APPROACHES TO ACTION

by

EUAN FINDLAY CARNIE

B.A., The University of British Columbia, 1988

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

in

THE FACULTY OF GRADUATE STUDIES

(Department of Philosophy)

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

January 1993

© Euan Findlay Carnie

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

(Signature)

Department of _PHILOSOPHY_

The University of British Columbia
Vancouver, Canada

Date _FEB 5 /93_

ABSTRACT

The objective of this thesis is to introduce recent philosophy of mind to philosophy of action in an attempt to evaluate the three currently dominant theories of action. These theories each concur with the general consensus in philosophy of action that the acts of an agent are causally derived. Yet, there is still dispute regarding the exact nature of the causation. This disagreement is reflected in the differences between the identity theory, functionalism, and eliminative materialism.

The identity theory, as described by Donald Davidson and Colin McGinn, states that materially instansiable beliefs and desires are the causes of action. However, problems arise from a lack of a priori validity, difficulties regarding the type of reduction or identification needed, and an unreasonable suppression of misrepresentation. There are also questions concerning the identity theory's account of learning, memory, and imagination. As a successful remedy has yet to appear, charges of theoretical stagnation have also gained momentum.

Nevertheless, faults in a theory do not automatically entail its elimination, and many feel that the predicted demise of the belief and desire model has been greatly exaggerated. For instance, functionalists like Daniel Dennett retain the belief and desire structure, but endeavor to avoid the ontological difficulties which beset the identity theory by functionally defining beliefs and desires. Unfortunately, this merely leads to indiscriminate ascriptions of belief and desire to subjects not normally associated with such attributes.

Functionalists, however, do tacitly recognize the legitimacy of a more empirical outlook in describing the causes of action. This empirical approach, referred to as eliminative materialism, is the basis of claims by theorists such as Paul Churchland that the belief and desire causal framework is fundamentally flawed, and hence, should be subject to an elimination similar to the abandonment of theories like alchemy, astrology, and caloric fluidity.

In light of the defects of the identity and functionalists theories, and enhanced by the appearance of a promising alternative which coalesces well with Churchland's conclusions (i.e.: connectionist parallel processing models of cognition), the balance now appears to be tipped in favour of the eliminativists. Yet, the question remains of just how much of an elimination of the ontological, practical, or semantic facets of the belief/desire model is warranted. While the final answer awaits confirmation, it seems evident that the solution will only be resolved by maintaining an eliminative approach towards the causes of action.

TABLE OF CONTENTS

ACKNOWLEDGEMENT

CHAPTER ONE

Three Approaches to the Causes of Action

Although it has been problematic for some philosophers,[1] there is currently widespread agreement[2] that actions are, in a straightforward sense, caused. The convention that agency and determinism are mutually exclusive is generally considered outmoded, and the path is now paved for a comprehensive causal explanation of action. Nevertheless, there is still disagreement about the exact nature of the causes of action. The disagreement is reflected in the three dominant approaches to action theory: the identity or reductive theory, eliminative materialism, and functionalism. This thesis, then, attempts to evaluate these choices.

The acceptance of causation's role in action is a good beginning, but from this point there is controversy regarding the best explanation of action's causal antecedents. From the outset, most solutions have centered on an identity or reductionist explanation of action's causes. This assumes that materially instantiable beliefs and desires are at the root of action's causes. Another popular answer has been to concentrate on a number of functionalist alternatives, thereby avoiding the ontological difficulties which beset belief and desire. And more recently, in marked contrast to its predecessors, eliminative materialism has denied that any reference to concepts like belief and desire is valuable because such an approach is suspect, particularly when compared to some of cognitive science's more promising alternatives.

While all the issues surrounding these arguments will be thoroughly examined later, a brief outline of what lies ahead will provide a useful framework. Each avenue will be explored through the arguments of some of its best known exponents. Thus, Donald Davidson and Colin McGinn represent the identity or reductionist theory, Paul Churchland the eliminative materialist agenda, and Daniel Dennett the functionalist initiative.

It was Davidson's causal version of an intuitively attractive, linguistically inspired model of action that first spurred interest in causal action theories. Along with McGinn, he claims that belief and desire are the reasons for an agent's action. These reasons create the agent's intention, which is significant, because it is the presence of intentionality which distinguishes action events from other more common events:

> Someone who acts with a certain intention acts for a
> reason; he has something in mind that he wants to promote
> or accomplish. A man who nails boards together with the
> intention of building a squirrel house must want to build
> a squirrel house, or think he ought to (no doubt for
> further reasons), and he must believe that by nailing the
> boards together he will advance his project.[3]

An action, then, is an intentional event, and since belief and desire are the basis of intention, they must also be the causes of action. Davidson and McGinn add that the rational features of action dictate that belief and desire be best defined as propositional attitudes - that is, logical premises about the world.

The advantage of the identity theory's explanation is that causally efficacious, propositional beliefs and desires solve two problems at once. First, belief and desire account for the causes of action, and secondly, in virtue of their propositional content,

they account for the reasons that rationalize action. For example, propositions such as, "I desire something to eat," and "I believe there is food in cupboard," act as premises in a rational syllogism with a conclusion like, "I intend to eat the food in the cupboard." In this way, belief and desire provide an elegant account of both action's rational and causal aspects.

Davidson's and McGinn's theory hinges upon belief and desire being causally efficacious, and this requires that belief and desire reduce to, or at least are identical with, the action's underlying physical causation. Yet, difficulties arise when such a reduction or identity is postulated. As explained more fully in the following chapters, commentators like Churchland, Dennett, and Cummins point out that the validity of the mental event supposition itself is suspect; that the proposed type-type or token-token reductions are problematic; and that even if a reduction were to succeed it would have serious problems in accounting for misrepresentation.

Eliminative materialists also attack belief and desire because they relegate these constructs to the dubious world of "folk psychology" (a psychology ostensibly on the brink of theoretical elimination):

> The propositional attitudes do not constitute an unbreachable barrier to the advancing tide of neuroscience. On the contrary, the principled displacement of folk psychology is not only richly possible; it represents one of the most intriguing theoretical displacements we can currently imagine.[4]

They state that the elimination of folk psychologically inspired reductionist and identity theories is necessary because such

explanations are inadequate, stagnant, and alienated from modern cognitive science.

Specifically, Churchland argues that the folk psychologically based identity and reductionist theories are unable to explain the causes of acts involving memory, imagination, and especially learning. For instance, how does folk psychology justify its propositional model of memory? As evidenced by artificial intelligence's difficulties with the frame problem, even the simplest acts of memory can require the retrieval of huge numbers of propositions. They also wonder how creative acts of imagination can be propositionally described. Since the imagination does not follow the dicta of propositional content, characterizing imaginative acts in terms of propositions appears redundant. Thirdly, they question how folk psychology explains learning in agents who have yet to acquire the cognitive skills needed for propositional manipulation. There are many examples of agents such as young children and some higher animals that lack the ability to manipulate propositions, but nonetheless display intentional action.

The fact that no folk psychology based theory of action has managed to successfully resolve these criticisms leads Churchland to conclude that such folk psychology provides an inadequate explanation of the causes of action. The perpetuation of this failure is the grist for Churchland's further assertion that folk psychology is stagnant. The stagnation is made all the more glaring in light of the comparatively massive strides made by more scientific theories of action. Connectionist artificial intelligence and computational neuroscience are good examples of this

progress. The alienation of the reductionist's propositions from science's research, coupled with its stagnant inadequacy appear to give Churchland a strong case for arguing that the folk psychological explanation is flawed. However it remains to be seen whether outright elimination is actually warranted.

Theorists such as Jeff Foss, Terence Horgan, and James Woodward all counter that the scientific evidence has yet to become conclusive. For these critics, cognitive science's explanation of action has not even reached folk psychology's sophistication. Hence, they hold that Churchland is overly optimistic in the claim that science will eventually displace reductive or identity explanations of action.

Hoping to stem the increasingly rhetorical exchanges between the reductionists and the eliminativists, Dennett's instrumental revision offers a less defensive reaction to Churchland's arguments. Rather than simply villainizing the eliminative position, he concedes that folk psychology has faults, but denies that they are fatal to the concepts of propositional causal action. He argues that propositions like belief and desire are still viable because they continue to be the only practical method of compre- hending an agent's actions:

> As many philosophers have observed, a feature of folk psychology that sets it apart from both folk physics and the academic physical sciences is that the explanations of actions citing beliefs and desires normally not only describe the provenance of actions, but at the same time defend them as reasonable under the circumstances. They are reason giving explanations, which make ineliminable allusion to the rationality of the agent.... I suggest that folk psychology might best be viewed as a rationalistic calculus of interpretation and prediction....[5]

For Dennett, the utility inherent in propositional action theory justifies the claim that folk psychological propositions are worth retaining.

Nevertheless, Dennett fully admits that instrumental explanations are merely pragmatic identifications of the causes of action. Instrumental action theory will always be somewhat tentative given its supposition that a correct prediction amounts to a correct explanation. Still, it remains a useful method simply because the more obscure, physical explanations of action's causes remain practically inaccessible in everyday situations.

Tentative assumptions of propositionally caused action, unfortunately, can prove enigmatic. For instance, "the intentional stance" is capable of ascribing belief and desire propositions to inanimate objects. Following Dennett's directions, the actions of a chair could be predicted by hypothesizing that the chair will not move because it desires to stay where it is, and believes that by not moving it will satisfy this desire. Given that the intentional stance's functional indiscriminancy permits such explanations, Dennett's revision of the folk psychological theory of action needs work.

Consequently, Dennett's failure to bolster folk psychology returns the investigation to the only remaining options. Either the faults of the folk psychologically based identity, reductive, and functional theories are ignored, or Churchland's eliminative/cognitive approach must be taken up. Before making a decision it should be noted that there is growing empirical evidence which supports a move like Churchland's. Thus a proper assessment requires some inquiry into cognitive science.

This brief overview, then, charts the terrain ahead of the more thorough investigation into the three primary approaches to explaining the causes of action. There is a broad based consensus that actions are caused, but the generic nature of action's causes remains in dispute. The identity, functional, and eliminative materialist action theories reflect this. Perhaps, in organizing the main arguments which fuel this dispute and directing them closely at their opponents, it is to be hoped that an understanding and a resolution may be somewhat closer.

CHAPTER ONE NOTES

1) Thomas Nagel, *The View From Nowhere* (Oxford: Oxford University Press, 1986) 7.

2) Donald Davidson, *Essays on Actions and Events* (Oxford: Clarendon Press, 1963-78) 48.

3) Davidson 83.

4) Paul Churchland, *A Neurocomputational Perspective* (Cambridge, MA.: The MIT Press/Bradford Books, 1989) 22.

5) Daniel Dennett, *The Intentional Stance* (Cambridge, MA.: MIT Press/Bradford Books, 1987) 48.

CHAPTER TWO

The Identity Theory

The two major innovations in relatively recent philosophy of action have been the identification of belief and desire as the reasons for acting, and the recognition that these constitute an intention to act. Allowing this understanding, identity and reductive theories then attempt to explain how the concepts of belief and desire might also be causally efficacious. A good example of such a reductive theory is Donald Davidson's theory of action.

