understood affair than it is today. Nevertheless, Feyerabend and Rorty (often cited pioneers of this theory) were able to provide some shelter for the nascence of eliminative materialism¹² with their seminal observations of, for Feyerabend, the presumptive strategy that most of the arguments levy against this theory; and for Rorty, the observation that presently we cannot know what science will discover in the future. Allow me to provide two quotes which well exhibit these insights:

1. A common feature of all the discussed arguments is this: they try to criticize a theory *before* this theory has been developed in sufficient detail to be able to show its power. And they make established modes of thinking and of expression the basis of this criticism. [Feyerabend, 1963.]
2. We have to wait and see what science does, and we may have to then modify our naive expectations...If future science tells us that the Fodorian notion of "mental representations with causal powers" is unreconstructible in terms of supervenience upon certain distinctive physiological microprocesses, then we may decide that the mental-physical distinction is as obsolete as Aristotle's sublunary-superlunary distinction. (But, of course,

that decision will no more hinder us from continuing to use folk psychology than Copernicus caused sailors to stop doing Ptolemaic navigation.) [Rorty, 1970.]

With great foresight, then, Feyerabend and Rorty were challenging the autonomy of the folk psychological idiom. For, with brand new empirical discoveries currently flourishing in the neurosciences, this idiom is being painstakingly criticized for what it might really be: a simple and effective, yet blunt, heuristic device which may no longer be able to carve out its own autonomous niche.

8.2 Eliminating the Obvious

Let us now turn to the explicit reasons why eliminative materialism vies for an outright elimination of the categories of FP rather than just a reduction of these categories to states of the brain and neural processes. Perhaps the most obvious reason is that if 'intentionality' and 'the propositional attitudes' (to name two such categories) can be shelved with other conceptual curios of the past (like: phlogiston, witches, and caloric fluid) then the eliminative materialist no longer has to treat these intrinsically paradoxical categories as problematic for his theory. Instead, he can offer new categories which better fit into our evolving picture of the world, and leave these philosophically vague terms in the past.

We will begin this endeavour with the conception of rationality in which FP plays the pivotal role; namely, the one you, me, her, him--pretty much *everybody*--has grown up with, tacitly accepts, and uses with varying degrees of success everyday. This conception

of rationality, described by our *idiom*, is comprised of a system of inference rules which govern its correct use and informs us when we are making mistakes. Ironically, however, the very mistakes which can be made here are mistakes *relative* to the accuracy of the idiom in the first place. And, since our present examination is questioning, among other things, these fundamental tenets, it seems like it would be a fool's wager to simply accept as bona fide the idiom which supports the causal efficacy of (purely mental) mental states: simply because this idiom is what we have been indoctrinated with and one whose accepted sets of inference rules currently reigns supreme.

The fact that we do indeed use this dualist idiom is (as I hope to have shown in chapter one) a result of much of the philosophy which has been promulgated by notorious figures in the past and reiterated down the generations. Once again, Feyerabend makes this point well when he compares what a reversal of philosophical fortunes might have yielded; yet he comes to the same conclusion. With respect to the choice of idiom which is in use he writes:

This is an irrelevant historical accident. Is it really believed that a vigorous propaganda campaign which makes everyone speak the materialist language will turn materialism into a correct doctrine? The choice of the language that is supposed to be the basis of criticism must be supported by better reasons. [Feyerabend, 1963, p. 4.]

We can then, I am claiming, leave behind the "argument from meaninglessness"²⁹ that many philosophers have tried to undermine eliminative materialism with. I say this because, and not to put too fine a point on it, the theories we use to characterize what real ontological items are is revisable. Therefore, when proponents of FP, most notably

functionalists, argue that the "intentional" aspects of internal states cannot be reduced to a naturalistic theory of our physical substrate, and this *because* it is precisely the abstract functional features of "intentionality" which make a person a person in the first place—and not the electro-chemistry of his brain—we can suspend our judgement. This is a prudent suggestion, because we are here interested not in the correct application of grammar from our own historical epoch, but in the underlying causes of the cognitive behaviour of our conspecifics.

Thus, the rationale behind the eliminativist program. And further, why the proponents of this view argue that an elimination of the propositional attitudes might clean up our cognitive ontology. Perhaps, then, this theory might be beginning to display some credibility. For, if we accept that FP is a theory, and moreover a "folk theory" (much like, say, folk physics, folk astronomy, or alchemy), then we might just entertain the thought that it very well could come to the same end as these erroneous conceptual models; that is, it might be reduced or eliminated in favour of a better, more advanced theory. Consider now, and in this light, P.M. Churchland's prediction that FP "will be eliminated, as false theories are, and the familiar ontology of common sense mental states will go the way of the Stoic pneumata, the alchemical essences, phlogiston, caloric, and the luminiferous aether" [1979, p. 114]. It is true that this approach to FP spells doom for the propositional attitudes. For if correct, talk of "autonomous mental states" or "intentional computation" might come to be viewed as just leftover idiomatic expressions from a pre-neuro-scientific time, which would, indeed, be short shrift for the attitudes.

8.3 Gathering Our Wits

Thus far in our investigation there has been little said about the positive claims of eliminative materialism. This is due mostly to the fact that, as a "reactionary thesis," eliminative materialism was not expressly designed to offer up a new idiom for scrutiny. Rather, and as its name implies, eliminative materialism is an argument *against* those categories which are said to name causally efficacious yet non-physical entities in their "physicalist" explanations of human behaviour and experience. For after all, the chief goal of the eliminativist program is to examine whether the categories of FP actually denote anything, despite what connotations these categories embrace. Moreover, I note, few positive claims have been made in support of the idiom which eliminative theory defers to because the creation of this idiom is currently in the works. Consider how people who are credited with "good minds" these days are often paid the compliment of being "cerebral," or less formally, "brains."

Despite the above disclaimers, however, there is some positive evidence in favour of eliminativism. Admittedly, this evidence comes from the neurosciences, and so the interpretation of these facts may still be dismissed as irrelevant by those committed to the ontological division of bodies and minds. For my own part, though, I am not so committed and thus think this division is an artificial one. Therefore, I have no qualms about including the following information as if it were, in fact, evidence that the human brain (and all, and only, its physical processes) is the sole producer of what has heretofore been characterized as "purely mental" mental states. This sentiment may appear to be over ambitious in scope, but as John Dewey once said "every great advance in science has issued from a new audacity of imagination." [Dewey, 1929.]

9. Rocket Surgeons and Brain Scientists

Despite the title of this section I will not be speaking of Rocket Surgeons, this is much too advanced. Rather, I will focus on some recent empirical discoveries made by the latter discipline; discoveries that would have challenged the imagination of even the most scientifically oriented philosopher of one hundred years ago. I want to be brief, here, and so I will provide just three small points of, what I consider to be, interest.

1. Anatomy: The exceedingly intricate anatomy of the brain is just beginning to be figured out. But even at this early stage it has been recognized that "The human brain is by far the most complex structure in the known universe" [Thompson, 1985, p. 1]. This emphatic claim is supported in part by the discoveries the Golgi³⁰ staining method has made possible. These discoveries show us that the brain is an organ made up of about one hundred billion individual, and self autonomous, nerve cells (neurons³¹) which, in each instance, can form up to about three thousand synapses with other cells. The resulting number of fibres has been estimated at about one hundred trillion (that's 100,000,000,000,000) individual information carrying strands. This is quite an enormous interactive capability considering that the estimated number of molecules in the physical universe is 10^{89} . Compared then to linguistic potential--the twenty six (that's 26) letters in the alphabet, digits from one to ten, and a few key punctuation marks that make up the constitutive elements³² of the English language--this number of potential neural connections between informationally encoded *symbols* (at least 100,000 different genes in each chromosome, in each cell) begins to make even the richness of the English language

look pale and wan.

Perhaps, then, Fodor's sentential "language of thought" may now really appear to be a paltry conception of how the brain communicates with itself. I say this because Fodor's conception of cognitive activity is modelled on a linguistic exchange; an exchange in which, presumably, one "language of thought" sentence *describes* a specific state of affairs, abstract or otherwise. By contrast, the neural model is predicated on a much more complex exchange; yet an exchange in which, presumably, one fundamental "bit" of awareness is *realized* in a neurophysiological process.

