

## REASONS, CAUSES, AND ELIMINATIVE MATERIALISM

REASONS, CAUSES, AND ELIMINATIVE MATERIALISM

By

NEIL CAMPBELL, B.A.

A Thesis

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Arts

McMaster University

(c) Copyright by Neil Campbell, February 1993

MASTER OF ARTS (1993)  
(Philosophy)

McMASTER UNIVERSITY  
Hamilton, Ontario

TITLE: Reasons, Causes, and Eliminative Materialism

AUTHOR: Neil Campbell, B.A. (University of Toronto)

SUPERVISOR: Professor Barry Allen

NUMBER OF PAGES: vi, 90

## ABSTRACT

This thesis is a criticism of the theory in the philosophy of mind known as "eliminative materialism". While this theory has been advocated by a number of philosophers, none have pressed its thesis harder than Paul Churchland. Consequently, Churchland's work in this area has been the focus of current debates in the philosophy of mind.

Although several philosophers have developed significant objections to Churchland's formulation of the thesis, Churchland has always been ready with a convincing reply. For this reason I propose to attack Churchland on as fundamental a level as seems possible, namely, by questioning his presupposition that folk-psychological explanation is a species of causal explanation. Without this presupposition there is little reason to expect Churchland's proposed theoretical elimination of folk psychology by neuroscience; for folk-psychological explanation (principally, reason-giving) must be characterized as a species of causal explanation if it is to be replaced by a better causal-explanatory theory of behaviour.

I argue that reason-giving explanations are not a species of causal explanation. By undermining one of the central presuppositions of eliminative materialism, I hope to

cast sufficient doubt on Churchland's thesis.

## ACKNOWLEDGEMENTS

I would like to thank Dr. Barry Allen and Dr. Jill LeBlanc for their helpful comments and guidance.

## TABLE OF CONTENTS

Introduction .....	1-7
Chapter I: Eliminate Materialism and the Causal Thesis .....	8-30
1. Eliminate Materialism .....	8
2. Reasons and Causes .....	17
3. Conclusions .....	28
Chapter II: Reasons as Epiphenomena .....	31-53
1. Anomalous Monism and Epiphenomenalism .....	32
2. Hess on Reasons and Causes .....	42
3. Conclusions .....	51
Chapter III: The Point of View Distinction .....	54-82
1. Reasons, Causes, and Cartesian Dualism .....	56
2. The Point of View Distinction .....	61
3. Thornton, Dennett, and Rationality .....	65
4. Two Problems: Self-Understanding and the Unity of Science .....	70
5. Conclusions .....	80
References .....	83-88
Works Cited .....	89-90

## Introduction

In recent philosophy a distinction between reasons and causes has been urged by a number of philosophers. In particular, they say that to cite a man's reasons for his intentional acts is not to give their causes .... If there is such a distinction, many philosophers seem to have ignored it. For example, the classical reconciliation of causal determinism with freedom of the will depends on the claim that acts of will cause intentional actions and are in turn caused by one's motives, desires, and beliefs; and motives, desires, and beliefs are mentioned in giving a man's reasons for acting.<sup>1</sup>

Since Donald Davidson's article "Actions, Reasons, and Causes," the conception of reasons as causes of action has become (as D.M. Armstrong has put it) "respectable" again.<sup>2</sup> The debate on reasons and causes has, to my mind, developed into a more interesting and significant one than ever before since the development of the thesis of eliminative materialism, particularly as it is formulated by Paul Churchland. This is because eliminative materialism includes the thesis that reasons are causes (henceforth "the causal thesis") as a tacit assumption which, as Kathleen Wilkes has pointed out, "needs but never gets defense."<sup>3</sup> What I propose to show is that eliminative materialism does indeed require the thesis that reasons are causes, but that this thesis cannot be supported. I will now briefly outline my approach to this problem.

In Chapter 1 I lay the groundwork for my discussion. This involves a fairly detailed account of Churchland's



eliminative materialism, its relation to the causal thesis, and two of the strongest arguments in favour of the view that reasons are causes.

Eliminative materialism is the radical thesis recently pressed by Paul Churchland that our common sense psychological framework is literally a theory and hence all of its entities, e.g., beliefs, desires, pains, etc., are theoretical entities posited to predict and explain human behaviour. According to Churchland, this theory, which he calls "folk psychology", provides a *mistaken* account of human action and so should be eliminated and replaced with another theory (neuroscience) which offers more accurate predictions of behaviour and offers greater explanatory power in terms of its account of human behaviour.

While many of Churchland's reasons for regarding beliefs and desires causally are tied to his model of the semantics of theoretical terms and his conception of the structure of theories generally, he characterizes them this way because he needs to. The thesis that folk psychology should be replaced by a better causal/explanatory theory requires the claim that folk-psychological explanation is a species of causal explanation, otherwise there would be little reason to expect a theoretical elimination. So the connection between eliminative materialism and the causal thesis is: Folk psychology can only be replaced by a *better* causal/explanatory theory if it is itself a

causal/explanatory theory of behaviour.

Once the thesis of eliminative materialism and its relation to the causal thesis are firmly established, I then explore two alternative accounts of the causal thesis. The first is Donald Davidson's, expressed in "Actions, Reasons, and Causes" and "Mental Events," and the second is Churchland's own argument, formulated in "The Logical Character of Action-Explanations."

Davidson's position is that reasons are causes because rationalizations explain actions by redescribing those actions in terms of their causes. Much of Davidson's elaboration and defense of this view (formulated within the context of his anomalous monism) involves a characterization of the relation between singular causal claims and causal laws. Davidson wants to regard reasons as causes yet at the same time insulate the mental from subsumption under strict deterministic laws. The key to his approach is to recognize that singular causal claims (for instance, "He started running because he believed that he was late") are not formulated in the terms in which the underlying laws are formulated. So mental events (having reasons) can cause actions without entailing any deterministic laws linking reason and action, because laws are not formulated in terms of "belief" or "desire"; i.e., "belief" and "desire" are not the terms in which laws are formally stated.

Once Davidson's position is described I turn to

Churchland's argument. Churchland tries to improve upon Davidson's approach by specifying the actual laws underlying action explanations. Churchland's approach, then, involves the development and defense of one such law, which he calls "L1". What Churchland proposes is that the capacity for a reason to explain an action is derived from the nomological character of the underlying law, L1.

Chapter 2 introduces and develops an objection to the causal thesis. I begin by offering a direct argument against Davidson's approach, and through a development of this objection also introduce problems for Churchland's argument.

As a point of departure I summarize and explain a debate from *Analysis* on Davidson's anomalous monism. The proposed attack on Davidson's thesis is that anomalous monism renders mental events (and that includes reasons, or having reasons) epiphenomenal (i.e., causally impotent). The point of contention is that a mental event can only cause a physical event by virtue of certain of its physical properties, but such properties do not rationalize actions, even though there is a token-identity between mental and physical. Tracing the debate, which runs primarily between Ted Honderich and Peter Smith, I conclude that Honderich's charge of epiphenomenalism is correct.

While this objection undermines Davidson's arguments for the causal thesis, it leaves Churchland's approach untouched. I therefore further develop Honderich's argument

with the aid of Peter Hess's account of action explanations, and show that depending upon whether one considers an event's mental or physical properties, one is tied to providing either a rationalization for the event or a causal explanation of the event. The idea is that the properties involved in a causal explanation perform the necessary causal work no matter how they are described, but the properties involved in rationalizations are *ascribed* to agents and hence, unlike physical properties, their capacity to explain events is dependent on a body of conventions. That is, mental properties cannot explain actions independently of conventions about propositional content (i.e., folk-psychological conventions) whereas the others (physical properties) can.

Since the differentiation between causal and reasoning explanations in Chapter 2 is characterized in terms of a differentiation of properties, it might seem that my arguments against the causal thesis require a commitment to some form of ontological dualism; i.e., it might seem that since the properties involved in causal explanations are characterized as "physical" properties, that mental properties are by comparison non-physical. The third chapter shows that this is not the case.

My starting point in Chapter 3 is to provide an alternative understanding of Descartes' model of explanation for human action. I start with Descartes because his account

of the person is paradigmatic of ontological dualism. What I argue is that Descartes' dualism of mind and body as it is represented in the Meditations and The Passions of the Soul need not imply an ontological dualism, but can in fact be seen to represent an explanatory dualism. I will suggest, then, that Descartes mistakenly posited a dualism of substances in his account of human action where there is in fact only a shift in models of explanation. Descartes takes the shift from causal explanation (which he employs in *res extensa*) to reason-giving (in *res cogitans*) to indicate a shift in ontology.

My discussion of Descartes is intended as a stepping-stone to introduce a point made by Mark Thornton. Recognizing, as Thornton did, that Descartes' dualism is simply the result of a mistaken attempt to combine two different models of explanation, I propose that we can likewise avoid an ontological confusion by recognizing that the difference between reasons and causes is to be accounted for in terms of an explanatory parallelism. Far from being the result of an ontological dualism, I suggest that the difference between causal and reason-giving explanations is the result of what Thornton calls a dualism of points of view.

According to Thornton, rationalizations are "first-person" explanations whereas causal explanations are "third-person". The basic idea here is that reason-giving

explanations are formulated from the first-person point of view of the subject. This is why rationalizations model themselves on the process of theoretical and practical reasoning engaged in by the agent, as opposed to focusing on publicly observable physical properties. Once Thornton's dualism of points of view is explained I then defend the resulting distinction between rationalization and causal explanation from several possible objections, draw the conclusions such a distinction has for the causal thesis, and explain its consequences for eliminative materialism.

## CHAPTER I ELIMINATIVE MATERIALISM AND THE CAUSAL THESIS

In this first chapter I will provide an account of Churchland's eliminative materialism and explain its relation to the causal thesis. This will consist primarily in explaining (1) how Churchland assimilates the semantics of folk psychology to the semantics of theories; (2) how he shows that folk-psychological explanation follows the same pattern as scientific explanation; and (3) how he shows that folk-psychological explanations involve causal generalizations. I will then outline Davidson's and Churchland's arguments in favour of the causal thesis.

### 1. Eliminative Materialism

Eliminative materialists doubt that there could (as was proposed by the identity theorists) be a neat one-to-one mapping or reduction of common sense (folk) psychological concepts to neurological concepts. The reason we should not expect a neat series of identity statements (or intertheoretic reduction) linking mental and physical states is that our framework of common sense psychological concepts

constitute a *mistaken* theory of behaviour:

As the eliminative materialist sees it, the one-to-one match-ups will not be found, and our common-sense [sic.] psychological framework will not enjoy an intertheoretic reduction, *because our common-sense psychological framework is a false and radically misleading conception of the causes of human behaviour and the nature of cognitive activity.*<sup>1</sup>

Consequently, with the development of neuroscience, we can expect that our common sense framework (folk psychology, henceforth "FP") will be eliminated rather than reduced.

Churchland appeals to several historical examples of successful elimination in order to lend his view some plausibility. In each case the ontology of the older theory is simply replaced by that of a newer, superior theory. One such example is the elimination of theories of heat built around the substance "caloric". In the Eighteenth and Nineteenth Centuries, it was believed that heat was a fluid substance called caloric. There developed a body of theory which explained the thermal behaviour of objects, such as heating, cooling, melting, etc., by appeal to the motion of this fluid. When the new "corpuscular/kinetic theory of matter and heat" was developed, it showed that heat is not a substance at all, but is simply the rapid motion of the molecules constituting the heated object itself. Given the greater explanatory power of the kinetic theory, and its ability to produce more accurate predictions of the behaviour of heated objects, caloric theory and its ontology were eliminated.<sup>2</sup>

So runs Churchland's example, which he repeats with



such concepts as phlogiston and witchcraft. In each case, we see a more primitive folk theory, ontology and all, replaced with a more respectably scientific theory with greater explanatory power. According to Churchland, this will be the case with FP. For, although more enduring and complex, FP is, after all, a folk theory akin to caloric theory, and will eventually be replaced by a mature neuroscience. To make a case for the folk-status of FP, Churchland appeals to a list of "explanatory, predictive, and manipulative failures."<sup>3</sup> Since FP cannot explain such phenomena as, for example, sleep, learning, memory, differences in intelligence, and mental illness, it should be replaced with a stronger theory that can.<sup>4</sup> Churchland attempts to illustrate the theoretical status of FP, first, by showing that folk-psychological explanation follows the same pattern as scientific explanation generally, and second, by showing that the terms of FP are defined in the same manner as theoretical terms. Hence he suggests that the semantics of FP are constructed in the same manner as the semantics of theories.

Churchland begins his argument by showing how the laws constituting a theory predict and explain events according to the deductive-nomological (D-N) model. On this model the explanation of an event is the conclusion of a valid deductive argument whose premises (explanans) contain particular facts and at least one law.<sup>5</sup> As Hempel has pointed out, the explanation shows how the explanandum (the fact to

be explained) is to be expected, given the relevant laws and particular circumstances.<sup>6</sup>

With regard to the semantics of theories, Churchland formulates what he calls "the network theory of meaning." The meaning of theoretical terms are not, in Churchland's view, formulated by means of single definitions (e.g., "the *electron* is the unit of electricity"), but holistically, by the way the terms are embedded in the network of theoretical principles constituting the theory in which they figure.<sup>7</sup> For instance, as Churchland puts it, "To fully understand the expression 'electric field' is to be familiar with the network of theoretical principles in which that expression appears. Collectively, they tell us what an electric field is and what it does."<sup>8</sup> Thus, the meanings of theoretical terms are "fixed by the set of laws/principles/generalizations in which they figure."<sup>9</sup>

Churchland's argument for the theoretical status of FP turns on his ability to assimilate both the explanatory function and the semantics of FP to those of theories generally, as they have been described above. He begins by citing several common sense folk-psychological generalizations:

- 1) Persons tend to feel pain at points of recent bodily damage.
- 2) Persons who feel a sudden sharp pain tend to wince.
- 3) Persons who want that *P*, and believe that *Q* would be sufficient to bring about *P*, and have no conflicting wants or preferred strategies, will try to bring it about that *Q*.<sup>10</sup>

According to Churchland, these and other generalizations or

laws collectively constitute FP. He argues that folk-psychological explanation follows the D-N model by illustrating how such laws play a role in the explanation of human behaviour. By showing that explanations involving such laws follow the D-N model, Churchland hopes to demonstrate the theoretical status of FP. Here is his example:

"Why did Michael wince slightly when he first sat down to the meeting?"