In "Actions, Reasons, and Causes" (1963), Davidson states that the explanation of action requires the revealing of a primary reason for an act. A primary reason amounts to possessing a belief and pro-attitude towards a particular action:

> Whenever someone does something for a reason, he can be characterized as a) having some sort of pro-attitude towards actions of a certain kind, and b) believing (or knowing, perceiving, noticing, remembering) that his action is of that kind.[1]

To put it more technically:

> R is a primary reason why an agent performed the action A under the description d only if R consists of a pro-attitude of the agent towards actions with a certain property, and a belief of the agent that A, under the description d, has that property.[2]

Therefore, Davidson claims that when an agent acts, the reasons for the action are best explained by the possession of the appropriate beliefs and pro-attitudes (or desires).

Once this assertion is established, Davidson posits an elegant reduction. He claims that if belief and desire can serve as the reasons for action, they can also serve as the causes of action:

> If causal explanations are "wholly irrelevant to the understanding we seek" of human action then we are without an analysis of the "because" in "He did it because...," where we go on to name a reason.[3]

Thus, his reduction:

A primary reason for an action is its cause.[4]

In one fell swoop, the primary reasons for action also become the primary causes of action.

Davidson continues his explanation of action by claiming that it is the possession of the appropriate reasons for action that makes up the intention to act:

> To know a primary reason why someone acted as he did is to know an intention with which the action is done. If I turn left at the fork because I want to get to Katmandu, my intention in turning left is to get to Katmandu.[5]

For Davidson, beliefs (or "knowing, perceiving, noticing, remembering") and pro-attitudes (or "desires, wantings, urges, promptings, and a great variety of moral views, aesthetic principles, economic prejudices, social conventions and public and private values,"[6] explain reasons. Reasons, in turn, rationalize intentions which then ultimately define actions. To put it very simply, as well as causing action, beliefs and desires are the reasons which explain intentional events like action.

The obvious advantage of Davidson's propositional theory is that it admits reasoned, intentional explanations into the relatively straightforward ontology of causation. He illustrates the theory's effectiveness with the example of his turning on a light and inadvertently alerting a prowler.[7] Upon entering his darkened home, Davidson flips a switch and turns on the light. The light's illumination warns a prowler who subsequently escapes. From a strictly causal standpoint, it appears that Davidson is the cause of the escape.

But to attribute that to Davidson as his action seems intuitively unacceptable. Similarly counter-intuitive situations had led earlier action theorists to search outside of causal determinism for descriptions of intentional action.[8] Nonetheless, under Davidson's causal theory it is possible to explain his action without implicating him in the escape of the prowler.

If he were to be blamed for the act of 'alerting the prowler,' it would need to be demonstrated that the cause of the prowler's escape was Davidson's intention to aid the burglar. It would have to be proven that he possessed the appropriate belief and desire pre-requisite to aiding the prowler. In the example, it is clear that he never intended to warn the prowler because Davidson did not possess, a) any beliefs regarding the prowler, nor b) any pro-attitudes towards the prowler.

It is evident that he intended nothing more than to turn on the light. Therefore, under Davidson's theory, he would be exonerated from alerting the burglar, and barring negligence, not responsible for aiding the escape. The alerting of the prowler is

simply an alternative description of the consequences of his action
- not a description of the action itself.

Davidson's identity theory is more or less the currently
accepted causal explanation of action in contemporary philosophy:
rational acts are defined as intentionally caused since they are a
product of belief and desire propositions. In fact, the theory is
so popular that the only significant alteration to his thesis has
been Colin McGinn's reductive elaboration in "Action and its
Explanation" (1979).

McGinn works backwards from Davidson's conclusions, filling in
many of the gaps. He begins by claiming that following the
dictates of commonsense, behavior must be divided into that which
agents do and that which happens to them. That which they do is
regarded as action. That which happens to them is simply clas-
sified as events in the world. The question then becomes, "What is
it about agency that makes what it causes an action?" McGinn
argues that the distinguishing feature is that the actions of an
agent are done intentionally:

> It seems clear at once that the concept of intention is
> integral to the notion of agency. For surely, if a piece
> of behavior is intentional, it is something the agent
> does - it is an action.[9]

Of course, McGinn admits that agents can also do things
unintentionally, such as spilling an ink bottle. Though it is
something the agent causes (accidentally), it is not an action
because such an event does not involve intention in the right way.
He points out that spilling the ink bottle is but one description
among many alternative descriptions of some core intentional act

such as, 'reaching for the pen.' The unintentional is usually an alternative description of some central intentional action, "in the immediate offing."[10]

For this reason, McGinn holds that it is mandatory for any action ascription that the action be described in some way that makes it true to say it was intentional. If the act cannot be described under some description which makes it truly intentional, then the act is not an action - it is an event. Naturally, because any event is potentially describable as an action, he accepts that the correctness of the ascription pivots upon the truthfulness of the description. To prevent untruthfulness, he cautions that an event can only be an action if its description can be substituted into an intentional sentence and the truth results. Hence, a reflex will not be an action under McGinn's view because such a behavior will not truthfully satisfy an intentional sentence.

It becomes obvious that the identification of intentional sentences is vital. The clue to identifying intentional sentences is to look for sentences that describe a behavior as acting for a reason:

> Acting intentionally seems intimately related to acting for a reason. When I empty my glass intentionally I do it for a reason, perhaps so I can refill it with vodka; but if my knee jerks reflexively no reason can be ascribed to me.[11]

McGinn facilitates the recognition of intentional descriptions by stating that the best conception of reason is to conceive of it as a combination of belief and desire. Thus, when an agent acts for a reason, it is because there is a desire which gives purpose to the action and a belief which attends to the means of satisfying

that purpose. Belief and desire can be seen, according to him, as the premises in a sort of practical syllogism:

> The components of an agent's reason can be construed as premises in a piece of practical reasoning. Practical reasoning is reasoning about what to do, and it involves taking account of one's desires and one's beliefs about certain courses of action.[12]

For example, if 'raising my arm' is described as an action it should be possible substitute into a true intentional syllogism premises like, a) I desire to reach an object, b) I believe that my arm can reach the object, and therefore, c) I raise my arm.

The doubt remains, however, whether such an analysis of action is really an adequate explanation of action. McGinn, like Davidson, claims that the analysis is sufficient because belief and desire descriptions are always a part of sentences that begin like, "I moved my arm because ...." It is just such sentences, according to them, which rationalize action:

> When we come to know an agent's reason for acting we learn what it was about the action, given his beliefs and desires, that made it appeal to him; and we learn that he acted because it thus appealed to him. This type of explanation may be called rationalization.[13]

Building on the foundation of rationalization, McGinn goes on to state that:

> Rationalization introduces an element of justification: knowing the agent's reason we appreciate why, from his point of view, what he did was a rational thing to do; we see of which practical syllogism his action was the 'conclusion'.[14]

The process of action justification relies on the propositional content of the relevant beliefs and desires. The propositional content explains *why* a particular action took place:

> An action is presented as reasonable by dint of its relation, as described, to desires and beliefs possessed of specific propositional contents.[15]

The content of beliefs and desires such as, "I desire to reach the object," and "I believe I can reach the object," justifies why I reached for the object.

The justifying role of belief and desire is fundamental to explaining why I raise my arm, but there is more. We may believe and desire that an action should take place, but still not actually perform the act. I might have good reason to raise my arm, but until I am actually caused to raise my arm, I will not perform the action. It must be known just what triggers action. McGinn argues that as belief and desire are the best candidates to explain *why*, they are the best choice to explain *what* triggers an action:

> When one event occurs because of another the natural and obvious account of the relation involved is that it is causation. And the case of reasons and actions seems to invite just such an account: reasons cause actions, and rationalization is a species of causal explanation. Here then is the second role of rationalizing propositional attitudes: they cause, and causally explain, actions. Reasons are rational causes. The practical syllogism is therefore not merely a means of reconstructing practical reasoning; it is also a sketch for a certain special sort of causal explanation.[16]

Basically, then, a syllogistic combination of belief and desire is what causes an agent's rational action.

In sum, McGinn and Davidson explain action in terms of belief and desire. They claim that belief and desire are the causes of action, and given that such propositions are the best conception of reason, these constructs also justify why an action occurs. The capacity of belief and desire to justify action stems from the content inherent in propositions like, "I believe that...," and "I desire that...." For the reductionist and identity theories, the propositional explanation of action is sufficient to explain how and why an action take place.

In every respect, Davidson's and McGinn's explanation is a good example of a commonsense theory about the causes of action. It seems able to account for action's causation and rationality. Additionally, the theory is intuitively attractive. It is ironic, then, that such a plausible theory is criticized as merely a good illustration of the radically misleading way in which commonsense psychology leads theorists to explain the causes of action.

CHAPTER TWO NOTES

1)   Donald Davidson, "Actions, Reasons, and Causes," in *Essays on Actions and Events* (Oxford: Clarendon Press, 1963-78) 3.

2)   Davidson 5.

3)   Davidson 11.

4)   Davidson 12.

5)   Davidson 7.

6)   Davidson 3-4.

7)   Davidson 4-6.

8)   Earnest LePore, Brian P. McLaughlin, *Actions, Reasons, Causes, and Intentions* in Lepore, E., McLaughlin, B. (eds.) *Actions and Events: Perspectives on the Philosophy of Donald Davidson* (Oxford: Basil Blackwell Ltd., 1985) 6.

9)   Colin McGinn, "Action and its Explanation," in *Philosophical Problems in Psychology*, ed. Neil Bolton (London: Methuen & Co., 1979) 20.

10)   McGinn 21.

11)   McGinn 23.

12)   McGinn 23.

13)   McGinn 25.

14)   McGinn 25.

15)   McGinn 25.

16)   McGinn 26.

CHAPTER THREE

The Identity Theory Questioned

Davidson and McGinn hold that the causes of an agent's action are the agent's beliefs and desires. For them, belief and desire are causally efficacious propositions that in virtue of their propositional content are also able to rationalize action. On the identity theory's reductive outlook, the underlying premise is that mental events like belief and desire are theoretically reducible to, or at least identical with, physical brain events. It is by dint of this physical identity that those mental states which are otherwise purely logical are able to display causal efficacy.

However, there are a number of difficulties with this approach to action. For instance, the initial assumption that there are mental events (at all) appears to rest upon an invalid supposition that mental events are distinct from brain events. Additionally, the theoretical reduction of mental events to physical brain events is itself problematic. And even if such a reduction were possible, a 1:1 correlation between brain events and metal events creates serious difficulties in accounting for the representational mistakes which inevitably arise during the exercise of agency.

The first problem the identity theory faces is the validity of the claim that mental events such as belief and desire are distinct from brain events. That is, in light of the fact that the mental event/brain event distinction rests upon an a priori assumption, validity is suspect. According to the identity theorists, mental events differ from brain events because they exhibit properties that distinguish them from the material brain events. However,

since this is grounded upon an introspective observation, it appears insufficient to sustain the distinction.

At first glance, the mental event distinction in Davidson's identity theory appears feasible. As Churchland casts Davidson's underlying premise:

>   1. My mental states are introspectively known by me as states of my conscious self.
>   2. My brain states are not introspectively known by me as states of my conscious self.
> Therefore, by Leibniz' Law (that numerically identical things must have exactly the same properties),
>   3. My mental states are not identical with my brain states.[1]

However, the argument's *intensional* fallacy is made apparent in the following parallel:

>   (1) Temperature is known by me, by tactile sensing, as a feature of material objects.
>   (2) Mean molecular kinetic energy is not known by me, by tactile sensing, as a feature of material objects.
>   (3) Therefore, temperature does not equal mean kinetic energy.[2]

The problem is that the observed properties are inappropriately held to be attributes of the scrutinized subject when in fact they are more correctly attributable to the observer.