2. Brain Function: There can be no doubt that the brain is at least the master control area for somatic functions. Indeed, it has been discovered that the autonomic nerves (the ones which connect to the heart, stomach, respiratory system, etc.), are connected at the other end to the hypothalamus. This could be construed as a trivial "physical" point. However, it has also been evidenced that the hypothalamus is the master control area for emotion, because:

Electrical stimulation in one part of the hypothalamus can elicit full-blown rage and attack behaviour in humans and other animals. Stimulation of a closely adjacent region of the hypothalamus produces feelings of intense pleasure. [Thompson, p. 16]

This discovery seems, to me at least, to lend plausibility to the claim that it is not an "abstract relation between thoughts" which is responsible for the creation of "mental" states, but rather, a very real and empirically detectible (see: point 3.) stimulation of certain

parts of the brain which is responsible for causing not only bodily behaviour, but for causing the experience which has been often described as a subject's "intentional" self motivation.

Another interesting discovery I wish to note regards the "qualia" of a perceived image. When a photon (light ray) bombards a rod this retinal structure changes its actual shape. The information which is produced from this phenomenon is then called "all-trans-retinal" and this physical process occurs in a few pico-seconds (pico = one billionth). And so, the basic process here is the speedy conversion of a photon--a light quale in philosophical terms--into atomic motion [Kuffler, et al., p. 248]. (Perhaps this discovery will remind the reader of how heat, formerly thought to be a substance, was discovered to be nothing more than the motion of the particles which made up the heated body itself.) Furthermore, there are millions of rods in the retina (some are activated by red, some yellow, etc.) and so the image that they collectively produce, and then send down the optic nerve, is already possessed of a certain character. Thus, and for example, "what it feels like" to see the colour azure is already, and in large part, "decided" before this information ever reaches the brain proper and becomes what is often referred to by FP as an intentional *mental state*.

3. Interpretation: Admittedly, the above information may still be considered trivial with regards to our discussion because, some might say, we still cannot look at someone's brain and know what they are thinking. However, this rebuttal is no longer as durable as it once was. For with the new techniques being developed in the neurosciences using PET (Positron-Emission Tomography) scanners, the functional neuroanatomy of the human

brain, and how this relates to a subject's "thoughts," is beginning to be revealed. If one looks at the images these machines produce one can see, and easily delineate, the different parts of the brain that are active in such varied tasks as require: visual, auditory, cognitive, memory, and motor faculties. [see: P.S. Churchland, 1986, p. 218-221.] Thus it may still be beyond the reaches of science to tell whether an observed subject is thinking 'cat on the mat' or 'cat on a hot tin roof' but with the overwhelming explosion of new discoveries being made, this may one day be in our reach. (Consider how unimaginable space travel was to the denizens of the 1890's.) This positive outlook on neural discovery is well articulated by Paul Churchland when he notes "In short, the greatest theoretical synthesis in the history of the human race is currently in our hands, and parts of it already provide searching descriptions and explanations of human sensory input, neural activity, and motor control" [1989, pp. 8-9].

Thus I hope to have provided some positive evidence for the materialist position I am advocating in this paper. I certainly do not think it is definitive nor resolute proof that some form of materialism is *ultimately* correct--but this has never been my goal. Rather, what I hope to have presented with the above arguments is that materialist explanations, whether "reductionist" or "eliminative" (although I think the latter is the logical outcome of the former in most cases), are not impoverished theories created by over zealous thinkers, but are instead, serious contenders for enlightenment in the philosophy of mind.

Now as a final word from this chapter, and an introduction to the next one--where we shall consider the persistent denial of materialist theses on the grounds that mental items like "intentionality" or "the propositional attitudes" are irreducibly non-physical--please indulge me in an anecdote. It is brief in any case.

10. The *Reductio ad Solutio*

Here, then, is my "reductive/eliminative" anecdote which, I think, sheds some light on this difficult subject. For, if you still have doubts about whether it is the propositional attitudes or some other "purely mental" mental state which *cause* cognitive behaviour or whether, ultimately, it is a strictly physical cause--i.e., you're not sure whether to buy this eliminativist sophistry or not--then please consider "A Telephone Conversation."

For the sake of argument, let's say that you are having second thoughts about how plausible the eliminative materialist account of mental states is. You decide that if you could ask me a few difficult questions about this thesis proposal, some really *wicked* ones, like: does this mean that emotions are to be reduced to just electro-chemico-neurophysiological states of the physical brain, does language then not really mean anything, how can the propositional attitudes be construed as causally inert, and--or, just simply--are mental states just brain states (with no metaphysical proviso)? So, you pick up your phone, you dial my number, and proceed to *fire away* as it were. Now, I am supposing that you have a workable notion of how your telephone reduces your *intentional* speech to electrical processes; that is, how your phone-cord delivers information from the electronic gizmo inside my hook, to electrical information inside yours; and that once there, onto the speaker in your ear-piece where it is then re-modulated and amplified. And furthermore, that this workable notion has something to do with electrons flowing through several thin strands of utterly un-fantastic copper wire (the "mysterious" causality of a digital phone may be an even better analogue). Now, if *my* propositional attitudes can be reduced to, or redefined as, physical processes--I mean here the "intentional" and

emotionally charged answers I respond to your questions with (via electrons flowing through to your telephone: and you then become aware of my *feelings, desires, beliefs*, et al.)—then, simply put, why can't yours? For, *ex fortiori*, and to entertain a physiological point, even when we communicate our hopes and fears face to face there is a medium between us: the environment (my internal states; which of course are external to yours) and your own internal relations (or "propositional attitudes"); and this medium is, namely, the activity of your own sensory modalities.

12. Onward

Perhaps this entire chapter has been in vain. For there might still be those who obstinately refuse to accept that our idiom is fallible when it comes to describing "mental" phenomena. And thus, they may claim that even though the mind is most certainly in the head, there is, nevertheless, an "irreducibly subjective" quality about internal states which defies explanation in physical terms. Now to be sure, there seemingly is a unique character about witnessing the world from the subjective point of view (unique, perhaps, because a "subjective" perspective is the only way in our idiom to describe 'the act of witnessing'). Nevertheless, I contend, this does not entail that that which is subjective must also be irremediably non-physical. We will thus move on to an investigation of the claim that subjective priority is a necessary component of any legitimate theory of cognition; because if this claim can be shown to be inconclusive, there may just be a plausible way to "wake from their dogmatic slumbers," the folks who, on the present account, have been lulled by the dualist mantra.

THE PERSISTENCE OF THE MANTRA

1. Tuning

Thus we have arrived at perhaps the most difficult task any materialist thesis faces: namely, to describe how subjective experience might be explained, in physical terms, without ignoring the "what it is like" aspect of such experience. This is indeed a difficult task; because the very nature of "subjectivity," as defined by our common idiom, is that it is a decidedly mental phenomenon which has, due to this characterization, no relevant description in physical terms. Despite this orthodoxy (what I am referring to as "the Mantra"), however, there might just be a way to describe how the brain produces what we have heretofore called 'the mind' which does not, ultimately, rely on some ineffable property, or at least abstract process, to describe this subjective phenomenon. My suggestion then, to begin, is that we do not forget the revealed errors of intuition and past metaphysical judgments we have observed up to this point, for this will remind us of the positive developmental process our theories are thankfully a part of.

Now in the pursuit of a materialist finale, I would like to consider two notable authors who stand fast against the possibility of physicalism. Frank Jackson and John Searle are these two authors, and they have been selected because they deliver their arguments with the savvy and persuasiveness we must seek to undermine in the following pages if we

are to have any hope of preparing our credulity for the conclusion which will ensue.

2. Jackson: Some Epistemic Concerns

Frank Jackson's "knowledge argument" is a fine example of how certain intuitions about the irreducibility of the subjective, and the epistemic necessity of this "non-physical" phenomena, can be expressed convincingly, and in so doing, wield a serious challenge against the thesis of physicalism. In his essay "What Mary Didn't Know," (1982) Jackson tells the story of a brilliant thinker who has been confined to a colourless room since birth. As we find out, all of the staggering intellectual resources that this thinker has at her command have been learned through lectures relayed on black-and-white television screen and the regular appearance of words on the printed page; at any rate, her life has been wholly absent of colour. Jackson further tells us that:

She knows all the physical facts about us and our environment, in a wide sense of 'physical' which includes everything in *completed* physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon all of this, including of course functional roles. [Jackson, 1982.]

