ELIMINATIVE MATERIALISM FROM A NORMATIVE NATURALIST PERSPECTIVE:

HOW WE OUGHT TO REGARD THE PROSPECT OF LOSING OUR MINDS

By

James Anthony Michael Wolfe, B. A.

A thesis submitted to the
Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for degree of

Master of Arts

Department of Philosophy

Carleton University
Ottawa, Ontario
April, 1998

ABSTRACT

Eliminative materialism is the thesis that our commonsense conception of psychological phenomena constitutes a radically false theory, a theory that is so fundamentally defective that both its principles and ontology will be displaced by completed neuroscience. Although this thesis has served as the focus for a number of positive and negative commentaries, most philosophers agree that only the future developments of empirical psychology can answer the question of whether or not it is correct. Yet with the appropriate *normative naturalist* principles in hand, Stephen Stich suspects that we might be able to determine whether it would be *rational* to accept the eliminativists' thesis. I take up Stich's proposal, and do so by appealing to various aspects of a particular normative naturalist enterprise — that of Robert McCauley's. However, whereas Stich is ultimately sceptical of normative naturalism's capacity to resolve the controversy over eliminativism, and resigns himself to a social constructivist stance when handling the ontological indeterminacy of folk psychological posits, I argue to the contrary: normative naturalism does show promise for resolving such issues, and for this and other reasons, a social constructivist account of how ontological indeterminacy is settled is anything *but* necessary.

ACKNOWLEDGMENTS

I owe my deepest thanks to Nicole, whose encouragement was present from start to finish.

Thanks to my parents, Bill and Heather, for their unwavering support of all my scholastic endeavours.

Robert McCauley patiently listened to and thoughtfully addressed all of my questions. For this I am greatly appreciative.

And finally, my thanks to Andrew Brook. It is without question that his good judgment and guidance made this project better than it would have otherwise been.

James Anthony Michael Wolfe

April 30th, 1998

CONTENTS

INTRODUCTION

Eliminative materialism is the thesis that our commonsense conception of psychological phenomena — what philosophers often call *folk psychology* — constitutes a radically false theory, a theory that is so fundamentally defective that both its principles and ontology will be displaced by completed neuroscience. According to this thesis, the idea of mental states distinguished in terms of their propositional content ought to be *eliminated* from a serious ontology of the mind-brain. The claim, in short, is that the propositional attitudes we allude to when explaining, predicting, and describing each others' behaviour *do not exist*: "like witches, phlogiston, and caloric fluid, or perhaps like the gods of ancient religions" these mental states are said to be "fictional posits of a badly mistaken theory" (Stich 1996: 3).

Not surprisingly, eliminativism has attracted a significant amount of attention, particularly since Paul Churchland's influential rendition of the thesis (1981). And while most philosophers of mind have strong convictions about eliminativism, most also agree that only the future developments of empirical psychology will definitively settle the controversy it inspires.

However, Stephen Stich (1996) has suggested that *normative naturalism* may provide a means by which to determine whether it would be *rational* to accept the eliminativists' thesis. For he imagines that a normative naturalist strategy might be deployed in such way as to produce some *principles of rational ontological inference or decision-making*. Such principles could inform us as to how we *ought* to regard the

entities posited by a theory once it is recognized that the theory is mistaken. In other words, the principles would yield a decision as to whether it would be rational to view the entities of a mistaken theory as existent or nonexistent.

It is Stich's supposition that a careful analysis of actual cases in the history of science wherein such decisions were in fact made could serve as a source for which to extract these principles. The first step might involve dividing a number of historical cases into the following two groupings: (i) theories that were judged to be incorrect and yet the entities therein were still taken to exist (e.g., planets, atoms, brains); and (ii) theories that were judged to be incorrect and their corresponding entities non-existent (e.g., caloric, phlogiston, the alchemist's 'spirits'). One could then search for prominent similarities and differences between the two groupings with the hope that certain features will emerge from most or all of the cases in one group and not from the other. These features, in turn, could be incorporated into principles that detail when it is rational to preserve items in the ontology of a mistaken theory and when such items should be 'eliminated'. With such principles in hand, we could then determine what we *ought* to do about the entities posited by a given theory once it was thought that that theory was mistaken — i.e., whether or not we should continue to conceive of its entities as existent.

It is here that we see the significance of this proposal to the current controversy over eliminativism, for *if*, at the eliminativists' insistence, we concede that folk psychology is a seriously mistaken theory, such principles could tell us whether or not we ought to retain the items that compose folk psychology's ontology (namely, the propositional attitudes). The principles would instruct us as to whether it would be

rational to accept what the eliminativist demands that we accept: that the entities signified by the terms of our folk psychological framework do not exist.

One particular normative naturalist project that appears to accord well with Stich's proposal (in that it shows promise of generating the aforementioned principles) is that developed by Robert McCauley. Indeed, of the various normative naturalist positions currently available it is McCauley's research into intertheoretic relations and the co-evolution of theories that commands a preponderance of Stich's praise. However, for a variety of reasons, Stich ultimately relinquishes the hope that McCauley's work, or for that matter, the fruits of any normative naturalist enterprise, might tell us whether or not we should accept the eliminativists' thesis as rational. And it is from this juncture that Stich begins to take up the chore of constructing his new position, a position firmly entrenched in the tradition of social constructivism.

I take up Stich's strategy, and do so by appealing to various aspects of Robert McCauley's research. However, whereas Stich is ultimately sceptical of normative naturalism's capacity to resolve the controversy over eliminativism, and resigns himself to a social constructivist approach when handling the ontological indeterminacy of folk psychological posits, I argue to the contrary: normative naturalism does show promise for resolving such issues, and for this and other reasons, a social constructivist account of how ontological indeterminacy is settled is anything *but* necessary.

~ Chapter 1 ~

Eliminative Materialism and the Future of Folk Psychology

Imagine setting out on a sunny afternoon for a drive in the country. While negotiating your way out of the city you notice another driver signaling a lane change and so you respond by decreasing your car's speed in order to let him in. Afterwards, you hear a siren and see the flashing lights of an ambulance in your rearview mirror so you pull over to the side of the road. And as you approach the city limits, several pedestrians ahead of you put their arms out at a crosswalk so you come to a full stop. You deal with each of these obstacles with such ease and proficiency that you are almost unaware of your actions. Indeed, throughout much of the drive the bulk of your thoughts were on where you might stop to have lunch, and why yesterday's tune-up did not resolve that incessant knocking coming from your car's engine. If asked about the details of your drive upon reaching your final destination you would probably find it difficult to provide a precise account. Nevertheless, the task has been accomplished — you have arrived safely and can now relax and enjoy your surroundings.

The successful completion of the preceding scenario is largely the result of an individual's expectations of others and of others' expectations of the individual. When I drive through a green light it is my *expectation* that those drivers who intersect my path will have stopped at their red light. When pedestrians motion at a crosswalk it is, if not their expectation, their *hope* that I will come to a stop. A closer look reveals that these expectations arise from attributing beliefs and desires to others. When driving through a green light I (often tacitly) attribute several beliefs and desires to those drivers who intersect my path. For example, I generally credit them with holding the *belief* that stopping at red lights will satisfy their *desire* to avoid accidents. I might also credit them with holding the *belief* that *not* driving through red lights can secure their *desire* to evade traffic charges. Similarly, pedestrians might credit me with holding the *belief* that stopping at crosswalks is the correct course of action since it best accords with my *desire* not to inflict injury on innocent bystanders.

Attributing beliefs and desires to others not only (or even primarily) serves as a means to justify our expectations. For in the act of ascribing beliefs and desires to another we are also offering an *explanation* — often simply to ourselves — of their behaviour. Why do the drivers that intersect my path stop when their light is red? Because they wish to avoid an accident and believe that stopping will enhance the probability of fulfilling this wish. This same approach can also be used to *predict* another's behaviour. For instance, the pedestrians *could* have surmised that as my car drew closer to the crosswalk I would depress the brake pedal, and then after a few moments, transfer my foot back to the accelerator — and in all of this they would

probably be quite right. To be sure, offering explanations and predictions of this sort is an entirely common practice. We continually explain and predict — with a degree of accuracy that is nothing short of remarkable — an immense array of complex human behaviour by attributing beliefs and desires to others. Yet because this talent is so thoroughly entrenched — and often tacitly employed — in our daily activities it is generally overlooked.

However, during the 1960's philosophers started to show a keen interest in our commonsense explanations and predictions of behaviour. They began to regard these explanations and predictions as particular instances of the more basic principles or law-like generalizations that seemed to govern our everyday behaviour (e.g., Hempel: 1965, Alston: 1967). In the decade that followed many went on to claim that these law-like generalizations were in fact part of a larger *network* of principles that constituted a commonsense *theory* of mental phenomena (e.g., Lewis: 1972). The central postulates upon which this theory was said to rest were *mental* states — states such as belief, desire, and intention (what philosophers have collectively labeled 'the propositional attitudes', since each expresses a distinct 'attitude' toward a particular proposition). And because ordinary people had a general understanding of these mental states and the principles that subsume them — as our mastery of explaining and predicting others' behaviour attests — the theory came to be known as 'folk psychology'. Thus it was alleged that we were all conversant in, and able to administer, a rudimentary psychological theory. We were all, whether we realized it or not, theory-using folk psychologists.

During this period (i.e., the 1960's/early 1970's) very few philosophers questioned the existence of minds or mental states such as belief and desire. However, the same cannot be accurately said of prior decades. Throughout the 1930's, 1940's and early 1950's scientific (i.e., psychological, methodological) behaviourism — the view, broadly speaking, that behaviour is fundamental to understanding mental phenomena — was enthusiastically received by many prominent figures in the philosophical community. Lead by the research of B. F. Skinner, C. L. Hull, and E. C. Tolman, the scientific behaviourist movement departed from the introspectionist tradition by redefining the central task of psychology as the explanation and prediction of behaviour where, crucially, "to explain behavior is to provide a 'functional analysis' of it, i.e., to specify the independent variables (stimuli) of which the behavior (response) is lawfully a function" (Marras 1995: 67). These variables were believed to be exhaustively specifiable by the experimental procedures of the natural sciences, and thus, introspectible, internal states of consciousness (what are in sum generally thought to compose the mind) were *excluded* from the proper domain of psychology. Although some behaviourists were prepared to admit internal neurophysiological conditions as 'intervening' variables, others, such as Skinner, took the more radical position in which environmental variables were the sole consideration (for, Skinner argued, any hypothetical inner states were, at bottom, functions of past and present environmental conditions). What united all psychological behaviourists was the absence of metaphysical claims: minds and mental entities *might* exist, but all agreed that this need not be presumed in psychological experiment or theorizing (Lycan 1990: 4).

Metaphysical claims were, in contrast, present in philosophical behaviourism. Philosophical behaviourism, in its most optimistic rendition (namely, that espoused by the logical positivists), asserted that statements containing mentalistic expressions "have the same meaning as, and are thus translatable into, some set of publicly verifiable (confirmable, testable) statements describing behavioral and bodily processes and dispositions [to behave] (including verbal-behavioral dispositions)" (Marras 1995: 67). Thus, for example, the expression 'Edmund is in pain' *does not mean anything* about Edmund's putative inner life or any episode taking place within him, but rather only that Edmund is "behaving in a wincing-and-groaning way or is disposed so to behave (in that he would so behave were something not keeping him from doing so)" (Lycan 1990: 4). As a result of the reductionist concerns expressed by the logical positivist thesis of physicalism and the unity of science, logical behaviourism (as it was often called) was "a corollary of the thesis that psychology is ultimately (via a behavioristic analysis) reducible to physics, and that all of its statements, like those of physics, are expressible in a strictly extensional language" (op. cit. 67). Accordingly, questions pertaining to the nature of the mind as traditionally construed (i.e., wherein mentalistic terms such as belief and desire are taken to signify some publicly *un*verifiable internal state), were, strictly speaking, irrelevant to the study of human psychology.

Further versions of philosophical behaviourism can be traced to Ryle (1949) and Wittgenstein (1953), both of which — albeit for different reasons — do not carry the reductionist implications of logical behaviourism. An attractive feature of these, and indeed all variations of philosophical behaviourism was that one could avert the perennial

problems associated with Cartesian Dualism. For in addition to solving the methodological problem of intersubjective verification, philosophical behaviourism "dispensed with immaterial Cartesian egos and ghostly nonphysical events, writing them off as ontological excrescences" (Lycan 1990: 5). Descartes' mind-body problem was no longer a concern, since philosophical behaviourism posited no immaterial, nonspatial causes of behaviour. In short, it raised no scientific mysteries concerning the intervention of Cartesian substances since it sanctioned no such intervention (ibid. 5).

However, many found the total repudiation of the inner unsettling, and throughout the 1950's and 1960's philosophical behaviourism as a whole underwent severe criticism. Among the most cogent points raised by those unconvinced of behaviourism was the seemingly inescapable fact that any conscious person fully knows that he or she experiences and can introspect inner mental episodes that are neither accompanied by characteristic behaviour nor merely serve as static hypothetical data of how he or she would behave if subject to a given stimulation (ibid. 5). Furthermore, many argued that nothing prevented the possibility that two people might differ psychologically despite a total similarity of their actual and counterfactual behaviour, as in, for example, a Lockean case of 'inverted spectrum'. Indeed, a creature might exhibit all the appropriate stimulus-response relations and yet lack mentality entirely (ibid. 5; see also: Fodor & Block: 1972). And finally, it became increasingly apparent that the logical behaviourist's behavioural analyses of mental ascriptions were adequate "only so long as one makes substantive assumptions about the rest of the subject's *mentality*", and were hence judged to be either circular or radically incomplete as analyses of the mental in general (ibid. 5; see also:

Churchland 1988: 23-5).[1] Thus philosophical behaviourism (as well as the sort of scientific behaviourism championed by Skinner) was, for all intents and purposes, abandoned, largely in favour of middle-ground materialist alternatives. What was noteworthy of these new positions was that they all accommodated at least some genuinely inner (i.e., mental) states and yet avoided, from the fact of materialism, the aforementioned objections raised against dualism (the identity theory and functionalism were the leading replacement candidates of the time). Accordingly, the overwhelming majority of philosophers again took the existence of minds and mental states to be obvious, and thus resumed their task of uncovering the *nature* of mind and the states therein.

Yet in the late 1960's/early 1970's a radical view — known as Eliminative Materialism (or simply 'eliminativism') — began to emerge in the philosophical community. Like the many materialist positions of the time, this doctrine took a straightforward physicalist stance on the nature of minds, viz., that 'minds' were functioning brains. It was the eliminativists' view concerning the nature of mental states, however, that was the source of controversy. According to the eliminativist thesis, the idea of mental states distinguished in terms of their propositional content ought to be *eliminated* from a serious ontology of the mind-brain since they were "relics of an outmoded 'folk theory' of human psychology" (Hannan 1993: 166). The claim, in other words, was that the propositional attitudes (such as belief and desire) we allude to when explaining, predicting, and describing each others' behaviour *do not exist*: "like witches, phlogiston, and caloric fluid, or perhaps like the gods of ancient religions" these mental

states were said to be "fictional posits of a badly mistaken theory" (Stich 1996: 3). And it was primarily for this reason that eliminativists, in a vein similar to their behaviourist predecessors, touted the utter irrelevancy of propositional attitude psychology and the necessity of a search for a legitimate, *scientific* account of the mind-brain. However, it was also on this count that eliminativism made a fundamental split from behaviourism. For rather than insist, with behaviourists, that external environmental conditions alone would furnish the correct account of human psychology, eliminativists, in their expectation that future neuroscientific research would wholly disclose the mysteries of the mental, took a pronounced turn inward.

Richard Rorty (1965, 1970) is responsible for much of the ground on which present-day eliminativism resides.[2] In attacking the arguments put forth by the mind-body dualists of his time, Rorty aimed to strengthen the notion that 'sensations' — what are generally regarded as close relations of the more familiar propositional attitudes (i.e., beliefs and desires) — could indeed be brain processes and thus analyzed in a strictly materialistic (i.e., scientific) idiom. What distinguished Rorty as an eliminativist was the manner by which he set out to accomplish this task — namely, by defending what he termed the 'disappearance' form of the identity theory. Unlike the traditional form of the identity theory, which foresees a numeric identification between mental states and brain processes (which would thereby preserve the ontological status of the former), Rorty's disappearance form, in anticipating profound discontinuities between the two, ultimately sought to impugn the existence of mental states (specifically sensations). In light of this, and in keeping with the foregoing (and admittedly cursory) historical remarks, the

following will chart Rorty's significant contributions to what would ultimately become contemporary eliminativism.

1.1: The Rise of Eliminativism

The basic strategy that underlies Rorty's reply to those who claim that empirical inquiry could not identify brain-processes with sensations is well known. For it exploits the simple fact that the classifications of linguistic expressions that form the foundation of the dualists' argument are classifications of a language "which is as it is because it is the language spoken at a given stage of empirical inquiry" (Rorty 1965: 17). Hence we should, the argument proceeds, expect that the sort of empirical results that would show brain-processes and sensations to be identical would, in all likelihood, bring about changes in our way of speaking that would render current classifications obsolete. Accordingly, to argue against the identity theory on the basis of the way we talk now would be like "arguing against an assertion that supernatural phenomena are identical with certain natural phenomena on the basis of the way in which superstitious people talk" (ibid. 18). The point being is that there is simply no method of classifying linguistic expressions that confers results that are guaranteed to remain intact despite the future discoveries of empirical inquiry.

Rorty expresses this familiar sentiment by defending the *disappearance* form of the identity theory — a form that reinterprets the relation between sensations and brain-processes in a rather unorthodox way. Traditionally, the identity theory holds the

relation in question as one of numeric or 'strict' identity, which can be expressed as follows:

$$(x)\,(y)\,[(x = y) \supset (F)\,(Fx \equiv Fy)].$$

By accepting this relation one maintains that anything truly predicated of (x) must also be truly predicated of (y) (and vice-versa). However, critics of the identity theory note that to comply with the said relation one must be prepared to admit, for instance, that physical processes such as brain-processes can be 'dim' or 'fading' or 'false', and correspondingly, that mental phenomena such as after images are 'publicly observable' or 'physical' or 'spatially located' (Cornman 1962: 490; quoted in Rorty 1965: 18)). In other words, with respect to the case of mental states and brain-processes, the identity theory forces one into making what philosophers call 'category mistakes' — the statements that ensue when properties sensibly predicated of one category become insensible when predicated of another. Consider the meaninglessness, for example, of statements such as 'My C-fibers are true' or 'My belief-that-the-sun-is-a-star is located in the temporal lobe of my left cerebral hemisphere' (Churchland 1988: 30). Since the identity theory would seem to demand that we accept these and other statements of the same sort, the critic concludes that the theory is seriously mistaken.

But what, Rorty asks, can the relation of identity mean if it is not one of 'strict' identity? For it would appear as though anything otherwise would be mere correlation (which dualists are quite willing to concede). Rorty suggests that two forms of the identity theory arise in response to this dilemma. The first, which Rorty calls the

translation form, grasps the first horn by attempting to show that the odd sounding expressions such as those mentioned above do not involve category mistakes, and that this can be shown by the appropriate translations into 'topic neutral' language (Rorty 1965: 19). The second, Rorty's *disappearance* form, grasps the second horn by maintaining that the relation is not a strict identity, but rather the sort of relation that holds between "existent entities and non-existent entities when reference to the latter once served (some of) the purposes presently served by reference to the former — the sort of relation that holds, e.g., between 'quantity of caloric fluid' and 'mean kinetic energy of molecules'" (ibid. 19). So it would seem as though Rorty interprets the identity relation as grounded in at least part of the referring role of the two propositions in question. Accordingly, he maintains that while there is an intuitive sense of 'same' wherein what used to be called 'quantity of caloric fluid' is the *same thing* as what is now called a certain mean kinetic energy of molecules, there is "no reason to think that all features truly predicated of the one may be sensibly predicated of the other" (ibid. 19).

Consequently, the disappearance form of the theory considers it a mistake to assume that "X's are nothing but Y's" entails "All attributes meaningfully predicable of X's are meaningfully predicated of Y's", for, as Rorty claims, such an assumption "would forbid us ever to express the results of scientific inquiry in terms of...'cross-category identity'" (ibid. 19). While the verb in statements like "Zeus's thunderbolts are discharges of static electricity" or "Demoniacal possession is a form of hallucinatory psychosis" might seem as though it were the 'is' of identity, it is clear that it does not express *strict* identity. This being the case, the disappearance form of the identity theory

recommends that we regard such statements as elliptical; as in, for example, "*What people used to call* 'demoniacal possession' is a form of hallucinatory psychosis," wherein the relation at issue is *strict* identity. For since there is no reason why "What people call X" should be in the same Rylean category as "X", Rorty contends that there is no need to claim, as the translation form must, that topic-neutral statements of "X" are possible (Rorty 1965: 19).

On the surface, one may notice what appears to be a serious flaw in the disappearance form of the identity theory. For when we say "What people call 'caloric fluid' is nothing but the motion of molecules" or "What people call 'witches' are nothing but psychotic women" we are generally prepared to say there is *no such thing* as caloric fluid, and similarly, *no such thing* as witches. However, to say "What people call 'sensations' are nothing but brain-processes" entails that there is no such thing as sensations seems, at least to most, ridiculous. Again, it is often said that such a statement embodies a category mistake; viz., the speaker is expressing a conceptual confusion: 'sensations' and 'brain-processes' are entities that belong in separate categories and must be treated as such. Rorty recognizes this common reproach and sets out to discredit it by way of analogy.