In the second argument the property of being detectable by tactile sensing is incorrectly regarded as a property of temperature. This leads to a mistaken distinction between mean molecular kinetic energy and temperature. Likewise, the property of being introspectively detectable is incorrectly applied to mental events (instead of the observer), mistakenly leading to the conclusion that mental events are distinct from brain events. Therefore, just

as there is no validity in distinguishing between mean molecular kinetic energy and temperature, there appears no validity in distinguishing between mental events and brain events.

Even if the distinction were valid, there are still serious difficulties in any potential reduction or identity between the two sets of events. The most straightforward reduction is to simply assert that for every type of mental event there is a corresponding type of brain event. Thus, for any given belief or desire proposition, there is a corresponding physical representation individuated somewhere within the material of the brain. This form of reduction or identity is known as a type-type relation.

Yet, such a strong form of reductionism is problematic. Given the vast spectrum of possible mental events, it is simply too optimistic to hope that somewhere in the brain there is a physical correlate for every type of mental representation. The physical logistics become overwhelming:

> According to the type-type identity theory, for every mentalistic term of predicate "M", there is some predicate "P" expressible in the vocabulary of the physical sciences such that a creature is M if and only if it is P. In symbols:
> (1) (x) (Mx = Px)
> This is reductionism with a vengeance, taking on the burden of replacing, in principle, all mentalistic predicates with co-extensive predicates composed truth-functionally from the predicates of physics. It is now widely agreed to be hopelessly too strong a demand. Believing that cats eat fish is, intuitively, a functional state that might be variously implemented physically, so there is no reason to suppose the commonality referred to on the left-hand side of (1) can be reliably picked out by any predicate, however complex, of physics.[3]

The demand that every type of mental event be matched by some type of physical event is too stringent.

As a consequence, many identity theorists have abandoned the familiar type-type identity for a simpler token-token identity. In order to place a more realistic burden upon the physical side of things, mental event types are broken down into more elementary mental tokens. These smaller mental tokens maintain rules of composition so that they can be recombined to form larger mental representations, but in their most basic form they offer a way out of the stringency inherent in their bigger type-type representational siblings.

Unfortunately, because such tokens can be instantiated in any number of material mediums (i.e.: electrically in connectionist artificial intelligence, or chemically in computational neurobiology) it seems extremely unlikely that they will prove identical or reducible to any one physical system:

> In general, any of various different events/states can properly be considered 'the cause,' at a given time, of a particular effect; normally a particular event/state will receive this designation on the basis of contextually specific parameters, involving such matters as the pragmatics of explanation. Consequently, if one considers only the context-independent features of the causal relation, it is very likely that for any creature C with a non-empty mentality set M(C), there will be numerous sets of physico-chemical events which qualify as physical causal isomorphs of M(C)....The upshot, then, is that any creature C with a non-empty set M(C) is likely to have several – even many – distinct physical causal isomorphs. And this is the principle premise of our argument against the token identity theory.[4]

In effect, there are simply too many physical systems that can manifest mental tokens for there to be much chance that only one particular physical token will prove identical. The sheer proliferation of physical tokens makes Davidson's token-token identity a poor prospect.

The prospects do not improve when it is considered that the biggest difficulty facing a successful 1:1 correlation between mental events and brain events is its accounting for the misrepresentations that inevitably occur in agent's actions. If mental and brain states were identical or directly reducible, then the linkage between the two should admit of no content variation.

Yet, there are many examples of agents mentally representing x while the brain mistakenly physically represents y. The daily inconvenience of incorrect change is probably one of the most common occurrences. Agents frequently mentally compute the change for a purchase and then mistakenly physically represent a different sum. When reminded of the mistake the agent immediately admits of the error, acknowledges that they knew the correct change amount, and usually blames the misrepresentation on a distraction. But, in the identity theorist's scheme where mental representations are identical to the brain's physical representations, there is actually no explanation of an agent mentally representing x while producing an action that indicates the brain was representing y:

> Under ideal conditions, the system must be infallible. Confronting whiteness must be nomically sufficient and necessary for the occurrence of the w-feature in precepts. Hence the transduced proximal stimulus, plus knowledge, plus nonpsychological laws of nature must entail (not just make highly probable; not just reliably indicate) that there is whiteness out there. Remember the "when and only when." "When": If a cat occurs and the c-pattern doesn't occur, then the possibility exists that only orange cats, or only Graycat, excite the pattern. "Only when": if the c-pattern occurs sometimes when it is a dog out there, then there is no principled reason not to say that the c-pattern represents CAT-or-DOG. Thus, the concept CAT must reduce to concepts that apply to simple perceptual properties - i.e., to proximal stimuli.

> Good luck. The literature since Descartes is littered with bankrupt programs that found this price too high.[5]

In reality, agents misrepresent constantly. However, if mental events were truly identical with or reducible to brain events, such misrepresentations should be impossible. The fact that such error is possible contradicts the Davidsonian concept of a direct, immutable link between mental events and brain events.

It should be noted that idealizing away from error will not solve the issue. The sort of resources required to quash error within a reductionist theory of representation are unrealistic. Its not just that Laplacean difficulties arise, or that too stringent a demand is again placed upon the physical side of cognition. Of greater concern is that idealized representation prevents mental events from entering the realm of practical agency. In real world situations agency depends on a certain amount of inaccuracy. It often uses representational shortcuts which are oversimplifications or presumptions in order to cope with a constantly changing environment. Room for error appears to be a significant aspect of a representational agency. An idealized identity theory's denial of error creates more pragmatic problems then are theoretically solved.

The internal inconsistencies of the identity theory are clear. The relatively commonsense idea that if mental events are to be causally efficacious, they must be identical with, or reducible to, brain events culminates in problems of intensional fallacy, difficulties in reduction, and inabilities to deal with misrepresentation. Worst of all, these issues undermine the

identity theory's explanation of action at just the time it faces
its greatest challenge: eliminative materialism.

o

CHAPTER THREE NOTES

1)   Paul Churchland, *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind* (Cambridge, MA.: The MIT Press/Bradford Books, 1984) 32.

2)   Paul Churchland, *A Neurocomputational Perspective* (Cambridge, MA.: The MIT Press/Bradford Books, 1989) 59.

3)   Daniel Dennett, *The Intentional Stance* (Cambridge, MA.: MIT Press/Bradford Books, 1987) 66.

4)   Terence Horgan, Michael Tye, M. 'Against the Token Identity Theory,' eds. Earnest LePore, Brain McLaughlin, *Actions and Events: Perspectives on the Philosophy of Donald Davidson* (Oxford: Basil Blackwell Ltd 1985) 438-439.

5)   Robert Cummins, *Meaning and Mental Representation* (Cambridge, MA.: MIT Press/Bradford Books, 1991) 48-49.

CHAPTER FOUR

The Eliminative Challenge

The proposal that Davidson's reductive theory should be eliminated from philosophy of action originates with theorists like Quine, Rorty, and Fayerabend. Its main current advocate is Paul Churchland, an ardent materialist who states in "Eliminative Materialism and Propositional Attitudes," (1981) that theories like Davidson's are expendable because they are based on a false "commonsense or folk psychology":

> 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 be displaced, rather than smoothly reduced, by a completed neuroscience.[1]

He argues that folk psychological concepts seriously mislead action theory, and therefore, need to be replaced by more scientific explanations.

Churchland begins his case that folk psychology is eliminable by asserting that folk psychology is an abstraction, and should not be assumed to be a natural starting point for explaining action. He claims that folk psychology is theoretical because it draws upon on abstract laws for explanation. As in other abstract theories, these laws provide content to the folk psychological concepts. There are many examples of such theoretical laws: 'x fears that p, where x desires that not p,' or 'x believes that p and x desires that p, therefore, x is pleased that p.'

In this manner, concepts such as belief and desire derive their meaning from their respective positions within the relations stipulated by the laws:

> This approach entails that the semantics of the terms in our familiar mentalistic vocabulary is to be understood in the same manner as the semantics of theoretical terms generally: the meaning of any theoretical term is fixed or constituted by the network of laws in which it figures.[2]

Belief and desire are concepts which owe their existence to the rules which define their properties.

Churchland cites the puzzles of introspection, intentionality, and other minds as further evidence that folk psychology is theoretical. These philosophical dilemmas can be resolved once it is recognized that they are a product of folk psychology's abstractions. That is, the commonsense conviction that there are other minds, when seen from a theoretical perspective, is no longer a mere generalizing from a single case. It is an abstract hypothesis meant to explain how other agents act.

Similarly, introspective judgments lose their incorrigibility or any other troublesome status under the theoretical perspective:

> Introspective judgments about one's own case turn out not to have any special status or integrity anyway. On the present view, an introspective judgment is just an instance of an acquired habit of conceptual response to one's internal states, and the integrity of any particular response is always contingent on the integrity of the acquired conceptual framework (theory) in which the response is framed. Accordingly, one's introspective certainty that one's mind is the seat of beliefs and desires may be as badly misplaced as was the classical man's visual certainty that the star-flecked sphere of the heavens turns daily.[3]

Introspection is essentially a theoretical axiom about how agents understand themselves. Intentionality is another product of the theorized relations proffered by folk psychology. The teleology inherent in intentionality is one way to explain action's rationality.

However, propositional attitudes provide the best evidence for Churchland that folk psychology is theoretical. Propositional attitudes are structurally similar to the propositions in other abstract theories. When physics postulates the proposition, "x has a mass n," it is structurally symmetrical to the propositional attitude, "x has a belief P," or "y has a desire P." Plus, the propositions of folk psychology can be combined into laws like, "belief plus desire equals intention," that are structurally parallel to laws like, "force divided by mass equals acceleration:"

> The "propositional attitudes," as Russell calls them, form the systematic core of folk psychology; and their uniqueness and anomalous logical properties have inspired some to see here a fundamental contrast with anything that mere physical phenomena might conceivably display. The key to this matter lies again in the theoretical nature of folk psychology. The intentionality of mental states here emerges not as a mystery of nature, but as a structural feature of the concepts of folk psychology. Ironically, these same structural features reveal the very close affinity that folk psychology bears to theories in the physical sciences.[4]

Having offered his argument that folk psychology is a theory, Churchland attempts to illustrate why folk psychology is false. He claims that: 1) folk psychology suffers many serious explanatory failures; 2) due to its inability to rectify these failures folk psychology is stagnant; and 3) cognitive science has yet to uncover any of the speculative conceptions folk psychology insists exist.

The explanatory failures of folk psychology are not immediately obvious because the theory appears to be reasonably successful in explaining ordinary action. However, once beyond the periphery of simple prediction, the failures of folk theory loom large. For example, it is unable to explain more complex actions involving memory, imagination, and learning.

Folk psychology has little to say about acts involving memory because the efficient retrieval of propositions remains to be explained. The problem is that intentional propositions are bulky constructs. They require cumbersome subject-predicate structures along with some account of their relations to the world. Into the bargain, experience with the frame problem demonstrates that even the simplest actions can entail reference to vast numbers of propositions. Subsequently, propositional retrieval is a slow and awkward process.

The problem of unwieldiness became apparent when early propositional task forces used computers to operate belief and desire based models of memory. Upon running the propositional programs, the computers became hopelessly bogged down in the enormous task of addressing and retrieving propositions.5 The proposed programs were unacceptably slow. Propositional unwieldiness continues to deny folk psychology a meaningful explanation of actions involving memory.