The story gets a little more imaginable when Mary is let out of her room and experiences colour for the first time. Jackson's opinion, and indeed the common intuition, is that Mary will *learn* something new when she experiences the colour of, say, an apple. However--and here's the rub--if Mary has genuinely learned something new (and as we have noted, she "knows all the physical facts..."), then obviously physicalism cannot be the entire story;

for there is some detail which physicalism cannot explain: namely, what *sensing* red is like to an observer. Therefore, because Jackson supposes of the physicalist thesis, that "complete physical knowledge is complete knowledge simpliciter" [*Ibid.*], then his argument seems to be directed at providing the reasons why complete physical knowledge is not complete knowledge simpliciter. If this project seems to be a little confused, with reference to what certain terms mean, then I propose that we have already sensed a problem with Jackson's theory. Let us see what happens when this problem is exploited.

3. Levine: On Equivocation

One objection to Jackson's resolute claims can be found in Joseph Levine's article "On Leaving Out What It's Like" [Levine, 1989]. Here, Levine argues that Jackson is relying on the equivocation between one's knowledge of an event and the actual existence of this event; or as Levine says of Jackson "[he is] using epistemological premises to support a metaphysical conclusion" [*Ibid.*]. What this objection amounts to, when unpacked, is that Jackson has conflated two senses of the term 'physical information' and used the resulting ambiguity to explain how Mary's case shows that because some information is not physical information, physicalism must be false. Now in order to clarify Levine's argument let's get clear what these two senses of "physical information" are. Levine writes:

In one sense this ["physical information"] might mean information expressed in terms used in the physical sciences. In another sense it might mean information about physical facts, processes, etc..[And moreover]...It is only

in the second sense that any reasonable physicalist is committed to the claim that all information is physical information. [Levine, 1997 p. 545.]

Levine's point here, of course, is that Jackson leads us to believe that Mary gains both types of information upon leaving her room and has thus learned some new detail about the world which is not physical, and therefore, physicalism is false.

There is, however, a problem with Jackson's argument. For consider, Jackson has built into his melodrama the fact that Mary already has all of the physical information possible (read: "complete knowledge simpliciter"), and so any new information she gains cannot, on either pain of contradiction or the failure of physicalism, have any bearing on the body of physical knowledge she possesses. However, this is where the ambiguity of Jackson's argument has its effect. For, and to be sure, there is new information that Mary gains upon witnessing colour for the first time; yet, although the information that Mary gains is not of the first type (because it is not literally articulated in terms used by the physical sciences), this is not equal to the proposition that Mary does not gain physical information of the second type--for she does learn something new about the world. Hence, what Jackson has hoped to show with his scenario is that because Mary has gained new information which is not "physical information" (read: "physical information of the first type"), then the information she has gained must be non-physical. However, it does not follow that because Mary has failed to acquire information of the first type (i.e., information "expressed in physical terms") that the information she does gain must necessarily be non-physical. Quite to the contrary, Mary does gain "physical information" (read: "...of the second type.") because she learns what it is like to experience a physical

object; namely, an apple.

Thus Jackson's argument against physicalism appears suspect here because he has equated the epistemic access one has to an object with the actual (or physical) reality of the same object. His argument, then, trades on the ambiguity between two different senses of "physical information." This is a very useful insight because it makes us aware that much of our knowledge, or "physical information," about the world is not gleaned from knowing what it is like to, for example: experience sublimation directly, occupy the state of an electron, experience someone else's pain, or to undergo the process of photosynthesis from the first person perspective--yet, and to once again be sure, all of these phenomena are certainly "expressed in terms used in the physical sciences." I suggest, then, that what we have seen so far of Jackson's revealed realism about qualia and his ensuing argument against physicalism do not yet provide sufficient premises for the acceptance of his conclusion, which is, of course, that physicalism is false.

Furthermore, I think that the premises of Jackson's argument can be shown to endorse the speculative, and often eschewed, notion of "epiphenomenalism" (more below), because their positive alignment with the priority of subjective experience also certifies this experience as non-physical. Therefore, let us now look at another critical approach to Jackson's "quality" argument, as this will further illuminate the reasons why the assumption of subjective priority might not, after all, provide such a safe haven for metaphysical keepsakes like qualia.

4. Qualia Schmalia: Dennett's response

With all due respect to Mary, Dennett begins his criticism of Jackson's argument by questioning whether it is even possible to imagine what "knowing everything there is to know about the physical nature of the world" [Jackson, 1982] would be. Dennett concludes that this is such a "preposterously immense [task, that] you can't even try" [Dennett, 1991, p. 399]. Indeed, it seems rather implausible that even the most astute and sympathetic reader of Jackson's argument could honestly conceptualize anything close to "knowing everything." Perhaps, then, Jackson's premise that "she [Mary] has *all* the physical information" can once again, yet in a different manner, be revealed as a suspect claim; because this state of affairs is, although mentionable, not even remotely imaginable.

Allow me to clarify the claim above. It is true that we can invent a name, or utter a sentence, which presumes to denote a certain state of affairs (much like the practice of our superstitious ancestors). However, this does not mean that what we describe in our idiom has, just because of its description, any real ontological importance (despite its cultural lore). For example, think of the Id, a flying pig, or "Hamlet." Therefore, even though it is possible to imagine that Mary knows an awful lot about physics, and even that she is the leading expert on every current physical hypothesis in the contemporary canon--and especially insightful when it comes to the neurophysiology of colour vision--this is not equal to knowing "all the physical facts." It seems correct to say, then, that the vast knowledge we can presently imagine Mary to possess is simply inconsistent with knowing *all* the implications of having *all* the facts of a completed physics would be; because "knowing all the facts" entails that one has experienced them. Why? Arguably, because "knowledge" requires at least "understanding" and understanding, in this case, requires sensory familiarity with the phenomena.

I think Dennett's criticism is warranted, then, when he says that Jackson has fallen prey to "a classic provoker of Philosophers' Syndrome, that is: mistaking a failure of imagination for an insight into necessity" [Dennett, p. 401]. If we do accept Dennett's criticism, here, then we should be wary of Jackson's methodology. For it seems disingenuous to suppose that just because Jackson can frame his hypothesis about "physicalism" in a set of sentences which assume to carry the semantical weight of describing an absolute state of affairs--whatever that would be--that he can then go on to "discover" that this state is not absolute after all.

Of course one might not be moved by Dennett's argument; because regardless of whether Mary truly knows everything about physics or not, she would still, without a doubt, learn something new about colour upon seeing it for the first time. However, this rebuttal, I suggest, would be to return to the ambiguity of Jackson's original premises; because Jackson's premises would have us believe that Mary *could* know everything about colour without actually knowing what it is like to observe it. Hence the logical structure of Dennett's criticism is valid. (We will investigate the problem of "subjectivity" below. See: sections 5 and 6.)

4.1 Epiphenomenalism

The above criticism of Jackson's epistemic conjuring trick is thus arguably solid; yet Dennett goes on to out perform it with his exploitation of Jackson's commitments³³ to "epiphenomenalism." To begin, the charge of ambiguity is again levelled. Yet levelled against those theories which employ this otiose term *epiphenomena* in the attempt to save

certain intuitive appearances. However, part of the problem with the use of this term is that it is widely assumed to have a fixed meaning in the philosophical vernacular. Yet as we shall soon see, if this were the case then those who are sympathetic to the use of this term to describe mental states are either bottom-line dualists (and thus on my account "incoherentists") or simply whimsical about the importance of logical consistency.

So let's get it straight. To be consistent, Epiphenomena must define one--and only one--of two possibilities when discussing the reality of mental states. Either, it must describe the by-products of brain activity as non-physical properties, or, it must describe an "effect" which itself has no effects in the physical world. The first definition names things like the *sound* of a steam whistle: the noise is empirically verifiable, yet it plays no causal role in the operation of the whistle. The second definition names "things" like *phenomenal experiences*: these are purported to be non-physical qualities (qualia) which are, because of their *rare* constitution, empirically undetectable.

Now, let's return to Jackson's 'knowledge argument'. Jackson has said that Mary (our brilliant scientist) already has possession of all the physical facts about colour perception, but nevertheless *learns* something new upon leaving her confinement. Hence, if Jackson wishes to avoid for his argument the charge of inconsistency, then he must say that the detail Mary has learned is epiphenomenal in character; which he does. Here's the problem. If Jackson is employing this term in the first sense, then Mary would have known all about the effects of her "reactive disposition" with respect to her neurological abilities for sensing the colour red, and thus, would not learn anything new--she would just experience something new. Yet, if Jackson is employing this term in the second (philosophical) sense, then Jackson must admit that the experience of sensing red could not

itself be considered a physical effect, and thus, could not cause any alteration to the state of her brain and central nervous system.