Imagine, Rorty asks, a primitive tribe that holds the view that illnesses are caused by demons — a different demon for each type of illness. If asked for further information about these demons, it is said that certain members of the tribe (namely witch doctors) have the ability, after ingesting a special kind of mushroom, to see and identify each type of (intangible) demon on or near a given patient. Moreover, these witch doctors have

come to recognize that, for example, a blue demon accompanies epileptics, a red one accompanies those who suffer from pneumonia, and so on. Witch doctors also know that, for example, red demons dislike a certain type of mold that they administer to patients who have pneumonia.

Upon encountering this tribe we would probably be tempted to tell them that there are no demons, and that illnesses are caused by germs, viruses, and the like. We would also probably explain that the witch doctors were not seeing demons, but were merely having hallucinations. In all of this we would be quite right, but, Rorty asks, would we be right on *empirical* grounds? What empirical criteria, built into the demon-talk of the tribe, go unsatisfied? What predictions that the *tribesmen* make fail to come true? If there are none, then a sophisticated witch doctor might insist that all modern science can do is show "(1) that the presence of demons is constantly correlated with that of germs, viruses, and the like, and (2) that eating certain mushrooms sometimes makes people think that they see things that aren't really there" (Rorty 1965: 21). Clearly, this would not be sufficient to show that there are no demons. At best, it shows that if we disregard demons, then "(a) a simpler account of the cause and cure of disease and (b) a simpler account of why people make the perceptual reports they do" can be provided (ibid. 21).

Indeed, there is not much else to say to a sophisticated witch doctor except that the simplicity of the accounts that follow if we disregard demons *is an excellent reason* for saying there are no demons (ibid. 21). While demon-discourse is one way of describing and predicting phenomena, there are other ways that seem preferable. Of

course, we could, at the witch doctor's request, attach demon-discourse to modern science by saying, first, that diseases are caused by the co-presence of demons and germs (each being necessary, but neither a sufficient, condition) and second, that the witch doctors (unlike drunkards and the mentally ill) really do see intangible beings. Doing so would retain all the predictive and explanatory advantages of modern science, for we would know as much about the cause and cure of disease, and about hallucinations, as we did before. Yet we would also have the additional encumbrance of dealing with problems that we did not have before: the problem of why demons are visible only to witch doctors and the problem of why germs cannot cause diseases by themselves (Rorty 1965: 21). We avoid both these problems in saying that demons do not exist. Any remaining objections to this use of Occam's Razor could only be met by citing the practical benefits obtained by use of the Razor in the past.

The identity theorist's claim is that sensations may be to the future advances of psychological science as demons are to current science. In other words, just as we are now inclined to deny that there are demons, future science may want to deny that there are sensations. Rorty notes that the only obstacle to replacing sensation-discourse with brain-discourse seems to be that sensation-statements have a reporting as well as an explanatory function (ibid. 21). However, the demon case illustrates that "the discovery of a new way of explaining the phenomena previously explained by reference to a certain sort of entity, *combined with a new account of what is being reported by observation-statements about that sort of entity*, may give good reason for saying that there are no entities of that sort" (ibid. 21); emphasis in original). The absurdity of saying "Nobody

has ever felt pain" is no greater than saying "Nobody has ever seen a demon" *if*, as Rorty states,

> ...we have a suitable answer to the question "What *was* I reporting when I said I felt pain?". To this question, [a scientist] of the future may reply "You were reporting the occurrence of a certain brain-process, and it would make life simpler for us if you would, in the future, *say* 'My C-fibers are firing' instead of saying 'I'm in pain'". In saying so, he has as good a prima facie case as the scientist who answers the witch doctor's question "What *was* I reporting when I reported a demon?" with "You were reporting the content of your hallucination, and it would make life simpler if, in the future, you would describe your experiences in those terms" [Rorty 1965: 22].

Whether by this quote Rorty truly intends to suggest that future science will, in keeping with his example, precisely match instances of pain with the firing of C-fibers is uncertain (although it would seem so, and the identity theory in its conventional form certainly does maintain something like this). If so, then current opinion almost universally rejects the view. Early functionalists, such as Putnam (1967) and Fodor (1968), soon recognized that the identity theory as such (namely, as a theory of *types* or *kinds* of mental items) carried disconcerting implications. For from the view that specific mental states (e.g., pain) are always and everywhere characterized by specific neurophysiological characterizations (e.g., firing C-fibers), it follows that "a creature of any species (earthly or science-fiction) could be in pain only if that creature *had* c-fibers and they were firing" (Lycan 1990: 7). But what grounds does the identity theorist have for imposing such a biological constraint? The answer is none, as there are no compelling reasons why we should suppose that an organism *must* be made from the same chemical

materials as ourselves in order that it have what can accurately be recognized as pain. In short, functionalists noticed that the identity theory, in focusing exclusively on the attributes of human neurophysiology, fell prey to species chauvinism. And to make matters worse for the identity theorist, preliminary research of the relevant neurosciences indicates that even *within* the human population some mental phenomena seem to have multiple neurophysiological realizations, making the identity theorist's convenient one-to-one match-up all the more unlikely.

While Rorty omitted a response to this particular charge, he did acknowledge certain disanalogies between the demon-case and the sensation-case. The first of these is that there is no easy way to fill in the blank in "What people called 'demons' are nothing but _____" since neither 'hallucinatory contents' nor 'germs' will suffice. We must, therefore, distinguish the observational and explanatory roles of 'demon':

> We need to say something like "What people who reported seeing demons were reporting was simply the content of their hallucinations," and also something like "What people explained by reference to demons can be explained better by reference to germs, viruses, etc." [Rorty 1965: 22].

Because of this need for a relatively complex account of how we are to get along without reference to demons, we cannot *identify* "What we called 'demons'" with anything, so instead, we simply deny their existence. However, in the case of sensations, we *can* give a relatively simple account of how to get along in the future. Statements about brain-processes can appropriate both the explanatory and the reporting functions of statements about sensations. Thus, we are prepared to identify "What we called 'sensations'" with

brain-processes, and to say "What we called 'sensations' turn out to be nothing but brain-processes" (Rorty 1965: 22).

Rorty notes that the pragmatic consequences of the proposed reduction in both the demon-case and sensation-case are the same — i.e., that we should stop asking questions about the causal and/or spatial-temporal relationships holding between the 'reduced' entities (demons, sensations) and the rest of the universe, and replace these with questions about the relationships holding between certain other entities (germs, hallucinatory experiences, brain-processes) and the rest of the universe (ibid. 22). As it turns out, for reasons just mentioned, the proposed reduction is put in the form of a denial of existence in one case, and of identification in the other. But, as Rorty points out, "There are no demons" and "What people called 'sensations' are nothing but brain-processes" can both be "equally well paraphrased as 'Elimination of the referring use of the expression in question ('demon', 'sensation') from our language would leave our ability to describe and predict undiminished'" (ibid. 22).

Yet the claim that there might be no such thing as 'sensations' — even in light of the preceding analogy — continues to astonish most. Rorty acknowledges this sentiment, and proceeds to explain that the reason the claim is received as such is because the elimination of the referring use of 'sensation' from our language would — *currently* — be highly impractical (ibid. 23). Nonetheless, Rorty insists that the statement "What people call 'sensations' might turn out to be brain-processes" is entirely sensible and unconfused. And even though it may seem intuitively implausible, and even if the thought of eliminating sensation-discourse seems fundamentally opposed to practicality, it should

not, Rorty contends "blind us to the facts that (a) entities referred to by expressions in one Rylean category may also be referred to by expressions in another, (b) expressions in the first category may drop out of the language once this identity of reference is realized, and (c) the thesis in question is a natural way of expressing the result of this realization in the case of 'sensation' and 'brain-process'" (Rorty 1965: 38). And this is all that Rorty aims to show with the disappearance form of the identity theory.

Clearly (b) offers the most conspicuous indication of the eliminativist undertones that accompany Rorty's disappearance form of the identity theory. For the implication of (b) is that once the identity between sensations and brain-processes has been secured, talk of sensations might be eliminated in favour of the advantages (both practical and empirical) afforded by brain-discourse. Yet it is perhaps because of the era in which his overall position was modeled that Rorty almost seems to understate the point. One must recognize that his primary objective was to *defend* materialism against the charge that its thesis — that thoughts might be nothing but brain-processes — was somehow conceptually confused and hence nonsensical. For he thought that if it could be demonstrated that this charge, in its various manifestations, was unfounded, then the materialists' thesis was vindicated: i.e., that it *was* reasonable to *suppose* (at the very least) that thoughts and sensations could be material processes and hence analyzed as such. Rorty did not, therefore, launch a serious *offensive* attack against the sort of sensation-discourse that folk psychology subsumes; indeed, he continued to hold the view that an identification of the posits of the common idiom with those of a future materialistic psychology would be realized, and that the question of elimination would be

largely a matter of the feasibility of changing linguistic conventions.[3] It might very well turn out that, unlike demon-discourse, linguistic conventions are too ingrained to allow the elimination of sensation-discourse, in which case we would be left with only an 'in-principle' elimination. It is precisely with respect to the last point that contemporary eliminativism is most markedly differentiated from Rorty's disappearance form of the identity theory. For it has become, as we shall see in the section to follow, the hallmark of contemporary eliminativism to expose the *faults* of folk psychology in order to illustrate the *futility* of seeking an identity between thoughts and brain-processes and the *need* for a wholly non-mentalistic psychology. Hence, replacing the common idiom with a suitably materialistic alternative is not simply an option, but a necessity.

1.2: The Contemporary Case for Eliminativism

Paul Churchland is considered by most to be the leading advocate of present-day eliminativism. At the outset of his now well-known essay "Eliminative Materialism and the Propositional Attitudes" (1981), Churchland sets forth what has since become the paradigmatic articulation of eliminativism:

> 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 [Churchland 1981: 67].

This statement alone provides a sense of the contrast between Churchland's eliminativism and the eliminativist connotations present in Rorty's disappearance form of the identity theory. For two crucial consequences, absent from Rorty's position, follow from Churchland's contention that our folk psychological account of mental phenomena is profoundly mistaken: (1) that the principles and posits (e.g., beliefs, desires,) that are taken to compose folk psychology's ontology will most assuredly be eliminated by a 'legitimately' scientific account of the mind-brain; and hence (2) that the prospect of identifying (and thereby reducing) thoughts to brain-processes is, given folk psychology's serious defects, unattainable.

Churchland appeals to three key arguments in support of this thesis. The first of these deals with folk psychology's explanatory power (Churchland 1981: 73). It is said that our commonsense conception of the mind fails to convincingly explain — and in many cases even *address* — a vast array of mental phenomena. Cognitive disorders, such as hemi-neglect and alexia without agraphia, as well as more pervasive phenomena, such as creative imagination, memory, disparities in intelligence, and the psychological function of sleep are currently unaccountable within the framework of folk psychology. Perhaps the most damaging of these purported deficiencies is folk psychology's inability to account for the learning process — that is, the manner in which we come to acquire the ability to administer the propositional attitudes in our daily explanations and predictions of others' behaviour. Since, as we have seen, the propositional attitudes are the central postulates of folk psychology, this shortcoming would appear quite harmful.

Although Churchland's second argument also questions folk psychology's explanatory competence, it does so by focusing directly on the historical progress of our commonsense conception of the mind (Churchland 1981: 74-5). Churchland observes that the applicability of folk psychology has been dramatically reduced over the course of time. Whereas primitive cultures often explained natural phenomena by attributing intentional states thereto (e.g., the 'fury' of the sea, the 'anger' of the storm), we now limit the use of such concepts to the behaviour of higher animals. Furthermore, it is said that even within the confines of the mental, folk psychology's development has been entirely insignificant. Indeed, in what has since become a somewhat notorious declaration, Churchland states that "the FP [folk psychology] of the Greeks is essentially the FP we use today" (ibid. 74). While it may be true that perfect theories need not evolve, folk psychology, as indicated above, is depicted as being far from perfect. Thus it is argued that this tremendously long period of stagnation serves as a good reason for questioning the integrity of folk psychology's central categories.

The final argument used in support of the eliminativist thesis measures the degree of theoretical integration exhibited by folk psychology (ibid. 75-6). This argument rests on the assumption that coherence with other theories is a virtue that every theory should strive for. Again, it is alleged that folk psychology performs poorly in this regard. The sum total of the physical sciences — which includes an assortment of theories pertaining to a number of fields (e.g., evolution; biology; physiology; chemistry; physics; etc.) — forms an undeniably impressive synthesis. Further, as this aggregate of science continues to develop and cohere, so does our ability to explain nature's curiosities, including the

realm of the mental — a domain traditionally attended to by folk psychology. Yet because the propositional attitudes are unable to integrate with the rest of science, folk psychology remains in isolation (interestingly, the same can be said of phonology and colour perception).

While conceding that each of the foregoing failings does not individually act as an incontestable confirmation of the eliminativists' thesis, when taken together, Churchland maintains, they clearly demonstrate the likelihood of folk psychology's future eradication.

Prima facie, each of Churchland's three arguments for elimination seems distinct. However, they all rely on two implicit presuppositions. The first of these maintains that folk psychology is in fact a proto-scientific *empirical* theory of mental phenomena. If folk psychology *is* such a theory, then it is simply another alternative in competition with other scientific-psychological accounts of the mind-brain, and hence, vulnerable to the same stringent criterion by which all theories are evaluated. Conversely, if folk psychology *is not* a theory, then Churchland's arguments are considerably weakened (for why, if folk psychology is not an empirical theory, should it be chastised for explanatory failures or an inability to cohere with the rest of the empirical sciences?). The second presupposition is that the meanings of the theoretical terms that denote the properties or entities posited by folk psychology are determined by virtue of the laws, principles and generalizations in which they figure (what philosophers call 'network semantics' or 'semantic holism'). The following example, borrowed from Churchland, will help to clarify this thesis. Electromagnetic theory postulates the existence of electric charges,

electric force fields, and magnetic force fields. The laws of electromagnetic theory specify how these things relate to one another and to the relevant observable phenomena. According to semantic holism, then, to fully understand the expression 'electric field' is to be "familiar with the network of theoretical principles in which the expression appears" (Churchland 1988: 56). Hence, theoretical terms do not, in general, get their meanings from single, explicit definitions stating conditions necessary and sufficient for their application. Rather, they are "implicitly defined by the network of principles that embed them" (ibid. 56).

Since folk psychology's elimination is urged largely as a consequence of the three arguments outlined above, Churchland's success in establishing its theoretical status is self-admittedly imperative. However, it is also worthwhile to note some less obvious incentives for accomplishing the said task. First, it would serve to guard against anti-eliminativist attacks of the sort that claim that folk psychology is partly or entirely *non*-theoretical (at least, as it often claimed, if we interpret the word 'theory' along the lines of classical usage), and as such, serves either partially or completely *non*-theoretical purposes (what can collectively be called 'anti-theory arguments'). Anti-eliminativist arguments of this type contend that folk psychology is *not* prone to the same methods of assessment as are empirical theories in general, simply because folk psychology is *not* a theory — at least in the normal sense (rather it is, for instance, a socially constructed device used solely for its pragmatic utility — an argument that, one will recall, is similar to that launched by the dualists against Rorty; for a contemporary example, see: Hannan [1993]). And second, it would serve to guard against anti-eliminativist attacks that find

their force by contesting network semantics (what can be collectively called 'anti-holism arguments'). These arguments can take the form of either a direct attack on semantic holism, or, by virtue of propounding an alternative thesis (e.g., semantic atomism, semantic molecularism), an indirect attack. For example, if semantic atomism is the correct account of meaning, then the eliminativists' thesis would be considerably weakened. For the falsity of folk psychology would *not* entail that the entities posited therein were nonexistent — rather, it might simply indicate that the terminology used for describing the entities is seriously misleading or mistaken. With this said, one can plainly see the multifarious motives that lurk behind Churchland's efforts to confirm folk psychology's standing as an empirical theory of mental phenomena. The following section will examine the grounds Churchland appeals to in support of this position.

1.3: Is Folk Psychology a Theory?

As we noted at the outset, it is a remarkable (yet often overlooked) fact that most people are able to explain and predict the behaviour of others with a high degree of success. Recall also that these explanations and predictions are generally made by attributing beliefs, desires, fears, hopes, intentions and so forth to a given subject(s). Acknowledging this fact, Churchland proceeds with the contention that such explanations presuppose laws — "rough and ready ones, at least" — that link the *explanatory conditions* to the *behaviour explained* (Churchland 1981: 68). But how exactly, one might ask, do such laws figure in our commonsense explanations of behaviour? Churchland's reply resides

within what philosophers call the 'deductive-nomological model of explanation'. An example will help clarify this conception.

Suppose I drop a sugar-cube in a glass of water. After some time, the sugar-cube 'disappears', bringing a look of astonishment to the face of my friend. "The sugar-cube disappeared!" he cries; "How is that possible?" "Well," I say, "it's soluble." Certainly not the brightest person around, my friend responds: "Huh?" A little frustrated, I respond: "All soluble objects dissolve in water." Hopefully, this would supply my friend with the sort of answer he was looking for. However, if this was still not enough to allay his confusion, I could assemble my explanation as follows:

1. All soluble objects dissolve in water.
2. My sugar-cube is a soluble object.
3. My sugar-cube is placed in a glass of water.
4. My sugar-cube dissolved.

One can see that the first three propositions (the *explanans*) deductively entail the fourth proposition — the statement of the event or state of affairs to be explained (the *explanandum*). Altogether it forms a valid deductive *argument* (while it may sound strange, virtually all of our explanations and predictions, when pressed, take the form of an argument). However, what is most important about the foregoing example is that it contains a *nomological* statement (namely, 1) — a law of nature that expresses the patterns to which nature adheres (Churchland 1988: 57). With the addition of some initial conditions (i.e., 2-3) we explain an event or state of affairs by deducing its description

from a law of nature — hence the name 'deductive-nomological model of explanation' (Churchland 1988: 57-8).

Churchland asserts that it is precisely this form of explanation that our commonsense explanations and prediction of behaviour take. Hence the claim that we all maintain an understanding of the laws of folk psychology; we all, as Churchland states, "share a tacit command of an integrated body of lore concerning the law-like relations holding among external circumstances, internal states, and overt behavior" (Churchland 1981: 69). So in review, then, we can see that Churchland holds the laws of a given theory as performing two functions: (1) they give sense to the theoretical terms they contain (semantic holism); and (2) they serve an explanatory and predictive function (via the deductive-nomological model of explanation).

If we agree with Churchland on (1), then the semantics of our mentalistic vocabulary should be discernible in the same manner as are the semantics of theoretical terms in general. Thus, to fully understand the meaning of 'pain', one must have knowledge of the set of folk psychological laws or generalizations in which it resides. In other words, to know the meaning of 'pain' is to know, for example, such relevant generalizations as "People tend to feel pain at points of recent bodily damage"; "People in pain tend to want to relieve that pain"; "People who feel a sudden sharp pain tend to wince"; and so on. According to Churchland, this account carries a certain straightforward plausibility, since after all, "who would say that someone understands the meaning of the term 'pain' if he has no idea that pain is caused by bodily damage, that people hate it, or that it causes distress, wincing, moaning, and avoidance behaviour?" (op. cit. 59).

(Incidentally, we can also see how generalizations such as those above, in combination with relevant initial conditions, figure into our commonsense explanations and predictions of behaviour as suggested by (2).) Yet even admitting the 'intuitive plausibility' of Churchland's explication of the meaning of mental state terms, it remains far from clear *how* the mental states *themselves* gain the perennially peculiar quality of 'intentionality'. However, as we will see in what follows, Churchland maintains that the recognition of folk psychology's theoretical status, in combination with a holistic account of meaning, provides an entirely plausible explanation of how the propositional attitudes acquire intentionality.

1.4: The Intentionality of the Propositional Attitudes

According to Churchland, the fact of folk psychology's standing as a theory entails that the semantics of its terms be understood in the same manner as the semantics of theoretical terms in general, i.e., wherein the meaning of a term is "fixed or constituted by the network of laws in which it figures" (Churchland 1981: 69-70). Obviously, we must take this claim with caution, since it remains far from certain that semantic holism is the only, or best, position with which to apprehend the semantics of our theoretical terms. However, if we grant this, Churchland insists that the intentionality normally accorded to mental states (which has long been interpreted as the profoundly enigmatic 'mark of the mental') will arise as a purely structural feature of folk psychology's concepts (ibid. 69-70). Indeed, according to Churchland, the intentionality of the propositional attitudes emerges not as mysterious anomaly of the mental, but as a product of folk psychology's

theoretical standing — in much the same manner as scientific theories discharge various 'attitudes'. Churchland offers a comparison that brings this last point to bear.

We may conceive much of the conceptual framework of physical science as being composed of certain 'numerical attitudes', as in, for example, 'has a mass (kg) of *n*'; 'has a velocity of *n*'; 'has a temperature (k) of *n*'; and so on. What is noteworthy about such statements is that they are predicate-forming expressions: "when one substitutes a singular term for a number into the place held by '*n*', a determinate predicate results" (Churchland 1981: 70). Moreover, the relations that are manifest between these 'numerical attitudes' are in fact relations between the numbers therein. Further, the argument place that takes the singular terms for numbers is open to quantification. The cumulative result of these traits allows one to express generalizations regarding the law-like relations that hold between the numerical attitudes in nature (ibid. 70). For example:

(i) (*x*) (*f*) (*m*) [((*x* has a mass *m*) & (*x* suffers a net force of *f*))

⊃ (*x* accelerates at *f*/*m*)].

Now consider the propositional attitudes of folk psychology: 'believes that *p*'; 'desires that *p*'; 'fears that *p*'; and so on. These statements are also predicate-forming expressions, for when one substitutes a proposition into the position held by '*p*', a determinate predicate results, as in, for example, 'believes that [James is tall]'; 'desires that [a cold drink is in the refrigerator]'; etc. Moreover, relations between these 'propositional attitudes' are in fact relations between the propositions therein (e.g., relations such as entailment, equivalence, and mutual inconsistency). Further, the

argument place that takes the singular terms for propositions is open to quantification. Again, the cumulative result of these traits allows us to express generalizations regarding the law-like relations that hold between the propositional attitudes (Churchland 1981: 71). For example:

(ii) (x) (p) [(x fears that p) $\supset$ (x desires that $\neg p$)];

(iii) (x) (p) [(x hopes that p) & (x discovers that p) $\supset$ (x is pleased that p)];

(iv) (x) (p) (q) [((x believes that p) & (x believes that (if p then q))

$\supset$ (barring confusion, distraction, etc., x believes that q)].