"Because he felt a sudden sharp pain."

"I see. And why did he feel pain?"

"Because he sat on the tack I placed on his chair."<sup>11</sup>

When we press this explanation, according to Churchland, we will uncover the first two background laws listed above. These two laws play a role in two deductive arguments, one on the heels of the other, all of which explains the event according to the D-N model.<sup>12</sup> Thus folk-psychological explanation follows the same model as scientific explanation generally.

Finally, Churchland draws on a structural analogy between FP and physics in order to show how the propositional attitudes derive their meaning in terms of the theoretical network thesis mentioned above, as well as to illustrate that the structure of FP is the same as the structure of paradigmatically physical theories. He compares a list of folk-psychological attitudes with numerical attitudes:

#### Propositional attitudes

...believes that  $P$   
 ...desires that  $P$   
 ...fears that  $P$

#### Numerical attitudes

...has a length  $m$  of  $n$   
 ...has a velocity  $m/s$  of  $n$   
 ...has a charge  $c$  of  $n$ <sup>13</sup>

Churchland identifies three similarities. First, in each case, a specific term is required to formulate a complete predicate; in the first list the term is a proposition, and in the second it is a number. Second:

just as the relations between numbers (for example, being twice as large as  $n$ ), can also characterize relations between numerical *attitudes* (for example, my weight is twice your weight); so do the relations between propositions (for example, logical consistency, entailment) also characterize the relations between propositional *attitudes* (for example, my belief is inconsistent with your belief).<sup>14</sup>

Third, and most importantly, the relations among certain kinds of propositional attitudes, like numerical attitudes, hold universally; that is, they are lawlike.<sup>15</sup> Although laws of physics exploit numerical rather than logical relations, its laws share the same structure as folk-psychological laws. As evidence, Churchland provides the following comparisons:

If  $x$  hopes that  $P$ , and  $x$  discovers that  $P$ , then  $x$  is pleased that  $P$ .

If  $x$  has a mass of  $M$ , and  $x$  suffers a net force of  $F$ , then  $x$  has an acceleration of  $F/M$ .

If  $x$  believes that  $P$ , and  $x$  believes that (if  $P$ , then  $Q$ ), then, barring confusion, distraction, and so on,  $x$  will believe that  $Q$ .

If  $x$  has a pressure of  $P$ , and  $x$  has a volume of  $V$ , and  $x$  has a mass of  $u$ , then, barring very high pressure or density,  $x$  has a temperature of  $PV/uR$ .<sup>16</sup>

The first conclusion Churchland draws from these comparisons is that:

since meaning arises from an item's place in a network of assumptions, and from the resulting conceptual role that item plays in the system's ongoing inferential economy, therefore our mental states can have the propositional contents they do because of nothing more than their intricate *relational* features.<sup>17</sup>

This means that terms like "the belief that  $p$ " or "the desire that  $q$ " derive their meaning in the same fashion as terms

like "electric field". For the relational features which Churchland takes to provide propositional content are exactly analogous to the network of principles in which a theoretical term is embedded. Churchland continues:

The second lesson concerns the very close structural analogies that obtain between the concepts and laws of folk psychology, and the concepts and laws of other theories. The emergence of these parallels coheres closely with the view ... that folk psychology is literally a theory.<sup>18</sup>

While Churchland's view of this thesis has remained virtually unchanged since Matter and Consciousness, he does offer one important alteration in a recent paper entitled "Folk Psychology and the Explanation of Human Behavior." He now thinks that the deductive-nomological model is inadequate both in the sciences and elsewhere. Churchland's reason for giving up the D-N model is described by him as follows: "My diagnosis of its failings ... locates the basic problem in its attempt to represent knowledge and understanding by sets of sentences. In this the framers of the D-N model were resting on the basic assumptions of folk psychology."<sup>19</sup> While Churchland does not elaborate on this, I take his meaning to be something like this. In order for some kind of scientific model of explanation to replace folk-psychological explanation, that scientific model cannot itself be formulated in folk-psychological terms. Because the D-N model formulates its explanations in terms of deductive arguments linking sentences, those explanations depend upon FP, for FP gives those sentences their meaning. Since the D-N model depends in this way on FP it is in no position to replace it.

Churchland needs a conception of scientific explanation which does not depend upon FP if scientific explanations are to replace folk-psychological explanations, otherwise the thesis of eliminative materialism would undermine itself.

One of the central claims about which Churchland has not changed his mind, however, is that the generalizations of FP are *causal* in nature. He recalls some of our previous examples:

- 1) A person who suffers severe bodily damage will feel pain.
- 2) A person who suffers a sudden sharp pain will wince.
- 3) A person who is angry will tend to be impatient.<sup>20</sup>

According to Churchland, all such generalizations fall into one of two categories. The first group (constituting the core of FP), consists of generalizations involving fully intentional concepts such as belief and desire, while the second group involves quasi-intentional concepts such as pain, fear and so on. (Churchland presumably calls generalizations of the first group "fully intentional" because they express propositional attitudes, and the others "quasi-intentional" because they express sensations and emotions. While this division of folk-psychological concepts may be questionable, we can grant Churchland this point as there is little riding upon this distinction.) While, in Churchland's view, those generalizations (such as 1 and 2 above) in the second category are "transparently" causal, generalizations in the first group are far less obviously so.<sup>21</sup> This is because generalizations of the first category,

particularly those involving belief and desire, are not as easily incorporated into the causal, stimulus-response schema characterizing the generalizations of the second group. Churchland acknowledges the apparent failure for belief and desire to fall under a clearly deterministic schema as the frequent point of attack against the theoretical view, but he insists that rationalization (formulated in terms of belief and desire) is nevertheless a species of causal explanation:

It won't do, then, to insist that the generalizations of folk psychology are on the whole nonempirical or noncausal in character. The bulk of them, and I mean thousands upon thousands of them, are transparently causal or nomological. The best one can hope to argue is that there is a central core of folk-psychological concepts whose explanatory role is somehow *discontinuous* with that of their fellows. The propositional attitudes, especially belief and desire, are perennial candidates for such a nonempirical role, for explanations in their terms typically display the explanandum event as "rational." What shall we say of explanations in terms of beliefs and desires? ... We should tell essentially the same causal/explanatory story ....<sup>22</sup>

The above quotation reveals that the causal thesis is central to Churchland's position. For the thesis that FP should be replaced by a *better* causal/explanatory theory such as neuroscience first requires the claim that the central part of folk-psychological explanation (reason-giving) is a species of causal explanation. In the following chapters I will argue against the causal thesis, thereby removing one of the central claims of eliminative materialism, and thus insulating FP from potential elimination by neuroscience. First, however, we need to be clear on what exactly the causal thesis is. In the remainder of this chapter I should like to outline what I take to be the causal thesis by

considering the views of Donald Davidson, as articulated in "Actions, Reasons, and Causes" and "Mental Events," and then Churchland's own paper "The Logical Character of Action-Explanations." These three papers embody the strongest arguments in favour of the causal thesis.

## 2. Reasons and Causes

Davidson first stated the causal thesis in his article "Actions, Reasons, and Causes." There he makes two claims, the most important of which for me is the second:

C1. *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.<sup>23</sup>

C2. A primary reason for an action is its cause.<sup>24</sup>

Davidson notes that from the characterization of a primary reason one can construct a syllogism from which the action follows as a rational thing to do. Thus a primary reason can (by appeal to the standards of rationality) *justify* an action. Since causal explanations do not justify the events they explain (in the sense of showing an action to be rational), it is tempting to conclude that reasons cannot be causes. However, according to Davidson, even if we grant that reasons, unlike causes, can justify actions, it does not follow that the explanation in which the reason figures is also not causal.<sup>25</sup>

Davidson considers a familiar Wittgensteinian line



according to which reasons provide explanations of events by redescribing them and fitting them into familiar patterns. This should not be misconstrued to mean that by fitting an event into a particular pattern we thereby understand *how* a reason explains an action:

➤ Talk of patterns and contexts does not answer the question of how reasons explain actions, since the relevant pattern or context contains both reason and action. One way we can explain an event is by placing it in the context of its cause; cause and effect form the sort of pattern that explains the effect, in a sense of "explain" that we understand as well as any. If reason and action illustrate a different pattern of explanation, that pattern must be identified.<sup>26</sup>

According to Davidson, then, when we offer a rationalization we explain an action by redescribing it in terms of its cause.

The most enlightening and productive discussion in Davidson's defense of the causal thesis, in my opinion, is his exchange with Hart and Honore. They claim that a reason cannot fulfill Hume's criteria for being a cause. According to Hume, "we may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second."<sup>27</sup> Hart and Honore take this (not implausibly) to mean that a singular causal statement implies a generalization or law, but since explanations of human behaviour do not appear to entail strict deterministic laws, it would seem that reasons cannot be causes: "The statement that one person did something because, for example, another threatened him, carries no implication or covert assertion that if the circumstances

were repeated the same action would follow."<sup>28</sup>

While Davidson agrees with the premise that any so-called generalizations connecting reasons and actions are unlikely candidates for general laws on the basis of which accurate predictions of human action could be made, he disagrees with the relationship Hart and Honore posit between a singular causal statement and a causal law. Hart and Honore seem to think that every true statement of causality instantiates a law, which, in Davidson's view, is false. Davidson's characterization of the relation between singular causal claims and causal laws is nicely summarized in "Mental Events" as follows:

The principle of the nomological character of causality must be read carefully: it says that when events are related as cause and effect, they have true descriptions that instantiate a law. It does not say that every true singular statement of causality instantiates a law.<sup>29</sup>

Davidson thinks that singular causal statements need not be couched in the same concepts in which laws are formulated. He develops this conception of causality in "Actions, Reasons, and Causes" by drawing attention to an ambiguity in Hume's thesis about causes:

... Hume's claim, as quoted above, is ambiguous. It may mean that "*A* caused *B*" entails some particular law involving the predicates used in the descriptions "*A*" and "*B*", or it may mean that "*A* caused *B*" entails that there exists a causal law instantiated by some true descriptions of *A* and *B*.<sup>30</sup>

Davidson believes that the second, weaker version of Hume's thesis is the only correct one, and that it is well suited to rationalizations. He supports this claim by means of the following example:

Suppose a hurricane, which is reported on page 5 of Tuesday's *Times*, causes a catastrophe, which is reported on page 13 of Wednesday's *Tribune*. Then the event reported on page 5 of Tuesday's *Times* caused the event reported on page 13 of Wednesday's *Tribune*. Should we look for a law relating events of these *kinds*? It is only slightly less ridiculous to look for a law relating hurricanes and catastrophes.<sup>31</sup>

In Davidson's view, we have only the vaguest ideas of causal laws, whether we speak of human action, of rocks breaking windows, or of hurricanes destroying bridges.<sup>32</sup> His response to Hart and Honore, then, turns on his removal of the apparent asymmetry between (seemingly) straight-forward causal generalizations (such as "rocks will tend to break windows if thrown at them with sufficient velocity") and folk-psychological generalizations. In either case, such generalizations only provide us with evidence of a causal law covering the situation at hand, not the terms in which we can expect that law to be formally stated.<sup>33</sup>

Davidson's conclusion is that the "laws whose existence is required if reasons are causes of actions do not, we may be sure, deal in the concepts in which rationalizations must deal."<sup>34</sup> A reason may be a cause provided that there is some true description of the events at hand which instantiates a law, even though such a description will not be in terms of beliefs or desires but will more likely be couched in neurological terms. For just as a hurricane can cause a bridge to collapse without the underlying laws being formulated in terms of bridges and hurricanes, a rationalization can explain an action without

the underlying law mentioning either belief or desire. Given this characterization of causality, it is not necessary to sharpen folk-psychological generalizations into laws of behaviour in order to defend the causal thesis and, hence, Hart and Honore's objection loses its force.

In "Mental Events," Davidson connects this claim to his so-called anomalous monism. He wants to reconcile three principles he holds to be true, yet which seem to lead to a contradiction. They are:

1. At least some mental events interact causally with physical events.
2. Events related as cause and effect fall under strict deterministic laws.
3. There are no strict deterministic laws on the basis of which mental events can be predicted and explained.<sup>35</sup>

Davidson is able to reconcile these three principles by means of his theory of token-identity between mental and physical, and a weak interpretation of (2), the same weak interpretation of Hume's thesis we saw in the reply to Hart and Honore.

Davidson's theory of token-identity proposes that although for every mental event there is a true physical description, there can be no identity established between kinds of mental events and kinds of physical events. So while every pain is identical with a specific neural event, there is not a uniform neural description for the class of pains as a *kind*. Davidson's position is nicely described by Quine in Pursuit of Truth:

Physicalistic explanation of neural events and states go blithely forward with

no intrusion of mental laws or intensional concepts. What are irreducibly mental are ways of grouping them: grouping a lot of respectably physical perceptions as perceptions that  $p$ , and grouping a lot of respectably physical belief instances as the belief that  $p$ .<sup>36</sup>

Although you and I may share the same pain (as a type), we probably do not share the same physical state, although there are, presumably, physical states on which our pains supervene.