(As an aside, and with respect to the propositional attitudes, it would thus seem that if the subjective experience of sensing red were epiphenomenal in *either sense*, but which nevertheless could cause beliefs like 'I am sensing red', then beliefs too must be epiphenomenal (in either sense). And, to make the point all the more regrettable for Jackson, if beliefs were epiphenomenal then there would be no way that they could play a causal role in the exchange of ideas, as this exchange relies on communication; which is always, unless you are a telepath, embedded in an empirically detectable medium--that is why it is communication.)

Therefore, if qualia (or anything else for that matter) are epiphenomenal in the first sense, then Jackson's argument is no threat to materialism and can be discharged from the problem list. But, if Jackson conceives of qualia in the second sense then, by definition, they could never have any effects in the world at all; everything would occur exactly the same way whether or not they were included, and thus, they are ontologically insignificant. Therefore, it is my contention, that to continue with any more talk of epiphenomenalism as being a refuge for qualia, or subjective experience, is not only moot, but ridiculous.

There is yet more to be said about how certain arguments which support "qualia realism" pose a challenge to any thesis of physicalism. Unfortunately, however, it has begun to look as if Frank Jackson is taking all the criticism for philosophers en masse who are sympathetic to any account which endorses the ontological realism of metaphysical phenomena. Thus I will make my apologies, here, and continue under the assumption that my criticisms about (excessive) metaphysical realism, heretofore and henceforth, extend to

anyone who describes subjective experience being in any way-shape-or-form, necessarily unavailable to the possible reaches of (broadly construed) physical theory.

5. P.M. Churchland

In an essay titled "Knowing Qualia: A Reply to Jackson" Churchland further questions the commitments Jackson must make to support his "knowledge argument." We will pass over the section where Churchland proves, canonically (that is, by the rigid application of the formal rules of logic), that Jackson's argument is supported by an equivocation of the term "knows."¹⁸ Rather, the specific part of Churchland's essay which I think is more pertinent to our discussion is the final section, headed, "Converting a Third-Person Account into a First Person Account." This section is, respectively, not aimed directly at Jackson's argument, but rather, and perhaps more menacingly, at the ideology which supports it. Churchland writes:

That ideology includes a domain of properties--the qualia of subjective experience--that are held to be metaphysically distinct from the objective physical properties addressed by orthodox science. [*Ibid.*]

Now so we do not confuse "that ideology" with some kind of *foreign* conceptual scheme, allow me to remark that the purported domain over which this ideology operates is exactly the framework of contemporary folk psychology. Accordingly, Churchland proposes that we recognize that this idiom is what "[w]e all of us, as children, were taught to use" [*Ibid.*]. It would follow, then, that we have relatively little choice in how we think of

subjective experience and other mental phenomena; and this restriction, I find, is prohibitive and irritating.

However, there are remedies for this irritation. Churchland has suggested that we simply accept the limitations of our current conceptual scheme which is intuitively committed to the dualist idiom; but that we accept these limitations not as epistemic absolutes, but as revealed flaws in our conceptual scheme. To be sure, what Churchland is offering us is a new way of understanding how (what we call in our idiom) "subjective mental states" do not, of necessity, require a bridge law to span the infamous "explanatory gap." Instead, he advises, we should look towards the young, yet ever developing conceptual resources of neuroscience; for here there is at least no ontological confusion.

Admittedly, neuroscience may presently appear to lack a certain explanatory capacity (as Churchland says "in providing a purely 'third person account' of mind"), but this too may be overcome. To begin, and to challenge the autonomy our idiom has presumed with respect to subjective explanation, Churchland suggests that:

What makes an account a "first-person account" is not the content of that account, but the fact that one has learned to use it as the vehicle of spontaneous conceptualization in introspection and self-description.
[Churchland, *Ibid.*]

For some this approach would certainly take some getting used to. This, I suspect, is because it is hard to imagine what giving a description of one's internal states would be like if it was not couched in the vocabulary of our familiar idiom--an idiom which is committed to explaining "internal" states as necessarily "mental" states. Moreover, Churchland's

approach may be avoided because it makes one have the dawning (and perhaps uncomfortable) realization that one's thoughts may not be purely of one's own design--in so far as they can be construed as the products of electro-chemical neural activity, and not, the result of some kind of intrinsic mind-stuff.

Now perhaps you may be asking yourself what this "new" idiom would *feel* like to use, and so I provide a possibility as Churchland sees it:

Given a deep and practised familiarity with the developing idioms of cognitive neurobiology, we might learn to discriminate by introspection the coding vectors³⁵ in our internal axonal pathways, the activation patterns across salient neural populations, and myriad other things besides. [*Ibid.*]

And so once again, yet for explicit reasons, this new idiom may appear strange, awkward, and certainly counter-intuitive. However, *that* type of response, if we have learned our lesson from history, has never offered sound reasoning for the expulsion of a potentially great theory. Indeed, just as so many other theories throughout our conceptual history have risen from the ashes of scorn to become adored enlightenment, so too might there be the potential for philosophy, when coupled with neuroscience, to provide us with, as Paul Churchland predicts, "a grand reconstruction and expansion of our subjective consciousness...that, unlike folk psychology, is at least equal to the kinematical and dynamical intricacies of the world within" [*Ibid.*]. Now admittedly, this may be waving the banner of materialism a little too vigorously, but surely something has to be done to distract the mind's eye from its orthodox glaze.

5.1. An Aside About Prerogative

To recall, the term 'impetus' has been formally, and legitimately, eliminated from the scientific canon because this term has been shown to denote an ontologically insignificant force. Nevertheless, this term continues to be used by those who express their ideas about force and motion in what we might call "folk physics." Therefore, even though 'impetus' has been redefined as an anachronism, it still enjoys "metaphorical" clout in folk theories about the causes of sustained motion. By analogy, then, this is perhaps what has unwittingly occurred in the philosophy of mind, because metaphysical categories (like "mental substance") have largely been eschewed as naive antiquities as their purported denotata no longer pick out any relevant phenomena. Despite this, mental categories continue to be used in our idiom. Thus, even though *this* ontologically distinct realm has been revealed as an anachronism, another metaphysical realm is still deferred to as naming a place from which "intentional" clout draws its potency. This ambiguity, I conclude, is a result of the inculcation of the dualist idiom into our conception of ontology. And this conception is one which, because of its strong ties with the priority of the subjective, has tacitly accepted that 'experience' names an ontologically distinct class of existence.

It may become apparent, then, that even if science were to prove materialism correct--and it was in fact correct--this would not entail that everybody would accept it immediately and start expressing themselves in the categories that this new and corrected ontology countenanced. What is more likely, if this discovery were to be made, is that many people involved with the description of subjective phenomena would accept the new knowledge of the brain as playing an intriguing role, but that nevertheless, these discoveries

only point to the fact that there still must be something else constitutive of subjectivity. The reason for this, perhaps, would be because the idiom in which we have learned to speak about "subjectivity" has been in place for millenia; long before there was any rational explanation of how a poorly understood bodily organ could organize and exchange information. Therefore, a good way of describing how our idiomatic understanding of mental phenomena is still viable might be to suggest that there were heretofore undiscovered connections occurring in this organ which, though freshly discovered, are still metaphysical in their constitution. Let us see if this is a defensible position.

6. Searle: Rediscovery of the Idiom

This section aims at appraising Searle's model. Before we begin, however, I think it wise to provide a definition of a term which Searle makes frequent use of. This should have the clarifying effect of dispelling any ambiguity over how this term is applied in the discussion which follows. For, as we saw above with Jackson's argument, when ambiguity is overlooked in this realm of discourse virtually any hypothesis can be tendered, whether perspicuous or not.

6.1 Emergentism

The doctrine of emergence is an explanatory category—a *theory*—which is used to describe how certain properties of whole objects, properties which are not deducible from the constituent elements of their physical substrate, nevertheless become manifest under

observation. Thus the "wetness" of water is a resulting emergent property from the combination of hydrogen and oxygen. The relation of emergence has been expressed otherwise by using a familiar altitude metaphor in order to describe, as Nagel puts it, the "consequent occurrence of properties at 'higher' levels of organization which are not predictable from the properties found at 'lower' levels" [1961, pp. 366-367]. To put it very simply, 'emergentism' names the process by which certain objective properties alluded to in the cliché, "the whole is greater than the sum of its parts," get realized.