As we can well imagine, (ii-iv) are a small fraction of the vast body of laws that compose our commonsense psychological framework. Yet Churchland contends that this should not blind us to the fact that what is taking place in folk psychological laws such as (ii-iv) is the very same thing as what is taking place in scientific laws such as (i), viz., that

> ...the abstract relations holding in the domain of certain abstract objects — numbers...or propositions — are drawn upon to help us state the empirical regularities that hold between real states and objects, such as between temperatures and pressures, forces and accelerations, interacting momenta,...and between various types of mental states [Churchland 1988: 65].

In other words, the conceptual framework of folk psychology is, as Churchland insists, exploiting an intellectual strategy that is standard in many of our conceptual endeavours, which is yet another indication of folk psychology's standing as a theory.

To be sure, it is difficult to ignore the resemblance between the preceding characterization of the structure of folk psychology and that of scientific theories, which is likely why Churchland is lead to adamantly declares that "not only is folk psychology a theory, it is so *obviously* a theory that it must be held a major mystery why it has taken until the last half of the twentieth century for philosophers to realize it. The structural features of folk psychology parallel perfectly those of mathematical physics; the only difference lies in the respective domain of abstract entities they exploit — numbers in the case of physics, and propositions in the case of psychology" (Churchland 1981: 71). It need only be added that numbers and propositions are not uniquely subject to this sort of exploitation; the relations that hold among various other abstract entities (e.g., vectors, sets, groups, matrices, etc.) are also used in the same fashion (Churchland 1989: 229).

In defense of the anti-theoretical thesis, some have been inclined to argue that the laws of folk psychology (such as ii-iv) are not *really* causal/explanatory laws, but rather hold some other, less empirical status (e.g., that of normative principles, or rules of language, or analytic truths). While such characterizations of commonsense psychological precepts may be helpful to certain (usually philosophical) endeavours, Churchland contends that the laws remain first and foremost causal/explanatory in nature, particularly when they are examined as they should be — alongside the daily activities of the folk. This point becomes more evident in what follows.

Churchland notes that the concepts of folk psychology can be separated into two broad categories — those that are wholly intentional (e.g., beliefs and desires), and those that are non- or quasi-intentional (e.g., fear, grief, pain, hunger) (ibid. 227). With regard to

this latter group, consider the following everyday generalizations: "A person who suffers severe bodily damage will feel pain"; "A person who is denied food will feel hunger"; "A person who is angry tends to be impatient". It is hard to deny the causal/explanatory character displayed by this kind of generalization. For these generalizations are continuously used by the folk to support simple explanations and predictions regarding *empirical* states and events. Hence, to interpret *all* the generalizations administered by folk psychology as non-causal and non-empirical in nature seems plainly inaccurate (and yet this is, in a sense, what some proponents of the anti- or non-theoretical thesis would seem obliged to accept).

Those who wish to advance the non-theoretical thesis might, alternatively, insist that the intentional concepts of the *first* group (i.e., beliefs and desires) are entirely discontinuous with those of the second; primarily because, as it is often argued, they fulfill non-empirical, normative ends — as in, for example, 'what it means to be rational'. However, we must recognize that any normative functions fulfilled by folk psychology are not incompatible with its standing as an empirical theory. This point will emerge with greater clarity in the final chapter. Yet for the moment it might prove interesting to briefly survey how Churchland responds to this anti-eliminativist tactic.

First, Churchland notes that the fact that the intentional core of folk psychology exhibits certain semantic relations among propositions is no basis for claiming anything uniquely normative about our commonsense psychological framework. An example illustrates this point: although the regularities ascribed by the classical gas law are predicated on arithmetic relations, it does not imply (as history has shown) anything

essentially *normative* about the classical gas law (Churchland 1981: 82). Any normative intuitions that might be sensed when dealing with folk psychology are apt to be due to the fact that we happen to *value* most of the patterns ascribed by the system; but we do not value all of them — for instance:

(v) (*x*) (*p*) [((*x* desires with all his heart that *p*) & (*x* learns that ~*p*))

⊃ (barring unusual strength of character, *x* is shattered that ~*p*)].

Second, Churchland insists that it is clear that the laws of folk psychology ascribe a very narrow and truncated sort of rationality, a sort of rationality that seems anything but 'ideal' (ibid. 83). However, he notes that this should really come as no surprise, since we have yet to develop a lucid, finished account of what 'ideal' rationality is. And third, we must grant that even if our best current notions of rationality are based on the propositional framework of folk psychology, there is no guarantee that this framework will be sufficient for a deeper, more accurate account of cognitive virtue (ibid. 83). It remains far from clear why the basic parameters of intellectual excellence should be discovered at the level of propositional attitudes; for we must remember that language is something *learned* by a structure that is predisposed to do so. Indeed, language use is but one of the *many* capacities afforded by the brain, so it would seem odd, at the very least, that a theory that models cognitive activity on elements of human language should be ideal.[4]

Churchland suggests that this commonplace anti-eliminativist rejoinder likely results from an inability to appreciate the fact that theoretical concepts can serve both

theoretical *and* non-theoretical (e.g., normative) purposes. Yet Churchland insists that we must still recognize that "whatever else humans do with the concepts for the propositional attitudes [e.g., using them to delineate the meaning of rationality], they do use them successfully to predict the future behavior of others" (Churchland 1989: 228). And this being the case, one must grant that any generalization that allows us to predict one empirical state or event on the basis of another, logically distinct, empirical state or event *has* to be empirical in character (ibid. 228). Moreover, it is clear that one's abilities are not limited to prediction alone; one can also *manipulate* and *control* (two functions that are utterly typical of all theories) the behaviour of others by selectively governing the information they receive. In essence, one can steer the cognitive states of another. How such a feat could be accomplished *without* an understanding of the objective empirical regularities that connect the internal states and the overt behaviours of others is something that advocates of the non-theoretical thesis must explain if their position is to be taken seriously (ibid. 228).

Finally, it might be argued that since the laws of folk psychology require incessant qualification via disparate *ceteris paribus* clauses they are surely, strictly speaking, false (or at the very least trivial); hence, we are justified in treating folk psychology as altogether non-theoretical. Objections of this sort seem to forget that it is not Churchland's aim to defend the laws of folk psychology as true or complete (indeed, he avows the exact opposite). Churchland need not vindicate the integrity of folk psychology and the laws it administers in order to show that it is in fact a theory; many historic systems (e.g., Ptolemy's geocentric model of heavenly movement) were, and

continue to be, outright empirical theories, even though they are laden with false laws and suffer from an overall lack of completeness — and on this count, Churchland holds folk psychology to be no different: "folk psychology may be a fairly ramshackle theory, but a theory it remains" (Churchland 1988: 231). In short, the fact that the laws of folk psychology are at times vague and suffer from explanatory failure is no reason to assume that it is not a theory.

This concludes our examination of Churchland's reasons for according theoretical status to folk psychology, and to be sure, he has constructed a strong case in support of the folk-theory-theory perspective. The parallels between the form and function of folk psychology and scientific theories would seem virtually impossible to disregard. While this observation in and of itself may not qualify as conclusive proof of folk psychology's theoretical status, it does demand that those who champion the anti-theory thesis answer several pressing questions. In the meantime, it would appear that Churchland's case for folk psychology's standing as an empirical theory of mental phenomena is convincing enough to fulfill its role in his three arguments for eliminativism.

However, before proceeding any further, we should pause to take note of an important point: although positing an internally represented theory (i.e., folk psychology) is the dominant strategy for explaining our cognitive capacity to explain and predict other's behaviour, it is not the *only* strategy. Recently, a number of individuals have proposed an alternative method with which to understand these abilities — simulation theory.[5] While variations in the details abound, all simulation theorists *deny* that an internally represented folk psychological theory plays a central role when we

exercise our power to explain and predict the behaviour of others. Rather, it is argued that we use a special sort of mental simulation whereby we substitute ourselves as a model for the person whose current or future behaviour is to be accounted for. One way this might be achieved is by taking our practical decision-making system 'off-line'. Broadly speaking, off-line simulation involves feeding what we *imagine* to be the relevant beliefs and desires of another into our decision-making system which, after the appropriate processing, outputs a *hypothetical* course of action. If your decision-making system is similar to that of the person whose behaviour you wish to explain and your imaginary beliefs and desires resemble their actual beliefs and desires, then your hypothetical decision should be similar to their actual decision. Hence, instead of applying a folk psychological theory to predict and explain a person's behaviour, one is using (part of) their own cognitive mechanism as a model for (part of) another's cognitive mechanism (Stich 1996: 140). And as is the case with folk psychology, this whole process may be largely unconscious, thus making us only aware of the end result (i.e., the prediction or explanation) (ibid. 140).

The prospect that we operate in the manner simulation theory describes is undoubtedly an intriguing one. If it turns out that some version of the simulation theory is in fact correct, the impact on cognitive science and the philosophy of mind will be profound. However, as it stands, there is a great deal of straightening out to be done. There have been a host of arguments directed against simulation theory and for the moment, at least, the folk-theory-theory continues to be the strategy of choice for philosophers.

1.5: What's Ahead

From its first presentation Churchland's essay has inspired a myriad of impassioned responses against eliminativism.[6] The bulk of this work has been directed at impugning the three key arguments for elimination set out in section 1.2. Rather than sifting through this vast assemblage of critical analyses, the following chapters will examine eliminativism from a somewhat singular vantage point: I will concede that the arguments in support of the eliminativists' thesis raise several legitimate concerns regarding the integrity of folk psychology. However, I will *not* concede that these shortcomings necessarily *entail* that the eliminativists' thesis is correct (i.e., that folk psychology *will* eventually be displaced by neuroscience). Similarly, I will also *not* concede that the absence of such entailment renders the eliminativists' thesis false. Thus I will take it that *neither* the arguments for *nor* against eliminativism are conclusive.

It is likely that some will find this stance metaphysically unsatisfying since it does not aim to definitively resolve the ontological indeterminacy of the propositional attitudes. However, current opinion within the philosophy of mind seems to be in keeping with this position, as all major parties (e.g., functionalists, instrumentalists, and even eliminativists) appear to agree that the question of whether the entities described by our commonsense psychological framework truly exist is at bottom an empirical issue that will only be settled by the future discoveries of the relevant sciences.

Yet this does not mean we should cease thinking about eliminativism and the issues it raises. Indeed, while most agree that the question of folk psychology's fate ultimately rests on the proceedings of science, most also agree that there are certain ways

in which one *ought* to regard eliminativism — ways that are, it is argued, plainly preferable to others. It is with respect to such perspectives that the following chapters are designed. More precisely, it is my belief that normative naturalism, in providing detached but penetrating insight, serves as the most illuminating standpoint from which to examine eliminativism. The most intriguing feature of a normative naturalist approach is the prospect that it might be able to tell us, by virtue of the normative principles produced therein, whether we *ought* to accept the eliminativists' conclusion.

The thought of using normative naturalism in this manner first came to my attention in Stephen Stich's recent book *Deconstructing the Mind* (1996). As we shall see in the next chapter, Stich, at least initially, regards the possibility that normative naturalism might ameliorate the indeterminacy regarding what one ought to conclude about the eliminativists' thesis as quite encouraging. By extracting normative principles of ontological inference from a purely descriptive (i.e., naturalist) examination of past ontological decisions in science one could, Stich imagines, be in a position to determine whether or not it would be *rational* to accept the eliminativists' thesis. In other words, one could determine whether or not it would be *rational* to accept the claim that states distinguished in terms of their propositional content do not exist. One particular normative naturalist project that appears to accord well with Stich's objectives (in that it shows promise of generating the aforementioned principles) is that proposed by Robert McCauley. Indeed, of the various normative naturalist positions currently available it is McCauley's research into intertheoretic relations and the co-evolution of theories that commands a preponderance of Stich's praise. However, for a variety of reasons, Stich

ultimately relinquishes the hope that McCauley's work, or for that matter, the fruits of any normative naturalist enterprise, might tell us whether or not we should accept the eliminativists' thesis as rational. And it is from this juncture that Stich begins to take up the chore of constructing his new position, a position firmly entrenched in the tradition of both social constructivism and pragmatism.

It is with respect to these final points that my position takes its most marked departure from Stich's. For I maintain that (1) Stich's scepticism about normative naturalism in general, and its significance for eliminativism in particular, is overblown; and (2) that his social constructivism, in addition to bearing thoroughly disconcerting implications, seems as if it were either an overreaction to the concerns that inspire (1) or a function of what appears to be an underlying affinity for radical naturalism (of course it could also be a combination of both). In light of this conviction, I have organized the remaining chapters as follows. Chapter two is primarily devoted to outlining how Stich suggests normative naturalism ought to be applied to eliminativism, and how McCauley's particular normative naturalist undertaking might fit the bill. Chapter three begins by describing a specific sort of moderate normative naturalism and its relation to traditionalist and radical naturalist positions when pursuing epistemology's central goals. In this we will see the unique benefits that only a moderate form of naturalism can provide. Upon finishing this, the focus will be redirected to Stich's scepticism of normative naturalism as well as the social constructivism it inspires. The final portion of the chapter will offer a response to Stich as well as a general prospectus on normative naturalism and its relevance to eliminativism.

[1] Perhaps the most obvious assumption in proffering behavioural analyses is the presupposition of at least a minimum of rationality within a given subject. Yet even still, in documenting the 'characteristic' behavioural responses of a particular mental ascription it soon became clear that an indefinite or even infinite number of conditionals could potentially override the said ascription and thereby change the resultant behaviour (making it thoroughly 'uncharacteristic').

[2] Although others, such as Feyerabend (1963, 1970, 1975), played prominent roles in the advancement of eliminativism.

[3] Rorty, like most identity theorists of the era (i.e., 1950's-60's), was not so bold as to claim that empirical inquiry *will most certainly* identify sensations with brain-processes; the claim, rather, was that it was *entirely possible* that empirical inquiry might do so (see Rorty [1965]: 17-18; ft.nt. 1).

[4] The theme of this, the final of Churchland's responses, will reappear with greater clarity in the final chapter.

[5] For examples of arguments for and against simulation theory, see Goldman (1989, 1992), Gordon (1986), Caruthers & Smith (eds) (1996) and Stone & Davies (eds) (1995).

[6] See, for example, McCauley (ed) (1996); S. Christensen & D. Turner (eds) (1993); and *Mind and Language*, Vol. 8 (1993).

~ Chapter 2 ~

Normative Naturalism as a Means of Adjudication

Those who follow the literature on eliminativism will most certainly welcome Stephen Stich's recent book *Deconstructing the Mind* (1996). Stich has an exceptional talent for sweeping aside philosophical clutter and bringing genuine issues into crystalline focus. As one might imagine, this talent often leads to some startling revelations — for both Stich and his readers. Indeed, those familiar with Stich's earlier work[1] will quickly discover that his position regarding eliminativism and the future of folk psychology has undergone a dramatic transformation. Once a principal defender of the eliminativist project, Stich now contends that even the most favourable formulation of the eliminativists' premises do not sustain the desired conclusion (i.e., that our folk psychological theory of mental phenomena is so radically false that its ontology will eventually be displaced, rather than reduced, by completed neuroscience) (ibid. 3-63). Yet Stich recognizes that this argumentative deficiency does not necessarily preclude the possibility that the eliminativists' conclusion is still correct. It is in light of this that Stich turns to *normative naturalism* with the anticipation that such an approach could deliver

some principles of rational ontological inference. Stich imagines that with such principles in hand, one would be in a position to determine, given the appropriate evidence, whether or not it is *rational* to accept the eliminativists' ontological conclusion. Clearly, either outcome would have a tremendous impact on, and perhaps even resolve, the current controversy surrounding eliminativism. However, after completing his investigation of the normative naturalist strategy, Stich remains doubtful that the method will produce principles of a kind that would yield a definitive judgment. This scepticism, in turn, serves as a catalyst for the formation of his revised view regarding eliminativism; a view that is, in light of its admittedly pragmatic basis, a radical departure from his earlier work.

It would seem sensible to begin with an examination of the motivation behind Stich's investigation of normative naturalism. Hence, the chapter will be primarily devoted to outlining one particular normative naturalist approach — that espoused by Robert McCauley. McCauley's work is noteworthy for two reasons. First, it is aimed sharply at the Churchlands and their brand of eliminativism; and second, Stich openly regards McCauley's work as offering a significant illustration of how a normative naturalist could generate the type of principles of rational ontological inference mentioned above (Stich 1996: 65). Upon completing our survey of McCauley's research, we should be cognizant of its relevance to the issue at hand — eliminativism and the future of folk psychology. For we will see that, among other things, McCauley's work strongly indicates that folk psychology *will not* be displaced by future neuroscience. Yet the manner by which this conclusion is derived makes it absolutely clear that the ontology of our commonsense psychological framework is not invulnerable; indeed, the possibility

remains that folk psychology might be displaced by another theory operating at the same level of analysis. In examining all of this we should gain a better understanding of whether a normative naturalist approach such as McCauley's is apt to furnish the principles Stich seeks. And with this accomplished, we will be well equipped to take up Stich's lingering scepticism about normative naturalism in general, and McCauley's proposal in particular, since this, along with Stich's newfound pragmatism, will be the focus of the next chapter.

2.1: Ontological Determinations in the Face of Theoretical Falsity

Prior to the sceptical commentaries of, for instance, Kuhn (1962) and Feyerabend (1970), a certain view of science commanded widespread popular and academic assent. This view, fostered in part by the seemingly continuous flow of scientific triumphs that pervaded the nineteenth and twentieth centuries, depicted the sciences as directed towards a noble goal: the attainment of truth (Kitcher 1993: 3). Granted, there were impassioned disputes over what the 'attainment of truth' entailed (since some thought of truth as The Truth, while others thought of truth as an approximation to The Truth), yet all proponents of the view agreed that the "discovery of truth was valued both for its own sake and for the power [it] would confer upon us" (ibid. 3). They also agreed that science had been uniquely successful in realizing this goal. For even in the midst of scientific mistakes and false steps, they all saw a trend that accompanied each successive generation of scientists and theories: the gradual accumulation of The Truth (or alternatively, the gradual accumulation of better and better approximations to The Truth).

The explanation for this predominantly progressive trend credited the use of the Scientific Method. Again, there was disagreement on the details of the method, but as Philip Kitcher has remarked, all concurred on some essential points:

> There are objective canons of evaluation of scientific claims; by and large, scientists (at least since the seventeenth century) have been tacitly aware of these canons and have applied them in assessing novel or controversial ideas; methodologists should articulate the canons, thus helping to forestall possible misapplications and extend the scope of scientific method into areas where human inquiry typically falters; in short, science is a "clearing of rationality in a jungle of muddle, prejudice, and superstition" [Kitcher 1993: 3].

Indeed, some even went so far as to regard science as the pinnacle of human achievement not because of its innumerable accomplishments, but *solely* by virtue of the fact that — in success or failure — its practice was thoroughly informed by reason (ibid. 4).

Interestingly, many of the most ardent champions of the preceding view of science were not practitioners but philosophers. For instance, Rudolph Carnap, Gustav Hempel, Ernest Nagel, and Hans Reichenbach all shared the conviction that the succession of theories in the physical sciences constituted a progression, and that the achievements of earlier theories were retained in later theories (ibid. 5). Moreover, while conceding that there was no algorithm for generating new hypotheses, all agreed that once hypotheses did arise, there were incontestable principles for their proper assessment (ibid. 5). Hence they endeavoured, along with the rest of the logical positivists, to uncover and transcribe the logic of confirmation, the logical structure of theories, and the logic of explanations; in

short, they attempted to lay down the canons that were taken to be both characteristic of 'good' science and tacitly employed by practitioners therein on a daily basis.

However, as the previous references to Kuhn and Feyerabend suggest, a steady barrage of critiques beginning in the late 1950's provided compelling evidence that such a view of science was, to borrow Kitcher's expression, "smug, uninformed, unhistorical, and analytically shallow" (Kitcher 1993: 5). Many critics challenged the notions of progress and rationality that were central to those who held the view. Questions such as 'Can we legitimately view truth as a goal of science?'; 'Is it meaningful to talk of approximations to the truth or to see science as converging on the truth about nature?'; 'Can we expect to attain even part of the truth about nature?'; and 'Does the historical record show even the accumulation of truths about observable phenomena?' cast serious doubt on beliefs that had formerly enjoyed almost universal acceptance (ibid. 6). This was especially true of the belief that each successive generation of scientists brought contributions to an ever-growing repository of truth.

As Kuhn (1962) so eloquently illustrated, the history of scientific progress bore little resemblance to the description sanctioned by advocates of the view. An honest examination revealed that the dynamics that framed each successive generation of scientists and theories were considerably more complex than was previously thought, and thus, anything *but* a relatively steady procession towards the truth. It also became clear that, contrary to what the logical positivists imagined, incorporating the achievements and eliminating the errors of past theories was not always a cut-and-dried procedure. For the tacitly employed canons of scientific rationality (which together composed the Scientific

Method) that the logical positivists strove to extract and uphold did not always endorse the past decisions and resultant triumphs and failures of scientists. Feyerabend (1970, 1975) accentuated this point in his exposition of how Galileo's defense of Copernicanism violated philosophers' favourite rules of good reasoning. This, in turn, led to Feyerabend's declaration that, aside from 'Anything goes', there were *no* standards of good reasoning that bound all scientists at all times, and thus, *no* maxim that definitively united all of the decisions found in the history of science (Kitcher 1993: 7). And in the absence of such standards, the possibility of settling the discrepancies that almost inevitably arise between successive theories in science was seriously hampered.