This version of the identity theory, coupled with Davidson's interpretation of the relation between a singular causal claim and a causal law, allows him to formulate the principle of the anomalism of the mental (there are no strict deterministic laws on the basis of which mental events can be predicted and explained).

According to this principle, a reason can cause an action, but there will not be a law connecting reasons and actions because beliefs and desires are not concepts which figure in the underlying causal law. Davidson calls the generalizations of FP which point to an underlying law supposedly couched in neurological terms "heteronomic."<sup>37</sup> He calls them heteronomic because there is a shift required in our vocabulary as we move from the empirical generalization (say, people who feel a sudden sharp pain tend to wince) to the underlying neurological law. We require this shift because, according to Davidson, a generalization can be made precise only if it points to a law which draws its concepts from a closed system.<sup>38</sup> A closed system is one which can yield "a standardized, unique description of every physical

event couched in a vocabulary amenable to law."<sup>39</sup> While an ideal physics can satisfy this criterion, FP cannot because belief and desire do not form a closed system. The reason FP does not constitute a closed system is, for Davidson, founded in the holism of the mental. Since beliefs and desires cause behaviour only when mediated by other beliefs and desires "without limit,"<sup>40</sup> it is impossible to formulate law-like statements linking them. So since folk-psychological generalizations can be made precise only by shifting to a physical vocabulary, we cannot have psychophysical laws. In this way, then, it is possible for rationalizations to be causal claims without falling under strict deterministic laws.

Having briefly taken account of Davidson's formulation and defense of the causal thesis, I would now like to investigate Churchland's view.

Churchland seems to agree with the spirit of Davidson's arguments for the causal thesis, but he thinks that Davidson falls short when it comes to specifying the actual laws underlying action explanations. In "The Logical Character of Action-Explanations," Churchland attempts to construct such a law (or law-sketch), to refine it in the face of possible objections and counter-examples, and then to determine whether it can support counterfactual conditionals in order to show that it has nomic status. Since this paper, like Matter and Consciousness, mentions the D-N model of

explanation, given Churchland's later remarks in "Folk Psychology and the Explanation of Human Behavior," I suggest placing more emphasis on the nomic as opposed to the deductive aspect of the explanations he offers here.

Churchland proposes the following Law (L1), where A is an action and O the intended goal reached by A-ing:

X A-ed because he wanted O iff:

- (1) X wanted O, and
- (2) X believed (judged, saw) that A-ing was a way for him to achieve O under those circumstances [these are unspecified], and
- (3) there was no action believed by X to be a way for him to bring about O, under those circumstances, which X judged to be as preferable to him as, or more preferable to him than, A-ing, and
- (4) X had no other want (or set of them) which, under the circumstances, overrode his want O, and
- (5) X knew how to A, and
- (6) X was able to A.<sup>41</sup>

Churchland notes an interesting feature of explanations falling under L1, in fact, the same feature Davidson cites as evidence against the causal thesis. The conditions (1 through 6) form the premises of a practical argument leading to a conclusion in favour of performing the action in question. The conclusion appears reasonable-in-the-light-of the conditions.<sup>42</sup> This feature, which, as has already been suggested, characterizes an action as a "rational" one, has been seized upon by critics of the causal thesis. Churchland's discussion of this point is intended to defuse the claim that the explanatory power of an explanation falling under L1 is derived from this reasonable-in-the-

light-of relation. While this reasonable-in-the-light-of feature may be helpful for picking out what Churchland calls "full-blooded" actions, like Davidson, he thinks that it is not "the unique explanatory relation instanced in action-explanations."<sup>43</sup>

Churchland characterizes a full-blooded action as follows:

An event-description of the general form "X A-ed" is a description of a full-blooded action if and only if there exists an explanans deductively entailing "X A-ed" such that (a) from the wants, beliefs, preferences, and whatnot ascribed to X in the explanans, a valid practical argument yields a conclusion in favor of A-ing, and (b) the explanans contains but one law [or law sketch], a law which is part of the common-sense theoretical framework and which embodies the corresponding "reasonable-in-the-light-of" relation between the wants, preferences, and whatnot ascribed in its antecedent and the action mentioned in the consequent, and (c) the explanans meets all the standard D-N criteria for explanatory adequacy. [Churchland never actually specifies what these criteria are beyond conditions (a) and (b) above, but I suspect he is thinking of the requirement that the terms in a valid D-N explanation should be subject to the law of substitutivity of identity.]<sup>44</sup>

According to Churchland, this characterization of full-blooded actions accounts for a curious opacity which he notes in connection with condition (5) in L1. Actions are full-blooded only under a certain description. Consider explaining the event E, "bringing water to a boil," by appeal to a desire to make a cup of coffee. This event is identical with the event "bringing the water's vapor pressure to Pa," but when we describe the event in this second way, the desire to make a cup of coffee seems curiously inappropriate to its explanation (assuming that the agent doesn't know he is raising the water's vapor pressure).<sup>45</sup> We lose the reasonable-in-the-light-of relation between explanans and



explanandum that holds under the first description:

Clearly then, there is something wrong with saying that X raised the water's vapor pressure to Pa because he wanted to make a cup of coffee, because he did not "raise the vapor pressure to Pa" in the full-blooded sense relevant [since he didn't know he was raising the water's vapor pressure and so didn't intend it]. That is, the question of *why he did that* is curiously inappropriate to his having raised the water's vapor pressure to Pa in a way that it is not inappropriate to his having brought it to the boil. On the other hand, we do feel that X's desire for a cup of coffee is in some way explanatory of his having raised the water's vapor pressure to Pa.<sup>46</sup>

Churchland suggests that we redescribe E as E' which includes a statement of identity between the two descriptions of the event ("boiling water"="raising the water's vapor pressure"). Thus, E' is the conjunction of E and the identity statement, "to bring water to the boil is to raise water's vapor pressure to Pa."<sup>47</sup> This substitution preserves the D-N relation between explanans and explanandum but the event ceases to be full-blooded because the reasonable-in-the-light-of relation is lost with this substitution.<sup>48</sup> This explains why reference to a desire to make a cup of coffee seems inappropriate yet is nevertheless seen to have *something* to do with the event. Churchland's conclusion is that the *real* explanatory force of an explanation of either E or E' is derived from the nomic character of L1. The reasonable-in-the-light-of relation illustrated in full-blooded actions (as in the case of E) is an interesting "extra-nomic relation," but is not "the unique explanatory relation instanced in action-explanations."<sup>49</sup>

Finally, Churchland considers William Dray's theory of "rational explanation," which places far more emphasis

than Churchland on this "extra-nomic" feature of action explanations. According to Dray (who denies the validity of the causal thesis), providing an action explanation is to show that the act in question was a reasonable thing to do. This is achieved by outlining the agent's beliefs and desires.<sup>50</sup> The reasonable-in-the-light-of feature which characterizes full-blooded actions is the essence of Dravian action explanations, and it is by means of this relation that rationalization provides an explanation of human action. In Dray's view, deducibility of the explanandum from empirical laws (such as L1) is neither necessary nor sufficient for the explanation of an action. It is not necessary because subsumption under a law is not the aim of such explanations, and it is not sufficient because empirical deducibility does not necessarily reveal an action as a reasonable one (as the example with E' has just shown).<sup>51</sup> While, as Hempel charges on purely logical grounds, there is an important sense in which the Dravian model fails to explain why one event rather than another occurred,<sup>52</sup> there is also, as Churchland concedes, an important sense in which it does explain the action: it shows that the action was rational.

In Churchland's view, there are two central difficulties with the Dravian account. First, since Dray must deny the relevance of L1, he must argue that conditions (1) through (4) are sufficient for explaining X's action (otherwise, empirical deducibility would be a necessary

condition for explanation).<sup>53</sup> This has the drawback of opening Dray to the sorts of counter-examples (5) and (6) were developed to avoid.<sup>54</sup> Second, if Dray were to include (5) and (6) as necessary conditions, he would have to limit himself to those actions characterized as full-blooded.<sup>55</sup> For, as we have just seen, the Dravian model cannot explain events such as E' because there is no reasonable-in-the-light-of relation between explanans and explanandum. Dray would have to insist that events of this kind require another model of explanation, and isolating full-blooded explanations as a distinct species of explanation, Churchland thinks, has the bad effect of "multiplying the types of explanation supposed to obtain in our conceptual tool shed."<sup>56</sup> Churchland's view avoids both of these difficulties by supposing that the explanatory power of full-blooded actions is derived from the underlying law (L1). Thus we do not need to distinguish between different species of explanation and, unlike the Dravian view, we preserve explanatory power even when we change the descriptions under which events are considered.

### 3. Conclusions

To conclude, I would like to point out the important similarities between Davidson and Churchland regarding the causal thesis, since these themes will be discussed in the

following chapter. First, both authors agree that rationalization is a species of causal explanation constituting a singular causal claim. Such causal claims, they agree, point to an underlying law which is not necessarily formulated in the same terms as the singular claim. For Churchland, the underlying law is folk-psychological, whereas for Davidson, it is presumably a physical law. Second, both Churchland and Davidson point out the "reasonable-in-the-light-of" relation that typically holds between the explanans and explanandum of action explanations. Although this feature of action explanations has been exploited in arguments against the causal thesis, both authors agree that it is not the source of the explanatory power of action explanations; the explanatory nature of such descriptions is actually derived from the underlying law. Finally, both stress how events are described and the effect this has on the "reasonable-in-the-light-of" feature of action explanations.

In the next chapter I will, by means of a critical evaluation of Davidson's argument for the causal thesis, show that the reasonable-in-the-light-of relation is in fact the relation by means of which rationalizations explain actions. I will argue that depending upon how we describe an action we are tied either to offering a rationalization of the action (to which the reasonable-in-the-light-of relation is central) or to providing a causal explanation of the

event.

## CHAPTER II REASONS AS EPIPHENOMENA

In Chapter 1 I discussed the effect of how events are described on certain features of an explanation of those events, in particular, the effect on the reasonable-in-the-light-of relation between explanans and explanandum. In this chapter I will argue that to speak of an event under a particular description is to pick out certain properties of that event and to ignore others, and furthermore, that depending upon which properties of an event are highlighted the explanation of that event can take on a different form. I will introduce this proposal by summarizing and evaluating a debate on Davidson's anomalous monism. I discuss this debate because it involves the proposal that events are causally efficacious only by virtue of possessing certain physical properties. This means that mental events are causally efficacious only if they possess the relevant physical properties. These properties, I will suggest, can only be picked out by considering mental events under their physical descriptions. Since mental events do not enter into causal relations with other events by virtue of their mental properties, those properties are epiphenomenal.

The relevance of this debate to the causal thesis is as follows. If mental events are causally impotent, and reasons are among such events, then reasons cannot be causes. I will begin then, with an account of the debate on anomalous monism. The debate, which took place in *Analysis* between 1981 and 1984, centers precisely on this charge that Davidson's anomalous monism renders mental events epiphenomenal. The protagonists of the debate are Peter Hess and Ted Honderich, who make the charge, and Peter Smith, who defends the Davidsonian view.

#### 1. Anomalous Monism and Epiphenomenalism

Since Hess's own argument in "Actions, Reasons, and Humean Causes" follows the same pattern as Honderich's, and Honderich's is more explicit, I will begin with Honderich's article, "The Argument for Anomalous Monism," for it includes, though it does not mention, Hess's position.

Honderich begins by pointing out a tension in Davidson's position. On the one hand, Davidson does not differentiate the properties constituting events: "The theory under discussion [anomalous monism] is silent about processes, states, and attributes if these differ from individual events."<sup>1</sup> His suggestion in "Mental Events" that causality "deals with events in extension and is therefore blind to the mental-physical dichotomy"<sup>2</sup> further illustrates

an apparent disinterest in distinguishing among the several different properties of events. On the other hand, to speak of events as "under a description" is nevertheless to do just that; that is, to pick out certain properties of an event and ignore others. As Honderich puts it: "To say things are not in lawlike connection under certain descriptions is to say that certain of their properties are not in lawlike connection, or, perhaps, that the things are not in lawlike connection in virtue of certain of their properties."<sup>3</sup> Given this, Honderich draws the not unreasonable conclusion that Davidson is bound to say that certain properties of an event as opposed to others are relevant to the causal efficacy of that event. Honderich's example is that the fact that the pears were green and French was not relevant to their causing the pointer to move to the two-pound mark when placed on the scale:

"Something weighing two pounds being put on the scale caused the pointer to move to the two-pound mark" does not entail that the events are in lawlike connection under the same descriptions. However, it *does* follow from any statement that the event of the pears' being put on the scale caused the pointer to move to the two-pound mark, and the statement that it did so in virtue of only certain properties, that events were in lawlike connection by way of those properties.<sup>4</sup>

Honderich calls the principle that events are in lawlike connection only by virtue of certain properties the "Principle of the Nomological Character of Causally-Relevant Properties."<sup>5</sup>

Given this principle, Honderich asks what we should say about any given mental event; i.e., which properties are



responsible for the efficacy of a mental event? He identifies two possibilities: It is either the event's mental properties (i.e., some characteristics of that event under the "mental" description), or it is by virtue of some physical properties that event has that it is causally efficacious. These physical properties will presumably be those neurological properties or processes on which the mental properties supervene.

As Honderich proposes, Davidson's belief in the efficacy of the mental (seen in the first claim of anomalous monism: at least some mental events interact causally with physical events), suggests the first answer. This is further supported by Davidson's conviction in "the efficacy of thought and purpose in the material world."<sup>6</sup> Surely such phrases are intended to suggest that mental events cause physical events. Indeed, it is precisely this efficacy of the mental that distinguishes Davidson's position from epiphenomenalism. The Principle of the Nomological Character of Causally-Relevant Properties, however, creates a tension between the first and third claims of anomalous monism (at least some mental events interact causally with physical events; there are no strict deterministic laws on the basis of which mental events can be predicted and explained). If we accept that it is the mental as mental which is causally efficacious, it seems that we are forced to deny the third claim of anomalous monism and hence give up anomalous

monism.<sup>7</sup> For if it were the mental as mental which is efficacious, then the causally relevant properties would be those mental properties picked out under the mental description of an event. This means that one could formulate laws employing mental concepts and, therefore, the mental falls under strict deterministic laws.