Creative imagination is an even greater problem for folk theorists. The difficulty arises because acts of creative imagination are not confined by the content of intentional propositions. Propositional content is rarely reflected in actions

involving imagination, making the propositional explanation of imaginative acts often seem redundant.

Acts of learning are the biggest area of difficulty. Given that folk psychological explanations of learning refer to nothing other than propositions, and assuming that infants and animals learn without the cognitive skills pre-requisite to propositional manipulation, folk theories can not explicate their learning:

> One particularly outstanding mystery is the nature of the learning process itself, especially where it involves large-scale conceptual change, and especially as it appears in its pre-linguistic or entirely non-linguistic form (as in infants and animals), which is by far the most common form in nature. Folk psychology is faced with special difficulties here, since its conception of learning as the manipulation and storage of propositional attitudes founders on the fact that how to formulate, manipulate, and store a rich fabric of propositional attitudes is itself something that is learned, and is only one among many acquired cognitive skills. Folk psychology would thus appear constitutionally incapable of even addressing this most basic of mysteries.[6]

Folk psychology comprehends learning solely as a process of formulating, manipulating, and storing propositions about the world. Unfortunately, many agents lack these abilities, leaving how they learn unaddressed by folk psychology.

Even if it is denied that the explanatory failures logically entail elimination, they at least indicate that some revision is necessary. Yet, there has been very little evolution in folk psychology's basic dictates. The ancient Greeks knew virtually as much about action as is currently known. A Greek of two thousand years ago would have no difficulty in following modern action

theory. In light of the obvious inadequacies that persist, the stagnation is particularly incriminating. As Churchland puts it:

> Perfect theories, perhaps, have no need to evolve. But folk psychology is profoundly imperfect. Its failure to develop its resources and extend its range of success is therefore darkly curious, and one must query the integrity of its basic categories. To use Imre Lakato's terms, folk psychology is a stagnant or degenerating research program, and has been for millennia.[7]

As a consequence of folk psychology's inability to advance the understanding of action's causes, other theories have seized the initiative. For example, when action is now analyzed from a neurological level, its descriptions of parallel processing and connectionist networks surpass anything offered by folk psychology. Such alternatives, despite their newness, provide far greater insight into the causes of action.

Worse still, folk psychology is not even a part of this growing body of scientific knowledge. Concepts such as propositional belief and desire are not present in the vocabulary of science. Of course, a reduction of these concepts to a scientific ontology may still be possible, but being an eliminativist, Churchland has little confidence in such a maneuver:

> Its intentional categories stand magnificently alone, without visible prospect of reduction to that larger corpus. A successful reduction cannot be ruled out, in my view, but folk psychology's explanatory impotence and long stagnation inspire little faith that its categories will find themselves neatly reflected in the framework of neuroscience. On the contrary, one is reminded of how alchemy must have looked as elemental chemistry was taking form, how Aristotlean cosmology must have looked as classical mechanics was being articulated, or how the vitalist conception of life must have looked as organic chemistry marched on.[8]

He concludes that by dint of folk psychology's omissions, stagnation, and scientific alienation, it is clearly on the verge of elimination.

Churchland's final tactic is to substantiate eliminativism by attacking two classic defenses of folk psychology. Both defenses follow either one of two types of functionalism. The first type states that the folk psychological explanation of action is in eliminable because it is an "ideal" which serves as a normative model of action. Even though the ideal functionalist readily grants that the model may never be actualized, the functionalist insists that folk psychology can still serve as a standard of comparison. As such, paradigmatic folk psychological explanations are not subject to the constraints Churchland imposes.

The second form of functionalism defends the folk psychological explanation of action on the basis that it is an "abstract" explanatory construction, instantiable in any number of mediums. On this view, folk psychology is simply a set of theoretical relations or properties which may be called upon to provide an explanation of action, regardless of scientific concerns. Hence, the abstract functionalist denies that Churchland can eliminate folk psychology purely on the grounds that science does not appear to support concepts such as belief and desire.

Churchland's initial reply is to the abstract functional defense. He attempts to undermine it by giving a functionally abstract defense of alchemy. Churchland's counter-example demonstrates how an abstract functionalist argument permits theories to evade honest criticism.

An alchemist can contrive an abstract functional defense of alchemy and the four fundamental alchemistic spirits by claiming that it is the alchemistic spirits which provide gold its unique properties. The alchemist claims that mercury, sulphur, yellow arsenic, and sal ammoniac lend gold it's metallicness, yellowness, luster, and pliability. He states that these properties are essential to gold, regardless of gold's periodic status. Thus, if it is accepted that gold possesses these properties, and if it is accepted that only the alchemistic spirits can provide these features, the alchemist stands a good chance of concocting a logical defense of alchemy. The functionally abstract roles of alchemistic spirits can persist despite the advent of atomic physics, assuming that the first premises of alchemy's defense continue to go unchallenged.

Of course, alchemy is false; but when given a functional defense it is disconcerting in its persuasiveness. By translating the alchemistic spirits into abstract functional states which logically account for a given set of properties, it is possible to be misled into holding that alchemy is a realistic theory about the manufacture of gold. Drawing on the numerous parallels between a functionally abstract defense of alchemy and a functionally abstract defense of folk psychology, Churchland inductively concludes that functionally abstract folk psychology is as misleading as alchemy:

> In summary, when confronted with explanatory impotence, stagnant history, and systematic isolation of the intentional idioms of folk psychology, it is not an adequate or responsive defense to insist that these idioms are abstract, functional, and irreducible in character. For one thing, this same defense could have

been mounted with comparable plausibility no matter what haywire network of internal states folklore had ascribed to us. And for another, the defense assumes essentially what is at issue: it assumes that it is the intentional idioms of folk psychology, plus or minus a bit, that express the important features shared by all cognitive systems.[9]

Dealing abstract functionalism this blow, Churchland returns to ideal functionalism. He has three objections to the use of folk psychology as a functional ideal. First, folk psychology's fortuitous use of the logical relations within the practical syllogism does not justify folk psychology being treated as an ideal. For instance, while the classical gas law also makes use of logical relations, it is not treated as an ideal since it displays recognized limitations - the realm of relatively non-extreme temperatures and pressures. There is no reason to assume that even if folk psychology were proven useful within a particular operational sphere, such logic would elevate it beyond other limitations.

Two, it is impossible for a defense of folk psychology to substantiate the claim that the relations set out in folk psychology are ideal. After all, since it is not known what a perfectly rational relation is, it cannot be verified that the rational relations in folk psychology are perfect. Even without defining perfect rationality, Churchland argues that folk psychological rationale hardly appears ideal. It pales in comparison to the logic displayed in modern mathematics and contemporary science. Churchland believes that folk psychology's rationality is minimal and at most confined to the truncated conception of what is popularly understood.

Thirdly, if it is replied that folk psychology has only to capture sufficient rationality to satisfactorily explain action, the fact that its rationality must be expressed exclusively in terms of language-inspired, sentence-like propositions will preclude a satisfactory explanation. Churchland feels the handicap is a result of the superficiality which language driven descriptions of action display:

> Even if our current conception of rationality - and more generally, of cognitive virtue - is largely constituted within the sentential/propositional framework of folk psychology, there is no guarantee that this framework is adequate to the deeper and more accurate account of cognitive virtue which is clearly needed. Even if we concede the categorical integrity of folk psychology, at least as applied to language-using humans, it remains far from clear that the basic parameters of intellectual virtue are to be found at the categorical level comprehended by the propositional attitudes.[10]

Obviously, Churchland is convinced that the language inspired propositionality simply skims the surface of a comprehensive understanding of action.

In final analysis, Churchland states that folk psychology deserves no special status as an explanation since it is only one among many abstract theories about the causes of action. And as abstract theories go, Churchland argues that folk psychology does not fair well:

> A serious advance in our appreciation of cognitive virtue would thus 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).[11]

Folk psychology only continues to pervade action explanations because of its social entrenchment. The entrenchment itself entails nothing, because under scrutiny folk psychological theory crumbles.

For philosophy of action, then, the demise of folk psychology would appear to have serious repercussions. Orthodox reductive theories, with their dependence upon folk psychology for concepts such as belief and desire would be deeply eroded. Realistically, the loss of folk psychology would effectively eliminate reductive and identity theories from the spectrum of plausible explanations of action's causes.

CHAPTER FOUR NOTES

1) Paul Churchland, *A Neurocomputational Perspective* (Cambridge, MA.: The MIT Press/Bradford Books, 1989) 1.

2) Churchland 3.

3) Churchland 3.

4) Churchland 3-4.

5) Daniel Dennett, *The Intentional Stance* (Cambridge, MA.: MIT Press/Bradford Books, 1987) 229.

6) Churchland 7.

7) Churchland 8.

8) Churchland 9.

9) Churchland 14.

10) Churchland 16.

11) Churchland 16.

CHAPTER FIVE

Eliminative Materialism Challenged

Not surprisingly, Churchland's eliminative approach to action faces many critics. Among the more skeptical of the eliminative treatise are Terence Horgan and James Woodward in, "Folk Psychology Is Here To Stay," (1985) and Jeffrey Foss in, "A Materialist's Misgivings About Eliminativist Materialism," (1983). The main thrust of these critics is that Churchland exaggerates both the failings of folk psychology and the advances of cognitive science.

According to Churchland, folk psychology inadequately explains acts involving memory, imagination, and learning; it is stagnated by its inability to rectify these difficulties; and it is becoming increasingly alienated from cognitive science. For Churchland, these faults make the elimination of the folk psychological explanation of action inevitable. Foss, Horgan and Woodward deny, however, that the demise of folk psychology is at hand.

They initially take exception to Churchland's conclusion that folk psychology fails to adequately explain complex acts involving memory, imagination, and learning. Foss admits that folk psychology has few specifics concerning the causes behind such action, but he quickly points out that cognitive science itself has little to offer. As cognitive science is yet to develop its own comprehensive theories of these acts, Churchland has no grounds for implying that cognitive science is in a better position. Even if folk psychology's explanation is not perfect, it is at least the equal of cognitive science.

Horgan and Woodward provide a slightly more positive defense of folk psychology. Though they admit that folk psychology does not directly discuss many types of complex action, there are many theories founded upon it which do try to explain such matters:

> While folk psychology itself may have little to say about the matters Churchland mentions, theories based on concepts deriving from folk psychology have a good deal to say about them. For example, cognitive psychologists have developed extensive and detailed theories about visual perception, memory, and learning that employ concepts recognizably like the folk psychological concepts of belief, desire, judgment, etc.[1]

They contend that there are many good theories based on folk psychology's precepts which deal more directly with imagination, memory, and learning.

Churchland's second conclusion is that folk psychology is stagnant because it continues to rely upon deficient propositional descriptions of action. Foss, however, sees nothing deficient in folk psychology characterizing action in terms of propositions like, "I believe x," and "I desire y." He concedes that while agents such as infants and animals may not employ propositions in their deliberation, this alone does not condemn folk psychology.