However, there is a danger with explaining emergentism this way; one which can cause confusion. For the fact that certain observable properties arise unpredictably from their constituent elements does not entail that these "emergent properties" have any intrinsic ontological status of their own. Now although I do not take any credit for confusing anybody, it is precisely this ontological conclusion which has been arrived at, and argued for, by many contemporary thinkers. Despite this support for the ontological reading of emergentism, however, it seems prudent to me to recognize that this type of reification--when and if it is tried--is perfectly analogous with the naive conclusions arrived at by certain predecessors who, as we have seen, thought that 'heat' was an actual substance; that is, until it was discovered that this emergent property could be fully explained in the "lower" level vocabulary of statistical mechanics.

So what then is the correct application of emergence? Well, on the original account:

To say of a given property that it is "emergent"³⁶ is to attribute to it a character which the property may possess relative to one theory or body of assumptions but may not possess relative to some other theory. Accordingly, the doctrine of emergence...must be understood as stating certain *logical* facts about formal relations between statements rather than any experimental

or even "metaphysical" facts about some allegedly "inherent" traits of *properties* of objects. [Nagel, p. 369.]

This is a lengthy quote, but worthwhile because it clarifies that "emergent properties" are not to be construed as the same "things" as the properties which many forms of dualism certify. Let us now see how the notion of emergence can be used to support claims of dual-ontological existence.

6.2 The reification of theory:

Searle calls emergent properties "system features." And the ones which cannot be deduced from the base constituents alone he calls "causally emergent system features...[as these]...have to be explained in terms of the causal interactions *among [my emphasis]* the elements" [Searle, 1992, p. 111]. We can thus infer, *mutatis mutandis*, how the finer points of emergentism can wreak havoc on any physicalist conception of what mentality is. For intuitively, it seems quite absurd to say that mental states can be directly identified with individual neurons. However, it seem less absurd, or maybe even downright plausible, to say that a set of neurons--maybe the entire set?--give rise to something else, and that this "emergent" something else is precisely what is commonly referred to as a mental state. This is the assumption Searle has taken for granted with regard to his explanation of mental phenomena in terms of 'emergent properties'. Yet because he has reified these properties as having an autonomous nature, and then assumed ontological identity, it becomes obvious that he has conflated the idea of property emergence with the idea of ontological genesis.

To be sure, an emergent property is not the kind of thing which has its own intrinsic nature; rather, it is an observable outcome, or characteristic, of the relation between two or more discreet entities.

Consider now Searle's seductive first words of chapter one (1992).

The famous mind-body problem, the source of so much controversy over the past two millennia, has a simple solution...Here it is: Mental phenomena are caused by neurophysiological processes in the brain and are themselves features of the brain...I call it "biological naturalism." Mental events and processes are as much part of our biological natural history as digestion, mitosis, meiosis, or enzyme secretion. [1992, p. 1.]

At first pass this articulation seems good. Searle has located the mind in the brain and has explained it as being a product of neurophysiological processes. One might thus conclude that Searle intends to go on and explain how mental states can be reduced to physical states, for he claims (on page four) that "consciousness, in short, is a biological feature of human and certain animal brains," and further (on page 84) that "The main aim of this chapter [chapt. 4] is to locate consciousness within our overall 'scientific' conception of the world." These claims seem to gesture at a physicalist conclusion.

However, and problematically, Searle also wishes to defend the position that mental states are an "irreducibly subjective" (non-physical³⁷) system feature of a properly functioning brain. I will spell out the problem: science does not countenance an ontology which is non-physical. And thus despite the fact that Searle believes himself to be arguing consistently, there does seem to be a problem. The problem, to lay it out, is that Searle wishes neither to endorse: behaviourism, the I.D. theory, functionalism,³⁸ materialism,

nor any form of the reductionist program; and furthermore, he believes that all forms of dualism are profoundly mistaken. Nevertheless he maintains that mental states are both features (read: properties) of brains, and, that they are not physical states. This is where the "rediscovery of the idiom" occurs. For once again we can find the theme of *irreducibility of the subjective* throughout his book (and specifically the title of section III chapter 5), and discover that the way Searle wishes to substantiate this subjective obstinacy is by the employment of the concept of emergentism. However, as we have already familiarized ourselves with what this relation between mind and brain amounts to, let us now investigate whether Searle is correct in his claims that 1) mental emergentism is irreducible to its physical substrate, and 2) that he (Searle) is not a dualist.

We shall begin by noting that for Searle the relation between mind and brain is causal and that this causality runs both ways; that is, that one's brain states cause one's mental states and that the reverse of this is true as well. Moreover, as we shall recall, Searle wishes to stand fast on his position that mental states are certainly part of reality yet do indeed have an independent metaphysical status (they are non-physical).

With respect to Searle's two claims, then, we can see how he has been influenced by Jackson's argument.³⁹ We shall not return again to the epistemic aspect of the above argument, however, because Searle is committed to making an ontological point. Indeed, Searle accepts that when a subject (say you or me) experiences pain, there is an order of ontological reality present in the world that has no magnitude, but which is nevertheless caused by the neurophysiological processes occurring in someone's brain.

To deal with this confusion smartly let us identify Searle's tacit dualism. By saying that (subjective) mental states are irreducible to physical states because of their unique

ontological constitution is to posit a separate realm of existence; and to posit such a realm is to endorse dualism. But why, if this is the correct characterization of Searle's position, is he so confident that he has got it right? (Especially since he has claimed that dualism is profoundly mistaken.) We have seen how Jackson's argument failed to make the metaphysical point it was designed to, and we have seen how Searle's program fails when "emergent properties" are suspected to lack the ontological status which he ascribes to them, and so to merely cling to the position that there is a real ontological division between subjective mental states and objective physical states at this stage of our argument would, perhaps, beg the original question. What we need, then, is a way to describe the necessary immediacy of subjective experience which can also explain why this phenomenon need not be construed as non-physical. We shall thus turn to this task presently.

7. The "Might" of Materialism

Perhaps the best way to commence with one's articulation of what might be called "subjective materialism" is with a bit of humility. For after all, is it not possible that one's intuitions about the underlying nature of subjectivity could turn out to be false? I think we have already asked this question and admitted that intuition is a corrigible epistemic device, and as such, needs to be continually augmented and revised for it to retain its efficacy on the one hand, and to remain consistent with the evidence of new discovery on the other.

For indeed, and given what we now know about the sheer complexity of the brain, if subjectivity were the final arbiter it is often celebrated as being, then one might think that this "unified" epistemic modality should be able to "observe" what is actually going on in

the brain; and not just the resulting effect these myriad neural combinations produce. However, if subjectivity were so equipped (and was in fact the unified entity it is often supposed to be) then one's awareness of internal phenomena would be apprised of not only all of the executive decisions one's brain makes with respect to bodily functions (like: 'blink now', 'breath now', 'beat the heart', 'produce more red blood cells', 'etc. '), but also, this awareness would be able to accurately chart the billions of individual synapses occurring at each and every second, each and every day, sleeping or awake, during one's entire life. Now obviously, this is a "preposterously immense task"--reminiscent of Jackson's hypothesis that someone could know all the physical facts (including functional roles), yet learn something new about subjectivity. Hence, it might just be reasonable to agree that the production of subjective awareness is only one of the manifold functions the brain is engaged in; one which has been selected for because of its prudent ignorance of superfluous details.

If this is true, then subjectivity *does* fit the description of being an emergent property of our neural hardware; but certainly not emergent in Searle's "ontological" sense. Rather, subjectivity might be construed as the resulting, yet unpredictable, property of the combination of empirically detectable components--much like, say, the resulting unpredictability⁴⁰ of the "wetness" of the combination of hydrogen and oxygen. Interestingly, the unpredictable properties of water can be likened, in inter-theoretical terms, as an "irreducible" property of this observed substance; yet this kind of "irreducibility" only marks an epistemological failing, rather than any type of dualist ontology in which the irreducible property figures. So, if we can make a connection between the irreducible properties of one theory, with irreducible properties of another

(viz., water and mental phenomena), then we might see how this irreducibility does not itself account for the fact that this property must be a non-physical entity. Hence, just because we haven't yet developed a model which can explain how the tactile property of subjectivity becomes, in Searle's terms, an "emergent system feature," this alone does not guarantee that this feature is an extraneous element of the most complex combination of parts we have heretofore tried to understand.