Given this difficulty, Honderich investigates the second alternative: that a mental event, by virtue of its physical properties, has the causal powers it does. While this preserves the third claim of anomalous monism (for we still have a heteronomic claim), it surrenders the first. For the mental descriptions or properties seem to have nothing to do with the efficacy of the event, since it is whatever physical properties that underlie them which do the necessary causal work. Hence we have the epiphenomenalism of the mental.

Finally, Honderich looks to supervenience for possible help against these difficulties. I think his reason for doing this is to cut short the sort of reply Smith made to Hess's "Actions, Reasons, and Humean Causes." In Smith's response to Hess's version of the article, he appeals to supervenience and counterfactual inferences in order to block the charge of epiphenomenalism. Since Honderich does not define exactly what he means by an epiphenomenal property, I shall use Hess's formulation:

A property P is epiphenomenal with respect to the relationship between an event C and its effect E iff

- (i) P is a property of C;
- (ii) It is not the case that C would not have caused E had it not had property P.<sup>8</sup>

Smith's response, which is perhaps what Honderich is addressing, is as follows:

[B]y the supervenience assumption, C cannot lack property P while remaining the same in all law-engaging physical respects ... So, an event which lacks P must be an event with some different physical characteristics from those actually possessed by C, and on a plausible view about event identity, this will necessarily be a distinct individual event. But if this is right, then it will be vacuously true that C would not have caused E had it not had property P (for it would not then have existed, and so couldn't cause anything!).<sup>9</sup>

This response, I believe, is reflected in Honderich's suggestion that "since the event as mental supervenes on the event as physical, the event as mental is efficacious with respect to the action."<sup>10</sup> Supervenience notwithstanding, Honderich seems to think that since the connection between mental and physical descriptions is anomalous (otherwise there would be psychophysical laws), we cannot have the efficacy of the mental.<sup>11</sup> For it is only by virtue of those physical properties underlying the mental event that that event is causally efficacious.

Peter Smith's reply to Honderich in "Bad News for Anomalous Monism?" focuses on Honderich's claim that it is an "accident" that a mental event has the causally efficacious physical features it has. Smith's strategy is to suggest that it is no accident that mental events have these causally salient features because "the physical state which is the belief is partially *identified* as the state which has the right physical properties to cause the action."<sup>12</sup>

He begins by asking how one would pick out the physical event with which a mental event is identical, given the mental specification. Since we cannot here appeal to an identity relation from the identity theorist, one must identify the belief functionally with the physical state which is:

(a) causally dependent on those antecedents which folk psychology recognizes as explaining *X*'s belief (such as perceptual input), and such that (b) it is causally involved in the production of such behaviour as folk psychology interprets as actions done because *X* believes that *p*.<sup>13</sup>

Thus a connection (though not a nomological one) is made between mental and physical descriptions which does not entail psychophysical laws, and hence insulates anomalous monism from Honderich's argument.

Smith is concerned, though, that his argument might seem circular. One might suggest that fixing which actions were done because of a given belief presupposes some causal claims about beliefs. For in the absence of psychophysical laws, one must appeal to physical laws to establish causal claims, which has the effect of trapping us in a purely physical vocabulary with no means of applying the relevant physical laws to belief and action unless we can independently identify a belief with a particular physical event.

Smith's response to this objection is that it simply is not true that we need to appeal to physical laws to establish causal claims. All that is necessary to know that dropping a plate will cause it to break, as Davidson himself

has said, is a rule of thumb, not the actual relevant laws. Similarly, FP provides rough generalizations for particular causal claims linking beliefs and desires to actions, and this is sufficient for identifying mental and physical properties in the manner suggested.<sup>14</sup> Thus Smith believes he can show that it is no accident that a mental event has the physical features it has, and with this he blocks the charge that anomalous monism, when examined carefully, cannot be distinguished from epiphenomenalism.

It will not take too keen a mind to recognize that Smith is using some sleight of hand here, for his response is question-begging. In his attempt to determine whether propositional attitudes are causally efficacious, he presupposes that the generalizations of FP are causal claims, which is precisely the point at issue.

In a response to Smith's defense of anomalous monism Honderich ignores the question-begging means Smith employs to identify mental and physical properties; instead he insists that Smith has misread his argument. Honderich claims that all he meant by saying that the identity between the mental and the physical is accidental is that the connection is not lawlike: "it is nomologically inessential, to the event's being the physical event that it was, that it was the mental event that it was."<sup>15</sup> Therefore it is inessential to its effect that it was the mental event that it was, and this amounts to the epiphenomenalism of the mental. The anomalous

monist is bound to accept this conclusion because of his denial of psychophysical laws. So all of Smith's labours have been beside the point:

The Anomalous Monist, having identified in Smith's way the physical event he has in mind, will presumably persist in the denial of lawlike connection noted above. It therefore remains as inessential as ever to a certain physical event's being as it is, and having the causal connections it does, that it is a mental event to which ordinary belief ... assigns the same causal connections. *That* is the objection.<sup>16</sup>

In "Anomalous Monism and Epiphenomenalism: A Reply to Honderich," Smith lays out what he takes to be Honderich's argument:

- (A) The anomalous monist is tied to saying that it is by virtue of its physical properties that a mental event causes an action.
- (B) There are no nomological links between an event's mental and physical properties.
- (C) It follows that there are no nomological links between a mental event's mental properties and its physical properties by virtue of which it is causally efficacious.
- (D) Therefore, it follows that it is an accident that a mental event, by virtue of its mental properties, should have the causal power FP attributes to it.<sup>17</sup>

Smith then sets out what he takes to be the doctrine of epiphenomenalism:

- (E) Mental events do not cause physical events at all.<sup>18</sup>

Smith simply denies the truth of (E), as the anomalous monist has a firm conviction that mental events can and do cause physical events. He then suggests that Honderich's charge of epiphenomenalism against anomalous monism is confused:

For it obviously can't follow from the thought that "it is inessential" to a given physical effect that its physical cause "was the mental event it was" that the mental event in question lacks all causal powers -- for by hypothesis the mental event is identical to a physical event *with* such powers.<sup>19</sup>

Given this identity, Smith thinks that it simply makes no

sense to suggest that a given mental event lacks all causal powers. And given Smith's account of the connection between mental and physical properties, which he likens to the connection between an event's being a hurricane and its having properties by virtue of which it can destroy buildings,<sup>20</sup> he believes he blocks Honderich's argument at stage (C).

In a final reply, "Smith and the Champion of Mauve," Honderich alleges that Smith cannot distinguish himself from the Champion of Mauve, who believes that although it is by virtue of the fleece lining that his mauve slippers keep his feet warm, and that although their being mauve is not nomically necessary to their being fleece, one should not "underrate" the connection between being mauve and keeping his feet warm, for the fleece slippers are *identical* to the mauve slippers.<sup>21</sup> Smith is as misguided as the Champion of Mauve because he believes that although it is by virtue of whatever physical properties a mental event has (as opposed to its mental properties), that it is causally efficacious, and that although there are no psychophysical laws, he nevertheless believes that mental events cause physical events.

Smith's policy here, in which he persists, is to go on saying that an event which has a mental property is an event which also has a physical property that is causal with respect to an action ... The Anomalous Monist can be as wedded as he wants to the proposition that *of course* a mental event in his sense causes a physical event. By way of that truth he is no nearer getting mental efficacy than the Champion of Mauve is to getting mauvish efficacy by going on saying that it *is* the mauve slippers that keep him really warm.<sup>22</sup>

Although Smith has not yet had opportunity to respond to Honderich on this point, he could suggest that Honderich is equivocating on the term "identity" between the case of Smith and the Champion of Mauve. Surely Smith implies a stronger sense of identity between the mental and physical characteristics of an event than can be secured between the mentioned properties of the slippers. While we might want to say that whatever is mental is also physical, we surely do not want to say that whatever is mauve is fleece. While there is some intuitive appeal to this suggestion, we must ask ourselves if Smith's use of "identity" is different enough from Honderich's to determine if Honderich is equivocating on "identity". It seems that the only way Smith could achieve this would be to conceive of the identity between mental and physical descriptions in terms of a reductive definition, analogous to "lightning"="rapid electrical discharge." But clearly Davidson is opposed to such a conception of identity since he refers to his own position as a brand of non-reductive materialism:

Anomalous monism shows an ontological bias only in that it allows the possibility that not all events are mental, while insisting that all events are physical. Such a bland monism, unbuttressed by correlating laws or conceptual economies, does not seem to merit the term "reductionism"; in any case it is not apt to inspire the nothing-but reflex ("*Conceiving the Art of the Fugue* was nothing but a complex neural event", and so forth).<sup>23</sup>

So it seems that Smith cannot support a disanalogy between himself and the Champion of Mauve, and is therefore unable to side-step Honderich's reductio. We can say in the light of the above discussion that Davidson's anomalous monism does



render mental events epiphenomenal. To the extent that reference to a mental event involves mention of that event's mental properties, that event is not causally efficacious. This means that reasons, since they pick out mental properties, are not causally efficacious and, therefore, cannot be causes.

## 2. Hess on Reasons and Causes

Several years after these articles were published, Peter Hess, in his book Thought and Experience, offers an analysis of action explanations which, in my view, illustrates the validity of Honderich's arguments against anomalous monism, and provides us with a means to distinguish causal explanations and rationalizations as different species of explanation. While this further distinction, which has already found expression in the Dravian view discussed in the first chapter, will not be fully articulated until the third chapter, I hope to lay some groundwork here in my discussion of Hess's account of action explanations.

Hess examines the following scenario: A mountain climber shouts the words: "There is going to be an avalanche," which, it so happens, is sufficient to cause an avalanche which sweeps him away to his death. In a similar fashion as Honderich, Hess asks what features of the mountaineer's utterance we should consider in formulating a

causal explanation of the event just described. According to Hess, we can rule out such considerations as the fact that the sentence was in English, that it was indicative, that it possessed certain semantic and syntactic properties, and instead focus on the utterance's acoustical properties. For surely it was due to the fact that the utterance was at such and such a time and place, that it involved sound waves of such and such cycles per second, etc., that it was instrumental in causing the avalanche. These physical properties, in Hess's view, do not describe the utterance in the same way as those features which will not figure in a causal explanation of the events at hand. This is because these physical properties exist no matter how we describe or measure them, and hence function independently of any interpretation, whereas the others require an ascriptive judgement involving an interpretation in accord with a body of conventions, namely, conventions about propositional content, e.g., what the mountaineer's utterance meant. It is only in the context of the speaker's intentions, rules governing shouts like the one in question, and the speaker's knowledge of those rules that it makes sense to ascribe propositional content to that utterance.<sup>24</sup>

In light of these considerations, Hess investigates an episode of human behaviour and asks whether its explanation is a causal one. He takes the following explanation: "Mary started to run because she realized that

her appointment with the bank manager was only 5 minutes away." Hess claims that to the extent that this explanation involves an ascriptive judgement (analogous to the case of the avalanche), i.e., that Mary entertained and accepted a certain proposition, this explanation does not pick out features which are causally relevant, and hence, the explanation does not appear to be a causal one.<sup>25</sup> But surely there are some properties of Mary's belief which caused her to run. Certainly, otherwise Mary would not be running. These will be whatever physical properties of the central nervous system physically realize the belief in question. According to Hess, although we can admit that such properties probably enter causal relations, we can still deny that the explanation for Mary's running is a causal one. To explain how, Hess calls upon the concepts of referential transparency and opacity.

Hess recalls the standard view that causal explanations are referentially transparent. This means that the truth-value of a causal explanation remains constant no matter how one refers to the elements involved in the explanation.<sup>26</sup> This seems to be an implicit assumption in Davidson's account of action explanations, given his indifference to specific descriptions, and is most certainly presupposed by Churchland. What this referential transparency shows is that truth-value is a function of the facts the explanation refers to (recall our example of the

physical features of the mountaineer's utterance) and not how it refers to them.<sup>27</sup> Is the explanation for Mary's running of this form? Hess invites us to assume that the bank manager is also Mary's long-lost brother. Subjecting the explanation to the appropriate substitution we arrive at the following: "Mary started to run because she realized that her appointment with her long-lost brother was only 5 minutes away." Since Mary doesn't know the identity of the bank manager, this statement is false, and therefore, the explanation is referentially opaque.<sup>28</sup> Hess takes this to mean that the explanation for Mary's running is not a causal explanation. This is because it involves intentional objects which tell us something about Mary's reasons for running, rather than about causal processes.

Somebody might claim that this explanation will not suffice unless we assume that there is a causal relation underlying the belief in question and the action it explains, but in that case it involves a causal claim. Recalling my discussion of Churchland from Chapter 1, we can see that Hess's example here is of the same form as Churchland's example, "boiling water"="raising water's vapor pressure." Churchland's claim that the full-blooded explanation (corresponding to Hess's first explanation for Mary's running) is explanatory by virtue of the underlying law is very close to the one Hess anticipates above, except that Hess has a physical law in mind.