That is, Foss claims that sentences such as, "I believe x," are only meant as a descriptive approximation of the events antecedent to action. Folk psychology only uses propositional descriptions of the causes of action because they enable agents to *talk* about action causation:

> Animals and pre-verbal children do see, hear, perceive, remember, want, avoid, and so on. What is wrong with expressing what it is they want, avoid, and so on, by means of language - why not, in fact, use sentences? We

> may know, for example, that the little girl won't go into the yard because she sees the dog, and is afraid of dogs - though she is pre-verbal and has never spoken or thought a word in her life. How can we otherwise express what is going on here, and in any case why should we try to find some other means of expression? As long as we do not picture sentences in her head, there is no harm in saying that she sees that *there is a dog in the yard.*[2]

Folk psychology is not trying to achieve a perfect representation of action's causes - it merely asks that the propositions facilitate the expression of an act's causation.

Horgan and Woodward add that although the linguistically inspired propositional concept of action is old, it is by no means stagnant. Their position contrasts Churchland's claim that folk psychology has advanced so little in the last two thousand years that an ancient Greek would still be able to understand contemporary action ascriptions. Horgan and Woodward counter that folk psychological development is clearly evident:

> For example, it is plausible to conjecture that Europeans in the 18th or 19th centuries were much more likely to explain human behavior in terms of character types with enduring personality traits than 20th century Europeans, who often appeal instead to 'situational' factors. (Certainly the difference is dramatically evident in 18th and 20th century literature; contrast, say, Jane Austen and John Barth.) Another example of empirically progressive change, perhaps, is the greater willingness, in contemporary culture, to appeal to unconscious beliefs and motivations.[3]

Horgan and Woodward also make a good point when they add that even if Churchland is correct about folk psychological progress being slow, moderate progress by itself is not sufficient justification for the elimination of a theory. After all, a gradual evolution is the not equivalent to being radically misleading.

Foss, Horgan, Woodward close their criticism of Churchland by attacking his final conclusion that since folk psychology cannot be reduced to cognitive science, it should be eliminated. Foss claims that just as chronomatic concepts like escapement mechanism and hairspring are meaningful despite not appearing in the categories of atomic physics, folk psychological concepts like propositional belief and desire are also meaningful despite not appearing in the classifications of cognitive science:

> Eliminativist materialists delight in pointing out the failure of folk psychological kinds like hearing and remembering to match new-fangled neurophysical kinds like neuron and hypothalamus. But the mismatch may be no more telling than the failure of chronomatic kinds like escapement mechanism to match physical kinds like iron - yet no one hesitates to reduce a clock to physics, much less deny the existence of the escapement mechanism.[4]

Foss concludes from his analogy of folk psychology and chronomatics, that a) there is no direct logical necessity for the concepts of one theory to perfectly match the postulates of another theory, and b) it does not follow that the concepts of one theory must be radically misleading and rejected simply because there are no logical connections between it and alternative theories.

Horgan and Woodward entirely agree with Foss' assessment. They state that the only conclusion which may be mustered from the fact that cognitive science and folk psychology do not mesh is simply that folk psychology and cognitive science are not amenable:

> But even if folk psychology cannot be reduced to lower level theories, and even if lower level theories provide a marvelous account of the nature and behavior of *homosapiens*, it simply does not follow that folk psychology is radically false, or that humans do not undergo the intentional events it posits. Churchland's

eliminative materialism is not the only viable naturalistic alternative to reductive materialism.[5]

They deny that Churchland is justified in drawing the far stronger conclusion that if folk psychology and cognitive science do not match, folk psychology must be eliminated.

To underscore their point, Horgan and Woodward cite Donald Davidson as a causal action theorist who actively supports the scientific view of action in addition to the folk psychological outlook. In "Agency" (1971), Davidson states that while folk psychology's intentional propositions are a necessary ingredient of the propositions of action, material causation is also central to agency. Moreover, he holds that these two features of action will never be directly reduced, despite being identical, because intentionality is completely semantic and causality is entirely extensional.

Thus, in Davidson's identity theory of action, folk psychology's propositions will never completely dissolve into the materialism of cognitive science:

> Davidson advocates a thesis which asserts that every concrete mental event is identical to some concrete neurological event, but which does not assert (indeed, even denies) that there are systematic bridge laws linking mental event-types, or properties, with neurological event-types. He calls this view anomalous monism; it is a form of monism because it posits psychological identities, and it is "anomalous" because it rejects reductive bridge laws (or reductive type-type identities).[6]

Horgan and Woodward take Davidson's theory as an illustration of a scheme that eschews the type of reduction that Churchland demands of folk psychology. Therefore, they claim that Churchland's demand

that folk psychology be eliminated because it does not comply with such a reduction is inconclusive.

Horgan, Woodward, and Foss ultimately assail Churchland for basing the eliminative attack of folk psychology upon the promissory note that cognitive science can provide a superior explanation of the causes of action. Cognitive science has not delivered on this promise. Hence, folk psychology is not logically obliged to be reducible to cognitive science in order to maintain its integrity. Furthermore, they contend that folk psychology is not necessarily the stagnant and superficial explanatory failure that Churchland portrays. They argue that he is too extreme in concluding that folk psychology is fatally flawed and irretrievably bound for elimination.

Where does this leave folk psychologically based theories of causal action? As their arguments attest, Churchland's critics concede that folk psychological propositional theory suffers difficulties. Yet, they do not accept the induction that the flaws necessarily leads to elimination. Nevertheless, the acknowledgment of folk psychology's faults is significant.

It is clear, in view of the eliminative arguments, that some alteration is mandatory if folk psychology's role in action theory is to remain tenable. Whether this means that folk psychology becomes more integrated with cognitive science or whether cognitive science replaces much of it depends on the form of revision which the folk psychologists offer.

CHAPTER FIVE NOTES

1)   Terence Horgan, James Woodward, "Folk Psychology Is Here To
     Stay," *The Philosophical Review*, XCIV, No. 2 (1985): 200.

2)   Jeff Foss, "A Materialist's Misgivings About Eliminative
     Materialism," *Canadian Journal of Philosophy*, Supp. Vol. 11
     (1983): 120.

3)   Horgan 201.

4)   Foss 121-122.

5)   Horgan 203-204.

6)   Horgan 206.

CHAPTER SIX

Functional Revisions

In an attempt to revitalize folk psychologically based causal action theories, Daniel Dennett takes the criticisms leveled at Churchland's eliminative materialism as an opportunity to offer an instrumentalist alternative. As a revision of the original abstract and ideal functionalisms, Dennett's book *The Intentional Stance* (1987) is an intriguing folk psychologically based response to Churchland's eliminative initiative. Like Foss, Dennett strives to sustain folk psychology by exploiting its inherent utility.

Dennett realizes from the start that paradoxes suffuse the folk psychological explanation of action's causes. At times, its attributions are encroached upon by subjectivity, cultural relativity, and Quinian indeterminacies of translation. At other times, folk psychological attributions of seem to be about as straight forward as, "speaking prose and counting beans."[1] In order to clear up this apparent contradiction, Dennett states that it must be recognized that folk psychology has two distinct outlooks.

The first outlook Dennett terms the "realist" view. This objective perspective holds that agent's experience belief and desire something in the same way to experiencing pain or stress. Belief and desire are considered physical aspects of the body and are simply an internal matter of fact - "something like suffering a virus."[2] To have a belief or desire on this view amounts to the brain being in a particular physical state.

On the other hand, the second, more "interpretive" outlook assumes that the possession of a belief or desire is like enjoying a talent or exhibiting a personal style. The attribution includes more than a simple physical fact. Consequently, the subjective identification of a belief relies on a certain amount of interpretation on behalf of the attributor.

Having noted that there are two distinct folk psychological viewpoints, Dennett asserts that this dichotomy has lead to much of the confusion surrounding folk psychology's attributions. Nevertheless, this divergence can be straddled and the confusion removed. The solution lies in understanding that while belief and desire may be completely objective phenomenon, their detection can only be achieved from a subjective vantage point.

Therefore, while belief and desire may be physical attributes of the body, their identification necessitates an interpretive strategy:

> My thesis will be that while belief is a perfectly objective phenomenon (that apparently makes me a realist), it can only be discerned only from the point of view of one who adopts a certain predictive strategy (that apparently makes me an interpretationist).[3]

For Dennett, the strategy is to interpret at least some of the events surrounding an agent as caused by the agent's beliefs and desires. When this interpretation produces reliable predictions about the agent's behavior, it can be reasonably assumed that it is an accurate explanation of the causes of the agent's action.

The starting point of Dennett's pragmatic approach is to treat the subject under scrutiny as a rational agent. That is, begin with the assumption that the subject actually experiences belief,

desire, and intentionality. Dennett then states that if the rational behavior of the subject becomes predictable in light of these assumptions, it may be deduced that it is correct to regard the subject as a "believer":

> To a first approximation, the intentional strategy consists of treating the object whose behavior you want to predict as a rational agent with beliefs and desires and other mental stages exhibiting what Brentano and others call *intentionality*.
> Then I will argue that any object – or as I shall say, any system – whose behavior is well predicted by this strategy is in the fullest sense of the word a believer. *What it is to be a true believer is to be an intentional system*, a system whose behavior is reliably and voluminously predictable via the intentional strategy.[4]

He labels this folk psychologically based, instrumental strategy the "intentional stance."

He notes that the intentional stance is actually very similar to other predictive strategies. For instance, astrology is an example of a similar strategy. It works on the premise that a birth date in conjunction with the movement of the planets is somehow indicative of the future. Hence, it parallels the intentional stance in its assumption that certain criteria are relevant to the prediction of behavior.

Dennett cites the physical stance as another example of a predictive strategy analogous to the intentional stance. To use the physical strategy, the assumption is made that the world is a physical constituency, where causality determines the outcome of all behavior, and indeed, everything else:

> Consider the physical strategy, or the physical stance; if you want to predict the behavior of a system,

determine its physical constitution (perhaps all the way
down to the microphysical level) and the physical nature
of the impingements upon it, and use your knowledge of
the laws of physics to predict the outcome of any input.
This is the grand and impractical strategy of Laplace for
predicting the entire future for everything in the
universe, but it has more modest, local, actually usable
versions.[5]

With this premise in place, when confronted with a question
concerning future actions and events, the relevant data is plugged
into a causal calculation and a prediction results.

As Dennett suggests, such Laplacean causal calculations are
inconvenient. They are hindered by the sheer impracticality of
enumerating every cause and effect relevant to a given situation.
Such impracticality precludes most agents from readily achieving
causal explanations. Therefore, agents find it a practical
necessity to devise alternative strategies that allow for their
limited resources, while still helping to make sense of the world.

It is within this practical realm that Dennett claims the
intentional stance exists. The intentional stance allows that it
is not always possible for agents to compile a complete causal
history for each action. It overcomes this deficiency by using the
assumption of belief and desire as a type of bridge that spans the
gaps left by an incomplete knowledge of a causal history:

Here is how it works: first you decide to treat the
object whose behavior is to be predicted as a rational
agent; then you figure out what beliefs that agent ought
to have, given its place in the world and its purpose.
Then you figure out what desire it ought to have, on the
same considerations, and finally you predict that this
rational agent will act to further its goals in light of
its beliefs. A little practical reasoning from the
chosen set of beliefs and desires will in many - but not
all - instances yield a decision about what the agent[6]
ought to do; that is what you predict the agent will do.

Born of practical necessity, Dennett claims that the strategy is frequently a helpful crutch. When alternative causal explanations prove inaccessible, he claims that pragmatism dictates that the intentional stance is adopted.