Currently, the difficulties associated with reconciling the achievements and errors of successive theories in science continue to resist simple resolutions. Some of these problems find expression within the debate over eliminativism. Reconsider, for instance, Paul Churchland's articulation of the eliminativist thesis:

> 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 [Churchland 1981: 67].

Setting aside the various arguments that are often used in support of this claim, one can ask a seemingly innocuous question: even if, as Churchland insists, folk psychology is to be succeeded by neuroscience, how are we to decide whether the entities posited therein (e.g., the propositional attitudes) exist? For, eliminativist inclinations notwithstanding,

the disclosure of a theory's falsity does *not* always entail the non-existence of its respective postulates — a point that is readily supported by a brief review of the history of science (e.g., modern astronomy views the Copernican theory of planetary movement as false and yet many Copernican theoretical entities, such as planets, enjoy a continued, unquestioned existence). If folk psychology's falsity will not suffice to settle this matter, what can? Moreover, how are we to decide whether or not the entities posited by *any* false theory exist?

Stich suggests that a potential answer to these questions may reside in what is commonly labeled the *normative naturalist* strategy (Stich 1996: 63-5). The central aim of the approach is to extract some normative principles of rational ontological inference or decision-making from actual historical cases in which such decisions have been made. The first step in deriving such principles might involve dividing a number of historical cases into the following two groupings: (i) theories that were judged to be incorrect and yet the entities therein were still taken to exist (e.g., planets, atoms, brains); and (ii) theories that were judged to be incorrect and their corresponding entities non-existent (e.g., caloric, phlogiston, the alchemist's 'spirits').

One could then search for prominent similarities and differences between the two groupings in the hope that certain features will emerge from most or all of the cases in one group and not from the other. These features could then be incorporated into principles that detail when it is rational to preserve items in the ontology of a mistaken theory and when such items should be 'eliminated'. With this done, the candidate principles can be evaluated by testing them against historical cases that were not included in either (i) or

(ii), and if the principles endorse most of the decisions that were in fact made, then we would have reason to believe the principles are sound. And even if there were a few instances in which the principles' judgment conflicted with the actual historical decisions, this would not immediately threaten the integrity of the principles, since the people involved in those historical decisions may have made the wrong (i.e., not rational) choice (Stich 1996: 64).[2]

According to Stich, there are very few normative naturalist endeavours akin to the preceding outline that show promise of generating principles of rational ontological inference strong enough to assist us in an assessment of the eliminativists' argument. As was intimated at the outset, Stich regards Robert McCauley's research into intertheoretic relations and the various 'levels of analysis' found in science as the most interesting. Therefore it warrants a closer look.

2.2: Intertheoretic Relations and the Co-evolution of Theories

Prior to presenting his own account of intertheoretic relations, McCauley offers an informative sketch of Paul Churchland's stance on the subject (McCauley 1986a: 69-71) Churchland conceives of intertheoretic relations along the lines of a basic continuum wherein one end is marked by theories that are thoroughly continuous with their successors, while the opposite end is marked by theories that are radically incommensurable. Filling out the continuum are numerous theory-pairings whose relations exhibit various degrees of compatibility. Correspondingly, the extent of reducibility between theories becomes "a direct function of (1) the number of the reduced

theory's propositions that are important both semantically and systematically that we can map onto the propositions of the reducing theory and (2) the ease with which they can be mapped" (McCauley 1986a: 69).

The success of a given reduction is thus judged on whether or not a faithful image of the reduced theory's central claims can be found within the framework of the reducing theory.[3] If a faithful image can be found without expending an undue amount of effort, then the ontology of the reduced theory is, at least roughly, preserved (e.g., the reduction of the laws of classical thermodynamics to the principles of statistical mechanics). Conversely, when very few or none of a theory's central claims map onto its potential successor, the theories are said to be radically incommensurable, which, in turn, usually leads to the wholesale displacement of the inferior theory's ontology (e.g., the 'elimination' of Stahl's system of chemistry and Darwin's theory of inheritance by their respective successors). And of course, this is precisely the fate the Churchlands predict for folk psychology, and indeed any other psychological theory bearing commitments to the propositional attitudes, upon the completion of neuroscience.

However, one should also recognize that at times, the disagreement between a theory and its proposed successor may persist, particularly when neither of the theories has the capacity to account for, or encompass, the other. Under such circumstances proponents of the clashing theories often seem to 'talk past one another' because of the vast dissimilarity of their conceptual frameworks, and traditional measures of theoretical excellenceare thus rendered ineffectual, since the disputants are likely to disagree on the relative importance of problems, the appropriateness of methods, and so forth.

Therefore, other considerations must be invoked in order to settle the matter (such as how well each of the theories coheres with the body of science), and, if they are addressing the same phenomena, one of the theories inevitably succumbs to elimination (McCauley 1986a: 70).

Elsewhere McCauley has remarked that Patricia Churchland's (1986) examination of the co-evolution of theories appears well-suited to the preceding account of intertheoretic relations since her work points to three co-evolutionary scenarios that are aptly differentiated in terms of their position on the aforementioned continuum (McCauley 1996: 24-8). The first of the said scenarios suggests that psychology and neuroscience might co-evolve in a fashion that results in an enhanced level of cohesion between the two. For Churchland concedes that if there is "theoretical give and take" then, in all likelihood, the two sciences will "knit themselves into one another" (op. cit. 374). McCauley observes that the metaphor of two sciences knitted into one another "implies an integration that is tight, orderly, and detailed" and hence most appropriately situated nearer the 'thoroughly continuous' endpoint of the continuum (op. cit. 24). However, the Churchlands remain steadfast in their insistence that such 'knitting' will never wholly satisfy the traditional demands of microreduction, and that throughout these co-evolutionary developments, neuroscience will invariably retain both explanatory and metaphysical priority. Yet it is still anticipated that intertheoretic integration would enable neuroscience to supply an equipotent image of psychological principles. McCauley is thus prompted to label this scenario 'co-evolution$_m$', since psychological

and neuroscientific theories are depicted as co-evolving in the direction of an *approximate* microreduction (i.e., *towards* the 'thoroughly continuous' endpoint).

The second co-evolutionary scenario suggests that a genuine integration of psychology and neuroscience, if it is indeed to be sought and attained, will most likely come *after* psychology's initial dismantling and subsequent reformation in conformity with the mandates of neuroscience. Accordingly, psychology and neuroscience are construed as moving, at least for the interim, in the *opposite* direction of co-evolution$_m$; that is, towards the 'radically incommensurable' endpoint of the continuum. Because this incipient process culminates with an effect typical of scientific revolutions (i.e., the eradication of substantial portions of psychology, such as suppositions about learning and memory), McCauley designates this scenario 'co-evolution$_s$'.

Clearly, co-evolution$_s$ underlies Paul Churchland's inaugural (i.e., 1981) rendition of eliminativism. Yet at times the Churchlands of recent seem to portray the foregoing scenarios as two stages of a single co-evolutionary process, the first of which attends to "the demolition of much current psychology [i.e., all that bear commitments to the propositional attitudes, including social and cognitive] via co-evolution$_s$", while the second ministers the "reconstruction of a neuroscientifically inspired psychology via co-evolution$_m$" (McCauley 1996: 26). Whether the Churchlands are in fact proposing an amendment to the eliminativist thesis is uncertain; however, the vital point, according to McCauley, is that both co-evolutionary scenarios interpret the relationship between theories as moving in one direction (as opposed to the other) on Churchland's continuum (ibid. 26). In light of this, it would seem necessary to ask: if a co-evolutionary shift in

either direction is possible, then what are the variables that decide the direction? McCauley claims that this question presses the aforementioned version of eliminativism just as forcefully as the original (i.e., elimination without a subsequent reconstruction of psychology), since it is still unclear what *precisely* determines a co-evolutionary shift in one direction as opposed to another. Although in fairness to the Churchlands, inquiries into intertheoretic relations (which are, relatively speaking, in their infancy) seem to indicate a greater degree of complexity than what was probably anticipated at first. Thus the fact that they have yet to form a direct reply may, for the moment at least, be excusable. Which brings us to the third and final co-evolutionary scenario: 'co-evolution$_p$' — that is, co-evolution as explanatory pluralism.

Although co-evolution$_p$ bears many similarities to co-evolution$_m$, there are some crucial differences. Whereas co-evolution$_m$ anticipates heightened intertheoretic integration largely guided by, and with a default preference for, neuroscience, co-evolution$_p$ "construes the process as preserving a diverse set of partially integrated yet semi-autonomous explanatory perspectives — where the non-negligible measure of analytical independence rests at each analytical level on the explanatory success and epistemic integrity of the theories and on the suggestiveness of empirical findings" (McCauley 1996: 27). In short, co-evolution$_m$ affirms unidirectional selection pressures (i.e., from the bottom [neuroscience] up), while co-evolution$_p$ respects the constraints imposed by the needs and demands of theories operating at higher levels of analysis (e.g., cognitive psychology).[4]

At a glance, this distinction may appear minor; however, it is actually an embodiment of a fundamental epistemological and metaphysical disagreement that has repeatedly divided scholars within the philosophy of science. For on the one hand, broadly speaking, if physicalists are not persuaded by co-evolution$_s$, they usually opt for co-evolution$_m$, since it "suggests a science unified in both theory and ontology that accords priority to the lower (i.e., physical) levels" (McCauley 1996: 27). On the other hand, philosophers with pragmatic leanings generally favour co-evolution$_p$, placing the explanatory resources afforded therein ahead of assurances of, and worries about, a unified science and metaphysical purity. Those who have followed the Churchlands' work will recognize that they have spent much of the last decade carefully balancing these two interests (i.e., an interest in a unified science and metaphysical purity versus an interest in enhanced explanatory resources). Yet currently, it would seem that the Churchlands' increasingly relaxed eliminativism (e.g., P. M. Churchland: 1993), in tandem with their emerging preference for co-evolution$_p$ (e.g., P. S. Churchland 1986: 363, 368, 373, 376; quoted by McCauley 1996: 28), indicate an escalating influence of pragmatic considerations in their thought.

Irrespective of the extent to which their pragmatic interests are gaining in import, Paul Churchland's continuum of intertheoretic relations, when combined with Patricia Churchland's research into the co-evolution of theories, forms an intuitively appealing model. This appeal is due in part, no doubt, to the model's relative simplicity. Moreover, whereas previous accounts were habitually confined to either classical microreduction (and in most cases, co-evolution$_m$; see Nagel: 1961) or to revolutionary

science (and in most cases, co-evolution$_s$; see Kuhn: 1970), the Churchlands' model accommodates a comprehensive *range* of intertheoretic relations, as well as three distinct co-evolutionary scenarios (and perhaps more than three if the Churchlands' ultimately admit to allowing combinations thereof). By gauging the practicability of mapping one theory's central claims onto another, and (by one manner or another) selecting the most appropriate co-evolutionary scenario, the model provides us with a generally straightforward plan for assessing specific cases in science. Yet McCauley insists that the Churchlands' model fails to capture the full complexity of both intertheoretic relations and the co-evolution of theories, and for this reason, demands revision. Such is the motivation behind McCauley's own model, which, aside from drawing a more robust (and presumably more accurate) picture of intertheoretic relations and the co-evolution of theories, carries profoundly negative repercussions for the eliminativists' project.

2.3: Interlevel Contexts and the Preservation of Folk Psychology

The crux of McCauley's proposal rests on the notion of 'levels of analysis' and the function of temporal factors when considering intertheoretic relations. Regarding the former, McCauley insists that there is a roughly hierarchical correspondence between levels of analysis in science and levels of organization in nature: "Broadly speaking, chemistry is a higher level of analysis than subatomic physics, since it concerns larger units and events which stand in causal relationships most economically described in chemical terms...Again, broadly speaking, biology is an even higher level, and psychology higher than that" (McCauley 1986a: 72). Furthermore, the altitude of a level of analysis

is said to be *inversely* proportional to the size of its respective domain of events, while also *directly* proportional to the complexity of the systems with which it deals. Thus, McCauley states that "higher-level sciences deal with increasingly restricted ranges of events having to do with increasingly organized physical systems [e.g., psychology]" (McCauley 1986a: 73), while lower-level sciences examine a wider range of events that pertain to comparatively simpler systems (e.g., molecular biology). Additionally, as one proceeds to lower levels of scientific analysis, one will recognize that the systems studied therein will have existed for a longer period of time; so, for example, whereas the particles studied by atomic physics "date from milliseconds after the Big Bang (approximately sixteen billion years ago, according to recent estimates)", the cultural arrangements that the socio-cultural sciences study "can be measured in thousands of years" (McCauley 1993: 4). Interestingly, and for McCauley crucially (as we shall see), practitioners of these upper-level sciences also tend to accord greater weight to *functional* ascriptions (as opposed to the purely *structural* considerations favoured by lower-level sciences) when dealing with the escalating intricacies exhibited by the systems in their fields (physiology is an excellent example of this partiality).

In addition to levels of analysis, one must also take temporal factors into account when evaluating intertheoretic relations. More precisely, McCauley stresses the importance of recognizing the contrast between relations that hold among theories at a *single* level of analysis *over time* (i.e., successional or *intralevel* contexts) and those that hold among theories at *different* levels of analysis at the *same time* (i.e., microreductive or *interlevel* contexts) (McCauley 1986a: 73). When mapping between theories in intralevel

contexts is fairly extensive (the 'thoroughly continuous' endpoint on Churchland's continuum), the new theory can be interpreted as *correcting* the old; it *explains* the older theory "in the sense that it offers a principled account of when and why [the older theory] fails" (McCauley 1986a: 74). Under such circumstances, the earlier theory's domain is most often regarded as a *special case* of the new theory and for which the old theory continues to suffice as a useful calculating heuristic (e.g., using Newton's laws of motion for certain calculations instead of the more complex laws of relativity) (McCauley 1996: 29). Hence, the transition is best understood as *evolutionary*, since the new theory *inherits* the evidence for the old (ibid. 29), and there is no talk of having categorically 'eliminated' the older theory's ontology:

> Generally, new theories retain old terms when possible. We have retained terms such as "planet," "evolution," and "gravity" and propositions about rectilinear inertial motion through numerous reinterpretations, because the effects of the reinterpretations taken individually were not especially severe...So long as the relevant changes have reasonably local effects that do not destroy larger conceptual patterns, to claim that incommensurability seriously threatens theory comparison is to overstate the case [McCauley 1986a: 74].

It is when intralevel relations evince profound discordance (the 'radically incommensurable' endpoint on Churchland's continuum) that the potential for ontological elimination arises. If it is essentially impossible to translate an older theory into its immediate successor, scientists generally decide to replace, and hence eliminate, the inferior theory and its ontology (such circumstances are akin to Kuhnian periods of revolutionary science) (ibid. 75). Compared to the foregoing scenario, such transitions are

(relatively speaking) rather abrupt. The examples that pervade the relevant literature in philosophy notwithstanding (e.g., phlogiston, caloric, impetus), the history of science indicates that this sort of theory-elimination is also extraordinarily *rare* (a point that many have come to realize; see, for instance, Thagard: 1992).

McCauley initiates a discernible split from Churchland with the further contention that *all* intralevel contexts *can* eventually result in the elimination of some theories (McCauley 1986a: 75-6). For such may be the final outcome when a current reigning theory is so greatly separated from its ancestral counterpart by successive generations of theories that mapping becomes both excessively convoluted and largely insubstantial. Consider, for instance, the abundance of incremental reinterpretations of 'natural motions' that divide Aristotelian and Newtonian mechanics. Some would insist that the cumulative effect of these transitions has been so significant that that which was originally conceived as 'natural motions' no longer exists (ibid. 75). Thus, in light of this inclusion, McCauley's account of *intra*level contexts opens theories and their respective ontologies to elimination from *either* relatively rapid, unforeseen scientific revolutions *or* prolonged, well documented evolutionary advances.[5]

However, it is McCauley's discussion of *inter*level contexts that signals the most dramatic departure from Churchland's continuum: "Interlevel contexts, by contrast, involve *no elimination* whatsoever. These are cross-scientific contexts where the goal is to associate theories that operate at different levels of analysis" (ibid. 76; emphasis added). Such activity is largely symbiotic (i.e., mutually advantageous); scientists look to theories at neighbouring levels in an effort to gain new insight into the phenomena at

issue. The discovery of conceptual ties (or lack thereof) between theories at different levels of analysis serves to support (or contest) the scientists' belief that they are dealing with a single explanandum from disparate perspectives. When such mapping between levels proves highly successful (i.e., extensive intertheoretic 'knitting'), the theories tend to heavily constrain one anothers' form (which is precisely the effect that co-evolution$_m$ envisions) (McCauley 1996: 30). Still, McCauley insists that, at best, only *partial* replaceability may result: "A theory at one level, well integrated with theories at adjacent levels, can (ideally) do some of their work under special circumstances" (McCauley 1986a: 76).

Notably, the foregoing applies to *both* upper-level and lower-level theories. A well integrated lower-level theory can — although typically at considerable computational expense — direct its conceptual resources in a manner that reproduces the predictive accomplishments of the upper-level theory; and analogously, the results of an upper-level theory will often conform precisely to the predictions of its lower-level counterpart (a characteristic that is entirely absent in both types of intralevel evolutionary cases). However, because of an admitted lack of fined-grainedness, upper-level theories will *at times* require correction from their lower-level correlate. But this is in stark contrast to *intra*level corrections, which always arise because "the earlier theory is wrong — by a little in evolutionary cases, by a lot in revolutionary ones" (op. cit. 31).

Moreover, the utility of most well-integrated upper-level theories in interlevel contexts far surpasses that of the replaced predecessor in intralevel evolutionary cases. For the conceptual resources of upper-level theories can be used in a manner that

efficiently yields what are often *invaluable* approximations of all lower-level variables (this can be seen, for instance, in the immense disparity in requisite computational effort between the classical and statistical solutions for simple problems about gases [an interlevel case] versus the comparatively diminutive disparity between classical mechanics and the mechanics of relativity for simple problems about motion [an intralevel case]). Yet in a sense, all upper-level theories in interlevel contexts possess a characteristic that is also present in most intralevel evolutionary predecessors. For upper-level theories invariably detail the regularities of a *subset* of the phenomena the lower-level theory encompasses, but which it has "neither the resources nor motivation to highlight" (McCauley 1996: 31)[6], and most intralevel evolutionary predecessors, recall, minister to a domain that is usually only a *part* of their successor's enlarged range.

However, when theories at corresponding levels exhibit fundamental incompatibilities, the basic distinction between intralevel and interlevel settings becomes plainly evident. In such cases, two theories at adjacent levels address some common explananda in ways that seem largely incommensurable — which is exactly how the Churchlands view the relationship between folk psychology and neuroscience. And as McCauley agrees, if all intertheoretic relations should indeed receive a unified treatment as traditional reductionists (such as the Churchlands) suggest, then it is perfectly reasonable to anticipate folk psychology's elimination. The problem, however, is that "neither the history of science, nor current scientific practice, not the scientific research the Churchlands champion, nor a concern for explanatory pluralism offers much reason to expect theory elimination in such settings" (ibid. 31). McCauley admits that in the

infancy of a science's history it is not always easy to distinguish levels of analysis, or, consequently, what would count as an intralevel as opposed to an interlevel context. The vital point, though, is that the history of science, especially the history of late-nineteenth and twentieth century science, offers *no* examples of large-scale interlevel theory elimination (particularly that of the wholesale variety the Churchlands' and co-evolution$_s$ envision) once an upper-level theory gains enough momentum to enjoy the accoutrements of other recognized sciences (e.g., characteristic research techniques and instruments, journals, university departments, professional societies and funding agencies) (McCauley 1996: 32). The reason for this, McCauley explains, is quite simple: mature sciences are principally defined by their theories, and more generally, by their research traditions, and thus, the elimination of an upper-level theory by a lower-level theory "may risk the elimination of the upper-level scientific enterprise!" (ibid. 32).

As previously mentioned, a motive for examining theories at adjacent levels of analysis — *especially* when intertheoretic connections are not plentiful — is to explore the possibility of using one science's successful problem-solving strategies as a means to "inspire research, provoke discoveries, and solve recalcitrant problems at another level" (ibid. 32). For monitoring the developments in theories at neighbouring levels very often serves as a productive heuristic of discovery, and this, to a large extent, is a direct effect of the two sciences having maintained a measure of independence from one another (as one will recall, this is the hallmark of co-evolution$_p$). Thus, a scarcity of intertheoretic links only serves to *strengthen* the autonomy of the upper-level science. Indeed, as Wimsatt remarks, "in interlevel reduction, the more difficult the translation becomes, the

more *irreplaceable* the upper-level is! It becomes the only practical way of handling the regularities it describes" (McCauley 1986a: 77; from Wimsatt 1976: 222; emphasis in original). Hence, radical incommensurability — in *inter*level contexts — *neither entails nor incites elimination* (op. cit. 76).

This final point poses a serious threat to the eliminativist thesis, since McCauley claims that the relationship of folk psychology to neuroscience is clearly one between theories *at different levels of analysis* (ibid. 79). Consider, for instance, the way in which folk psychology and neuroscience address many of the same phenomena. Folk psychology, like physiology and most other upper-level sciences, makes extensive use of functional ascriptions when dealing with the complex systems it subsumes. Neuroscience, on the other hand, is typical of lower-level science in that it tends to favour structural considerations. Each discipline brings different concepts and principles to bear, and are commonly interpreted as highlighting different aspects of a shared explanandum. Finally, it is clear that folk psychology deals with a wider range of variables, while neuroscience deals with a wider class of event — qualities that are fully anticipated if we take the theories to operate at contrasting levels of analysis. In short, the evidence overwhelmingly indicates that folk psychology and neuroscience are indeed situated in an interlevel context.