But the claim that there is a causal relation underlying the action is beside the point. According to Hess, we do not need to pay any attention to the physical characteristics of the belief in question (or, presumably, any folk-psychological laws) in order to have a cogent explanation. When we speak of actions, we are speaking of more than just observable bodily movements; we are rendering ascriptive judgements (i.e., we are attributing some kind of propositional content to the person in question), and this content does not figure among those properties which an event can be caused to have. This does not mean that these properties expressing propositional content are uncaused, it means that they are not intrinsic to the event in question and are therefore not the product of whatever causal processes produced the event in question. In order for events to be meaningful then, and presumably actions are among such events, they must be situated in a broader context involving an interpretation.<sup>29</sup>

To illustrate this, Hess asks us to imagine that a gust of wind miraculously arranges grains of sand in a way that spells out the sentence "This is a public beach." By citing the relevant causal relations and conditions enabling the wind to do this, we will not, in Hess's view, be compelled to regard the event as an action or see it as the expression of a belief. Hence we will have trouble assigning any propositional content to the grains of sand so long as we

are prevented from situating the event in an appropriate broader context in such a way that we can regard it as an action:

By restricting ourselves to a description of the causal conditions that produced the arrangement, we prevent ourselves from viewing this happening in a larger context which, due to its particular nature, might perhaps have justified us in judging the arrangement of sand particles as one which expresses a proposition which is true. Similarly, if we restrict our account of an action (or a belief) to the description of the causal factors that produced the publicly observable behaviour involved in that action ..., we prevent ourselves from seeing the phenomenon in question as an action (or as the acceptance of a belief).<sup>30</sup>

So, we cannot, by citing causes, render the appropriate ascriptive judgements required for rationalizations; for they lack the appropriate larger context, namely, the context of folk psychology. This prevents us from viewing the "because" of rationalization as a causal "because". Therefore, according to Hess, rationalizations function to explain behaviour in a different fashion than causal explanations:

When we explain why somebody acted as he did by referring to his intentions, decisions, and wishes, we are not offering a causal explanation. We are, instead, putting his action "in a certain light". We are depicting it as something which was a reasonable thing to do for an agent who formed such intentions and decisions and who entertained such wishes.<sup>31</sup>

With respect to Honderich's arguments concerning anomalous monism, we can now better appreciate why mental events are epiphenomenal. To the extent that mental events (having reasons) rationalize actions by situating them in a broader context, they cannot be considered causes and, consequently, cannot be causally efficacious. Since this view is contrary to Davidson's I would like to ask what means Davidson himself employs to identify a primary reason. If

these correspond to Hess's criteria, then Davidson's view is all the more problematic. Recall Davidson's first claim from "Actions, Reasons, and Causes":

C1. *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.<sup>32</sup>

The features Hess argues must figure in a rationalization are precisely those which will characterize a Davidsonian primary reason. It is by virtue of a certain propositional content (here represented by *d*) that we can identify a primary reason, and although that belief (i.e., the belief expressed in the primary reason) may have a physical description which figures in a causal relation, that physical description is not what makes *R* a reason. So if reasons are in fact characterized in the manner suggested by Hess, and it seems they must be, then Davidson is surely wedded to the view that reasons are not causes and, hence, to the denial of the efficacy of the mental.

Before I draw my conclusions, I would like briefly to consider an argument by Armstrong that in order for a reason to properly rationalize an action, it must cause it in the right sort of way.

Reasons must be linked to actions characterized as intentional ones, and according to Armstrong, intentional actions are necessarily tied to the notion of causation. This creates a difficulty for my distinction between causal explanations and rationalizations, and so the matter deserves

some attention.

In "Acting and Trying," Armstrong investigates what must be added to "A tried to do *P*" to make *P* an intentional action. His suggestion is that part of the answer involves considering the act of trying to be a cause of *P*, and furthermore, that this cause must follow a particular pattern.<sup>33</sup> His reason for suggesting this is to circumvent a problem raised by Chisholm.<sup>34</sup> Imagine that A wants to kill his uncle and sets out to do it with a great deal of excitement. On his way to commit the murder he fails to notice a pedestrian and runs him down. It so happens that the pedestrian is A's uncle. So A's attempt to kill his uncle does in fact cause his uncle's death, but the act cannot be viewed as an intentional one, for the action was not brought about in the right way:

How is a case like Chisholm's to be excluded? The reflection that naturally occurs is that, although A's attempt brought about the death of his uncle, it did not bring it about *in the right way*. The causal pattern in which A's attempt to bring about his uncle's death brings about that death is not the right sort of pattern. What would the right sort of pattern be?<sup>35</sup>

Armstrong's answer is that the causal chain which culminates in the action must follow the pattern of practical reasoning the agent develops as he proceeds with his action:

[I]f purposes, beliefs, etc. can be thought of as *causes*, then we can think of A acting as he does *because*, in the causal sense of "because", he has these objectives, these beliefs about the current and developing situation, these principles of reasoning and acting. *The pattern of the practical reasoning shadows out a pattern of operation of causal factors in A's mind.*<sup>36</sup>

This view is subject to two criticisms. First, if, as Hess has suggested, we have good reason to resist the idea



that reasons are causes, why should we look for such a symmetry between the flow of practical reasoning and of causal processes? Surely all we need is for there to be *some* causal account (indeed, there must be or the act would never have occurred), and it is altogether unclear why one should expect an undoubtedly complex network of causal relations to "shadow" the stages of practical reasoning. What would the relevant similarities be? How would one pick them out? Of course a defense of this criticism depends upon a further substantiation of the distinction between causal and reason-giving explanations which will follow in the next chapter.

Secondly, Armstrong's view seems to identify the inferential role of a belief with its causal role, and there is good reason to resist this identification. As Mark Thornton points out, inferential role depends upon referential mode (i.e., whether or not we draw certain inferences depends in part on how the propositions in question are referred to or described) whereas causal role does not. The causal role of my belief that I am marrying Jocasta and my belief that I am marrying my mother are the same, but certainly these two beliefs do not have the same inferential role.<sup>37</sup> Since inferential role plays an integral part in practical and theoretical reasoning, and given the difference that exists between the inferential and the causal roles of a belief, there is little reason to expect any significant relation of mirroring to exist between causal

explanations and rationalizations.

### 3. Conclusions

The conclusions I draw are first, that Davidson's anomalous monism and his characterization of a primary reason as a cause cannot be supported. Primary reasons are formulated and identified by means of propositional content, and it is by means of this content that they rationalize actions, i.e., they show the act in question to be a reasonable thing to do. This involves situating the action in a broader context involving an interpretation in the light of folk psychology, and makes no mention of causal relations. In fact, it seems that causal explanations function on a completely different level from rationalizations, and hence, perhaps, serve a different purpose.

Second, there is clearly an important sense in which it is true that mental events are epiphenomenal. Since reasons are not causally efficacious, they cannot be causes, except perhaps in the strained sense Honderich implies in "Smith and the Champion of Mauve." They might be described, as Frederick Stoutland has suggested elsewhere, as "oblique causes."<sup>38</sup>

This is where we once again make a connection with Paul Churchland. Recalling my remarks in Chapter 1, I suggested that the thesis of eliminative materialism does not

require Davidson's anomalous monism. What Churchland does require is the claim that folk-psychological explanations are causal explanations. Since Davidson and Churchland disagree on the question of whether there can be folk-psychological laws, the criticism of anomalous monism I have reviewed in this chapter leaves Churchland's causal thesis untouched. But Hess's discussion is more far-reaching than the debate reviewed in *Analysis*. He has distinguished two types of explanation, something which Churchland admits, in his discussion of Dravian explanations, would pose a problem for eliminative materialism. As I said in Chapter 1, if FP employs a different species of explanation from neuroscientific explanation (that is to say, folk-psychological explanations are something other than causal explanations), then there is little reason to think that FP is liable to replacement by neuroscience. My elaboration and substantiation of this claim follows in the next chapter.

In sum then, my discussion of the debate on Davidson's anomalous monism was intended to show that to speak of an event under a particular description is to pick out certain properties of an event and ignore others, and furthermore, that depending upon which properties of an event are highlighted, the explanation of that event can take on a different form. As Hess has shown, to limit oneself to the physical properties of an event is to be tied to providing a causal explanation of the events in question. To focus on the

mental properties of an event, however, is to render an ascriptive judgement and hence to locate that action within the context of folk-psychological explanation. It is to understand the action as a rational one and hence to rationalize it.

The suggestion that these two models of explanation depend upon a differentiation of properties might seem to lead to the conclusion that the distinction between causal explanations and rationalizations depends upon some kind of ontological claim i.e., it requires an ontological dualism of properties. In the following chapter I will show that this is not the case.

### CHAPTER III THE POINT OF VIEW DISTINCTION

Ted Honderich's closing remarks from the debate discussed in the previous chapter include the following interesting statement:

... Anomalous Monism, having denied psychophysical nomic connection, is so far indistinguishable from what we all hoped had been put to rest, which is to say mere psychophysical parallelism. Will this be followed by a revival of Pre-Established Harmony?<sup>1</sup>

This remark suggests a good deal of dissatisfaction with the theory of mind which results from the preceding criticisms of anomalous monism. I believe, however, that there is little to fear in the picture of mind which has resulted in Chapter 2. What Honderich has failed to recognize is that the parallelism between causal and reason-giving explanations arrived at in the last chapter is an explanatory parallelism and not an ontological parallelism. I would like to show that the distinction between causal explanations and rationalizations need not lapse into some form of ontological dualism by providing an analysis and critique of Descartes' account of action explanations. (I am taking Descartes' dualism as the model for our understanding of an ontological dualism of substances or properties; i.e., there are two kinds of things constituting the universe: non-physical, mental substances or properties, and physical, unthinking

substances or properties.) My suggestion will be that we can understand the Cartesian division between substances as a form of explanatory parallelism which is compatible with an ontological monism. I will argue that what motivated Descartes to make his ontological claim was simply his failure to recognize that he was attempting to combine two different explanatory models of human behaviour. As he unwittingly shifted explanatory stances in what he took to be one uniform explanation of behaviour, he was led to posit a dualism of substances and then connect the two. It is my contention that if we can recognize these two models of explanation and keep them separate, we thereby avoid the difficulties associated with ontological dualism, i.e., questions as to what sort of thing a mental property or substance is, (if and) how it can interact causally with physical substances or properties, etc.

With this discussion of Descartes I will introduce and explain Mark Thornton's point of view distinction, developed in Folk Psychology: An Introduction, and show that Thornton's dualism of points of view is the principle underlying the distinction made in Chapter 2 between causal explanations and rationalizations. By bringing this distinction into focus I will demonstrate the falsity of the causal thesis and draw its consequences for eliminative materialism. I will begin, then, with a brief account of Descartes' model of explanation for human action.

# 1. Reasons, Causes, and Cartesian Dualism

Although Descartes had the tendency to analyse human behaviour according to a rigorous causal-explanatory model, his philosophy of mind has been regarded as the obstacle to be overcome by more scientific theories of the person. For, as Flanagan points out in The Science of the Mind, Descartes' philosophy was highly motivated by scientific theory. This interest in scientific description can be seen in his work on mathematics, his interest in human anatomy, and his frequent analogies between the human body and mechanical objects such as watches and automata. Descartes' desire to understand human behaviour in terms of causal models of explanation is, as Flanagan observes, particularly well illustrated by his interest in the automata at the French Royal Gardens:

You may have seen in the grottos and fountains which are in our royal gardens that the simple force with which water moves in issuing from its source is sufficient to put into motion various machines and even to set various instruments playing or to make them pronounce words according to the varied disposition of the tubes which convey the water.<sup>2</sup>

Apparently there were many such complex machines which responded to passers-by who unwittingly triggered their movements by stepping on pressure-sensitive tiles. The movements of these automata were evidently quite life-like, and led Descartes to theorize about whether it would be possible to explain the behaviour of persons in a manner analogous to the explanation of the movements of these

robots. The actions of these automata could be explained in terms of a series of reflex arcs: deterministic three-term causal sequences involving a stimulus (someone stepping on a pressure-sensitive tile), a precise internal mechanism (the flow of water through a series of elaborate tubes), and a response (the resultant motion of the mechanical figure).<sup>3</sup> The analogy between the hydraulic processes governing the behaviour of these automata and the animal spirits, as well as some other basic physical similarities, gave Descartes reason to pursue this hypothesis.<sup>4</sup>

There are passages in the Meditations and The Passions of the Soul where Descartes does in fact explain some human behaviour in precisely these terms. In Meditation VI, when he compares the function of the human body to that of a clock, he writes:

In the same way, I might consider the body of a man as a kind of machine equipped with and made up of bones, nerves, muscles, veins, blood and skin in such a way that, even if there were no mind in it, it would still perform all the same movements as it now does in those cases where movement is not under the control of the will or, consequently, of the mind.<sup>5</sup>

In The Passions of the Soul, Descartes is even more explicit in his mechanistic treatment of the human body. He provides what could be regarded as primitive neurophysiological explanations for certain types of bodily behaviour, tracing the movement of the animal spirits through the brain, central nervous system, and limbs. While some behaviour requires reference to the mind, there are many motions of the body for which Descartes provides wholly physiological explanations.



For the most part these include behaviour such as breathing, eating, walking, habitual behaviour, and any other actions shared with lower animals:

[W]hen one of the pores [in the brain] is opened somewhat more or less than usual by an action of the sensory nerves, this brings about a change in the movement of the spirits and directs them to the muscles which serve to move the body in the way it is usually moved on the occasion of such an action. Thus every movement we make without any contribution from our will ... depends solely on the arrangement of our limbs and on the route which the spirits ... follow naturally in the brain, nerves and muscles. This occurs in the same way that the movement of a watch is produced merely by the strength of its spring and the configuration of its wheels.<sup>6</sup>

In Meditation VI it seems that Descartes extends his reflex model to more complex behaviour. There he posits a rigid relation between physical states of the pineal gland and mental states, which would seem to draw higher level cognitive functions such as acts of the will under coverage by the reflex model, for the relations between mind and pineal gland become deterministic. This deterministic relation is expressed in the following proposal: "My final observation is that any given movement occurring in the part of the brain that immediately affects the mind produces *just one* corresponding sensation" (my emphasis).<sup>7</sup> I am, of course, suggesting that one might interpret "just one" as meaning "just one type". While this passage makes explicit mention of sensations as opposed to actions, the rigid relationship that arises between pineal and mental states suggests that human action might also fall under the reflex model of explanation, particularly in the case of reactions to stimuli.