Dennett also elaborates on the ascription of belief and desire, and the application of the practical syllogism. Beginning with the position that it is best to ascribe the beliefs a subject ought to have, he states that it is important to understand how those beliefs come to be formed. For a subject to gain a belief, the subject must have exposure to situations that result in belief formation. This may seem straight forward, but it actually results in a serious restriction on the number of beliefs attributable.

For example, if an experience is of little interest or relevance, the agent often neglects to form a belief. Therefore, Dennett's first rule of belief ascription is drawn: "attribute as beliefs all the truths relevant to the system's interests." Dennett concedes that postulating too many beliefs is still a problem since many germane beliefs actually lack truth, and so, discarded. Additionally, even if a belief is both relevant and true, there is no guarantee that the belief will be remembered. Undaunted, he insists that enough dependable beliefs will be retained to maintain the explanatory value of the belief hypothesis.

Just as the rule for belief attribution stemmed from the more basic rule, "attribute all the beliefs the system ought to have," Dennett's rule of desire attribution is derived from the basic dictate, "attribute all the desires the system ought to have." His desire rule, then: "attribute desires for those things a system

believes to be the best means to other ends it desires." Again, this effectively limits the desire attributions to only those desires relevant or of interest to the subject in question.

At this point, he feels it is important to interject a warning regarding the employment of his rules of belief and desire attribution. He cautions that the necessity of expressing belief and desire within linguistic representation can have distorting effects:

> Language *enables* us to formulate highly specific desires, but it also forces us on occasion to commit ourselves to desires altogether more stringent in their conditions of satisfaction than anything we would otherwise have any reason to endeavor to satisfy.[7]

Dennett's example:

> "I'd like some baked beans, please."
> "Yes Sir. How many?"[8]

He continues:

> I dwell on this because it has a parallel in the realm of belief, where our linguistic environment is forever forcing us to give - or concede - precise verbal expression to convictions that lack the hard edges verbalization endows them with. By concentrating on the *results* of this social force, while ignoring its distorting effect, one can be easily misled into thinking that it is obvious that beliefs and desires are like sentences stored in the head.[9]

He cautions that if the linguistic over-explicitness is ignored, distortions can create the mistaken impression that the antecedents to action are merely sentences in the head.

Having delivered his caveat, Dennett addresses a final doubt about the assumption that intentionality plays such an integral

role in an agent's behavior. He argues that the use of intention

has become so habitual that the doubt is unwarranted:

> Do people actually use this strategy? Yes, all the time.
> There may someday be other strategies for attributing
> belief and desire and for predicting behavior, but this
> is the only one we all know now. And when does it work?
> It works with people nearly all the time. Why would it
> not be a good idea to allow individual Oxford colleges to
> create and grant academic degrees whenever they saw fit?
> The answer is a long story, but very easy to generate.
> And there would be widespread agreement about the major
> points. We have no difficulty thinking of the reasons
> people would have for acting in such ways as to give
> others reasons for acting in such ways as to give others
> reasons for...creating a circumstance we would not want.
> Our use of the intentional strategy is so habitual and
> effortless that the role it plays in shaping our
> expectations about people is easily overlooked.[10]

In sum, Dennett argues that since the intentional stance is such a

popular and practical everyday shortcut, it is reasonable to assume

that folk psychology is an adequate explanation of the causes of

action.

He concludes by considering why the intentional stance works

at all. First, he holds that the intentional stance works because

it operates on the premise that intentionality is real. Dennett

claims that it is real because of its natural advantage. The

process of natural selection, which selects for intentionally

conditioned agents, denotes intentionality as a survival feature.

Thus, intentionality is real because it is a consequence of a

natural reality.

Secondly, he claims that the intentional stance works because

of the physical structure of agents themselves. Having said this,

Dennett is the first to accept that enough is not yet understood

about the physical workings of agents to verify this claim. But,

he points out that while the definitive physical explanation of intentionality awaits final confirmation, two theories currently predominate.

The first candidate is behaviorism. However, Dennett dismisses it quickly. He claims that beliefs and desires are internal attributes, and the observations of behaviorism shed little light upon an agent's internal workings.

The second candidate, the language of thought, captures more of Dennett's support. The language of thought states that the actual physical processes working in the human brain will prove to be some form of natural language based symbol manipulation. He has some confidence in the language of thought because language is the only model (so far) of how the brain might organize and manipulate the masses of information it continually receives from the body's perceptions.

With this, Dennett's instrumental revision of folk psychology's account of causal action is complete. The revised folk psychology is founded upon the premise that intentionality is a useful explanation, since its propositional beliefs and desires produce reliable predictions of agents' actions. If Dennett's amendments prove durable, the intentional stance may be an auspicious revision of folk psychology's account of the causes of action.

CHAPTER SIX NOTES

1) Daniel Dennett, *The Intentional Stance* (Cambridge, MA.: MIT

    Press/Bradford Books, 1987) 14.

2) Dennett 14.

3) Dennett 15.

4) Dennett 15.

5) Dennett 16.

6) Dennett 17.

7) Dennett 20.

8) Dennett 20.

9) Dennett 21.

10) Dennett 21-22.

CHAPTER SEVEN

Functionalism Faulted

If Dennett's revision of folk psychology is successful, it would revitalize the claim that folk psychological propositions such as belief and desire have a place within causal action theory. However, there is some doubt about just how successful the intentional stance might really be. After all, the intentional stance pivots upon Dennett's claim that it provides a useful explanation of action when alternative explanations are unavailable. But how useful is his theory when it proposes to simply assume the potentially misleading premise that beliefs and desires are relevant to the explication of an agent's action?

More precisely, how reasonable is the suggestion that any subject of a behavioral investigation can be assumed to experience belief and desire? Surely there are some subjects that lack the necessary pre-requisites to intentionality. Yet, he appears to places no limits on the range of subjects for which the intentional strategy may be claimed a suitable approach. Therefore, it appears unavoidable that the intentional stance will license belief and desire based explanations of subjects not normally associated with intentionality.

For example, it is possible for Dennett's intentional stance to devise a folk psychological explanation of action for subjects such as plants, lightning, thermostats, and lecterns. As he admits:

> The strategy even works for plants. In a locale with late spring storms, you should plant apple varieties that

are particularly *cautious* about *concluding* that it is
spring - which is when they want to blossom, of course.
It even works for such inanimate and apparently
undesigned phenomena as lightning. An electrician once
explained to me how he worked out how to protect my
underground water pump from lightning damage: lightning,
he said, always wants to find the best way to ground, but
sometimes it gets tricked into taking second-best paths.
You can protect the pump by making another, better path
more *obvious* to the lightning.[1]

These examples of belief and desire being ascribed to clearly

inanimate objects undermine his argument that the intentional

stance is a reasonable revision of folk psychology.

Dennett's lectern example is a more detailed illustration of

the intentional stance contriving a dubious intentional explanation

of action. Were it needed to be explained why a lectern remains

sitting in the lecture hall, adhering to the guidelines set out by

the intentional stance, it would be correct to initially treat the

lectern as a rational agent. That is, it would be supposed that

the lectern sits in the lecture hall because this action is

consistent with its belief and desires. The supposition is that it

desires to stay in the hall, and remains there because it believes

doing so will satisfy this desire:

> For instance, it seems that the lectern in this lecture
> room can be construed as an intentional system, fully
> rational, believing that it is currently located at the
> center of the civilized world (as some of you may also
> think), and desiring above all else to remain at that
> center. What should such a rational agent so equipped
> with belief and desire do? Stay put, clearly, which is
> just what the lectern does.[2]

Given that the intentional stance permits such an errant

explanation of a lectern's "action," the doubts surrounding the

instrumental approach are only fueled.

How serious a charge is the intentional stance's vulnerability to indiscriminate application? Obviously, Dennett denies the charge. He argues that in the cases such as the lectern, the intentional stance is not the appropriate explanatory tool. It is inappropriate because it uncovers no new knowledge about lectern behavior that was not evident prior to the adoption of the intentional strategy:

> For one thing, the strategy does not recommend itself in this case, for we get no predictive power from it that we did not antecedently have. We already knew what the lectern was going to do - namely nothing - and tailored the beliefs and desires to fit in a quite unprincipled way. In the case of people or animals or computers, however, the situation is different. In these cases often the only strategy that is at all practical is the intentional strategy: it gives us predictive power we can get by no other method.[3]

Dennett's response is that the intentional stance is more legitimate when applied to explaining the actions of agents more complex than mere lecterns, plants, and thermostats.

He claims that complex agents legitimize the intentional stance since it is often the only practical method of explaining the actions in these cases. The causal complexity antecedent to some agent's actions often prohibits an inquirer from cataloguing sufficient causes to produce a reliable physical explanation. According to Dennett, the intentional stance then becomes a reasonable alternative.

However, Laplacean difficulties are contingent upon each investigator's fathoming of the causal history of a particular agent. Some investigators will be better than others at piecing together the causal antecedents, making the application of the

intentional stance inconsistent. But Dennett denies that this lead
to inconsistencies within the intentional stance's explanations.
The stance does not unhinge if it's application varies with every
user since the results are subject to objective scrutiny:

> The decision to adopt the intentional stance is free, but
> the facts about the success or failure of the stance,
> were one to adopt it, are perfectly objective.[4]

Still, an objective testing of the intentional stance does not rule
out indiscretions like the stance's explanation of lectern's
actions. As Dennett himself concedes, postulating certain belief
and desire propositions for a lectern yields an objectively
successful prediction. Since the intentional interpretation of the
lectern would meet Dennett's mandatory objective correctness,
Dennett's stance could not eradicate the intentional explanation of
the lectern's actions.

Thus arrives Dennett's second response to the intentional
stance's indiscrimination. He claims that there has only been the
admission that it is possible to view a lectern as a rational agent
in order to underline the intentional stance's logical approach to
belief and desire. But in reality, as a system grows in
complexity, such a possibility becomes less likely:

> The perverse claim remains: *all there is* to being a true
> believer is being a system whose behavior is reliably
> predictable via the intentional strategy, and hence *all
> there is* to really and truly believing that p (for any
> proposition p) is being an intentional system for which
> p occurs as a belief in the best (most predictive)
> interpretation. But once we turn our attention to the
> truly interesting and versatile intentional systems, we
> see that this apparently shallow and instrumentalistic
> criterion of belief puts a severe constraint on the

internal constitution of a genuine believer, and thus
yields a robust version of belief after all.[5]

For Dennett, complexity incurs limitations that curtail indiscrimi-
nate applications of his approach.

He uses a thermostat as an example of how the complexity of a
system actually places severe constraints upon the postulation of
belief and desire. Were a thermostat to be considered a realistic
agent, it would have to be endowed with features which enable it to
have an intentional capacity. For instance, it might have a lens
through which it could watch the reactions of the occupants of its
controlled environment to variations in the ambient temperature.
Perhaps, it might even have a voice box so that the thermostat
could ask the occupants about how warm or cold they were.

It becomes evident that in making the thermostat a more
realistic agent, one is also adding to the specialization of the
machine. The thermostat is no longer a bi-metalic spring that
simply contracts and expands. It now has the necessary sophistica-
tion to see and speak, and it uses these abilities to appropriately
regulate the temperature. Dennett argues that one consequence of
the thermostat's increased complexity is that it can only operate
within a very limited assortment of environments.