Hence, it seems imperative to determine whether co-evolution$_p$ appropriately describes the relationship between folk psychology and neuroscience (for given their prima facie incongruities, it seems unlikely that co-evolution$_m$ will prove an appropriate description — although it may be for cognitive psychology and neuroscience). Therefore,

the pivotal question must be: does folk psychology contribute to our knowledge? Or better yet: does folk psychology contain resources that may assist more systematic psychological theorizing? The answer is obvious, for there has been a wealth of knowledge purveyed by theories that make as much use of the propositional attitudes as does their originator, folk psychology (e.g., memory theory, perception theory, reasoning theory, etc.). Interestingly, it also looks like psychological science is "simultaneously employing and, ever so gradually, *transforming* familiar folk psychological notions" (McCauley 1996: 33) — a point that lends credence to the hypothesis that folk psychology is currently undergoing intralevel evolutionary processes. In light of this, it would appear that the Churchlands might be guilty of underestimating and undervaluing the role of folk psychology's conceptual resources in social psychological and cognitive theorizing. McCauley suspects the same, and attributes the Churchlands' neglect to their earlier aspirations of a unified account of intertheoretic relations and the accompanying propensity towards co-evolution$_s$.

The trouble, of course, is that the eliminations co-evolution$_s$ depicts are *exclusively* intralevel processes, and only so when "the levels in question concern scientific pursuits as well established as neuroscience and psychology and...when those levels are construed as thickly, i.e., as inclusively, as the distinction between those two sciences implies" (ibid. 33). These conditions were present in the theories and ontologies informing, for example, Stahl's account of combustion and Young's account of the propagation of light, both of which were replaced by new theories with new ontologies operating at the *same* level of analysis, and that were identified, both now and then, as *continuations* of the

established research traditions associated with those levels. However, in applying co-evolution, to *inter*level cases (i.e., folk psychology and neuroscience), the Churchlands have conflated "the dynamics of the co-evolution of theories at *different* levels of analysis with those of scientific revolutions, which are [without exception] *intra*level processes" (McCauley 1996: 34; emphasis added). Consequently, since folk psychology and neuroscience are situated in an interlevel context, we must conclude that, irrespective of the Churchlands' decree, there can be *no* elimination:

> The mistake *all* versions of eliminative materialism have made is to draw their eliminativist conclusions about the *inter*level relationship between psychology and neuroscience on the basis of an analysis appropriate to *intra*level contexts. The eliminativists *correctly* claim that theories at the two levels have important conceptual discontinuities, but they incorrectly conclude...that such incommensurability requires the elimination of one or the other. In *intra*level contexts during scientific revolutions such crises *do* require that sort of radical surgery, but in interlevel contexts such a measure would eliminate potentially important stimuli for scientific discovery [McCauley 1986a: 79].

This is not to deny that decisions at both lower and upper levels can influence theoretical developments at another level. Rather, McCauley's point is that these influences are reliably mediated by developments in the conceptual apparatus and research practices that are associated with the tradition of the level in question. The implication, of course, is that this consummate denial of the eliminativists' conclusion does not render folk psychology invulnerable. For while McCauley's model illustrates why lower-level sciences do not threaten the continued existence of folk psychology, it also suggests that

*intra*level perils may persist. Indeed, the intralevel elimination of folk psychology, resulting from either scientific revolution *or* evolutionary advances in science, remains a very real possibility — something that has not escaped the Churchlands' notice: "Even if an abstract or higher-level explanatory framework were somehow essential to grasping psychological phenomena, it would remain an open question whether our current [folk psychology] is the *correct* framework with which to meet this challenge" (P. M. & P. S. Churchland 1996: 225).[7] And the fact remains that neuroscience may play a significant role in fostering developments in what might eventually become folk psychology's intralevel successor.

It is of interest to note that this last point carries peculiar implications for those who have *defended* folk psychology's ontological status by citing its contributions to disciplines, such as cognitive psychology, that operate at the same level of analysis. For example, Horgan and Woodward have referred to the widespread use of folk psychological concepts in theories that detail such divers phenomena as visual perception, memory, and learning in their attempt to discredit the Churchlands' claim that folk psychology exhibits profound explanatory deficiencies (Horgan & Woodward 1985: 146-7).[8] While such an approach may help to weaken this specific allegation, it also seems to *accentuate* the prospect of folk psychology succumbing to intralevel elimination, particularly as a result of evolutionary advances within the relevant sciences. For the probability of intralevel evolutionary elimination will undoubtedly *increase* as folk psychology (or neuroscience, for that matter) continues to facilitate the development of its potential successors. Although such a process would likely take decades (if not

centuries) to complete, the fact remains that folk psychology may play a central role in bringing about its own elimination — a point that many ardent defenders of folk psychology fail to recognize, let alone address. Yet we should also realize that in the event of folk psychology's elimination by some other upper-level theory that (regardless of extent) implements intentional concepts, the Churchlands could claim no victory; indeed, they would simply have a new target.

However, before we can be assured of the preceding conclusions, we must ask: is it indeed the case, as McCauley maintains, that the history of science offers no examples of interlevel theory-elimination? What about, for instance, the apparent displacement of many intentional-concept upper-level psychological theories of cognitive disorders, such as schizophrenia, by lower-level neurophysiological accounts?

McCauley insists that these types of cases should not affect an analysis of intertheoretic relations in science largely because such upper-level explanations are not, strictly speaking, scientific.[9] It is argued that such explanations are so vague and underspecified that they rarely submit to empirical testing (a presumably essential characteristic of scientific theories), and when they are detailed enough to submit to such tests, they prove thoroughly inadequate. In short, McCauley agrees with Popper's position; viz., that such explanations are so far removed from paradigmatic examples of scientific theories that one can justifiably say that they hold some sort of 'fringe status'. However, this is not to suggest that these upper-level explanations cannot perform an important service to scientific inquiry. To be sure, while offering thoroughly mistaken explanations of cognitive disorders, these 'theories' often fulfill an indispensable function

in that they provide a means for efficiently *identifying* cases of the said phenomena, cases that can then be expedited to the relevant sciences.

It remains uncertain whether McCauley's response convincingly addresses the foregoing concern. For it seems as though the outcome turns on what sort of criteria is utilized for distinguishing scientific theories from non- or quasi- scientific theories. And on this count there is ample room for reasonable people to disagree (i.e., there is no single criteria that enjoys widespread acceptance). It is clear, however, that if there is a legitimate example of interlevel theory-elimination in science, such as that which pertains to schizophrenia, then McCauley cannot draw the sort of conclusion he does about eliminativism: i.e., we cannot be assured that neuroscience will not displace folk psychology.

2.4: Will Normative Naturalism Prevail?

There is certainly no shortage of impassioned defenders of folk psychology. Many of the rebuttals issued by this fellowship are motivated by an intuitive feeling that the posits of folk psychology, particularly beliefs and desires, are somehow essential to human experience, and hence, not open to ontological elimination. Of course the methods by which this sentiment has been articulated vary greatly. Some have argued that the propositional attitudes are necessary to the perception of one's self as a rational agent (e.g., Hannan 1993); others cite the indispensable pragmatic utility of the propositional attitudes to everyday life (e.g., Lahav 1992); and still others have defended (albeit indirectly) the ontological status of intentional concepts by citing their requisite use in

scientific psychology (e.g., Fodor 1987, 1990). The list could go on. The important point, however, is that they all, in one way or another, argue for the *necessity* of the propositional attitudes. And simply because it *lacks* this prevalent trait, McCauley's position occupies a unique place within the debate surrounding eliminativism.

McCauley's defense of folk psychology is notable for other reasons as well. First, it does not treat the propositional attitudes as exempt from the possibility of ontological elimination (many arguments in support of folk psychology do accord such immunity, often in virtue of some 'special standing' such as that described above). Second, the theory of intertheoretic relations and the accompanying view of the co-evolution of theories that informs the argument that ultimately shields folk psychology from the eliminativists' charge is thoroughly informed by comprehensive research into the history of science — a unique tactic, given the literature.

McCauley's model of intertheoretic relations and the co-evolution of theories also seems preferable to the Churchlands' alternative. Though more complex, McCauley's model still comes across as a reasonable outcome given the apparently incontrovertible circumstances (i.e., that there exist levels of organization in nature to which levels of analysis in science roughly correspond, and that this has an impact on the character of individual theories as well as their relations to one another). And it is precisely with respect to this point that the simplicity of the Churchlands' model would seem to impede its accuracy, particularly when one considers its inability to account for several prominent characteristics of intertheoretic relations and the co-evolution of theories. For instance, whereas the Churchlands' model does not account for the fact that scientific

revolutions are remarkably rare occurrences (and in fact suggests the opposite in order to reach the desired ontological conclusion), McCauley's model offers a clear and comprehensive explanation of *why* the phenomenon is so scarce. Furthermore, it is evident that some sciences make greater use of functional ascriptions when dealing with increasingly complex systems that concern a smaller domain of events, while other sciences favour structural considerations that involve comparatively fewer variables and offer wider application. While absent from the Churchlands model, we have seen that McCauley's notion of 'levels of analysis' easily integrates this fact into the larger scheme of intertheoretic relations.

Altogether, McCauley develops a formidable case against the Churchlands' thesis, and does so without forcing those with preferences for naturalism and scientific realism to abandon their convictions. Thus, in terms of its specific relevance to eliminativism, and its broader relevance to intertheoretic relations, McCauley's proposal is of substantial importance. However, while he does not explicitly say so, it is evident that Stich finds McCauley's resolution of the eliminativist controversy unconvincing. For Stich, while applauding McCauley's efforts, remains unwilling to reject the prospect that the eliminativists' conclusion — i.e., that *neuroscience* will displace folk psychology — might be valid.

Hence we turn to the possibility that McCauley's insights might be manipulated in such a way so as to confer the sort of principles of rational ontological inference Stich desires. At first glance, McCauley's proposal seems promising in this regard, for the manner by which he ultimately discounts the eliminativists' conclusion seems to have the

semblance of a normative judgment. Indeed, in light of McCauley's final determinations, we *could* derive one normative principle: no theory at one level of analysis has eliminated a theory at another level of analysis, therefore none should be thought to do so. With this principle in tow, many controversial claims regarding the elimination of one theory by another could be quickly resolved by simply ascertaining — via the appropriate criteria[10] — whether or not the theories in question reside at the same level of analysis. If they do not, as is the case with folk psychology and neuroscience, then the claim of elimination can be immediately discarded — since, in light of our *normative naturalist* principle, it would be *irrational* to accept such a claim.

However, it is evident that such a principle does not address the eliminativists' conclusion in its entirety; nor does it, for that matter, address Stich's primary concern. For the eliminativist thesis, as noted at the outset, states that folk psychology is a *false* theory that, along with its ontology, will eventually be displaced by neuroscience (Churchland 1981: 67). Clearly, the foregoing principle applies only to the *latter* portion of the eliminativists' thesis, since it simply tells us that it would be irrational to think that *neuroscience* will displace folk psychology. It says *nothing* about folk psychology's truth or falsity. Nor does it say anything about what might become of folk psychological posits, such as beliefs and desires, *if* it should turn out that folk psychology is seriously mistaken. And this final point — that is, how to decide whether or not the *entities* postulated by *any* false theory exist — is a question that is of the utmost importance to Stich. In light of this, one can understand why Stich expresses dissatisfaction with the often imprecise manner with which McCauley differentiates the 'elimination' of theories

and the 'elimination' of their ontology (or individual entities therein) (Stich 1996: 65), since this distinction would seem to play a vital role in answering the preceding question. Whether McCauley's model can be readjusted so as to correct this particular deficiency remains to be seen.

It will turn out, however, that this shortcoming is almost trivial, since, as we shall see in the next chapter, Stich harbours even deeper reservations about normative naturalism in general; reservations that are in fact *so* deep that one is led to wonder whether *any* normative naturalist undertaking — McCauley's included — could *ever* impart the type of principles of rational ontological inference that Stich seeks.

---

[1] The most prominent example being: *From Folk Psychology to Cognitive Science: The Case Against Belief* (1983).

[2] For an engaging (albeit somewhat speculative) discussion of the conditions that led Darwin and Huxley to make the rational decision in favour of common descent, and conversely, that which led Fitzroy, Owen, and Wilberforce to make the irrational choice, see: (Kitcher [1993]: 180-82).

[3] Notably, as McCauley has pointed out, this does not necessitate either deductive relations between the two theories or preserving the truth of the reduced theory's claims. See: McCauley (1986a: 69).

[4] 'Levels' of analysis and the theories therein will be discussed in greater detail in the section to follow.

[5] Yet McCauley admits that while elimination "is surely not a principled consequence of [evolutionary] scientific change, it is certainly a real one" (McCauley [1986a]: 76).

[6] This, according to McCauley, is a direct consequence of the lower-level theory's greater generality and finer-grain.

[7] Strangely, the Churchlands see this point as somehow counting against McCauley's position; however, McCauley has willingly admitted folk psychology's susceptibility to intralevel elimination (see Churchland [1996]: 224-5 and McCauley [1986a]: 80).

[8] Others, such as Dennett (1987), have also questioned the integrity of eliminativism by citing the successes, owed in a large part to the use of folk psychological concepts, of cognitive psychology. However, Dennett has no interest in defending the ontological status of folk psychology, and hence, the intralevel elimination of folk psychology does not affect his critique.

[9] This reply comes from personal discussions with McCauley.

[10] The criteria, recall, are outlined at the start of section 2.3. For a more comprehensive discussion, see: McCauley (1993): 2-5.

~ Chapter 3 ~

Reconciling the Normative and the Naturalistic

It is worthwhile to begin with a more detailed account of the rationale behind normative naturalism before examining why Stich is doubtful that such an approach can resolve the controversy eliminativism inspires. Proceeding in this fashion seems to demand at least a brief survey of the basal tension that exists between many traditional epistemologists and their radically naturalistic counterparts. This is so because both contingents, in light of their seemingly irreconcilable perspectives, generally deem any epistemological proposal bearing the badge of *both* the normative (as conventionally construed) and the naturalistic as misconceived from the start. For traditionalists are internalists about norms; viz., they view it as necessary that the believer have access, in some sense or another, to the justificational support for their beliefs. In other words, the individual must be able to access, or bring to awareness, the justification underlying their beliefs. Traditionalists tend to explicate such conceptions of epistemic justification in evaluative terms, which is why they take the project of naturalising epistemology to be in error. Radical naturalisers, on the other hand, are externalists about norms; viz., norms are thought to be

realized in terms of certain physical processes, which is why they are said to require examination from a thoroughly naturalistic standpoint. Accordingly, the norms of which traditionalists speak are, strictly speaking, non-existent, which is part of the reason why some radical naturalisers (such as the early Quine) insist that we abandon the normative mission of traditional epistemology and place scientific psychological study in its place.

Hilary Kornblith (1994) provides a straightforward approach for understanding the foregoing division. Consider, Kornblith asks, the following questions: (1) How ought we to arrive at our beliefs?; (2) How do we arrive at our beliefs?; and (3) Are the processes by which we do arrive at our beliefs the ones by which we ought to arrive at our beliefs? (ibid. 1). The traditional epistemological view advises philosophers to occupy themselves with answering (1) while psychologists contemporaneously engage in research aimed at answering (2). Each is urged to pursue their respective questions independently, and once both answers are found, they are compared with one another, thus producing an answer for (3). A central presupposition of this view is that the answer to (1) need not have anything to do with the answer to (2); viz., how we *ought* to arrive at our beliefs may have nothing to do with how we *actually* arrive at our beliefs (and vice versa) (ibid. 2). For since the normative questions philosophers ask are taken to be completely independent of the purely descriptive questions psychologists ask, the subsequent answers, it is claimed, are apt to have nothing to do with each other (it is here that we arrive at the familiar traditionalist claim "you can't derive an 'ought' from an 'is'").

The history of the traditional epistemological view is largely the history of the foundationalist program (Kornblith 1994: 4). Radical foundationalists, such as Descartes, strive to disclose a type of knowledge (i.e., a class of beliefs) about which it is impossible to be wrong (i.e., beliefs that are certain in some sense — e.g., they are indubitable, incorrigible, infallible, etc.). Typically, such beliefs are taken to be non-inferential and can be either *a priori* (if one is a rationalist, such as Descartes) or *a posteriori* (if one is an empiricist, such as Hume). Contemporary foundationalists typically endorse *modest* foundationalism — i.e., the view that foundational beliefs need not be certain, but must, in some manner, be non-inferentially justified. Of course, opinion as to what non-inferential justification is vary considerably; the three prominent accounts of non-inferential justification are (a) self-justification, whereby it is said that foundational beliefs can justify themselves (b) justification via non-belief, non-propositional experiences, such as sensory experience, and (c) justification by a non-belief reliable origin of a belief, such as memory, perceptual, and introspective processes. Foundationalists maintain that such beliefs are used to justify the rest of our inferential beliefs; hence, "in addition to identifying those beliefs that would serve as the foundation of knowledge, foundationalists [attempt] to show how foundational beliefs provide us with good reason [i.e., standards of justification] for adopting the remainder of our beliefs" (ibid. 4). It is the two-tiered internalist conception of knowledge and justification as foundational that is at the core of many disagreements between traditionalists and naturalists.

Naturalistic approaches to epistemology can vary considerably from one another, but all insist that question (1) cannot be answered independently of question (2). In other

words, questions about how we actually arrive at our beliefs *do* bear on questions about how we ought to arrive at them. *How* the purely descriptive projects of psychology are thought to affect inquiry into the normative assessment of belief acquisition is what differentiates naturalists from one another. Radical or exhaustive naturalisers, such as Quine, argue for, to borrow Kornblith's terminology, the 'strong replacement thesis'. According to this thesis, to understand how we ought to acquire our beliefs is to understand how we actually acquire them, viz., the norms of belief acquisition *are natural processes*. Evolution is commonly used in support of this thesis: since believing truths enhances our survival and reproductive value, natural selection must guarantee that our cognitive mechanisms have an innate predisposition for believing truths (Kornblith 1994: 5; see also: Quine [1969]). Knowledge is thus depicted not in terms of meeting some abstract criteria, but as a necessary by-product of evolution. Traditional epistemological questions should therefore be *replaced* by psychological questions. Given this approach epistemology "or something like it" would simply "[fall] into place as a chapter of psychology and hence of natural science" (Quine 1969: 101).

Needless to say, traditionalists are at odds with Quine's proposal. For among other things, it urges epistemologists to abandon the traditional view (which, as we noted, typically includes the foundationalist program[1]) and take up the descriptive practices of psychology. Those who do otherwise, i.e., proceed along the lines of epistemology's historical standing as an autonomous enterprise, would seem to risk the chance of having their work pigeon-holed as illegitimate (that is, if they could still find employment!).

There is hope, however, for those who wish to preserve the autonomy of epistemology — moderate naturalism. Moderate forms of a naturalised epistemology propound a weaker version of the replacement thesis in that they acknowledge some epistemological questions as legitimately distinct in content from questions of descriptive psychology (Kornblith 1993: 7). Consequently, epistemology is not just about 'the facts'. Again, moderate forms vary according to *which* questions they regard as distinctly epistemological. However, broadly speaking, moderate naturalists accept that the norms by which we assess the truth and justification of our beliefs cannot be reduced to natural processes (Brook & Stainton: forthcoming). Hence, questions that pertain to norms fall, as one would expect, within the domain of epistemology.

Crucial to the sort of moderate naturalism that will be presented in the pages to follow, however, is the expectation that psychological research concerned with the mechanisms by which we acquire and relate our beliefs *can and is very likely to play a significant role in the shaping of at least some of our norms*. Put more generally, while both epistemic and moral norms are forged by a number of factors (many of which are clearly non-scientific), their contours can also be influenced — sometimes markedly — by developments in the physical sciences. What follows from this is that if epistemologists take the study of our norms as one of their central tasks, they *must* become more responsive to the developments within the relevant (i.e., cognitive) sciences. This entails, among other things, that epistemologists attend to questions that are not normally a part of their repertoire; questions such as 'What mechanisms will yield beliefs that satisfy our norms of truth and justification?', 'When mechanisms do not yield justified true beliefs,

under what circumstances can the agent be held responsible?', 'What circumstances will lead us astray no matter what we do?', and so on.[2]

At this point it should be evident that each of the positions mentioned above interprets the relationship between epistemology and psychology, and the relationship between values and facts, differently. The traditional view regards epistemology and psychology as separate, autonomous enterprises that go about their respective activities independently of one another. Here values and facts are taken to have nothing to do with each other. Radical naturalisers regard epistemology as a subcategory of, and hence continuous with, the encompassing field of psychology (thus the sort of epistemology traditionalists speak of is 'eliminated'). Under this view, values *just are* facts. Moderate naturalisers, in contrast, view epistemology and psychology as two separate fields that employ different methodologies in an effort to answer equally legitimate, but different, questions. Yet because of the intimate relation between their respective subject matter (i.e., how we ought to acquire beliefs versus how we actually acquire them), moderate naturalisers insist that epistemologists and psychologists be ever mindful of each other's work. The necessity of this approach becomes even more apparent once we recognize that it is not always easy to distinguish a purely normative issue from a purely descriptive one. Indeed, questions of fact and value often so thoroughly interpenetrate and reconfigure one another that rendering a distinction is at times extremely difficult, if not impossible. Nevertheless, if, in our discussions of either, we ignore the impact of developments in one front on developments in the other, we will most assuredly confine ourselves to an impoverished and misleading account of the matters at hand.

In short, even though it is affirmed that neither epistemology nor psychology could ever co-opt the other, moderate naturalism still calls for a heightened level of interaction between the two. For it is expected that both fields stand to gain a great deal by carefully monitoring the progress of the other. (This type of relationship, one will recall, is generally what co-evolution$_p$ describes.)