Later on in The Passions, however, Descartes drops

this relation between the mind and the brain and hence, sets the explanatory limit for the application of the reflex model at the pineal gland. While the mind or soul does not fall under this model, it is clear that the physiological causal chains beginning with or terminating in the pineal gland remain deterministic. Descartes' reasons for setting the limits to the reflex model in this way are clearly tied to his conviction in the freedom of the soul with respect to its control of the passions:

But the will is by its nature so free that it can never be constrained .... And the activity of the soul consists entirely in the fact that simply by willing something it brings it about that the little gland to which it is closely joined moves in the manner required to produce the effect corresponding to this volition.<sup>8</sup>

In order to preserve the freedom of the soul, Descartes proposes an alternative relation between mind and brain which is far more idiosyncratic, even to the extent that there is no longer a one to one relation between brain states and mental states. With respect to behavioural reactions, in section 39 of The Passions, under the heading "How one and the same cause may excite different passions in different people," Descartes writes:

The same impression which the presence of a terrifying object forms on the gland, and which causes fear in some people, may excite courage and boldness in others. The reason for this is that brains are not all constituted in the same way. Thus the very same movement of the gland which in some excites fear, in others causes the spirits... to move the hands in self-defence....<sup>9</sup>

This passage shows conclusively that Descartes did not believe that there could be a type-type identity between brain (or pineal) states and mental states. Descartes

anticipated what Quine takes for granted: that our neural networks are idiosyncratic.<sup>10</sup> So although there is still a causal relationship between mind and body, this relation does not fall under the reflex model. The reflex model of explanation reaches its limit of application at the pineal gland so as to preserve the freedom (through anomaly) of the soul.

Here, then, is the general model of Descartes' account of human behaviour. In the case of behaviour which has its origins in the mind, Descartes is far from dropping his primitive neurophysiological descriptions of the body as an element central to its explanation. In fact, for the most part his explanation for such behaviour is formulated in terms of physical causal relations. The point of origin, though, is the soul, a different kind of thing from all the others in the explanation, since it is non-physical. Hence, Descartes maintains his reflex model until he has to make reference to the reason or the desire in the soul as the efficient cause of the behaviour in question.

I think it would be correct to say that the relevant physiological causal chain does not terminate in the pineal gland (or any other part of the brain), but could in principle be related to a potentially infinite series of other physical events both within and outside the body. Calling upon the distinctions made in the previous chapter, I suggest that where Descartes begins to refer to a different

kind of thing (i.e., an immaterial substance), he has simply shifted explanatory frameworks; he has given us a reason instead of a cause. Because he nevertheless regards the reference to the soul to be a continuation of the *same* causal explanation, he is forced to formulate an uneasy union between mind and body.

So while Descartes' dualism is traditionally regarded as a combination of different substances linked through causal relations, I think, given the distinction between reasons and causes in the second chapter, that what is really happening in Descartes' account of behaviour is that he is combining different kinds of explanation. The ontological shift actually reflects an explanatory shift. In his account of human action Descartes moves from offering a causal explanation to offering reasons for behaviour at a convenient juncture (i.e., where knowledge of the function and construction of the brain begin to fade into uncertainty).

## 2. The Point of View Distinction

While we have seen good reason to formulate this distinction between reasons and causes in terms of Hess's criteria (reasons are distinct from causes because the properties by means of which they explain events are ascribed to those events; they do not function independently of an interpretation of the events they explain), I want to suggest

that there is an even more fundamental distinction at work here. This distinction, however, is not an ontological one. My discussion of Descartes was intended as a stepping-stone to introduce this point. The two species of explanation that have been identified are the result of what Mark Thornton calls a dualism of points of view. According to Thornton, Descartes was attempting to combine two different points of view under one explanatory model, but because each point of view determines its own model of explanation, they cannot, as Honderich demands and Descartes attempted, be combined. In Thornton's view, rationalizations (or folk-psychological explanations) are developed according to the "first-person" point of view, whereas causal explanations take the "third-person" point of view. Thornton characterizes these two points of view as follows:

The other-person [third-person] approach sees unconscious or non-conscious mental events as the paradigm. These are internal events mediating between stimulus and response; consciousness is an added level of information-processing, a response to one's own basic level information-processing. The first-person approach sees conscious mental events as the paradigm. Unconscious mental events are not ruled out, but their claim to be *mental* depends on their similarity to conscious mental events.<sup>11</sup>

The two points of view, then, are well characterized in terms of the two varieties of knowledge one has of one's body. The first-person point of view is marked by the kind of awareness one has of one's own body as a conscious subject, and so has a predominantly self-descriptive role (for instance, the way my arm feels when it is broken), whereas the third-person point of view is epitomized by the kind of

knowledge one can have of one's body as it appears to others (looking at the results of an X-Ray, for example).<sup>12</sup> Causal explanations fall into the third-person category because as we have seen in the second chapter, the events they involve (neural configurations, for example) can be described only by taking the third-person perspective. This is supported by Hess's discussion of the message in the sand. To the extent that we provide a causal explanation for that event, we are tied to an analysis of its physical features as they are open to public observation, and this prevents us from interpreting the event as if it were produced by the conscious act of an agent.

Although Thornton, like Hess, is willing to make the concession that propositional attitudes are physically realized in the brain or central nervous system, he does not think that this means the "because" of intentional explanation is a causal "because". Like Hess, he characterizes it as the "because" of reason-giving. Thornton identifies three sorts of "because":

Rational: "He thought that the apples were cheap, *therefore*, he bought them" iff "He bought the apples *because* he thought they were cheap."

Causal: (Where C and E are events, and if C then E is a causal law) "If C then E, C *therefore*, E" iff "E *because* C and if C then E."

Logical: " $n^2=4$ , *therefore*,  $n=2$ " iff " $n=2$  *because*  $n^2=4$ ."<sup>13</sup>

The "because" of intentional explanation is a rational "because", and does not express a causal relation,

since rationalizations are of the first-person variety of explanation. This is due to the fact that reason-giving explanations take as their model the pattern of theoretical and practical reasoning engaged in by the agent in question. For, according to Thornton, the application of the categories of psychological explanation in reason-giving is simply the "corollary" of the agent's reasoning.<sup>14</sup> He illustrates this point by showing that the "therefore" of theoretical and practical reasoning can be transformed (as shown above) into the "because" of reason-giving: "That 'because' corresponds to 'therefore' is a familiar fact. We say 'because' when we take the conclusion first and proceed to the premises, we say 'therefore' when we take the premises first and proceed to the conclusion."<sup>15</sup> In Thornton's view then, "The distinctness of the rational 'because' is most easily seen in 'first-person' terms (as one would expect)."<sup>16</sup> Consider his following illustration:

Suppose that I announce that I am leaving at 4 p.m. for the airport. When asked why, I reply, "Because I want to be at the airport two hours before the plane leaves, the plane leaves at 6.45 p.m., and it takes 45 minutes to get to the airport." *Because*. Is it seriously contended that I am saying that I have this desire and I have a couple of beliefs and together these three states will cause me to leave at 4 p.m.? ... Precisely because human beings act for *reasons* (and believe and desire for reasons), once we know what the agent's reasons are we know the explanation of the agent's act.<sup>17</sup>

So rationalizations, on Thornton's account, are formulated from the first-person point of view. They take as their model the (actual or implied) pattern of theoretical and practical reasoning engaged in by the agent and hence are formulated

from the point of view of the subject. I say "implied" because it would be extravagant to demand that every intentional human action involve a piece of conscious reasoning. The point is rather that the *model* of explanation is based on the process of reasoning in general. The connection we might make between Hess's characterization of reason-giving and Thornton's approach is as follows. The mental properties which play a role in a rationalization are ascribed to an agent according to the first-person point of view. Since these properties are ascribed rather than inherent in the agent, we need not call upon an ontological dualism of properties to account for them. Hence, the differentiation between reasons and causes does not depend on an ontological commitment to any form of dualism.

### 3. Thornton, Dennett, and Rationality

The apparent similarity between Thornton's first and third-person points of view and Dennett's intentional and physical stances might lead one to regard Thornton's position as a kind of instrumentalism akin to Dennett's. For both authors regard FP as a kind of idealized theory and neither base the distinction between stances or points of view on an ontological dualism of substances or properties. I believe, however, that this is where the similarities end. For while I agree with Churchland that Dennett's attempt to insulate the



intentional stance from reduction to or replacement by neuroscience is unsuccessful, I think that Thornton's position avoids Dennett's pitfalls. The significant respect in which they differ, as we shall see, lies in their respective accounts of rationality. Before I explain this I would like first to show how Dennett's position plays into Churchland's hands.

In "Three Kinds of Intentional Psychology" Dennett proposes that we can be realists about propositional attitudes; i.e., we can say that there are such things as beliefs and desires because they are not merely heuristic devices *posited* to explain human behaviour. Dennett argues this by calling upon a distinction made by Reichenbach. He proposes that the concept of belief or desire is more like the concept of a center of gravity ("abstracta" or calculation-bound entities) than of a posited theoretical entity ("illata").<sup>18</sup> It would seem that grouping intentional concepts under the rubric of abstracta prevents their being reduced or eliminated since it is presumably only *posited* theoretical entities which are liable to elimination or reduction. Posited theoretical entities, for Dennett, serve a solely heuristic function, so if some other posited entity can do a better job explaining events than the first did, the first can be eliminated without entailing any revisions of our ontology. Since we can say that belief and desire are real (i.e., they are not posited heuristic devices) and can

provide accurate predictions of human behaviour, we should, according to Dennett, continue to exploit those concepts.

I agree with the criticisms Churchland makes of Dennett's appeal to this distinction between abstracta and illata. First, I think that Churchland is correct when he says "Instrumentalism is first and foremost an antirealist position ...."<sup>19</sup> Second, I, like Churchland, cannot see any relevant similarity between folk-psychological concepts and abstracta. Churchland expresses his confusion as follows:

The reality of equators, centers, and rotational axes I am happy to grant. They are all places or loci of some sort that are decisively specifiable by reference to the shape or behavior of the relevant concrete object. But the alleged similarity of these items to beliefs, desires, and other intentional states escapes me entirely. In what respects are they similar and why should they be grouped together in advance of the issue here at stake?<sup>20</sup>

While this confusion might be clarified by a further elaboration on Dennett's part, I suspect that this criticism is compounded by the fact that Churchland could easily point to his network model of the semantics of theories, which would suggest that there is no such distinction between posited and abstract entities, at least not in FP. Churchland's attempt to assimilate the semantics of FP to those of theories generally is intended precisely to show that beliefs and desires are posited theoretical entities. Furthermore, as Churchland correctly points out, the explanatory power of FP does not itself constitute a sufficient reason to warrant its concepts any special status (in terms of their ultimate irreducibility), for according to Churchland a great many theories which have already been

eliminated (e.g., theories of crystal spheres and alchemical essences) also had a capacity to explain and predict events, otherwise they would not have been adopted in the first place.

In my view, Thornton's position is untouched by these difficulties. This is because his account of rationality is sufficiently different from Dennett's to foster a disanalogy between the first-person point of view and the intentional stance. While Dennett appears to think that rationality is equally attributable (i.e., in an equally derivative sense) to humans and other species of animal, Thornton prefers to distinguish between two kinds of rationality. The first kind identified by Thornton corresponds to Dennett's model, which he describes as "basic rationality". Basic rationality is simply optimality of design. A system possessing basic rationality is hard-wired to have beliefs and desires (of a certain kind) which make a contribution to the survival of the organism. Dennett characterizes a basically rational system as one "whose behavior can be predicted by the method of attributing beliefs, desires, and rational acumen according to the following rough and ready principles":

- (1) A system's beliefs are those it *ought to have*, given its perceptual capacities, its epistemic needs, and its biography.
- (2) A system's desires are those it *ought to have*, given its biological needs and the most practicable means of satisfying them.
- (3) A system's behavior will consist of those acts that *it would be rational* for an agent with those beliefs and desires to perform.<sup>21</sup>

While this characterization of rationality is good for animals, Thornton suggests that it is insufficient when it comes to human beings. In our account of human behaviour we must also take into consideration the agent's desire to be rational as well as the desire to "have true beliefs and desires worth having."<sup>22</sup> This requires a stronger normative dimension than functions in basic rationality as it requires that the agent have the ability to evaluate beliefs and desires according to the norms of rationality:

The normativity of higher rationality is manifested in the application of standards to beliefs, desires, reasoning, and actions, and the appraisal of them in the light of those standards. The behaviour of the basically rational conforms to norms of rationality but it is not guided by them; their behaviour is not evaluated by them in the light of those norms. The behaviour of the higher rational is guided by those norms in the sense that it is viewed as open to critical appraisal in the light of those norms.<sup>23</sup>

The phrase "acting for a reason" is consequently ambiguous depending on whether the agent in question has higher or basic rationality.<sup>24</sup> If the act in question is only basically rational, then Dennett's model of explanation is as good as any, but if higher rationality is implied, then one must regard that action from the first-person point of view of the subject; i.e., one must consider how that action would (or did) appear, in the light of the agent's background of belief, desire, and knowledge, to that agent. Hence we arrive at Thornton's general characterization of an agent with higher rationality: "An agent who acts or believes for a reason must see the action or belief as reasonable in the light of his or her reasons."<sup>25</sup> It is because higher

rationality is intimately interwoven with the first-person point of view that rational explanations are to be viewed differently from causal explanations. And since folk-psychological explanations are constituted from the first-person point of view, they are not liable to replacement by a third-person theory, for first and third-person models of explanation are not in conflict.