THE FINAL CUT

1. The Stimulation of Memory

We are near the end of our investigation and so it is now time to recall what we set out in the beginning to accomplish. In chapter one I suggested that a large part of the confusion surrounding the mind-body problem could be avoided if it was realized that much of this confusion is a result of tacitly accepting the handed-down linguistic categories our idiom has been committed to since at least the seventeenth century. To this end, I endeavoured to explain how these distinctly "theoretical" categories might yield up their reified status when analyzed by the objective methods of science. Thus, I was of the opinion then, as I am now, that although subjective experience is indeed a real phenomenon and even characterized in our idiom as "mental phenomena," this semantic commitment does not entail that it must be non-physical.

In chapter two we saw how the notion of intertheoretic reduction provided a more comprehensive understanding of the underlying causes of many complex phenomena. The felicity of this approach to broadening our knowledge base was supported by many concrete examples which, even though unsuccessful with the assumed autonomy of "mental"

phenomena, nevertheless supported the view that formerly accepted truisms in our idiom could be shown to be false when more comprehensive and explanatorily ubiquitous theoretical models were discovered. This failure of reduction to capture and explain away the basic elements of mental phenomena was then suggested as the reason why many contemporary thinkers have continued to search for a theoretical model based on the posits of what has come to be described as folk psychology.

Chapter three was committed to first investigating the assumed rationality of accepting common sense articulations--made scientific--to see if these posits really corresponded to anything; besides, of course, the simple theories of our neurologically naive ancestors. It was then suspected that perhaps they did not. And, as a result of this suspicion, it was then queried whether these ancient theoretical categories might really be false categories. From here it was suggested that perhaps the more perspicuous method for understanding subjective phenomena was to look at the ever revealing findings of the new discoveries being made in the neurosciences. The selected evidence that was provided at the end of chapter three did not "prove" that the categories of folk psychology were in fact false, yet this evidence did provide rationale for thinking that this may be the case--because this evidence did prove that discoveries which were formerly contradictions in our idiom because they were *then* beyond our imagination are currently being made, and made often. Nevertheless, there remained a worry; and that was how the immediacy of the "subjective" mode seemed to imply that this experience was of a non-physical kind even if it was accepted that our idiom could be revised.

Then, in the last chapter, we examined why subjectivity was, because of its "emergent" nature, often considered to be a non-physical phenomenon. I would not say that

we proved this consideration to be ultimately erroneous, but I think we did establish that this possibility might actually be borne out of the further discovery of how one's brain produces a phenomenon which is itself not designed to immediately apprehend all of the subtle processes which it is engendered by. Finally, it was suggested that the reason this discovery is so elusive is because the brain is so utterly complex that the product of its function is not yet even remotely predictable. And moreover, because we cannot predict how the physical brain produces what it does from one perspective, this resulting product has struck our dumbfounded intellectual community as being "irreducibly" non-physical from another. Admittedly these are strong sentiments coming from yet another perspective which does not presume to have any ultimate authority on this issue; but this, I contend, is precisely what makes the materialist thesis a plausible one.

2. The End of the Line

Ultimately, and thus paradoxically, it seems rather pretentious to make any final and resolute claims about the mind-body problem—or ultimate claims about any kind of knowledge for that matter. Yet again, this sentiment is supported by my thesis. For, with all we have acquainted ourselves with regarding the revisability of both scientific and philosophical claims about reality, it would then seem that any position which argued, righteously, for a final truth, or the outright denial of possibility about the issues we have herein been considering, would be just as susceptible to revision as were previous theories which we have long since abandoned. And therefore, even perhaps our most cherished theories about what the mind really is could be revised, or even eliminated, in favour of

a materialist conclusion.

For it does seem hard to deny that there has been a unification of theory occurrent in our intellectual environment. And this unification has precipitated all of our technological advances. New discoveries are being made all the time with the resources of the natural sciences we have fostered, and this fecundity of information, invention and new theory, does not appear likely to let up any time soon. This is perhaps telling when considered against the relative stagnation of metaphysical discoveries; although here, I will not try to add any further insult to the injury which this field of endeavour has had to endure. Instead, I would like to admit that metaphysical pursuits are very interesting and challenging theoretical abstractions which do much to expand one's artistic vocabulary; yet, this astounding creative capacity of rhetoric in our idiom outstrips what is fundamentally real.

To be sure, by denying an explanation that is, ultimately, ineffable, the materialist thesis does hold out for an explanation that can have some kind of bearing in the observable physical world, and thus, an explanation that could one day be apprehended by a more sophisticated scientific community. Admittedly, though, even if the approach to the mind-body problem we have been examining in this essay is a tenable one, perhaps there is yet a thousand different ways to be wrong about materialism and we have only entertained the first few. Nevertheless we count on the fact that this theory will continue to be revised if philosophy is to adequately synthesize our evolving picture of cognition and subjectivity.

Notes

1. I include this endnote to clarify which type of materialism I here speak of, viz., a materialism which construes all substances, properties, processes, etc., as being ultimately definable from within a causally closed physical system. This is contrasted with, for example, "non-reductive" accounts. Furthermore, I have not included *Idealism* in my introduction, nor in my essay, as this view has largely been passed over and is not, in my humble opinion, a live option—even though it is a well formed and difficult theory to refute.

2. Plato is famous for, among other things, explaining that there is a universal "realm of forms (or ideas)" whose categories are more perfect than any mere mortals can become apprised of. This ancient "belief" has saturated western thought and can be found, in a derivative form, in much of the contemporary literature about the real ontological status of the mind. I provide a reference presently:

Socrates: ...[i]t really is proved to us that if we are ever going to have pure knowledge of anything, we must get rid of the body and survey things alone in themselves by means of the soul herself alone...and if we are thus purified and freed from the foolishness of the body, we shall probably be in the company of the pure, and through our very selves come to have knowledge of all that is unsullied—that is, I suppose, of truth; for it is, perhaps, not lawful for the impure to attain to that which is pure. [Plato, 1986, p. 75.]

Whether Plato's prediction that we will come to have pure knowledge of the truth only after we discard our mortal coil is certainly beyond the scope of this paper. I have included this reference only for the purpose of demonstrating how the concept of mind (or, *soul*) is no longer described in such blatantly metaphysical terminology, although the underlying epistemic commitments still appear, in many cases, to be intact.

3. Bertrand Russell (1940) has been credited with coining the phrase 'the propositional attitudes'. In short, this term refers to the attitude one can take towards any given proposition and usually, or paradigmatically, the propositional attitudes refer to 'beliefs',

'desires', 'hopes' and so forth. Moreover, as Russell noticed, these terms are usually followed by a "that-clause" (e.g., I believe *that it is sunny*"), and so they seem on the one hand to be a necessarily mental class (because of the embedded *attitude*, or *intentionality*, they display), while also manifesting a natural class of grammatical instances.

4. To be sure, the theory of Evolution does not countenance non-physical traits; nor non-physical anything for that matter.

5. Although, and as I have already alluded to, it could be argued on pain of cliché that this problem is just one more footnote to Plato.

6. There are, of course, problems with this construction. Quine's theory of *indeterminacy of translation* is legendarily difficult to account for; however, these finer points of Quine's canon obscure the present discussion and so will be left for another time.

7. This sentiment runs counter to Popper's notion that, logically speaking, evidence can only refute a theory. Popper may be correct about his view of induction. Yet on pragmatic grounds, and because I agree with Quine that "inductive expectation underlies all learning," [Quine, 1995, p.25], I respectfully disagree that this condition is as negative as Popper makes it out to be. I am more inclined to agree, both, with Quine's slightly loosened articulation that "An unfavourable outcome refutes the theory—the categorical. A favourable outcome leaves the theory for further consideration" [*Ibid.*, p. 26.], and, E. Nagel's claim that "Observation statements may on occasion formulate initial and boundary conditions for a theory or law; they may also be employed to confirm or refute theories and laws" [Nagel, 1961, p. 348].

8. For example: F. Jackson, (1982); T. Nagel, (1974); J. Searle, (1992); M. Tye, (1995); and D. Chalmers, (1996), *inter alia*.