It is with the aim of further defining (and at times defending) this sort of moderate naturalism that the chapter is organized along the lines of two general themes. The first of these (3.1—3.3) bears upon the difficulties associated with adopting either a radically naturalist position or an unwaveringly traditionalist position when pursuing epistemology's core projects (e.g., the search for norms, or standards, of rational acceptability with which to evaluate our claims, including that made by the eliminativist). Throughout this discussion we will see that moderate naturalism is in keeping with much of the normative naturalism developed in the previous chapter. Indeed, it is my intention to show that the advantages of combining various aspects of the two outweigh those afforded by any alternative. This will become particularly evident at points wherein a reply is made to those who contest, either directly or indirectly, the plausibility of the *very idea* of maintaining this sort of naturalism. For in showing that such allegations are misconceived and thereby misplaced, the desirability of this type of moderate normative naturalist position will be unmistakable.

With Stich's comments the allegations against normative naturalism continue, and in doing so, serve to motivate the development of the second theme. The difference here is that unlike traditionalists and many radical naturalisers, Stich does not reject the idea of

normative naturalism or deny that normative naturalist enterprises can yield lucrative results. Indeed, Stich maintains that many sophisticated and powerful strategies of reasoning and inquiry have been produced by such undertakings. However, he is — for a number of reasons — extremely sceptical that a normative naturalist project could be of any assistance in settling the particular controversy regarding the eliminativists' thesis. The grounds for this scepticism will be outlined (section 3.4) and then evaluated (section 3.5) alongside the (allegedly) inextricable social constructivism that is said to result. The final passages will contest Stich's assertion that normative naturalism is unlikely to furnish principles that would inform us of how we ought to regard the eliminativists' thesis. And by doing this, we will see that the current ontological indeterminacy pertaining to the propositional attitudes is not necessarily destined to resolve itself in the manner that Stich's social constructivism anticipates.

3.1: The Normative and the Naturalistic: A Tale of Two Cities

The Churchlands, in their campaign to bring the precision of the sciences to epistemology, provide a clear illustration why many traditional epistemologists are led to repudiate naturalism. For the manner in which the Churchlands ultimately expect to deliver precision to epistemology is simply to treat it as the part of scientific research concerned with human cognition. Thus the Churchlands advice to those who wish to bring the determinateness of scientific judgment to epistemological claims is, in short, to pursue the relevant sciences (which, of course, is also what Quine recommends) (McCauley 1988: 143). The Churchlands are far from being the sole proponents of this view; as McCauley

has commented, "enthusiasm for this project has increased in the past decade in part because of the startling progress in the...[cognitive] sciences. Progress in psychology, linguistics, cognitive anthropology, artificial intelligence (especially of the connectionist variety), and the various neurosciences has rendered increasing expanses of human cognitive activity accessible to systematic empirical investigation" (McCauley 1988: 143).

As we previously observed, the Churchlands — especially Paul — have repeatedly emphasized the particular importance of neuroscience to matters (such as those pertaining to rationality and the acquisition of knowledge) that have been historically perceived as the exclusive subjects of traditional (i.e., *normative*) epistemology.[3] The erosion of conventional epistemological boundaries is, according to both Churchlands, inevitable, since

> ...a serious advance in our appreciation of cognitive virtue would...seem to *require* that we go beyond [folk psychology], that we transcend the poverty of [folk psychology's] conception of rationality by transcending its propositional kinematics entirely, by developing a deeper and more general kinematics of cognitive activity, and by distinguishing within this new framework which of the kinematically possible modes of activity are to be valued and encouraged (as more efficient, reliable, productive, or whatever) [Churchland 1981: 84].

Thus eliminativism does not, Churchland purports, imply the end of our normative concerns; rather, it suggests that they will have to be "*reconstituted* at a more revealing level of understanding, the level that a matured neuroscience will provide" (Churchland 1981: 84; emphasis added). If Churchland is correct in claiming that

neuroscience can, in the sense indicated above, serve as a source for the development of norms, then it is clear that, given such a radical reinterpretation of norms, epistemologists would have no choice but to naturalise. Indeed, any epistemological speculation regarding cognitive virtue that does otherwise (i.e., that ignores the developments of neuroscience [and the rest of the cognitive sciences, for that matter]) would seem to risk the chance of thoroughgoing irrelevance (McCauley 1988: 143).

Of course all of this is, at bottom, a warning to those (traditional) epistemologists who rely extensively on the framework of folk psychology and who assume that the conceptual habits that evolved prior to systematic psychological theorizing are by and large adequate (recall that in support of this, the Churchlands have pointed out that (1) systematic scientific research has continually encroached on and usurped fields of inquiry that were once the sole possessions of folk theories, and (2) that many recent findings in the cognitive sciences seem to conflict with, and hence undermine, the presumptions made by folk psychology; see pp. 62-3). Yet in accepting the Churchlands' recommendation — viz., to carry out the project of naturalising epistemology by proceeding along the lines of a single (or number of) natural science(s) — one is also accepting the apparent predicament that comes with confronting the inherently *normative* aspect of the discipline. For it would seem — at least in light of the central task of the traditionalist's foundationalist program (i.e., to formulate general *a priori* principles for the *assessment* of our beliefs in an effort to distinguish those *meriting* the title of 'knowledge' from those of a lesser stature) — as though a normative dimension is inextricably tied to all epistemological endeavours. Yet by proceeding solely on the basis of developments

within the relevant natural sciences, it is difficult to imagine *what* could count as an explicitly *normative* epistemology.

While troubling for some, the prospect of jeopardizing the normative dimension of epistemology appears to be of little concern to the Churchlands.[4] For since, as we saw in the previous chapters, the question of whether folk psychology is the *correct* framework with which to grasp mental phenomena remains very much an open one, so too must be the question of whether normative epistemology, in drawing on the categorial level comprehended by folk psychology, is the *correct* framework with which to delineate the parameters of such notions as 'rational belief', 'truth', or 'knowledge' (Churchland 1979: 123; 1981: 82-4). The consequence, in short, is that the elimination of folk psychology's core postulates (i.e., the propositional attitudes) would also, presumably, entail the elimination of the normative features of epistemology that are delineated in terms of those postulates (McCauley 1988: 144). If, as the Churchlands anticipate, such circumstances are realized, what — given the absence of norms — will guide epistemologists in their efforts to appraise our beliefs? The short answer is empirical psychology. The long answer lies within several of the Churchlands sub-personal approaches to 'rationality' and 'knowledge', all of which accord well with recent findings in cognitive neurobiology and connectionist architecture and, of course, make none of the assumptions that inform the class of normative epistemologies they reject (e.g., see: P. M. [1995]; P.S. [1986]). Yet it is still difficult to say whether these unconventional measures of cognitive performance can actually fulfill a 'normative' purpose. Radical naturalisers, such as the Churchlands and Quine, certainly suggest that they can and will (recall how Quine uses

evolution in support of this). However, because scientific epistemology of this sort is in its infancy, no one really knows precisely how — or even *if* — the results of these studies are to assume a normative role.

One of the strongest opponents of such attempts to naturalise epistemology is Hilary Putnam. Putnam insists that if we were to explain all of our norms away we would be left with no remaining standards by which to assess competing explanatory claims — including those concerning the explanation of our norms. In light of this, Putnam states that "the elimination of the normative is attempted mental suicide" (Putnam 1983: 246), since without any normative standards there would be nothing we could be right or wrong about — all our arguments and assertions would be, as McCauley states, "all for naught" (McCauley 1988: 145). While Putnam is wary of aprioristic arguments about the necessary shape of epistemic concepts, McCauley has noted that "the alternative picture he offers portrays epistemology as inescapably normative, nonetheless" (ibid. 145). For Putnam maintains that *explanation is in itself* an epistemic notion, since all explanations presuppose values (Putnam 1983: 290-98). If, as Putnam suggests, explanation contains an ineluctably normative dimension, then one would presume that such mainstay epistemic notions as 'rationality' and 'knowledge' must as well. Accordingly, we should expect that any of the naturalisers' proposed *explanations* of these central concepts would just as surely presuppose further epistemic norms (op. cit. 145).

3.2: The Problem with an Exhaustively Naturalised Epistemology

As we shall see, McCauley is convinced that naturalist attempts to eliminate the normative outright, such as that suggested by Churchland, will face the following dilemma: they will either "prove incomplete or, if complete, then inimical to the progress of science" (McCauley 1988: 145). However, by constructing an alternative (i.e., *normative*-naturalist) proposal, McCauley hopes to demonstrate that neither this conclusion nor Putnam's should discourage advocates of a naturalised epistemology from pursuing their projects. His aim, in his own words, is not to defeat the project of naturalising epistemology, but rather to "help clarify what it does and does not amount to" (ibid. 145).

McCauley begins by examining Putnam's contention that explanation is a fundamentally epistemic notion. McCauley notes that two episodes in the history of science — the demise of geocentrism and that of creationism (and correspondingly, the failure of normative defenses thereof) — brought significant contributions to the evolution of new standards for acceptable scientific work (such as, for example, that, *ceteris paribus*, we should prefer theories that have an array of consequences that are specific enough to submit to extensive empirical tests over their competitors who do not [Kitcher: 1982]). Thus, these episodes had an effect on what would subsequently count as an acceptable *explanation*, suggesting that, among other things, good explanation in science should not appeal to the intervention of the supernatural. However, McCauley notes that it is clear that "this sort of *evolution* of new scientific standards is equivalent neither to the *abolition* nor to the *elimination* of all such [i.e., normative] standards" (McCauley

1988: 146). The failure of attempts to defend geocentrism and creationism on the basis of their centrality to prevailing normative positions *does* suggest that arguments of this sort are incapable of preserving what have proved (during the course of research) to be deficient theories. Yet this does not, as McCauley insists, support the claim that normative matters in general can be eliminated once and for all. Rather, what these cases show is that "*as a result of theoretical progress in science our normative commitments sometimes change* (both in the epistemological and in the moral sphere)" (ibid. 146; emphasis in original).

Simply put, the cases of geocentrism and creationism show that what were once *obviously* matters of great moral significance may, as a result of scientific developments, become less conspicuously so. At the moral level, it could be argued that homosexuality will ultimate prove to be an analogous case, for it would appear as though attitudes towards homosexuality stand a good chance of undergoing a similar sort of transformation as a result of current and future scientific developments.[5] For consider (1) the astonishing array of 'naturally occurring' sexual arrangements and relations that arise among individuals across a variety of species (e.g., the homosexual parings of doves discovered in California); (2) research on the value to the species of a gene for a kind of hypertrophic attraction to males; and (3) preliminary findings about small but consistent differences between certain brain structures of homosexual and heterosexual human males. Such research might, in time, significantly alter prevailing perceptions of the 'naturalness', if not the moral acceptability, of homosexual relations. At the epistemic level, consider the influence of the development of the modern synthetic theory of evolution on our

(in)tolerance for teleological forms of explanation in the natural sciences. This would seem to be an obvious instance wherein an inquiry into *natural processes* has had a derivative affect on the *standards* by which we evaluate certain claims.

However, as mentioned above, none of this demonstrates that matters of normative concern in general are candidates for elimination. For contrary to the implications of radical naturalism, what counts as a normative matter pertaining to science — be it moral or *epistemic* — is *not* always decided exclusively on the basis of considerations *internal* to science (incidentally, McCauley notes that the contrary position "covertly resurrects the logical empiricists' assumption that the standards for a satisfactory explanation can be formally specified once and for all" [McCauley 1988: 146]). A number of factors — scientific and non-scientific — can be involved in the determination of if, and to what extent, a given issue is normative. Some ardent creationists, for example, continue to regard the conflict between their convictions and the theory of evolution as one of profound moral significance that has yet to be settled — a fact that is entirely consistent with both a moderate form of normative naturalism and McCauley's own understanding of how the contours of normative issues are defined:

> What counts as normative within science depends, at least in part, on the relative prominence of the dissenting views (both within the scientific community and within the culture at large) that are critical of the dominant methods and the theoretical developments of some science. It also depends upon the kind of challenge that dissenters present to scientific standards and achievements. (Hence, it is much easier for us to understand how the origin of humanity still constitutes an issue of moral significance than it is for us to understand how the position of the earth does.) [In short], *the polemical*

> *context carries the major share of the responsibility for highlighting the normative issues* [be they moral or epistemic] [McCauley 1988: 146; emphasis added].

These comments also accord well with Paul Roth's contention that "we cannot clearly mark off what is normative from what is not" (Roth 1987: 32; quoted in McCauley 1988: 146). Roth arrives at this conclusion through the recognition that, as with the distinction between the conceptual and the empirical, the distinction between the descriptive (or explanatory) and the normative is "overwhelmingly mitigated by the sort of holism Quine proffers in semantics" (op. cit. 43) — a position to which, interestingly, all parties in the foregoing discussion (i.e., the Churchlands, Putnam, and McCauley) generally subscribe. If we accept this portrayal, then we must grant that holistic considerations constrain claims about justification no less than other claims within the web of belief (op. cit. 146). Accordingly, Roth argues that, depending on the conceptual and theoretical frameworks at hand, different issues will seem more plainly normative than others. One consequence of this view, then, is that what is interpreted as a normative issue, even in epistemology, is partly a function of the polemical context in which it resides (McCauley 1988: 146).

McCauley notes that this is not to argue that "any particular assumption is intrinsically normative come-what-may", but it does imply "*that all positions involve presuppositions (e.g., about methods) that function normatively* (in the same way that all positions involve conceptual presuppositions that function *within those positions* like analytic truths" (ibid. 147; emphasis in original). The question of *whether* the normative

dimensions of such assumptions are manifest usually depends upon the ferocity of relevant disputes. Typically, *which* normative dimensions are manifest depends upon the problems and the purposes of the disputants (McCauley 1988: 147).

What follows from this is that every new explanation of norms will itself presuppose further norms that will in turn require explanation. Thus the proposed encapsulation of the normative within the framework of our explanations — which the Churchlands foresee and which Putnam deems suicidal — will "surely remain forever incomplete" (ibid. 147). This is the first horn of the dilemma that confronts those who wish to exhaustively naturalise the normative.

Given the foregoing depiction of inquiry (i.e., wherein each systematic explanation of a normative consideration contains further normative assumptions), the relation of the normative and the explanatory emerges as one of a "perpetual and productive interaction where the borders between the two domains could shift (often only imperceptibly, perhaps) with each and every advance" (ibid. 147). The vital point, though, is that *the persistence of both the normative and explanatory dimensions in science is a prerequisite for its progress*. Hence, the elimination of the normative is antagonistic to the advance of science — which is the second horn of aforementioned dilemma. McCauley describes this predicament as follows:

> If the confluence of opinion about our explanatory successes ever completely overshadowed the normative assumptions that inform scientific work, it would signal the absence of disputes. The absence of disputes would indicate the cessation of theoretical competition, and the cessation of theoretical competition would put an end to any substantial scientific change. (It is

> indifferent...whether or not we regard that change as progressive.) When science ceases to change, science ceases. Only on the most naive forms of realism could science ever grind to such a halt — presumably, because it had finally gotten all of its descriptions right! [McCauley 1988: 147].

In light of this outcome one can see why McCauley is inclined to deem naturalist positions that anticipate the wholesale elimination of the normative as traditionally construed, such as that found in Paul Churchland's early work (wherein the disappearance of the normative is seen as a consequence of eliminating folk psychology in favour of some future neuroscientific account of cognitive functioning), as 'inimical to the progress of science' (a charge that would appear to be particularly distressing for both Churchlands considering their fondness for, and faith in, science in general). Furthermore, with respect to that which was said in the previous chapter, we can identify a further problem with the Churchlands' position *vis-à-vis* normative epistemology: even *if* the categorial framework of folk psychology, upon which normative epistemology is said to rest, were to succumb to elimination, it would *not* come at the hands of some future neuroscience, but rather those of some *intralevel* competitor. Accordingly, the Churchlands may be encumbered with the same difficulty as that which loomed at the end of section 2.3; namely, that if folk psychology is eventually displaced, it will come via some higher level counterpart (e.g., cognitive psychology), which would, in all likelihood, implement some of folk psychology's core (i.e., intentional) concepts. (For recall that, at least currently, the most promising intralevel successors to folk psychology all utilize intentional concepts to some extent [see p. 24].) Granted, this is not to say that

neuroscience could not play a role in *shaping* folk psychology's intralevel successor. Indeed, as McCauley states, "in all probability the neurosciences will share some responsibility in the crime [i.e., the elimination of folk psychology]"; yet nevertheless, "it is much more likely to be some descendant of contemporary cognitive psychology that will do the actual dirty work" (McCauley 1988: 149). Thus the point remains: as long as intentional concepts are preserved to some degree by folk psychology's intralevel successor, then the possibility of a *normative* epistemology couched in such concepts remains open.

3.3: The Normative Naturalist Alternative

Fortunately, neither the Churchlands' or Putnam's arguments ought to deter philosophers from actively investigating the relevance of the cognitive sciences to epistemology. For the considerations they raise do not, according to McCauley, undermine Quine's insight, viz., that once we are bereft of appeals to 'First Philosophy', it is the standards of those pursuits that contribute most obviously and most successfully to human flourishing in general, and to our knowledge about the world in particular, that offer the clearest and most defensible standards of rational justification (McCauley 1988: 150; see also: Quine [1969]). And it is with respect to such pursuits that McCauley holds the non-controversial assumption of science as deserving of a prominent — if not the *most* prominent — position, particularly when one considers that standards that are "*simultaneously external to, superior to, and independent of*" those of science simply do not exist" (McCauley 1988: 150; emphasis in original).

This last comment of McCauley's, motivated largely by the recommendations of Quine, provides a concise representation of what, above all, continues to divide many traditionalists and naturalists (be they radical or moderately normative). This division stems from the naturalists' insistence that the conventional epistemological project — both in terms of its content and approach — is misguided. In terms of its content, it is said to be misguided because it searches for nonexistent standards of rationality from which to judge the epistemic status of our (specifically scientific) beliefs (i.e., the standards to which the preceding quote of McCauley's refers). Naturalists often reinforce this conviction by emphasizing the fact that "three hundred years of epistemological inquiry has failed to reveal any set of principles that could claim the allegiance of a majority of even philosophers" (Giere 1989: 378). This misguided goal, in turn, is said to trigger the traditionalists' misguided approach: i.e., attempting to determine the aforementioned standards *a priori*. In light of all this confusion, most post-Quinian naturalists, including McCauley, are led to advise traditionalists to abandon their task of establishing a 'First Philosophy' (i.e., a class of "unconditionally accepted, non-scientific, truths upon which to justify science" [Gibson 1987: 62]). Such abandonment is crucial, the naturalist is apt to continue, since it is only *after* this is done that epistemology can proceed, and in doing so, be in a position to freely consult the relevant sciences in answering its central questions.

Along with the expectation of a future epistemology drawing on the resources of science comes a further claim that is commonplace among naturalists, viz., that the norms used to judge scientific beliefs *must* come (radical naturalisers such as the Churchlands)[6],

or are *very likely — though still not necessarily* — to come (normative naturalisers such as McCauley), from within science itself. Note the similarity here between McCauley's position and the specific sort of moderate naturalism that was outlined at the beginning of the chapter. Like McCauley's own position, the said variety of moderate naturalism goes somewhat further than that which is characteristic of the approach in general. This is so because moderate naturalists typically eschew the question of *where* our norms are to come from and instead focus on the cognitive mechanisms involved in their satisfaction (or lack thereof). However, while the variety of moderate naturalism I envision is equally concerned with mechanisms of belief acquisition, it also carries the additional conviction that the findings of such inquiries and/or the methodologies used in their attainment *might* serve as a source of norms in and of themselves. For if one accepts that scientific developments can influence the shape of our norms, one must also acknowledge at least the *possibility* that such developments could exert *so much influence* that the norms could be said to have come straight from science.

Nor should such a possibility be construed as far-fetched; indeed, we have good reason for thinking otherwise. Consider the basis of McCauley's preference for science — i.e., its (relative) determinateness — both in its findings and its methods. If, at McCauley's recommendation, we regard the findings and methods of our knowledge-seeking activities (when viewed in terms of their respective determinateness) as falling along a continuum, we can see that as one moves further away from the physical sciences, claims and methods become progressively less and less precise. Although not the only sort of justification, scientific justification does appear to have a distinct advantage over

other methods in that it yields comparatively determinate results in accordance with comparatively determinate methods.

Of course this is not meant to suggest that science could ever definitively justify itself (and thereby monopolize some portion of the truth). Yet it still seems as though normative positions — even that of Putnam's — must look to science at some point. For it is difficult to see how one would cash-out, for example, Putnam's depiction of truth as "an idealization of rational acceptability" (Putnam 1983: 200) without devoting a substantial amount of attention to *the standards of science in general* and *the findings of cognitive science in particular*. While granting that not all truths are the truths of science, the truths of science do seem to stand much nearer to the core of our concept of truth than do our less determinate, veracious claims (McCauley 1986b). McCauley realizes that assertions such as this carry an air of question begging since "whatever force [they possess] seems to be psychological only" (McCauley 1988: 150). However, McCauley avows that after all versions of foundationalism are found to be wanting, the accounts of rationality "whose begging of questions is least offensive" will be precisely those which attend most conscientiously to our best accounts of the relevant facts — and the relevant facts here are largely psychological (ibid. 150). Hence, developments within the cognitive sciences that deal with the acquisition, processing, and utilization of information would seem, at least *prima facie*, to provide an obvious means by which to augment the body of facts that inform our current epistemological projects (such as those that seek standards of rational acceptability). Of course, the question of whether a particular development would amount to just one of many factors (including those that might be non-scientific)

that often contribute to the forming of our norms, or whether it would take on a principal role by furnishing the norm itself, depends on the content of the development and the polemical context in which it figures.