The sophisticated thermostat would no longer be of all-purpose
use. It could not regulate an oven or a boiler since ovens and
boilers do not speak, and their reactions to temperature variations
are difficult to visually gauge. The complex thermostat would be
relegated to the specialized function of moderating the temperature
surrounding subjects which spoke and reacted visibly. These

requirements subsequently limit the type of belief and desire that could be potentially ascribed to the sophisticated thermostat:

> At that point we say this device (or animal or person) has beliefs about heat and about this very room, and so forth, not because of the system's actual location in, and operations on, the world, but because we cannot imagine another niche in which it could be placed where it would work.[6]

Additionally, the limitation on the belief and desire ascriptions makes these ascriptions more creditable.

So briefly, Dennett's argument is that as a system grows more complex it necessarily becomes more specialized. This makes any belief and desire ascription quite specific - a specificity which lends legitimacy to those beliefs and desires ascribed. The increasing legitimacy of specific belief and desire ascriptions applies to all systems including thermostats, computers, and human beings.

However, Dennett's argument that complex systems can be attributed specific beliefs and desires simply does not work. The problem is that sophistication does not entail specialization. It cannot be maintained that complex systems always perform specialized functions. Ironically, for Dennett, the best example of a sophisticated system which is extremely diverse in its functioning is a human being. It is possible to define a myriad of potential functions for the highly complex, yet extremely adaptable human animal. Thus, the belief and desire ascribed by the intentional stance could not be confirmed by reference to any specialized human function.

Therefore, the problem of indiscriminately ascribing functional beliefs and desires is not overcome. The general problem of simply assuming the presence of beliefs and desires without actually understanding the causation which underlies a behavioral subject remains. Devoid of causal corroboration, problems with indiscriminate ascription of functional belief and desire are a fact of life. As a result, Dennett's instrumental intentional stance appears unable to provide a substantial revision of folk psychologically based causal action theory.

CHAPTER SEVEN NOTES

1) Daniel Dennett, *The Intentional Stance* (Cambridge, MA.: MIT Press/Bradford Books, 1987) 22.

2) Dennett 23.

3) Dennett 23.

4) Dennett 24.

5) Dennett 29.

6) Dennett 31.

CHAPTER EIGHT

The Alternatives

Reviewing the conclusions so far, the conventional causal theories of action like Davidson's and McGinn's rest upon the supposition that belief and desire cause action. However, as eliminativists such as Churchland point out, a reliance upon folk psychological constructs like belief and desire is problematic. There are doubts concerning folk psychology's a priori validity; difficulties regarding the type of reduction or identification necessary; and an unrealistic suppression of misrepresentation. Furthermore, folk psychology provides an inadequate account of actions involving learning, imagination, and memory. As a result it appears stagnant, especially as it becomes increasingly alienated by modern cognitive science.

Churchland consequently demands the elimination of folk psychology from causal action theory. But, the eliminative demand is criticized by Foss, Horgan, and Woodward as being too extreme. They concede that folk psychology is flawed, but argue that its faults do not logically entail its elimination. They counter that Churchland's assertion that cognitive science will provide a better explanation is merely a leap of faith.

Yet, Dennett grants cognitive science a legitimate place in action theory. He makes many references to cognitive research in his own arguments, and is undaunted by cognitive science's threat to the traditional folk psychological concepts. Unfortunately, Dennett's compromise is undone by its lack of discrimination: his

instrumental functionalism ascribes belief and desire to subjects not normally attributed such capacities.

It therefore becomes evident that the standard folk psychologically based identity or functionalist theories will not sustain a causal explanation of action. Regardless of the objections of detractors that cognitive science is a leap of faith, it appears inevitable that a revision of causal action theory will have to reflect an increasingly eliminative approach. What are the ramifications of this development?

To begin with, it should be emphasized that an eliminative position includes more than a basic skepticism of folk psychology. Also bedded within its tenets is the conviction that the identity theorists, reductionists, and functionalists are, "missing the metaphysical boat."[1] For the eliminativists, it is obvious that cognitive science is making significant advances in understanding how the brain causes behavior like rational action. Of course, such progress often raises more questions than answers, but at least it progresses beyond the stubborn problems which haunt folk psychology. Folk psychologically based theories of action have lost their intuitive appeal not just because of internal problems, but also because cognitive science is beginning to offer worthwhile alternatives.

As William Allman's book *The Apprentices of Wonder* (1989) outlines, these alternatives have had a long gestation. Since the time of Leibniz and Babbage, early science originally attempted to explain action's causes in purely theoretical terms. The early theories used the contemporary a priori scientific truths to build

logical explanations of behavior. However, as recent philosophers like Quine have made clear, the explanation of behaviors like action will not be achieved solely by referring to a priori truths and logical reasonings. Thus, it is a research emphasis which characterizes modern cognitive science.

From a research standpoint, cognition has always been a good starting point for the study of behavior like action. The conventional scientific view of cognition starts with the idea that representing is a matter of logical symbol computation. The symbols have no intrinsic meaning in themselves, but display various derived values:

> In much the same way the letter e represents "energy" in the equation e=mc2 in physics, in the traditional cognitive science approach "ice cream" could be represented as a symbol - ic, say.... These mental symbols are processed in the mind by the rules of formal logic.[2]

The computational explanation assumes that the representation which causes things like action involves manipulating basic symbols according to a set of universally applicable logical rules.

Russell and Whitehead substantiated this approach with their proposal that the laws of natural numbers in mathematics and the formal laws of thought were actually one and the same system. This school, logical empiricism, holds that (as with all things) the workings of the brain and the causes of action can be explained in terms of symbolic propositions. Such an explanation starts with simple axioms, builds these into propositions, and then composes logical explanatory syllogisms. As such, logical empiricism and

its propositional attitudes is the scientific equivalent of philosophical action theories like Davidson's and McGinn's.

With the advent of innovations like cybernetics and Turing machines, the opportunity arose for a wholly mechanistic explanation of cognition and action's causes. The propositional reasoning which embodies the empiricist's concept of action's causes became mechanically realizable, leading to the conclusion that the brain is a mechanical processor very similar to a digital computer:

> Because such a system of universal symbols and rules can also be run as a computer program, machines can be used to mimic the processes that presumably go on in our minds. Not only could scientists test their models of mental operations on a computer but they could also incorporate such models in a machine to give it intelligence. To most scientists, this model of the mind is very appealing, and for the last thirty years the main thrust of cognitive science has been to discover the mind's "program."[3]

For the logical empiricists, the brain is the hardware and behaviors like action are simply a consequence of the software.

Nonetheless, it is now widely accepted that the early hope of the digital computer providing a useful model of cognition has not materialized. Even after twenty five years of research, digital computers still cannot attain the cognitive proficiency of 3 year old.[4] This is due to (as Churchland addressed earlier) digital computers proving too demanding on available resources. Hence, artificial intelligence's computational, propositional task forces have been permanently disbanded.[5]

It is probably reasonable to hold that the events antecedent to action involves some logical symbol manipulation, but according to many cognitive researchers this is more of a sideline. The

experimental evidence suggests that the laws of cognition are not
the laws of logic:

> The traditional model of the mind is based on the
> principle that, at its roots, the mind is an engine of
> logic, and if the mind does not always behave logically,
> the problem lies in the brain. But this model doesn't
> account for the way people actually think. "All the
> experimental evidence points to the fact that people
> aren't rational," says Rumelhart. "That is a simple fact
> that the rationalists refuse to accept."[6]

Perhaps this explains why logical puzzles are such a challenge -
they strike at the brain's Achilles heal: logical reasoning.

As a result, the most recent research has turned to another
computer spawned model of cognition - parallel processing. It does
away with the assumption that computational representation
necessarily precedes behavior. That is, sophisticated behaviors
like action emanate from layers of increasingly sophisticated
abilities. In a terrestrial system these layers might be (from top
to bottom) object avoidance, wondering, wexploring and map making.
The advantage of this is that parallel processors can reach
conclusions faster and more efficiently than the traditional
machines which have to sequentially identify an environment,
represent it, rationally plan, and then execute a task. In effect,
parallel processing does more with less, rendering the
computational digital computer model of cognition obsolete.

Perhaps the best developed parallel-type explanation of
cognition is connectionism. In this theory, it is the brain's vast
array of billions of interconnected cellular switches that makes it
cognitively powerful. Like the sound waves which emerge from
billions of molecules, the cognitive causes of action evolve from

the brain's billions of neurons. Therefore, in a connectionist's explanation of action's causes, the events antecedent to action are not part of a computational predicate calculus. The events which causes action are non-sequentially layered interactions of activity within the neural network of the brain.

Avoiding the highly technical aspects of cognitive science, Robert Cummins' book *Meaning and Mental Representation* (1989) provides a detailed account of how connectionist networks operate. The networks consist of nodes and links which control the transmission of electrical activity. Each node plays a very limited role, being simply a reservoir of activity, though each is interconnected by many links which act as communicators:

> At a given moment, a node is characterized by a variable representing its threshold. When the activation of a node exceeds its threshold, activation is propagated from that node to others along whatever links happen to exist. Links are weighted, the weights determined by the (relative) quantity of activation they carry. Connection weights may be positive or negative: that is, a given node can *inhibit* the nodes to which it is connected by a negatively weighted arc by subtracting from their activation. The state of the network at a time is given by pattern of activation at that time (i.e., the activation level at each node at that ime) and by the connection strengths.[7]

These patterns of activation follow a geometric (as opposed to linear) causal progression, making them surprisingly intricate and complex. Yet they can be quickly generated, easily manipulated, and efficiently stored. As such, the connectionists claim that if real-world information can be translated into these electrical states, the networks would meet the constraints behavior like action imposes upon representation. These patterns appear well

suited to the enormous job of information retrieval and manipula-
tion which precedes an act. Their suitability is further confirmed
by the ability of these complex ensembles to "learn" using a
process called back-propagation.

Back-propagative learning is a result of the architecture of
connectionist networks. The design of a network is such that when
an "input node" is stimulated by an influence external to the
network, its excitement is transmitted along the links to huge
numbers of neighboring nodes throughout the system:

> An input is given to the network by providing some
> specified level of activation to the nodes designated as
> input nodes. Activation then spreads throughout the
> network in a way determined by the antecedently existing
> pattern of activation and by the strengths of the
> connecting links. The *dynamics* of the system is given by
> differential equations that determine activation spread
> as a function of time. The "output" of the system is the
> state of that portion of the network designated as the
> output nodes when the entire network "settles down" into
> a steady state.[8]

It has been found that specific external influences will yield
specific network patterns. The size and shape of a particular
pattern can be retained by permanently altering the transmission
links that connect the nodes. Again, this is achieved by adjusting
or reweighting the carrying capacity of the links.

Learning itself starts at the final "output nodes" of the
network. The output nodes transmit activity back through the
network in response to the effects a pattern's generation is having
upon the external influence. If the current pattern is having the
correct effect, it is left unaltered. But, if the external effects

are unacceptable, the output nodes send responses back through the network that reweight the links and alter the pattern generation.

In this way, the patterning in the network is sensitive to the effects it is causing, and can adjust and correct itself as required. Self correction is the root of the connectionist's claim that these networks are learning. Some argue that connectionist models of learning are psychologically unrealistic9 - a point to which even the connectionists concede. But, according to network proponents, if information about the outside world can be converted into reliable patterns which are quickly generated, easily manipulated, and efficiently stored, in addition to being primitively self-correcting, then this system exhibits at least the fundamentals of a useful model of how the brain produces complex behavior.