9. It is standard practice in this literature to speak of different "levels" of explanation. This *altitude* metaphor is an expedient prosaic device which isolates certain realms of discourse. Thus, and by using this device, one can distinguish, e.g., Sociology, as being a "higher level" theory compared with the "lower level" theory, of say Psychology, which defines over the behavioral traits of an individual; and further "down," the neural states of an individual which (on some accounts) are the underlying causes of a person's behaviour.

10. As a gloss: think of Quine, 1960, for theoretical contingency; and with reference to "elimination," think of Kuhn's (1962) *paradigm shift*

11. An interesting point with reference to the imperfect "identity" of this reduction (and relevant to the forthcoming discussion on *eliminativism*) is made by Paul Churchland. But because I have not treated Nagel's explanation of "Gas Laws," I am obliged to cite it as an endnote (I will remind the reader that kinematics is the science of pure *motion*):

Strictly speaking, however, this identity [temperature = mean kinetic energy of constituent molecules] is true only for the temperature of a gas, where

simple particles are free to move in ballistic fashion. In a *solid*, temperature is realized differently, since the interconnected molecules are confined to a variety of vibrational motions. In a *plasma*, temperature is something else again, since a plasma has no constituent molecules; they, and their constituent atoms, have been ripped to pieces. And even a *vacuum* has a so called 'black-body' temperature--in the distribution of electromagnetic waves coursing through it. Here temperature has nothing to do with the kinetic energy of particles. [1994, p. 41.]

12. This is just one of many instances where "intuition" can be shown to be a faulty epistemic mechanism. I mention this here because intuition is often considered to provide good evidence, particularly in introspection, for establishing the truth of theoretical posits in the philosophy of mind. It will thus serve us well to remember this example in what follows.

13. To clarify, I understand *precision* to be functionally related to the degree of refinement in measurement, whereas *accuracy* is the exactness of results arising from careful effort.

14. There is also a version of the Identity Theory which is referred to as the Token Identity Theory. However, as my concerns lie elsewhere I will not be making reference to this theory. See Davidson (1984) for more on this.

15. Leibniz's Law (or the law of the identity of indiscernibles) says that if x and y are identical, that is they share all the same properties, then anything truly predicated of x must also be true of y, and vice versa. Hence, for any occurrence of x we should be able to replace this term with y and observe no difference in truth value. In logical notation:

$$(x)(y)[(x = y) \text{ iff } (F)(Fx = Fy)]$$

[see also endnote #23.]

16. I include "seemingly" here, as a clause, to alert the reader to my opinion that there is not one thing being reflective and another thing supporting the reflection.

17. I am aware that Galen, the 1st century Roman physician, recognized that the brain was an essential organ for the operation of motor skills, health, and rational behaviour. However, I am here referring to "neuroscience" as the pursuit of studying the brain and its functions on the macro- as well as the micro-physical level; hence my reading of the term *neuro-science*.

18. Seminal work on functionalism can be found in Putnam (1967), and Davidson, (1969).

19. I here provide an extended quote for the purpose of identifying Fodor as, indeed, a functionalist; but also, to support other claims I make (in the body of the paper, below) which exhibit Fodor's "desire" to both have his cake, and "identify" it too:

For, if Functionalism is true, then there is plausibly a *level of explanation* between common-sense belief/desire psychology, on the one hand, and neurological (circuit-theoretic; generally 'hard science') explanation on the other...Moreover, it was [sic] possible to tell a reasonable and aesthetically gratifying story about the relations *between* the levels: commonsense belief/desire explanations *reduce* to explanations articulated in terms of functional states (at least the true ones do) because, according to Functionalism, beliefs and desires *are* functional states. And, for each (true) psychological explanation, there will be a corresponding story, to be told in hard science terms, about how the functional states that it postulates are "realized" in the system under study. (Fodor, 1990, pp. 9-10.)

20. Gordon writes:

One problem with this conception of folk psychology is that mastery of its concepts would seem to demand a highly developed theoretical intellect and a methodological sophistication rivalling that of modern-day cognitive scientists...[and further, that this]...could be a sign that it is no theory at all, not, at least, in the accepted sense of (roughly) a system of laws implicitly defining a set of terms. (1995a, p.71.)

21. According to functionalist doctrine, in each instance of a mental state there must be a particular brain state in which this mental state has its genesis. However, and largely because of "multiple realizability," type-identical mental states do not have to be instantiated by type-identical brain states because mental states can be *multiply realized* by a variety of different physical systems.

22. Some may argue that it was Dewey and James that first dabbled in functionalism. However, authors such as Putnam (1964, 1967); Lewis (1966); and Fodor (1968), better characterize the version of functionalism I am speaking of in this essay.

23. Leibniz's "law of identity" (or, law of indiscernables) runs analogously thus: for any x and any y , if x is "identical" (i.e., '=') to y , then they share all of the same properties.
[See also endnote #15.]

24. For accuracy I here note that the hypothesis of "Multiple Realizability" defines over "physical systems" as not just other brains, but of artificial systems as well. This hypothesis is hotly debated by AI (artificial intelligence) theorists and other cognitive scientists but will not be further pursued at this time.

25. The notorious relation 'R' is described creatively by many authors; however, I think I can pare it down. In the literature it is frequently agreed that ("specific," or) "token"-mental states correspond to "token"-physical states. Thus let us say that M_1 corresponds to P_1 (M = mental state; P = physical state, etc.). However, and here is where relation 'R'

makes its specious debut, the "intentional" relation between two different mental states, say between M_1 and M_2 , does not correspond to the "material" relation between two different physical states, say P_1 and P_2 . Yet, if M_1 and P_1 are identical, and M_2 and P_2 are identical, then the relation $M_1:M_2$ and $P_1:P_2$ will, by definition, be identical as well. I will note further that relation 'R' corresponds solely to the above mentioned relation "between mental states" only.

This is a fine bit of dualist sophistry. To be sure, it allows for the identity relation between variables in one instance, yet nevertheless denies it in a second instance where the variables are "the same." [see endnotes 15 & 23 for more on Identity.]

26. Functionalism is compatible with many of the theories being currently entertained by A.I. (artificial intelligence) researchers whom suspect that a "mental state" could be supported, or (multiply) realized, in a variety of non-biological systems. This is interesting stuff, yet cannot be pursued further in this present essay.

27. Other versions of the "Representationalist" thesis which are tempting, but ones which, nevertheless, retain realist commitments to folk psychological categories can be found in (Dretske, 1995; and Tye, 1995); the former offers an externalist account of the content of the propositional attitudes; while the latter extrapolates on phenomenalism.

28. For historical accuracy I will note that "eliminative materialism" was originally referred to as the "disappearance identity theory" (by both Feyerabend (1963), and Rorty (1965), because, according to these authors, the mental terms of FP will eventually "disappear," and be replaced by more tangible theoretical entities.

29. The "argument from meaninglessness" claims that it is either false or meaningless to accept that eliminativism is true; because for it to have any relevance one must "believe" it, yet "beliefs" are precisely the phenomena which eliminativism eschews. Therefore, eliminativism is either false or comprised of a string of meaningless marks (sentences) on a piece of paper.

However, this argument is merely a case of "begging the question." For, it assumes that the idiom in which mental states are engendered with meaning is an accurate characterization of these "legitimate" phenomena, which, to recall, is precisely the point in contention.

30. Camillo Golgi (1843-1926) discovered the first method of staining individual neurons. Using a silver (Ag) compound, which was "sucked up" through the soma, dendrites, and/or axons, he was able to impregnate certain neurons with a dye which made it possible to observe the individual nervous units of the human brain. However, Golgi thought (1883) that neurons must fuse (or, "anastomose") in one continuous nerve net (or, "reticulum") for the nervous system to be able to do what it does. It was not until 1891 that the word "neuron" was used to denote a single, independent cell. [see: P.S. Churchland, 1986, p. 25-29.]

My purpose for including this reference is to demonstrate that pretty much prior to our own century (the twentieth), there was virtually no evidence that the brain was not just

one big processing unit. Perhaps, then, the etiology of functionalism's "black box" is apparent; and, when cast against the above theoretical progression of the neurosciences, forgiven for its naive explanans.

31. An interesting aspect of neuronal degeneration is that an average adult human being loses approximately 50,000 neurons per day due to cell death. This discovery bodes poorly for the I.D. theory which requires identical brain states to support identical mental states.

32. There are of course about seventy phonemes which are used in the articulation of the English language. However, these "significant sounds" result from the syntactical combination, and thus nomological relation, of individual letters.