#### 4. Two Problems: Self-Understanding and the Unity of Science

I would now like to consider two criticisms of Thornton's account of the explanation of human behaviour. The first suggests that since intentional explanation has its foundation in a process of projection and empathy from one's own case to the behaviour of others, it rests on a systematic understanding of oneself, and this understanding will literally constitute a theory. This theory, like any other, is subject to reduction or elimination by better theories. Thus, characterizing FP as predominantly first-person does not shelter it from potential elimination. The second criticism argues that the posited explanatory parallelism (first and third-person; reason-giving and causal) does not satisfy the principle of the unity of science. If the unity of science is a workable hypothesis, then all levels/types of explanation should cohere with all other levels/types in the sense of bearing mutual relations of reducibility and deducibility; i.e., all explanations should be: (1) *reducible*

to lower levels (physics), and (2) *deducible from lower levels*. Since FP does not satisfy these criteria, it ought to be eliminated.

The first criticism can be found in Churchland's "Folk Psychology and the Explanation of Human Behavior," where he criticizes Goldman's alternative account of the explanation of human action.<sup>26</sup> According to this model (which is similar to Thornton's), one uses oneself as a kind of simulation and extrapolates the behaviour of others on the basis of one's own case. In Churchland's view this is quite consistent with the view that FP is a folk theory, so long as one does not regard the empathic account as rendering a nomic framework for the explanation of an action unnecessary. According to Churchland, first-person experience is neither necessary nor sufficient for the understanding of human behaviour. It is not necessary because people who are blind or deaf can know that others have perceptual capacities beyond their own and can know what sort of contribution such capacities make to knowledge and behaviour. Also, people who have never experienced profound grief can nevertheless understand it and effectively predict the behaviour of those inflicted with it.<sup>27</sup> First-person knowledge is not sufficient because, while it may suggest predictions, first-person models are not themselves explanatory of behaviour. To illustrate this Churchland asks us to imagine that we have a miniature model of the universe which yields accurate

predictions of real events. The mere possession of this model and its capacity to provide predictions does not *explain* real events. In fact, it would be a mystery how the model itself works and so we would now have two things requiring explanation instead of just one.<sup>28</sup> In Churchland's view, self-understanding is similarly left a mystery by the empathic account.

If one is to be able to apprehend even the *first-person* intricacies at issue, then one must possess a conceptual framework that draws all the necessary distinctions, a framework that organizes the relevant categories into the appropriate structure, a framework whose taxonomy reflects at least the more obvious of the rough nomic regularities holding across its elements, even in the first-person case. Such a framework is already a theory.<sup>29</sup>

This is quite damaging to Thornton's position since such a theory is in principle subject to reduction or elimination.

I think that Thornton could respond to such a criticism by suggesting that it mistakenly presupposes two things. First (with respect to the mysteriousness of the first-person model), that Thornton's account amounts to a form of abstractionism which leads to insuperable epistemological problems both concerning one's own mind and other minds, and second, that the sense in which the first-person point of view constitutes a theory is the sense required by the thesis of eliminative materialism.

With regard to the first point, it seems that Churchland's characterization of the "empathic" account presupposes (in the absence of a nomic framework) that folk-psychological concepts are derived by means of a process of

abstraction: One notices similar events recurring in one's experience from which one can abstract the idea of a "sort" of thing which is then given a name, e.g., "pain". Churchland suggests, following Wittgenstein, that according to abstractionism, one's vocabulary for sensations would constitute a private language. Wittgenstein's objection to this approach is that in a private language there are no means to distinguish between correct and incorrect applications of a private sensation term and without such means a term is rendered meaningless. Therefore, a private language (or abstractionism) is impossible.<sup>30</sup>

Thornton makes it perfectly clear in the preface to his book that he does not advocate the thesis of abstractionism (nor does Goldman). For Thornton connects the self-ascriptive use of folk-psychological terms to "behavioural and other observable cues" in such a way that thoughts or sensations can be characterized as the "sort" of thought/sensation that would normally occur under certain conditions, and these conditions provide the necessary criteria to determine the proper application of a mental term. While the intrinsic character of thoughts or sensations so characterized is not thereby captured, it "indicates what is from a first-person point of view the intrinsic character of a mental event."<sup>31</sup> This means that the "what it is like to ..." (to use Nagel's phrase) have an experience of a certain sort, or to hold a belief of a certain kind, is not captured



under the behaviourist description. The qualia of sensations and the content of thoughts are only indirectly indicated by a behavioural description, the content (or intrinsic character) is only accessible to the person who is in the state thus specified:

The "sort" in question cannot be further *described* but is *known* [by acquaintance] to the person having the thought or sensation. (If the "sort" can be further described or identified it will be in *other* terms, e.g. [sic.] physical, not in psychological terms.) A person who is told that a sensation of red is the sort of "inner" state which comes from looking at red objects in standard conditions does not know what a sensation of red is, but someone who is told that *she* is in this sort of "inner" state (when *she* is) *does* know what a sensation of red is, i.e. knows its intrinsic nature.<sup>32</sup>

The point is then, that someone who has never felt pain, although he or she might use the word "pain" correctly, will not know what the state of pain is like unless he or she experiences it for him or herself. I could hit that person sharply over the head and, since I know that pain is typically associated with such physical events as sharp blows to the head, correctly tell the victim of my attack that he or she is now in pain and he or she would consequently know something about pain that he or she did not know before, i.e., its intrinsic nature. By tying our knowledge of mental states in this way to behavioural ones, Thornton avoids the position of abstractionism and hence the problem of a private language.

With regard to the second presupposition (that one's self-understanding from the first-person point of view requires a theory), I would like to investigate Thornton's discussion of theories. Thornton asks what is involved in

speaking about a scientific theory of behaviour:

If all that is meant is that (a) FP is (quasi-) *scientific* because it contains at least some causal generalizations, (b) it is *explanatory* because we use FP to explain people's actions, and (c) it is a *theory* because it is a conceptual framework involving a set of inter-related concepts which together constitute a way of viewing human beings ("Any network of concepts is a speculative assumption or *theory*:" Churchland 1984, p. 80), then *of course* folk psychology is a (quasi-) scientific explanatory theory of behaviour.<sup>33</sup>

To infer from this weak sense of *theory* that FP is a theory in the strong sense that (i) mental events are postulated entities used to explain behaviour, and (ii) that FP is in principle liable to replacement by some other theory, is, as Thornton points out, to commit the fallacy of equivocation.<sup>34</sup> For it is possible for FP to be a theory in the weak sense (i.e., a network of concepts used to explain human behaviour) without either (i) or (ii) being true (indeed, Thornton has convincing arguments to suggest that both these assumptions are false.) As Kathleen Wilkes points out, failure to distinguish between theories in the weak sense and theories in the strong sense (her criteria include how systematic each "theory" is), causes the term to lose its robustness: "everything becomes a theory, and that's equivalent to saying that nothing is."<sup>35</sup> And so it appears that Churchland's charge that first-person experience requires a theory is rendered harmless because the sense of "theory" this involves is not sufficiently strong. In fact, Churchland's willingness to refer to any network of concepts as a theory renders the term meaningless because there is no means to distinguish a real theory (with postulated entities) from any old network

of concepts used to explain events. I would now like to turn to the second objection.

As Thornton himself points out, Churchland is a firm believer in the unity of science. This belief is motivated by the drive for a coherent system of beliefs, which includes a coherence among theories. If certain theories or explanations are irreducible, this is a good sign that they should either be altered so that they cohere with the rest of science or eliminated and replaced with theories that do cohere. Since the dualism of points of view and its corresponding parallel models of explanation cannot be combined or reduced one to the other, an ultimate distinction between reasons and causes cannot satisfy the principle of the unity of science.

While Putnam and Oppenheim<sup>36</sup> present the thesis of the unity of science as a *working hypothesis* or *trend* in scientific explanation, as opposed to an actual state of science, they do offer a number of powerful examples as evidence for the thesis. The empirical evidence they offer in support of the principle (e.g., the proposed microreductions of major biological generalizations to elements of the DNA molecule), when coupled with Churchland's eliminativism, renders the principle of the unity of science a powerful thesis.

The proposed coherence of different levels of explanation (formulated in terms of relations of reducibility and deducibility) and the thesis of eliminativism would

appear to render Fodor's arguments against the unity of science ineffective.<sup>37</sup> His attempt to insulate FP from reduction on the grounds of something like Davidson's token physicalism plays right into Churchland's hands. The resulting anomalism of the mental is just the sort of evidence Churchland would point to to show that FP is a false theory and should be replaced by a better one that fits with the rest of science. Any attempt to show that mental events are in some way irreducible plays into the hands of the further thesis (added by Churchland) of eliminativism.

Thornton's own solution to this problem is simply not to believe in the unity of science, at least not as it is formulated above. He calls for a weaker conception of coherence which drops out the conditions of reducibility and deducibility:

Active coherence also obtains when it is clear from the conjunction of different explanatory schemes why the explanation of one phenomenon described in one way lies within the purview of one scheme, and of another phenomenon described in another way within the purview of another scheme. And that is precisely what we do find when we consider FP alongside neurophysiology, biology, and other sciences of the human person. If the third-person perspective of neurophysiology can give an explanation in terms of what is happening when, in first-person terms, we think, remember, experience,<sup>38</sup> desire, etc., then this is "active" coherence even if there is no reduction.

While I agree with what Thornton is proposing, I think that his argument needs reinforcement. For he can only offer an alternative version of coherence or of unity of science if he has independently given us arguments against Putnam and Oppenheim's version of the thesis. Such an argument can be found in chapter 2 of Alan Garfinkel's Forms of Explanation.

Garfinkel suggests that the Oppenheim-Putnam thesis can be captured in a pair of "slogans":

1. for every state, a microstate; and
2. for every microstate, a microexplanation.<sup>39</sup>

Assuming these two principles are true does not, according to Garfinkel, render macroexplanations reducible. He illustrates this with the help of an example. He compares the macroexplanation of the death of a rabbit with its microexplanation:

1. Macroexplanation: The cause of the death of the rabbit was that the fox population was high.
2. Microexplanation: Rabbit  $r$  was eaten because he passed through the capture space of fox  $f$ , at place  $p$ , at time  $t$  ....<sup>40</sup>

Garfinkel points out that the "objects" (i.e., what is explained) of these two explanations do not correspond. In the first explanation the object is *why the rabbit was eaten*, whereas in the second the object is *why the rabbit was eaten by fox  $f$ , at time  $t$* , etc. Garfinkel characterizes the second explanation as "hyperspecific", suggesting that it, unlike (1), does not lend itself to the pragmatic concerns of rabbits (assuming they have any).<sup>41</sup> For what a rabbit would want to know is not why his friend was eaten by fox  $f$ , but why he was eaten at all. As Garfinkel points out, even if the rabbit had not passed through the capture space of that particular fox, given that the fox population was high, it would probably have passed through the capture space of

another fox. Therefore, only the first explanation (in terms of the high fox population) is of any significance to rabbits.

That there exists a significant pragmatic dimension in explanations (for macroexplanations in particular), suggests that different levels of analysis are relative to a body of interests. What Garfinkel proposes is that it is usually the macro level of analysis which meets our interests in the explanation of human behaviour because upper level explanations are more stable under possible perturbations than lower level explanations:

In each case the stability of the upperlevel object under perturbations of the microstate demands an autonomous level of explanation appropriate to its own object. If, for example, the explanation of why my arm moved is that I was shaking hands with someone to whom I was being introduced, this explanation gives us what the underlying neurophysiology does not: a conception of what the allowable variation in the circumstances might have been.<sup>42</sup>

These possible perturbations or "contrast spaces" as Garfinkel also refers to them, are what concern us (and rabbits), for these explanations allow us to understand what might have happened had the circumstances been different.

So, given that levels of explanation are motivated by the interests of those proposing them, and that different levels of analysis of the same event actually have different objects, there is little reason to think that the unity of science as proposed by Putnam and Oppenheim is a workable hypothesis. Having provided independent grounds for rejecting the principle of the unity of science, we can now accept Thornton's weaker conception of coherence as an alternative

proposal. And this alternative model of coherence is not in conflict with our two models of explanation (causal explanations and rationalizations) because causal explanations and rationalizations can cohere in Thornton's sense without bearing mutual relations of reducibility and deducibility.

## 5. Conclusions

Having defended Thornton's account from two objections, I would now like to sum up and draw some conclusions. By conducting a brief examination of Descartes' account of action explanations, I showed that the explanatory parallelism characterized in the second chapter need not rest on an ontological distinction. Since causal explanations and rationalizations are distinguished by means of the properties each type of explanation involves, it would seem that the distinction rests on a dualism of properties. For since causal explanations involve physical properties, it might seem that mental properties are by comparison non-physical. What I tried to establish in this third chapter is that this distinction actually rests on what Mark Thornton calls a dualism of points of view. The difference in properties is not indicative of an ontological difference, but is instead a reflection of how we consider an event, which is not ontological because the "mental" properties are ascribed

according to the first-person point of view to the person in question, they are not non-physical properties somehow inherent in the agent. Depending on whether we consider human action from the first or third-person point of view, we are tied to a particular type of explanation: explanations from the first-person point of view are formulated in terms of higher rationality, while third-person explanations are causal.

Finally, I have shown that the explanatory parallelism between rationalizations and causal explanations is not open to significant objections from the principle of the unity of science. The principle of the unity of science (as formulated by Putnam and Oppenheim, implying reducibility and deducibility between levels of explanation) is not a workable hypothesis. Our interest in a coherent system of beliefs can instead be accounted for by Thornton's alternative proposal, namely that active coherence obtains when neurophysiology explains in its terms what is happening when, from the first-person point of view, we think, feel, etc.