What really encourages the connectionist's optimism is that the patterns of brain activity that have been observed in the cerebral hemispheres appear to work in the same way as the patterns in the connectionist networks. In the brain the patterns are regularly generated in response to external stimuli that arrive from the body's sensory apparatus at the input neurons. The resultant brain patterns which emanate from the neural network constitute the brain's response to the external influences. It is inferred that the ensembles then orchestrate the body's reaction to the external stimuli.

Most versions of connectionism, then, attempt to account for intelligent behavior without any reference to folk psychological constructs like belief and desire. Irrespective of which specific connectionist theory ultimately wins out, it is assumed that the

events which lead to action are neural patterns orchestrated by the brain's parallel processing cellular network. As an example, Cummins suggests one somewhat extreme version of connectionism:

> The requirement that cognitive behavior be conceived as behavior under an interpretation is dropped, and hence the implication that cognitive behavior is to be explained as the computation of representations is dropped. Cognitive performance, on this conception, becomes purely ostensive and comparative: To the extent that a system performs comparably to a normal adult human doing distinctively normal adult tasks, that system is "intelligent," or "cognitive." There is just nothing else to be said. It is behaviorism of the old school, with the following slight revision: Neurophysiology, which is where behaviorists thought talk of internal states belonged, is replaced by something slightly more abstract - a mathematical model of neural dynamics. The result is pretty much what radical, behaviorism envisaged all along: We have a psychology discovering "effects" - i.e., non-semantically characterized regularities in human and animal performance - and a nonpsychological explanation of the underlying mechanism.[10]

In this instance an action's explanation would consist of a comparative identification and a mathematical equation tracing its neurological causation.

Of course, such connectionism remains an experimental alternative to the more traditional explanations of action. Some have simply derided it as the latest scientific fad - like robotics or perceptrons. Yet the greatest strides in cognitive science have been in uncovering the basic mechanisms that operate our motor controls and the sensory apparatus that allows us to monitor these skills. From this research, it is amply demonstrated that belief and desire based intermediaries are not needed as a part of these abilities. This naturally raises the question of whether propositions like belief and desire are really necessary at the higher levels of agent operation - levels at which action is produced.[11]

Connectionism would seem to suggest that they are unnecessary. Such a response is hardly definitive. But, the final answer to this question will greatly depend on how folk psychologically inspired theorists like Davidson, McGinn, and Dennett anticipate the challenge. At present they merely attempt to make sense of folk psychological constructs more or less as they are popularly understood. But, the eliminativists recognize that it is no longer sufficient to simply rearrange in such a way that criticism is temporarily restrained. In the face of cognitive science's surging influence, the viability of the folk concepts must be rejustified. There may even be a continuum between reductionism and outright elimination12, but it will take an eliminative approach to establish this. It is possible that the relevence of belief and desire may be reconstituted by somehow tying them into the eliminativist's scientific agenda (see Coval 1992).

In spite of the objections by Foss, Horgan and Woodward, an audit of causal action theory grows all the more appropriate as burgeoning evidence accumulates on the side of the eliminative position. As cognitive science slowly uncovers the precise neurological causes of action, it charts an extremely rigorous map of the causation antecedent to action. The rigor of this causal blueprint appears to exceed anything currently offered by the identity or functional folk psychological action theories.

At the present time in philosophy, there is a widespread consensus that actions are caused. However, the specific causes which lead to action remain in dispute, as the three preeminent causal action theories each offer a different explanation of action's causes. Nevertheless, it would appear that from among the

identity, functional, and eliminative alternatives, it is the scientifically fostered eliminative approach which will prove most influential.

CHAPTER EIGHT NOTES

1) Paul Churchland, *A Neurocomputational Perspective* (Cambridge, MA: MIT Press/Bradford Books, 1989) 9.

2) William Allman, *Apprentices of Wonder: Inside the Neural Network Revolution* (New York: Bantam Books, 1989) 41-42.

3) Allman 42.

4) Allman 10.

5) Daniel Dennett, *The Intentional Stance* (Cambridge, MA: MIT Press/Bradford Books, 1987) 227-229.

6) Allman 192.

7) Robert Cummins, *Meaning and Mental Representation* (Cambridge, MA.: MIT Press/Bradford Books, 1991) 146.

8) Cummins 146

9) Dennett 231.

10) Cummins 155.

11) Dennett 228.

12) Paul Churchland, *Matter and Consciousness* (Cambridge, MA: MIT Press/Bradford Books, 1984) 48.

BIBLIOGRAPHY

Angel, L. (1989). *How to Build a Conscious Machine*. Boulder, CO: Westview Press.

Allman, W. (1989). *Apprentices of Wonder: Inside the Neural Network Revolution*. New York: Bantam Books.

Boden, M. (ed.) (1990). *The Philosophy of Artificial Intelligence*. Oxford: Oxford University Press.

Brand, M. (1984). *Intending and Acting*. Cambridge, MA.: MIT Press.

Bratman, M. (1985). "Davidson on Intentional Action," in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. LePore, E., McLaughlin, B. (eds). Oxford: Basil Blackwell Ltd.

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

--- (1979). "Eliminative Materialism and the Propositional Attitudes," *Journal of Philosophy*, 78: 67-90.

--- (1984). *Matter and Consciousness: A Contemporary Introduction to the Philosophy of Mind.* Cambridge, MA.: The MIT Press/Bradford Books.

--- (1985). "On the Speculative Nature of Our Self Conception: A Reply to Some Criticisms." *Canadian Journal of Philosophy,* Supp. Vol 11, 157-173.

--- (1989). *A Neurocomputational Perspective.* Cambridge, MA.: The MIT Press/Bradford Books.

Churchland, P.S. (1986). *Neurophilosophy: Toward a Unified Theory of Mind/Brain.* Cambridge, MA.: The MIT Press/Bradford Books.

--- (1986). "Replies to Comments." *Inquiry,* 29, 241-272.

Costa, M. (1987). "Causal Theories of Action." *Canadian Journal of Philosophy,* Vol. 17, No. 4, 831-854.

Coval, S.C., Smith, J.C. (1986). *Law and its Presuppositions: Actions, Agents and Rules.* London: Routledge & Kegan Paul.

Coval, S.C., Campbell, P.G. (1992). *Agency in Action.* Dordrecht: Kluwer Academic Publishers.

Cummins, R. (1991). *Meaning and Mental Representation.* Cambridge, MA.: MIT Press/Bradford Books.

Davidson, D. (1963-78). *Essays on Actions and Events*. Oxford: Clarendon Press.

Dennett, D. (1969). *Content and Consciousness*. London: Routledge & Kegan Paul.

--- (1978). *Brainstorms: Philosophical Essays on Mind and Psychology*. Sussex: Harvester Press.

--- (1984). *Elbow Room*. Cambridge, MA.: MIT Press/Bradford Books.

--- (1987). *The Intentional Stance*. Cambridge, MA.: MIT Press/Bradford Books.

Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, MA.: MIT Press.

--- (1988). *Explaining Behavior*. Cambridge, MA.: MIT Press/Bradford Books.

Duran, J. (1988). "Reductionism and the Naturalization of Epistemology." *Dialectica,* Vol. 42, No. 4, 295-306.

Fodor, J. (1976). *The Language of Thought*. Sussex: Harvester Press.

--- (1981b). *Representations*. Cambridge, MA.: MIT Press.

--- (1987). *Psychosemantics*. Cambridge, MA.: MIT Press.

Fodor, J., Pylyshyn, Z. (1988). "Connectionism and Cognitive Architecture." *Cognition* 28,: 3-71.

Foss, J. (1983). "A Materialist's Misgivings About Eliminative Materialism." *Canadian Journal of Philosophy*, Supp. Vol. 11, 105-133.

--- (1987). "Is The Mind-Body Problem Empirical?" *Canadian Journal of Philosophy*, Vol. 17, No. 3, 505-532.

Ginet, C. (1990). *On Action*. Cambridge: Cambridge University Press.

Grice, P. (1975). "Method in Philosophical Psychology (From the Banal to the Bizarre)" in *Proceedings and Addresses of the American Philosophic Asssociation*, XLVIII: 23-53.

Hobbes, T. (1651). *Leviathan*.

Honderich, T., Burnyeat, M. (eds.) (1979). *Philosophy As It Is*. Middlesex: Pelican Books.

Horgan, T., Tye, M. (1985). "Against the Token Identity Theory," in *Actions and Events: Perspectives on the Philosophy of*

Donald Davidson, LePore, E., McLaughlin, B. (eds.). Oxford: Basil Blackwell Ltd.

Horgan, T., Woodward, J. (1985). "Folk Psychology Is Here To Stay." *The Philosophical Review,* XCIV, No. 2, 197-226.

Hornsby, J. (1980). *Actions.* London: Routledge & Kegan Paul.

Hume, D. (1739). *A Treatise of Human Nature.*

--- (1748). *Enquiry Concerning Human Understanding.*

Kurzweil, R. (1990). *The Age of Intelligent Machines.* Cambridge, MA.: MIT Press.

LePore, E., McLaughlin, B. (eds.) (1985). *Actions and Events: Perspectives on the Philosophy of Donald Davidson.* Oxford: Basil Blackwell Ltd.

Locke, D. (1982). "Beliefs, Desires and Reasons for Action." *American Philosophical Quarterly,* Vol. 19, No. 3, 241-249.

McGinn, C. (1979). "Action and its Explanation," in Bolton, N. (ed.) (1979). *Philosophical Problems in Psychology,* pp. 20-42 London: Methuen & Co.

--- (1982). *The Character of Mind*. Oxford: Oxford University Press.

Nagel, T. (1986). *The View From Nowhere*. Oxford: Oxford University Press.

Putnam, H. (1988). *Representation and Reality*. Cambridge, MA.: MIT Press.

Quine, W.V. (1951). "Two Dogmas of Empiricism," in *The Philosophical Review*, 60, 20-43. Reprinted in Quine, W.V. (1963). *From A logical Point of View*. Cambridge, MA.: Harvard University Press.

--- (1990). *Pursuit of Truth*. Cambridge, MA.: Harvard University Press.

Rorty, R. (1965). "Mind-Body Identity, Privacy, and Catagories." *Review of Metaphysics,* Vol. 19. Reprinted in Rosenthal, D.M. (ed.) (1971). *Materialism and the Mind-Body Problem*. Englewood Cliffs, NJ: Prentice-Hall.

--- (1970). "In Defense of Eliminative Materialism." *Review of Metaphysics,* Vol. 24. Reprinted in Rosenthal, D.M. (ed.) (1970). *Materialism and the Mind-Body Problem*. Englewood, NJ: Prentice-Hall.

Rumelhart, D., McClelland, J., and the PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1 and 2. Cambridge, MA.: MIT Press/Bradford Books.

Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.

Silvers, S. (ed.) (1989). *Representation: Readings in the Philosophy of Mental Representation*. Dordrecht: Kluwer Academic Publishers.

Smart, J.J.C. (1963). *Philosophy and Scientific Realism*. London: Routledge & Kegan Paul.

Stich, S. (1983). *From Folk Psychology to Cognitive Science*. Cambridge, MA.: MIT Press/Bradford Books.

Vermazen, B., Hintikka, M. (eds.) (1985). *Essays on Davidson*. Oxford: Clarendon Press.

Wilson, G. (1985). "Davidson on Intentional Action," in *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, LePore, E., McLaughlin, B. (eds.). Oxford: Basil Blackwell Ltd.