Scientific contributions notwithstanding, the normative, in one sense, is ineliminable, since the persistence of the normative, as we noted earlier, will invariably preclude the possibility of ever *exhaustively* naturalising epistemology (for each of the radical naturalisers' explanations, recall, will presuppose further norms). However neither this nor, for that matter, Putnam's argument, demonstrates that our *current* epistemic norms or the categorial framework (based predominantly on folk psychology) in which they are posed must both persist. It is, rather, that normative discourse of *some sort* or another will endure. And this (i.e., the persistence of normative discussions) does *not* establish any principled autonomy for epistemology as traditionally construed (indeed, talk of epistemological autonomy seems even more unbefitting once one acknowledges that "just as surely as all debates about the facts presuppose norms, all normative discussions make assumptions about the facts" (McCauley 1988: 151).

In summary, then, McCauley argues that the normative will continue (within any account of scientific methodology) on the grounds that a complete elimination will render any exhaustively (i.e., radically) naturalised successor forever incomplete or inimical to the progress of science. In other words, contrary to what Churchland (1979) suggests, epistemology is *not* 'just about the facts' (McCauley: personal comm.). Yet crucially, the persistence of the normative is understood to have no bearing whatsoever on "the urgent need of contemporary epistemologists to carefully attend, henceforth, to developments in

the cognitive sciences and to take as their primary occupation the task of making sense of those findings" (McCauley 1988: 151). For from the fact that epistemology will never be *completely* co-extensive with science it follows neither "that it cannot nor should not become *much more* responsive [to the relevant sciences]" (ibid. 151; emphasis in original). Thus he argues not that we *must* naturalise, but rather that we *can* naturalise without forsaking normative commitments — and most importantly, have very compelling reasons for doing so.

Some traditionalists, such as Ellen Klein (1992), have argued that the 'middle-ground' alternative presented in McCauley's version of normative naturalism is unworkable. Among other misgivings, Klein alleges that McCauley's proposal implies that — given the difficulties associated with the traditional epistemological endeavour — naturalism is the *only* option. The problem, as Klein states, is that "McCauley cannot consistently maintain the belief that the choice between traditionalism and naturalism are the only choices open to epistemologists"; he must also acknowledge the possibility, for instance, that one may wish to affirm epistemological scepticism (Klein 1992: 292). In other words, McCauley cannot claim — as many stanch naturalists often do — that epistemologists have no choice but to naturalise. Klein admits, however, that this point does not pose a serious threat to McCauley's proposal, since he is not constrained to show that one *must* naturalise, but rather only that one must abandon traditionalism (because it is not until philosophers drop the quest for a 'First Philosophy' that epistemology can proceed along the lines McCauley imagines it should — namely, in direct consultation with the cognitive sciences]). Yet even on this count, Klein finds

McCauley unconvincing. Klein attributes two arguments to McCauley's purported efforts to urge the abandonment of traditionalism, the first of which is said to rely primarily on the historical failings of the enterprise. Klein's reply to this argument (and indeed to all arguments of this sort) is simply to note that "the empirical fact of the failure of all past epistemologies, even if true, holds little weight concerning prospects for the future" (Klein 1992: 293).

While this retort might coerce some naturalists into reformulating or recanting arguments that invoke the past failures of epistemology (although I tend to think it should not), it has no bearing on what McCauley proposes. This is so because Klein's comments rest on a misunderstanding of McCauley's position: McCauley *is not* required to show (1) that epistemologists must abandon traditionalism, or (correspondingly) (2) that naturalism is the only choice open to epistemologists. What McCauley urges is not nearly as strong as what Klein suggests. McCauley's recommendation that epistemology become 'more responsive' to the findings of the cognitive sciences carries no categorical force. As McCauley states, this recommendation is merely presented "in light of what I take to be a reasonable assessment of the relative promise of the various directions of research open to epistemologists"; and it is with respect to this that he holds that "the balance of the arguments over the past few decades indicates that the traditional project possesses some profound limitations and that during those same decades the findings of the relevant cognitive sciences seem to raise important questions about many of the central categorial assumptions of the traditional approach in epistemology" (McCauley 1992: 301).[7] Consequently, McCauley does not deny the possibility of pursuing the

traditional or even sceptical epistemological projects — epistemologists are free to do both. However, *he is* pessimistic that such avenues of inquiry will produce any deeper understanding of our intellectual and practical successes, or, for that matter, of the wide range of knowledge that we have at our command, including much of our tacit knowledge and our intuition (McCauley 1992: 302). While this may, practically speaking, amount to the claim that epistemologists who wish to pursue fruitful research paths must abandon traditionalism, it again carries no categorical force.

Klein credits McCauley with launching a second argument for the abandonment of traditionalism, one that impugns the epistemologist's search for a 'First Philosophy'. Although we know that, in light of what was previously said, McCauley has no such aim, it is worthwhile to survey Klein's confusion — for in doing so we can further clarify McCauley's position. Klein interprets McCauley's assertion that there are no standards that are 'simultaneously external to, superior to, and independent of' those of science as, in essence, a restatement of Quine's (1981) belief that there is no need for a 'super-scientific tribunal'. To this, Klein insists that "there is a need for a super-scientific...tribunal and the need is (among other places) precisely where Quine denies that there is any — determining the adequacy (justificatory status) of specific scientific claims, arguments and methodologies" (Klein 1992: 294). Again, Klein is misconstruing McCauley's position if she thinks that he categorically disavows extra-scientific deliberation about scientific endeavours. McCauley does, however, regard the metaphor of a super-scientific 'tribunal' as entirely inappropriate. As McCauley states, "neither epistemologists nor anyone else can sit as a 'super-scientific tribunal' possessing any

confidence either that their judgments can justifiably supersede the judgments of science or that their judgments boast epistemic virtues superior to those of science" (McCauley 1992: 300). For although McCauley agrees with, among others, Klein and Putnam that other values and considerations besides those of science shape our epistemic investigations and ideals, he agrees with Quine that none of these investigations or ideals lead to standards that either "supersede or ground or stand in authoritative judgment over the judgments of science" (ibid. 301). In short, even though McCauley views the search for a 'First Philosophy' as fruitless, and the related notion of a super-scientific tribunal as faulty, he is not required to show that epistemologists must give up the hope of either. It is, rather, that he simply regards epistemological projects that are uninformed by the findings of the cognitive sciences as needlessly impoverished.

Klein's comments are representative of the traditionalist's aversion to the prospect of utilizing a normative naturalist approach for pursuing epistemology's central projects. Yet we have seen that her allegations rest on a misunderstanding of what normative naturalism envisions (and hence, what it requires of epistemologists). Contrary to the arguments of Klein and many other traditionalists, it appears that epistemologists *can* make use of the developments in the relevant sciences without relinquishing their normative commitments. Moreover, it would seem as though they have compelling reasons for doing so, since such a course looks to be, at least prima facie, aptly suited for the tasks at hand (such as devising standards or principles of rational acceptability). Stich assents to this point, and indeed maintains that scientists have already employed such strategies with great success. Thus, Stich does not regard a normative naturalist approach

as problematic in and of itself. However, as we shall see in the section to follow, Stich *is* sceptical that such a strategy will produce principles of rational ontological inference that are strong enough and detailed enough to dictate what we ought to conclude about the entities invoked by a theory once we come to believe that the theory is seriously mistaken.

3.4: Pragmatic Obstacles to Normative Naturalist Solutions

Stich expresses two separate reasons for doubting the possibility that a normative naturalist approach— such as McCauley's — might serve to resolve the controversy surrounding the eliminativists' conclusion. The first of these proceeds from the purported normative status held by the principles normative naturalists sanction (for recall Stich's hope that normative naturalism might produce *rational* principles of ontological inference — principles that would direct us to conclusions we *ought* to accept). In contemplating how such principles are apt to acquire their normative force, Stich makes two significant observations (Stich 1996: 66).

First, it is a mistake to ask how *the* normative naturalist approach produces rational principles of inference, as there is no *single* normative naturalist approach, but rather several strategies that can be reasonably classified as such (McCauley's being but one of these). Given this, Stich concludes that it is quite probable that normative naturalists with dissimilar loyalties will disagree on whether certain principles possess normative status. And until a single normative naturalist strategy can prove unequivocally correct — a daunting task, to say the least — one will be unsure as to

which set of principles ought to be followed. Second, sometimes efforts to substantiate a particular normative naturalist enterprise can carry relativistic implications. For the principles endorsed therein are claimed to be *hypothetical imperatives*; that is, "they are principles we ought to follow *if* we have certain goals" (Stich 1996: 66; emphasis in original). The problem, of course, is that "if our goals are different, the normative-naturalist procedure may well sanction different principles" (ibid. 66). In other words, if the participants of a particular controversy do not share the same goals (which may be the case in the debate over eliminativism), normative naturalism might prescribe different conclusions to different people. Taken together, Stich contends that these observations justify at least some circumspection about normative naturalist projects.

The second reason for Stich's scepticism centres on what we can expect normative naturalism to disclose about patterns of ontological inference, and further, how this disclosure might prove ineffectual when assessing the eliminativists' argument: "even if we grant that the output of the normative-naturalist procedure will be principles that rational people ought to follow, it is important to realize that there is no *guarantee* that this approach will succeed in resolving our problem about what to conclude from the eliminativists' premises" (ibid.; emphasis in original). For, as Stich continues, it is entirely possible that there *just are no rational principles of ontological inference* to be found — or at least none that are strong enough and comprehensive enough to determine what one ought to conclude if the eliminativists' premises are in fact true (ibid. 66-7). Indeed, a thorough study of past cases in which ontological inferences were made might reveal that, in some or many cases, there is simply no correlation between the features of

the theories involved and the conclusion that was ultimately accepted. Yet decisions were still made. If normative principles remain silent when applied to such cases, then, Stich concludes, additional factors *must* have been involved in the deliberative process. And Stich maintains that it will be precisely these sorts of factors that will furnish an explanation of how ontological uncertainties are settled when appeals to normative principles falter. Stich goes on to present an admittedly tentative outline of four types of cases in which these extraneous factors are most likely to affect ontological decisions — cases that he feels a careful investigation of actual examples will corroborate (Stich 1996: 67-70).

(i) 'Don't cares'. There are some cases, Stich argues, in which "it just doesn't much matter to anyone whether the theorists conclude that some of the entities invoked in a mistaken theory really do exist, though their properties are quite different from those attributed to them by the old theory, or whether they conclude there are no such things — that some of the theoretical terms in the old theory didn't refer to anything — and that the phenomena that need explaining are best explained by invoking different theoretical entities" (Stich 1996: 67). In other words, the ontological decision is largely arbitrary. Stich believes that this is what might have happened with electrons, which were originally conceived not as particles but as units of quantity of electric charge, and is presently happening with genes, as modern molecular genetics discerns nothing that maintains all of the important properties originally attributed to them.

(ii) 'Implicit previous agreements'. In other cases, there may be an implicit agreement within the scientific community about an essential property or set of properties ascribed to a posited entity, and that if it turns out that nothing has these properties, then everyone will agree that the entity does not exist. Stich suggests that the historical conception of 'the aether' as a uniform and stationary framework for inertial systems comes close to illustrating such circumstances, for after Michelson and Morley proved that such a rendering was untenable, most scientists agreed there was no aether.

(iii) 'Social and political factors internal to the relevant science'. It is Stich's suspicion that in certain situations the personalities of scientists and the micropolitics of their respective communities exert some influence on ontological determinations. At times, for example, the prospect of receiving a grant or promotion may improve by announcing the discovery of a new entity or denying the existence of an entity that was formerly thought to exist. Either course may also serve to facilitate a scientist's personal desire to break from tradition and be perceived as a radical innovator. However, at other times, it may be more politically sagacious to conclude that, while a particular entity exists, one's colleagues were heretofore mistaken about some of the properties attributed to it — an avenue that is perhaps more attractive to scientists with more conservative dispositions. It is conceivable, for instance,

> ...that had Lavoisier wished to be viewed less as a radical innovator and more as a conservative, he might have retained a venerable old term rather than introducing a new one. And rather than maintaining that there is no such

> thing as phlogiston, he might instead have claimed that Priestly, Stahl, and earlier theorists were mistaken about lots of the properties they attributed to phlogiston. So if Lavoisier had a somewhat different personality, what we now call "oxygen" would be called "phlogiston" instead [Stich 1996: 68].[8]

Stich admits that he is in no position to definitively say *what* Lavoisier's actual temperament was, or if, in turn, it played a part in this particular case. However, the point of this invented scenario, according to Stich, is that it provides an indication of what he suspects is the case — that the personalities of scientists and the micropolitics of their communities can affect the outcome of situations such as these.

(iv) 'Broader social and political factors'. Currently, the social and political consequences of affirming the existence of witches in liberal democracies are, by and large, benign, as the practice of witchcraft (e.g., worshipping nature, etc.) has gained a not inconsiderable degree of legitimacy, while perennial myths have been debunked (e.g., that witches make a pact with the devil, cast evil spells, ride on broomsticks, etc.). However, for those who held a similar view of witches during the sixteenth-century (i.e., that they exist, but made no pact with the devil or caused any harm), it was probably more politically advantageous to deny the existence of witches altogether, in that such a stance might better prevent the torture or death of innocent women. Stich claims that similar social and political considerations are also intermixed in the debate over homosexuality:

> While no one denies that same sex sexual activity occurs, there is some reason to believe that many widely held beliefs about sexual preference, homosexuality, and homosexual people are false. This has led some to

> conclude that there is no such thing as homosexuality or that there are no homosexual people. Others...argue that this is exactly the wrong conclusion to draw. Rather, they insist, homosexuality and homosexuals do indeed exist, though many claims made about them are false. I think it is pretty clear that this dispute is in large measure a political dispute. Both sides agree that there is no kind or condition that has all the features commonly attributed to homosexuality. What is really in dispute is how hatred, prejudice, and discrimination in this area are best confronted and overcome [Stich 1996: 70].

According to Stich, the preceding case — and indeed all of the aforementioned case types — suggest that many of the most interesting and fervent ontological disputes lie outside the bounds of the normative naturalists' rational principles of inference. Consequently, if we wish to compile explanations of how cases of this sort were resolved in the past, and how others are likely to unfold in the future, Stich insists that we must look to any or all of the relevant personal, social, and political factors that exist in both the scientific community and the larger society of which it is a part. And it is in this way that Stich envisions a resolution to the current controversy over eliminativism. More precisely, Stich asserts that if the eliminativists' premises turn out to be true, then the current indecision over what ontological conclusion to draw about the entities of folk psychology will most likely resolve itself in the fashion set out in (iii) (Stich 1996: 71).

Interestingly, Stich claims to find support for this supposition from Patricia Churchland — is a leading advocate of the eliminativist thesis:

> Determining when the fit is close enough to claim identities between properties and entities of the old [theory] and those of the new is not a matter for formal criteria, and the decision is influenced by a variety of

> pragmatic and social considerations. The whim of the central investigators, the degree to which confusion will result from retention of the old terms, the desire to preserve or to break with past habits of thought, the related opportunities for publicizing the theory, cadging for grants, and attracting disciples all enter into decisions concerning whether to claim identities and therewith retention or whether to make the more radical claim of displacement [Stich 1996: 71; from P. S. Churchland 1986: 283-84].

Stich also suspects that (iv) is apt to play a role in what is finally concluded about the existence of intentional states; for one would presume that if scientists and scholars do agree on the nonexistence of such states, the impact on broader social and political issues will be momentous (Stich 1996: 71-2). For instance, Stich imagines that the chance of mounting a successful case for the increased funding of various forms of psychotherapy (which often rely on altering a patient's beliefs and desires) would be seriously impeded if prominent scientists and scholars maintained that such states were nonexistent. Correspondingly, those who wish to develop a case for the increased use of psychoactive drugs might find that such circumstances would substantially bolster their efforts. Whatever the case, one would expect, given the preceding explication, an upsurge of impassioned public and political discussions (concerned, no doubt, with issues such as those just mentioned) that could conceivably pressure scientists and scholars into recasting their conclusions in a less provoking manner — or even recanting them altogether.

Given Stich's account of ontological decision-making, it would appear that if folk psychology is judged to be profoundly mistaken, we should expect a rather lengthy period of indeterminacy while the relevant social and political deliberations run their

course. Eventually, this process will terminate in a broad consensus with respect to the ontological standing of the propositional attitudes, and "the conclusion that wins will have much the same status as the conclusions that were reached in the case of stars or atoms or phlogiston" (Stich 1996: 72). Accordingly, if we look upon historical conclusions such as these as being conclusions rational people ought to accept, then *whatever* conclusion is reached about the existential status of propositional attitudes should be perceived the same way. Stich admits that this position bears a strong resemblance to social constructivism, and even more so, to the sort of pragmatism championed by Peirce, Quine and Rorty — which, while thoroughly incongruous with his earlier eliminativism, is a ramification that the current Stich is entirely willing to concede (ibid. 72).

3.5: Conclusions

It is interesting to look at the relationship between Stich's view, wherein prolonged social and political debate play a central role in resolving the ontological indeterminacy surrounding folk psychological entities, and Paul Churchland's more recent portrayal of scientists' engaging in a similar style of discussion:

> Whether folk psychological categories will find some kinematical and dynamical role within the new framework remains a strictly open question. In all of this there is plenty of empirical evidence to mull over, and ample room for reasonable people to disagree. It is an exciting period of theoretical and empirical evaluation. It would be inauthentic not to enjoy it for what it is [P. M. Churchland 1993: 222].

In one sense, there is a striking kinship between these two viewpoints, in that both Stich and Churchland believe that an extended interval of vigorous deliberation will precede any formal decisions regarding the ontological status of folk psychological entities.[9] There is, however, a crucial difference: the purpose of Stich's account is to articulate his own belief as to how the current ontological dispute over the propositional attitudes is *most likely* to be resolved, while the preceding quote of Churchland's is an expression of his belief of how the said conflict *should* be resolved. In other words, whereas Stich contends that non-scientific factors are apt to determine the outcome, Churchland maintains that nothing other than empirical evidence should affect what is ultimately accepted (a belief, incidentally, that proceeds directly from Churchland's version of eliminativism, which is firmly embedded in scientific realism). And it is for this reason that Churchland, and indeed most any naturalist, will likely find Stich's prognosis unsettling — which brings us to the following criticism.

Stich anticipates that some will likely regard his position as anti-rational, nihilistic, and self-defeating. For if social and political factors determine what conclusion we ought to accept regarding the ontological status of intentional states, phlogiston, stars, and other entities, then what would prevent these factors from determining the answers to questions such as 'Does smoking cause cancer?' (Stich 1996: 73). If we are inclined to think that such socio-political factors *do* determine the answers to such questions, then it would seem that 'might makes right' and there can be no objective or rational inquiry at all. Indeed, by effectively corralling the relevant social and political forces, those who

oppose Stich could just as easily mount a campaign urging the rejection of the conclusion he advises us to accept!

Stich recognizes this line of attack, and attributes its motivation to a misunderstanding of the view he propounds (Stich 1996: 73-4). First, Stich reminds his detractors that he is not unconditionally sceptical of normative naturalism. Indeed, Stich esteems the sophisticated and powerful strategies of reasoning and inquiry scientists have devised over the centuries, strategies that have clearly been used with great success. However, Stich interprets the process of acquiring and deploying these strategies as largely a function of interpersonal exchanges between scientists and their apprentices, from one generation to the next. And Stich contends that the knowledge gained by this process, much like knowledge found in most other domains, is largely tacit — a point that, interestingly, Churchland has also defended (ibid. 74; Churchland 1993: 218). Stich suspects that such knowledge is made up of a diverse array of principles, and that many of these principles — when used for certain goals — carry a significant degree of normative clout. However, if it is one of the goals of normative naturalism to sort out and explicitly define this aggregate of principles, then, Stich contends, we have good reason to doubt that this goal will be attained.

Thus Stich claims he is not committed to denying the fact that many of the normative naturalists' principles *are* rich and detailed enough to provide instruction in many situations. It is, however, the *specific* possibility of normative naturalism furnishing principles of rational ontological inference "that are strong enough to dictate what we should conclude about the entities invoked by a theory when we come to believe

that the theory is seriously mistaken" that arouses Stich's scepticism (Stich 1996: 73). Consequently, Stich believes that it is unlikely that normative naturalism could tell us what we ought to conclude about the propositional attitudes if the eliminativists are indeed correct in their contention that folk psychology is in error. Yet in arguing that the principles endorsed by normative naturalism will not yield a determinate answer in *some* cases (such as the case of eliminativism), Stich insists that he is not constrained to the view that normative naturalist principles will *always* fail to supply determinate answers.

Yet even in light of this qualification Stich's position remains disconcerting. For most, the prospect that thoughts distinguished in terms of their propositional content might not exist is worrisome enough itself. To add that this uncertainty is apt to be settled by virtue of scientists 'cadging for grants', seeking promotions, or desiring to either break or keep with tradition makes it all the more so.

However, before we resign ourselves to this outcome, it may be worth the effort to contemplate again what Stich asks of normative naturalism — namely, principles of rational ontological inference or decision-making. As we saw in the previous chapter, it was Stich's supposition that a careful analysis of actual cases in the history of science wherein such decisions were made might provide a source from which to extract these principles. After completing this task, the principles could then be used to tell us what we *ought* to do about the entities posited by a given theory once it was thought that that theory was mistaken — i.e., whether or not we should continue to conceive of its entities as existent. In other words, the principles would detail when it is *rational* to preserve the items in the ontology of a mistaken theory and when such items should be 'eliminated'.

And it is here that we see the significance of this proposal to the current controversy over eliminativism, for *if*, at the eliminativists' insistence, we concede that folk psychology is a seriously mistaken theory, the principles would then tell us whether or not we ought to retain the items that compose folk psychology's ontology — namely, the propositional attitudes. The principles would instruct us as to whether it would be *rational* to accept what eliminativism demands that we accept: that the entities signified by the terms of our folk psychological framework do not exist.