The distinction I proposed between rationalizations and causal explanations has significant consequences for the thesis of eliminative materialism. If, as I have argued, reasons are not the causes of action, then the validity of eliminative materialism is doubtful since, as I demonstrated in Chapter 1, eliminative materialism requires the causal



thesis. While developments in neurophysiology will continue to open new ways of understanding human behaviour, neuroscience will not be in conflict with FP. Such developments can only change the *causal* story behind human action and this FP is silent about.

## REFERENCES

## NOTES FOR INTRODUCTION

1. Keith S. Donnellan, "Reasons and Causes," in Encyclopedia of Philosophy, vol. 7 (New York: The Macmillan Company and The Free Press, 1967), 85-86.
2. D.M. Armstrong, "Acting and Trying," in Armstrong, The Nature of Mind and Other Essays (St. Lucia: University of Queensland Press, 1980), 74.
3. Kathleen Wilkes, "The Brilliance of Common Sense Psychology," presented at McMaster (1988), 29.

## NOTES FOR CHAPTER ONE

1. Paul Churchland, Matter and Consciousness (Cambridge Mass.: MIT Press, 1984), 43.
2. Ibid., 43-44.
3. Ibid., 45.
4. Ibid., 45-47.
5. Ibid., 57.
6. Carl Hempel, "Laws and Their Role in Scientific Explanation," in The Philosophy of Science, ed. R. Boyd, P. Gasper, and J.D. Trout, (Cambridge Mass.: MIT Press, 1991), 301.
7. Churchland, Matter and Consciousness, 56.
8. Ibid., 56.
9. Ibid., 56.
10. Ibid., 58-59.
11. Ibid., 59.
12. Ibid., 59.
13. Ibid., 64.
14. Ibid., 64.
15. Ibid., 64.
16. Ibid., 64-65.
17. Ibid., 66.
18. Ibid., 66.
19. Paul Churchland, "Folk Psychology and the Explanation of Human Behaviour," in J.D. Greenwood, ed., The Future of Folk Psychology (Cambridge: Cambridge

University Press, 1991), 61.

20. Ibid., 53.
21. Ibid., 53.
22. Ibid., 53.
23. Donald Davidson, "Actions, Reasons, and Causes," in Davidson, Essays on Actions and Events (Oxford: Oxford University Press, 1980), 5.
24. Ibid., 12.
25. Ibid., 10.
26. Ibid., 10.
27. Ibid., 15.
28. Ibid., 15.
29. Donald Davidson, "Mental Events," in Davidson, Essays on Actions and Events, 215.
30. Davidson, "Actions, Reasons, and Causes," 16.
31. Ibid., 17.
32. Ibid., 16.
33. Ibid., 16.
34. Ibid., 17.
35. Donald Davidson, "Mental Events," in Davidson, Essays on Actions and Events, 208.
36. W.V. Quine, Pursuit of Truth (Cambridge Mass.: Harvard University Press, 1990), 71.
37. Davidson, "Mental Events," 219.
38. Ibid., 219.
39. Ibid., 223-224.
40. Ibid., 217.
41. Paul Churchland, "The Logical Character of Action Explanations," Philosophical Review 79 (1970), 221.
42. Ibid., 228.
43. Ibid., 228.
44. Ibid., 229.
45. Ibid., 220.
46. Ibid., 230.
47. Ibid., 230.
48. Ibid., 230.
49. Ibid., 228.
50. Ibid., 232.
51. Ibid., 232-233.
52. Hempel's suggestion is that the explanans of a Dravian explanation does not provide a set of conditions which are sufficient for our inferring the occurrence of the explanandum. See "Reasons and Covering Laws in Historical Explanation," in Sidney Hook, ed., Philosophy and History: A Symposium (New York: New York University Press, 1963), 143-163.
53. Churchland, "The Logical Character of Action-Explanations," 233.
54. Conditions (5) and (6) are built in primarily to illustrate the propriety of an action explanation for the case at hand. Churchland's example is that if Peter's finger

crooking were the result of a nervous paralysis, then it would be inappropriate to regard the movement of his finger as an action beckoning someone to join him from across the room. Only if the agent knows how to crook his finger and is able to do it (and conditions 1 through 4 are also met) does "X A-ed because he wanted O," express a true statement. Ibid., 218-220.

55. Ibid., 234.

56. Ibid., 234.

## NOTES FOR CHAPTER TWO

1. Donald Davidson, "Mental Events," in Davidson, Essays on Actions and Events, 210.

2. Ibid., 215.

3. Ted Honderich, "The Argument for Anomalous Monism," Analysis 42 (1982), 61.

4. Ibid., 62.

5. Ibid., 62.

6. Davidson, "Mental Events," 224-225.

7. Honderich, "The Argument for Anomalous Monism," 63.

8. Peter Hess, "Actions, Reasons, and Humean Causes," Analysis 40 (1981), 80.

9. Peter Smith, "Hess on Reasons and Causes," Analysis 41 (1981), 208-209.

10. Honderich, "The Argument for Anomalous Monism," 64.

11. Ibid., 64.

12. Peter Smith, "Bad News for Anomalous Monism?," Analysis 42 (1982), 222.

13. Ibid., 222.

14. Ibid., 223.

15. Ted Honderich, "Anomalous Monism: Reply to Smith," Analysis 43 (1983), 148.

16. Ibid., 148.

17. Peter Smith, "Anomalous Monism and Epiphenomenalism: A Reply to Honderich," Analysis 44 (1984), 83.

18. Ibid., 84.

19. Ibid., 84.

20. Ibid., 85.

21. Ted Honderich, "Smith and the Champion of Mauve," Analysis 44 (1984), 86.

22. Ibid., 88-89.

23. Davidson, "Mental Events," 214.

24. Peter Hess, Thought and Experience (Toronto: University of Toronto Press, 1988), 64-65.

25. Ibid., 66.

26. Ibid., 67.
27. Ibid., 67.
28. Ibid., 68.
29. Ibid., 70.
30. Ibid., 71.
31. Ibid., 72.
32. Donald Davidson, "Actions, Reasons, and Causes," in Davidson, Essays on Actions and Events, 5.
33. D.M. Armstrong, "Acting and Trying," in Armstrong, The Nature of Mind and Other Essays (St. Lucia: University of Queensland Press, 1980), 76.
34. See Roderick Chisholm, "The Descriptive Element in the Concept of Action," *Journal of Philosophy* 61 (1964), 615.
35. Armstrong, The Nature of Mind and Other Essays, 76-77.
36. Ibid., 79.
37. Mark Thornton, Folk Psychology: An Introduction (Toronto: Canadian Philosophical Monographs, 1989), 16-17.
38. Stoutland characterizes the oblique theory of causation as follows: "Attitudes cannot be causes because they play a role in the agent's practical reasoning, for attitudes are causes only because they are tokens of physical types. The most we could say is that the attitudes which played a role in the agent's reasoning also happened to be tokens of causal (physical) types. But then the caused act can in no sense be a *conclusion* of the agent's reasoning, for though the act which was caused may be the act which was the conclusion, its being caused was not explained by its being the conclusion." See Frederick Stoutland, "Oblique Causation and Reasons for Action," *Synthese* 43 (1984), 366, 351-367.

### NOTES FOR CHAPTER 3

1. Ted Honderich, "Smith and the Champion of Mauve," *Analysis* 44 (1984), 89.
2. Owen J. Flanagan, Jr., The Science of the Mind (Cambridge: MIT Press, 1984), 2.
3. Ibid., 2-3.
4. Ibid., 3-4.
5. Descartes, Meditations, in The Philosophical Writings of Descartes vol II, trans. by John Cottingham, Robert Stoothoff, and Dugald Murdoch, (Cambridge: Cambridge University Press, 1984), 58.
6. Descartes, Passions of the Soul, in Descartes: Selected Philosophical Writings, trans. by Cottingham et al., (Cambridge: Cambridge University Press, 1988), 224-225.
7. Meditations, in Cottingham et al. (1984), 60.
8. Passions, in Cottingham et al., (1988), 233.

9. Ibid., 232-233.
10. W.V. Quine, Pursuit of Truth (Cambridge Mass.: Harvard University Press, 1990), 44.
11. Mark Thornton, Folk Psychology: An Introduction (Toronto: Canadian Philosophical Monographs, 1989), 3.
12. Ibid., 44.
13. Ibid., 23-24.
14. Ibid., 23.
15. Ibid., 23.
16. Ibid., 24.
17. Ibid., 24.
18. Daniel Dennett, "Three Kinds of Intentional Psychology," in The Intentional Stance (Cambridge: MIT Press, 1987), 52-53.
19. Paul Churchland, "Folk Psychology and the Explanation of Human Behavior," in J.D. Greenwood, ed., The Future of Folk Psychology (Cambridge: Cambridge University Press, 1991), 66.
20. Ibid., 66.
21. Dennett, "Three Kinds of Intentional Psychology," 49.
22. Thornton, Folk Psychology: An Introduction, 30.
23. Ibid., 31.
24. Ibid., 31.
25. Ibid., 31.
26. For a current account of Goldman's position see Alvin I. Goldman, "The Psychology of Folk Psychology," forthcoming, (Cambridge: Cambridge University Press, 1992). Churchland is referring to Goldman's "Interpretation Psychologized," *Mind and Language* 4 (1989), 161-185.
27. Churchland, "Folk Psychology and the Explanation of human behavior," 59.
28. Ibid., 60.
29. Ibid., 60.
30. Paul Churchland, Matter and Consciousness, (Cambridge: MIT Press, 1984), 54-55.
31. Thornton, Folk Psychology: An Introduction, vii.
32. Ibid., vii-viii.
33. Ibid., 8.
34. Ibid., 8.
35. Kathleen Wilkes, "The Brilliance of Common Sense Psychology," presented at McMaster (1988), 7.
36. Paul Oppenheim and Hilary Putnam, "Unity of Science as a Working Hypothesis," in The Philosophy of Science ed. R. Boyd, P. Gasper, and J.D. Trout, (Cambridge: MIT Press, 1991), 405-428.
37. Jerry Fodor, "Special Sciences," in Boyd et al., The Philosophy of Science, 429-441.
38. Thornton, Folk Psychology: An Introduction, 48.
39. Alan Garfinkel, "Reductionism," in Boyd et al., The Philosophy of Science, 445.
40. Ibid., 446.

41. Ibid., 447-448.
42. Ibid., 452.

## WORKS CITED

- Armstrong, David Malet. The Nature of Mind and Other Essays. St. Lucia: University of Queensland Press, 1980.
- Boyd, R., P. Gasper, and J.D. Trout, ed. The Philosophy of Science. Cambridge: MIT Press, 1991.
- Chisholm, Roderick. "The Descriptive Element in the Concept of Action," *Journal of Philosophy* 61 (1964), 613-624.
- Churchland, Paul. Matter and Consciousness. Cambridge: MIT Press, 1984.
- . "The Logical Character of Action Explanations," *Philosophical Review* 79 (1970), 214-36.
- . "Folk Psychology and the Explanation of Human Behavior," in Greenwood (1991), 51-69.
- Davidson, Donald. Essays on Actions and Events. Oxford: Oxford University Press, 1980.
- . (1963). "Actions, Reasons, and Causes," in Davidson (1980), 3-19.
- . (1970b). "Mental Events," in Davidson (1980), 207-25.
- Dennett, Daniel. The Intentional Stance. Cambridge: MIT Press, 1987.
- Descartes, Rene. Meditations.
- . Passions of the Soul.
- Donnellan, Keith S. "Reasons and Causes," in Encyclopedia of Philosophy, vol. 7. New York: The Macmillan Company and The Free Press, 1967, 85-86.
- Flanagan, Owen J., Jr. The Science of the Mind. Cambridge: MIT Press, 1984.
- Garfinkel, Alan. "Reductionism," in Boyd et al., 443-459.
- Goldman, Alvin I. "The Psychology of Folk Psychology," forthcoming. Cambridge: Cambridge University Press, 1992.



- Greenwood, John, D., ed. The Future of Folk Psychology. Cambridge: Cambridge University Press, 1991.
- Hempel, Carl. "Reasons and Covering Laws in Historical Explanation," in Sidney Hook, ed., Philosophy and History: A Symposium (New York: New York University Press, 1963), 143-163.
- . "Laws and Their Role in Scientific Explanation," in Boyd et al., 299-315.
- Hess, Peter. "Actions, Reasons, and Humean Causes," *Analysis* vol. 40 No. 2 (1981), 77-81.
- Hess, Peter. Thought and Experience. Toronto: University of Toronto Press, 1988.
- Honderich, Ted. "The Argument for Anomalous Monism," *Analysis* vol. 42 No. 1 (1982), 59-64.
- . "Anomalous Monism: Reply to Smith," *Analysis* vol. 43 No. 3 (1983), 147-149.
- . "Smith and the Champion of Mauve," *Analysis* vol. 44 No. 2 (1984), 86-89.
- Oppenheim, Paul and Hilary Putnam, "Unity of Science as a Working Hypothesis," in Boyd et al., 405-428.
- Quine, W.V. Pursuit of Truth. Cambridge: Harvard University Press, 1990.
- Smith, Peter. "Hess on Reasons and Causes," *Analysis* vol. 41 No. 4 (1981), 206-209.
- . "Bad News for Anomalous Monism?" *Analysis* vol. 42 No. 4 (1982), 220-224.
- . "Anomalous Monism and Epiphenomenalism: A Reply to Honderich," *Analysis* vol. 44 No. 2 (1984), 83-86.
- Stoutland, Frederick. "Oblique Causation and Reasons for Action," *Synthese* 43 (1980), 351-367.
- Thornton, Mark. Folk Psychology: An Introduction. Toronto: Canadian Philosophical Monographs, 1989.
- Wilkes, Kathleen. "The Brilliance of Common Sense Psychology," presented at McMaster (1988).