33. Dennett writes: "...I am interested in directly considering the conclusion that Jackson himself draws from his example: visual experiences have qualia that are 'epiphenomenal'." [1991, p. 401.]

34. ...If this ambiguity is not already agreed to exist in Jackson's argument I fear nothing--besides, perhaps, reading Churchland's entire essay--which will make a difference to this epistemic confusion.

35. "Coding vectors" refer to the predetermined processes which occur in the exchange of neurotransmitters.

36. A example of how reification can occur because of grammar is to consider two ways in which a similar, but not identical, sentiment can be expressed:

1. To say of a given property that it is "emergent" is to...(no reification implied.)

2. To say of a given property that it is an "emergent" one is to...(reification implied.)

My purpose here is to exhibit how the reification of a proposition can logically imply a material conclusion.

37. Searle states, on page 94: "In the sense in which I am here using the term, 'subjective' refers to an ontological category, not to an epistemic mode." Unfortunately for Searle, I suggest, he has got this relation perfectly backwards.

38. Searle states: "If you are tempted to functionalism, I believe you do not need refutation, you need help" (p. 9).

39. Indeed, Searle writes: "In different ways the argument occurs in the work of...Frank Jackson (1982). I think the argument is decisive, though it is frequently misunderstood in ways that treat it as merely epistemic and not ontological" (p. 116-117).

40. For the purpose of clarifying that what I mean here by "unpredictable" is not a metaphysical relation, I provide the following quote from E. Nagel:

Oxygen has certain properties and Hydrogen has certain properties. They combine to form water, and the proportions in which they do this is fixed. Nothing that we know about Oxygen itself or in its combination with anything but Hydrogen could give us the least reason to suppose that it could combine with Hydrogen at all. Nothing that we know about Hydrogen by itself or in its combination with anything but Oxygen, could give us the least reason to expect that it would combine with Oxygen at all. And most of the chemical and physical properties of water have no known connection, either quantitative or qualitative, with those of Oxygen and Hydrogen. Here we have a clear instance where, so far as we can tell, the properties of a whole composed of two constituents could not have been predicted from a knowledge of those properties taken separately, or from this combined with a knowledge of the properties of other wholes which contain these constituents. [E. Nagel, 1961, p. 368.]

Word Count (Essay only): 27,745

Bibliography

Block, N; Flanagan, O; Guzeldere, G; eds. (1997) *The Nature of Consciousness: Philosophical Debates*. Cambridge MA: MIT Press.

Chalmers, D.J. (1996) *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Christensen, S.M. and Turner, D.R. eds. (1993) *Folk Psychology and the Philosophy of Mind*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

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

Churchland, P.M. (1984) *Matter and Consciousness*. Cambridge, MA: The MIT Press.

Churchland, P.M. (1989) *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science*. Cambridge, MA: MIT Press.

Churchland, P.M. "Folk Psychology and the Explanation of Human Behaviour," in *Folk Psychology and the Philosophy of Mind*. eds., Christensen & Turner (1993), NJ: Lawrence Erlbaum Associates, Publishers.

Churchland, P.M. (1995) *The Engine of Reason, the Seat of the Soul*. Cambridge, MA: MIT Press.

Churchland, P.S. (1986) *Neurophilosophy: Toward a Unified Science of the Mind/Brain*. Cambridge, MA: MIT Press.

Dennett, D.C. (1978) *Brainstorms: Philosophical Essays on Mind and Psychology*. Cambridge, MA: MIT Press.

Dennett, D.C. (1991) *Consciousness Explained*. Boston: Little, Brown and Company.

Descartes, Rene. (1979) *Meditations on First Philosophy*. Cambridge, MA: Hackett Publishing Company, INC.

Dretske, F. (1995) *Naturalizing the Mind*. Cambridge, MA: MIT Press.

Feyerabend, P. "Materialism and the Mind-Body Problem," in *Folk Psychology and the Philosophy of Mind*. eds., Christensen & Turner, (1993) NJ: Lawrence Erlbaum Associates, Publishers.

Feyerabend, P. (1963) "Mental Events and the Brain," in *Materialism and the Mind-Body Problem*. ed. Rosenthal, D. (1987) Cambridge, MA: Hackett Publishing Company, Inc.

Fodor, J. (1968) "Materialism," in *Materialism and the Mind-Body Problem*. ed. Rosenthal, D. (1987) Cambridge, MA: Hackett Publishing Company, Inc.

Fodor, J. (1987) *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.

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

Fodor, J. "The persistence of the Attitudes," in *Folk Psychology and the Philosophy of Mind*. eds., Christensen & Turner (1993), NJ: Lawrence Erlbaum Associates, Publishers.

Fodor, J. (1994) *The Elm and the Expert: Mentalese and Its Semantics*. Cambridge, MA: The MIT Press.

Goldman, A.I. "Interpretation Psychologized," in *Folk Psychology*. eds., Martin Davies and Tony Stone (1995), Oxford: Blackwell Publishers.

Gordon, R.M. "Folk Psychology as Simulation," in *Folk Psychology*. eds., Martin Davies and Tony Stone (1995), Oxford: Blackwell Publishers.

Gordon, R.M. "The Simulation Theory: Objections and Misconceptions," in *Folk Psychology*. eds., Martin Davies and Tony Stone (1995), Oxford: Blackwell Publishers.

Graham, G. "The Origins of Folk Psychology," in *Folk Psychology and the Philosophy of Mind*. eds., Christensen & Turner (1993), NJ: Lawrence Erlbaum Associates, Publishers.

Horst, S. (1990) *Symbols, Computation, and Intentionality*. Los Angeles, California: University of California Press.

Kitcher, P. (1982) *Abusing Science*. Cambridge, MA: MIT Press.

Kuffler, S.W.; Nicholls, J.G.; Martin, A.R. (1984) *From Neuron to Brain*. Sunderland, MA: Sinauer Associates Inc.

Levine, J. (1997) "Leaving Out What It Is Like," in *The Nature of Consciousness: Philosophical Debates*. Cambridge MA: MIT Press.

- McCauley, R.N. "Intertheoretic Relations and the Future of Folk Psychology," in *Folk Psychology and the Philosophy of Mind*. eds., Christensen & Turner (1993), NJ: Lawrence Erlbaum Associates, Publishers.
- Nagel, E. (1961) *The Structure of Science*. New York: Harcourt, Brace & World, Inc.
- Nagel, T. (1986) *The View From Nowhere*. New York: Oxford University Press.
- Nagel, T. (1974) "What Is it Like To Be A Bat?" in *The Nature of Consciousness*. eds., Ned Block (et al.) Cambridge, MA: The MIT Press, 1997.
- Pernier, J. & Howes, D. (1995) "'He Thinks He Knows': And More Developmental Evidence Against the Simulation (Role Taking) Theory," in *Folk Psychology*. eds., Martin Davies and Tony Stone (1995), Oxford: Blackwell Publishers.
- Putnam, H. (1967) "The Nature of Mental States," in *Materialism and the Mind-Body Problem*. ed. Rosenthal, D.(1987) Cambridge: Hackett Publishing Company, Inc.
- Plato. (1986) *The Dialogues of Plato*. New York: Bantam Books.
- Quine, W.V. (1960) *Word & Object*. Cambridge, MA: MIT Press.
- Quine, W.V. (1990) *Pursuit of Truth*. Cambridge, MA: MIT Press.
- Rorty, R. "Consciousness, Intentionality, and Pragmatism," in *Folk Psychology and the Philosophy of Mind*. eds., Christensen & Turner (1993), NJ: Lawrence Erlbaum Associates, Publishers.
- Rosenthal, D. M. ed., (1987) *Materialism and the Mind-Body Problem*. Cambridge: Hackett Publishing Company, Inc.
- Rosenthal, D. M. ed., (1991) *The Nature of Mind*. Oxford: Oxford university press.
- Searle, J.R. (1992) *The Rediscovery of the Mind*. Cambridge MA: MIT Press.
- Smart, J.J.C. (1962) "Sensations and Brain Processes," in *Materialism and the Mind-Body Problem*. ed. Rosenthal, D. (1987) Cambridge: Hackett Publishing Company, Inc.
- Tye, M. (1995) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press.
- Thompson, R.F. (1985) *The Brain: A Neuroscience Primer*. New York: W.H. Freeman and Company.
- Wittgenstein, L. (1953) *Philosophical Investigations*. (trans. by G.E.M. Anscombe),

Oxford: Basil Blackwell & Mott, Ltd.