As we now know, Stich ultimately concludes that the prospect of successfully carrying out a normative naturalist project such as this (viz., one that would yield the aforementioned principles) is extremely unlikely. Hence, the corresponding hope of determining how we ought to regard the eliminativists' conclusion is also apt to go unsatisfied.

In all of this, though, we must ask: is it reasonable to even *surmise* that a normative naturalist endeavour could produce the sort of principles Stich speaks of? Is it not possible, as Stich himself confesses, *that there simply are no principles of rational ontological inference*? And if this is the case, is it then not senseless to doubt that normative naturalism will produce that which cannot be produced? Granted, there is no question that *if* such principles exist, they would be of tremendous value. For it is difficult to imagine a more efficient method by which to resolve cases of ontological indeterminacy. However, it is equally difficult to imagine a class of principles that could settle every conceivable case of ontological indeterminacy in science. And yet this is what Stich desires of normative naturalism — principles that are at all times applicable to

all cases. A cursory look at just the *history* of science evidences an astonishing array of cases in which the ontological standing of vastly dissimilar theoretical entities was at issue; consider, for example, the following posits: alchemic 'spirits', planets, witches, electrons, phlogiston, atoms, caloric, and the aether. Each of these entities belonged to the ontology of a theory that was eventually found to be mistaken. Some of these entities found a place within the ontology of theories that superseded those that were in error while other entities were 'eliminated'. Each of these decisions — i.e., whether or not an entity was judged to exist — was made subsequent to a prolonged period of debate and discussion (usually within the scientific community, although not always).

If, however, the involved parties had the sort of principles Stich desires, none of this would have been necessary. Such lengthy periods of debate and discussion would have been superfluous, for the parties would have simply invoked the principles which would have then promptly detailed how a rational person ought to regard the ontological standing of each of the aforementioned entities. If this sounds too good to be true it probably is — i.e., there probably are no such principles. But even if such principles do exist, does it follow that, since normative naturalism is unlikely to produce them, all normative naturalist endeavours will yield unprofitable results? Clearly no — a point that, as we have seen, Stich readily admits. But from this does it follow that normative naturalism is unable, in some other manner, to resolve the controversy eliminativism inspires? Given the fact that he regards a social constructivist position as the preferable standpoint for handling the alleged ontological indeterminacy of the propositional attitudes, it certainly seems as though Stich thinks so. However, it remains far from clear

that this is the case — i.e., that a normative naturalist endeavour could not inform us of how we ought to regard the eliminativists' thesis in some other fashion. Hence, it remains far from clear *why* we should accept that the debate over the ontological standing of the propositional attitudes will resolve itself in the way that Stich's social constructivism describes.

In keeping with this is the ironical fact that Stich's social constructivism unwittingly lends *support* to the idea that normative naturalist strategies *are* used to settle *precisely those types of ontological controversies that are claimed to lie beyond its jurisdiction*. For reconsider what (ii) — i.e., 'Implicit previous agreements' — amounts to: scientists come to collectively agree that "some property or set of properties *are the essential ones* for some posited entity, and if it turns out that nothing has those properties, then everyone in the community would agree that the entity doesn't exist" (Stich 1996: 68; emphasis added). Admittedly, there is nothing overtly normative about the actual experiments that might indicate a missing property here or an additional one there. However, prior to using the results of experiments as a confirmation or refutation of an entity's existence, *scientists must collectively decide upon which properties are to be taken as essential and which are not*. And is it not the case that any decision as to which properties are to be regarded as 'essential' and which are not contains an ineluctable normative element? How else could scientists differentiate 'essential' properties from 'inessential' ones? Is there, for instance, some objective line, rule of thumb, or formula that is used to separate essential properties from inessential ones? Clearly no. What is determined to be essential and inessential is at least partly a function of what the

decision-maker(s) value. Hence, there must be a normative dimension involved in reaching the sort of implicit agreements Stich speaks of.

Given this, it would seem that Stich's depiction of implicit previous agreements depends on the fact that scientists (tacitly) invoke norms in reaching their consensus about what properties or set of properties are to be taken as essential. Thus, not only do the agreements follow largely as a result of principles bearing a significant degree of normative clout within the relevant scientific community, but that they do so in precisely those sorts of situations where Stich claims they are powerless. Aside from undermining Stich's own position, this point seems to corroborate what was surmised in our previous dealings with radical naturalism — namely, that at some point, normative interests invariably enter into scientific discussion.

Much of the previous chapter was aimed at illustrating how normative naturalism might yield some concrete determinations about the eliminativists' thesis. We saw how the results of one specific normative naturalist project — McCauley's examination of intertheoretic relations and the co-evolution of theories — seemed to directly affect at least part of the eliminativists' thesis (namely, that which contends that a mature *neuroscience* will displace folk psychology). Recall that the upshot of McCauley's model, in drawing on the notion of levels of analysis and temporal considerations in science, was that theory-elimination was an exclusively *intralevel* phenomena. Although this allowed for the possibility that advances in neuroscience could serve to influence the character of some upper-level theory that, in turn, might eventually displace folk psychology, the central point was left undiminished — i.e., that the replacement would

not come at the hands of neuroscience. Hence, it was decided that part of the eliminativists' thesis was, strictly speaking, in error.

It is not absolutely certain whether we could go on to regard this decision as the product of some underlying normative naturalist principle (such as 'no theory at one level of analysis has eliminated a theory at another level of analysis, therefore none should be thought to do so'). This is so because, among other reasons, it has yet to be seen if McCauley's account can convincingly answer all the objections levied against it (e.g., that there are in fact some cases of interlevel theory-elimination in science). But if it turns out that McCauley can refute these criticisms, we would clearly have an example of a normative naturalist principle that (1) would appear to provide determinate results, and (2) is directly applicable to a range of cases, including that which is the focal point of eliminativism — i.e., the anticipated displacement of folk psychology by a matured neuroscience. And even though a principle such as the foregoing would not be of the sort that Stich envisions, it would still serve to highlight a significant point — namely, that a normative naturalist strategy *can* produce principles that are strong enough to yield determinate results with respect to how we ought to regard a core element of the eliminativists' thesis. Thus, if a normative naturalist strategy can do this much, one would assume that we must allow for the possibility that such a strategy might also produce principles that detail how we ought to regard the eliminativists' thesis in its entirety. It is with this very prospect in mind that we draw to a close with some final observations.

As mentioned above, chapter two was primarily concerned with illustrating how a particular normative naturalist project might be of assistance in assessing the eliminativists' thesis. In a similar vein, the chapter at hand highlighted the advantages of adopting a normative naturalist approach when pursuing epistemology's diverse goals. Part of this involved a side-by-side comparison with the main alternatives — traditionalism and radical naturalism. While it may not be completely obvious at first, Stich's social constructivism is quite compatible with the sort of radical naturalism championed by Quine. For even though Stich *begins* with the hope that normative naturalism might yield some determinate principles with which to adjudicate instances of ontological indeterminacy, we must remember that he *ends* with the firm suspicion that no such principles are apt to be found — viz., that a normative naturalist strategy will fail to produce any definitive norms. This conclusion, of course, is *precisely* what a radical naturaliser would expect, since it is maintained that norms — as conventionally construed — simply do not exist. If Stich's social constructivism is indeed indicative of his support of radical naturalism, then, of course, he must deal with all of the same challenges levied against his compatriots in section 3.2.

In closing, then, we might again ask the question that has driven this inquiry from the outset: can normative naturalism tell us how we ought to regard the eliminativists' thesis? If the only way it can do so is by yielding the principles of rational ontological inference of the sort that Stich imagines, then it is unlikely. If, on the other hand, we adopt a moderately naturalist approach to epistemology and thereby remain responsive to developments within the burgeoning field of the cognitive sciences, we could soon find

ourselves with some definitive standards with which to evaluate our various claims — such as those made by the eliminativist.

---

[1] As we shall see in the pages to follow, Quine and many other naturalists also have independent reasons for questioning the idea of a foundational epistemology. Broadly speaking, these arguments either (1) attack the notion that there is a class of beliefs about which we cannot be wrong or (2) cite the perpetual failure of such projects in the history of philosophy.

[2] I borrow these examples from: (Brook & Stainton: forthcoming).

[3] See, for instance, P. M. Churchland: 1979, 1981: 84-90.

[4] This is especially true of Paul Churchland's early work. However, some of his recent efforts attempt to redefine 'normative' issues in terms of neural-computational models, which are (not surprisingly) radically unlike normative issues as traditionally conceived.

[5] I have developed this example, as well as the one that follows, through personal communication with McCauley.

[6] At least this is what Paul Churchland suggests in *Scientific Realism and the Plasticity of Mind* (1979).

[7] Stich (1990) has, incidentally, discussed such limitations and has also surveyed some of the relevant empirical findings.

[8] Stich was apprised of this imaginary scenario by Kim Sterelny, who, in turn, heard it from Elliot Sober.

[9] This appears to be what Churchland has been admitting as of late. And even this seemingly slight confession signals a strong change of heart from what was presented in his earlier (e.g., 1981) work.

REFERENCES

Alston, W. P. (1967). "Wants, actions and causal explanations" in *Intentionality, Minds and Perception*, Castaneda, H. N. (ed). Detroit: Random House.

Brook, A. & Stainton, R. S. (in progress). *Knowledge and Mind.*

Caruthers, P. & Smith, P. K. (eds). (1996). *Theories of theories of mind.* Cambridge: Cambridge University Press.

Cheyne, C. (1993). "Reduction, Elimination, and Firewalking" *Philosophy of Science* Vol. 60, 349-57.

Churchland, P. M. (1981). "Eliminative Materialism and the Propositional Attitudes" *The Journal of Philosophy* Vol. 78, 67-90.

_____. (1993). "Evaluating Our Self Conception" *Mind & Language* Vol. 8, 211-22.

_____. (1989). "Folk Psychology and the Explanation of Human Behavior". *Philosophical Perspectives (3): Philosophy of Mind and Action Theory.* Atascadero: Ridgeview Publishing.

_____. (1988). *Matter and Consciousness* (rev. ed.). Massachusetts: MIT Press.

_____. (1979). *Scientific Realism and the Plasticity of Mind.* Cambridge: Cambridge University Press.

_____. (1995). *The Engine of Reason, the Seat of the Soul.* Massachusetts: MIT Press.

Churchland, P. M. & Churchland, P. S. (1996). "McCauley's Demand for a Co-level Competitor". *The Churchlands and Their Critics*, McCauley, R. (ed). Massachusetts: Blackwell Publishers.

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

Clark, A. (1993). "The Varieties of Eliminativism: Sentential, Intentional and Catastrophic" *Mind & Language* Vol. 8, 223-33.

Davies, M. & Stone, T. (eds). (1995). *Folk Psychology and the Theories of Mind Debate: Core Readings*. Oxford: Basil Blackwell.

Dennett, D. C. (1987). *The Intentional Stance*. Massachusetts: MIT Press.

Feyerabend, P. (1970). "Against the Method" *Minnesota Studies in the Philosophy of Science* Vol. 4, 17-130.

_____. (1975). *Against Method*. London: Verso.

_____. (1963). "Materialism and the Mind-Body Problem". *Review of Metaphysics*, Vol. 17, 49-66. Reprinted in *Folk Psychology and the Philosophy of Mind*, Christensen, S. & Turner, D. (eds). Hillsdale: Lawrence Erlbaum Associates, 1993. Page references are to the latter.

Fodor, J. A. & Block, N. J. (1972). "What psychological states are not" *Philosophical Review* 81, 159-81.

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

_____. (1968). *Psychological Explanation*. Random House.

_____. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Massachusetts: MIT Press.

Gibson, R. F. (1987). "Quine on naturalism and epistemology" *Erkenntnis* Vol. 27, 57-78.

Giere, R. (1989). "Scientific rationality as instrumental rationality" *Studies in History and Philosophy of Science* Vol. 20, 377-384.

Goldman, A. (1992). "In Defense of the Simulation Theory" *Mind & Language* Vol. 7, 104-19.

_____. (1989). "Interpretation Psychologized", *Mind & Language*, Vol. 4, 161-85.

Gordon, R. (1986). "Folk Psychology as Simulation" *Mind & Language* Vol. 1, 158-71.

Hannan, B. (1993). "Don't Stop Believing: The Case Against Eliminative Materialism" *Mind & Language* Vol. 8, 165-79.

Hempel, C. G. (1965). *Aspects of Scientific Explanation*. New York: Oxford University Press.

Horgan, T. & Woodward, J. (1985). "Folk Psychology is Here to Stay" *The Philosophical Review* Vol. 94, 197-226. Reprinted in *Folk Psychology and the Philosophy of Mind*, Christensen, S. & Turner, D. (eds). Hillsdale, USA: Lawrence Erlbaum Associates, 1993. Page references are to the latter.

Kornblith, H. (1994). "What Is Naturalistic Epistemology?". *Naturalizing Epistemology*, Kornblith, H. (ed). Massachusetts: MIT Press.

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

_____. (1993). *The Advancement of Science*. New York: Oxford University Press.

Klein, E. R. (1992). "Is 'normative naturalism' an oxymoron?" *Philosophical Psychology* Vol. 5 (3), 289-97.

_____. (1992). "Normative naturalism undefended: a response to McCauley's reply" *Philosophical Psychology* Vol. 5 (3), 307-8.

Kuhn, T. (1970). *The Structure of Scientific* (2nd ed.). Chicago: The University of Chicago Press (1st ed. 1962).

Lahav, R. (1992). "The Amazing Predictive Power of Folk Psychology" *Australasian Journal of Philosophy* Vol. 70, 99-105.

Lycan, W. (ed). (1990). *Mind and Cognition*. Massachusetts: Basil Blackwell.

Marras, A. (1995). 'Behaviorism'. *The Cambridge Dictionary of Philosophy*, Audi, R. (ed). Massachusetts: Cambridge University Press.

McCauley, R. N. (1993). "Cross-Scientific Relations and the Study of the Emotions". To appear in: *The Emotion: Cultural and Biological Perspectives*, Shore, B. & Worthman, C. (eds).

_____. (1992). "Defending normative naturalism: a reply to Ellen Klein" *Philosophical Psychology* Vol. 5 (3), 299-305.

_____. (1988). "Epistemology in an Age of Cognitive Science" *Philosophical Psychology* Vol. 1 (2), 143-52.

_____. (1996). "Explanatory Pluralism and The Co-evolution of Theories in Science". *The Churchlands and Their Critics*, McCauley, R. N. (ed). Massachusetts: Blackwell Publishers.

_____. (1986a). "Intertheoretic relations and the Future of Folk Psychology" *Philosophy of Science* Vol. 53, 179-98. Reprinted in *Folk Psychology and the Philosophy of Mind*, Christensen, S. & Turner, D. (eds). Hillsdale: Lawrence Erlbaum Associates, 1993. Page references are to the latter.

_____. (1986b). "Truth, epistemic ideals, and the psychology of categorization". *PSA–1986 Volume 1*, Fine, A. & Machamer, P. (eds). East Lansing: Philosophy of Science Association.

Nagel, E. (1961). *The Structure of Science*. New York: Harcourt, Brace, & World.

Putnam, H. (1983). *Realism and Reason*. New York: Cambridge University Press.

_____. (1969). "The Nature of Mental States". Originally appeared in *Art, Mind and Religion* as "Psychological Predicates", Capitan, W. H. (ed). Pittsburgh: University of Pittsburgh Press.

Quine, W. V. (1969). "Epistemology naturalized". *Ontological Relativity and Other Essays*. New York: Columbia University Press.

_____. (1981). "Five Milestones of Empiricism". *Theories and Things*. Massachusetts: Harvard University Press.

Rorty, R. (1970). "In Defense of Eliminative Materialism" *Review of Metaphysics* Vol. 24, 112-121.

_____. (1965). "Mind-Body Identity, Privacy, and Categories" *Review of Metaphysics* Vol. 19, 24-54. Reprinted in *Folk Psychology and the Philosophy of Mind*, Christensen, S. & Turner, D. (eds). Hillsdale: Lawrence Erlbaum Associates, 1993. Page references are to the latter.

Roth, P. (1987). *Meaning and Method in the Social Sciences*. New York: Cornell University Press.

Ryle, G. (1949). *The Concept of Mind*. London: Hutchison.

Skinner, B. F. (1933). *Science and Human Behaviour*. New York: MacMillan Press.

Stich, S. (1996). *Deconstructing the Mind*. New York: Oxford University Press.

_____. (1983). *From Folk Psychology to Cognitive Science: The Case Against Belief*. Massachusetts: MIT Press.

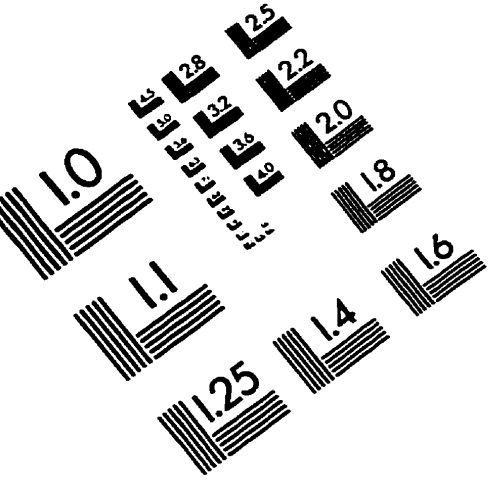
_____. (1990). *The Fragmentation of Reason*. Massachusetts: MIT Press.

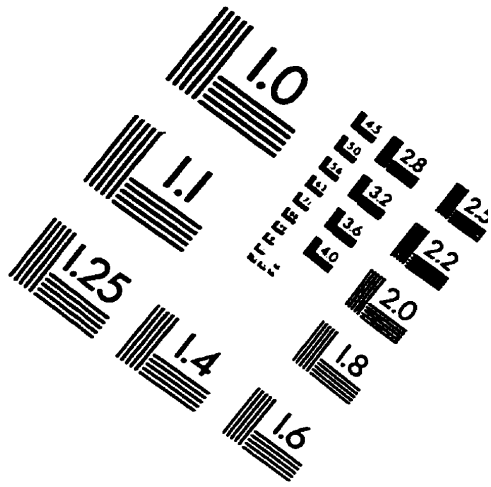
Thagard, P. (1992). *Conceptual Revolutions*. Princeton: Princeton University Press.

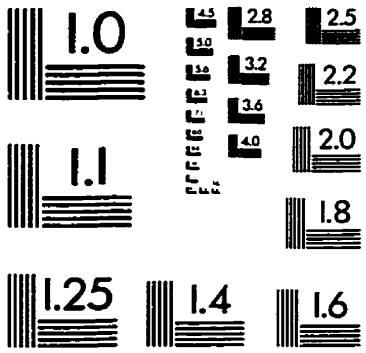
Wimsatt, W. (1976). "Reductionism, Levels of Organization, and the Mind-Body Problem". In *Consciousness and the Brain*, Globus, Maxwell, & Savodnik (eds). New York: Plenum Press.

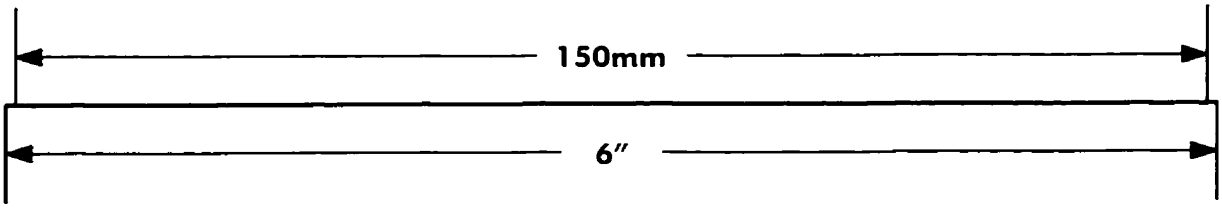
Wittgenstein, L. (1953). *Philosophical Investigations* (3rd ed), trans. by Anscombe, G. E. M. Englewood Cliffs: Prentice Hall (©Basil Blackwell 1958).

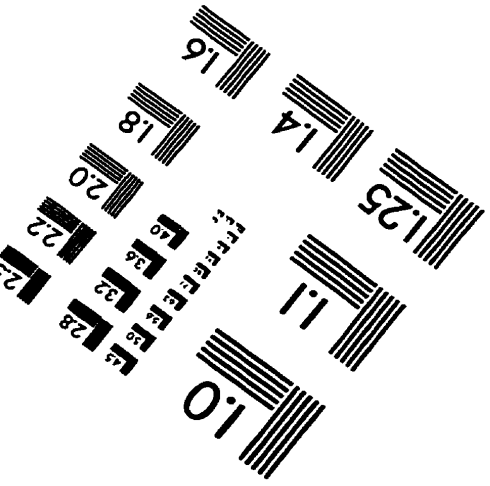
IMAGE EVALUATION
TEST TARGET (QA-3)

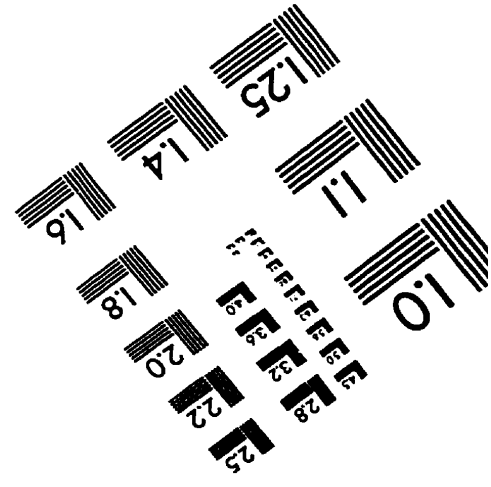












APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989