

CONTENTS

Abstract

Introduction

Chapter One

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Conclusion

Reference List

**Master Thesis**

**Mechanistic and Psychological explanations of behaviour :  
the problem of mental causation**

by

**Ronald Davies**

**Supervisor : Paul Taylor**

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

ABSTRACT  
CONTENTS

Abstract	ii
Introduction	1
Chapter One	3
Chapter Two	13
Chapter Three	21
Chapter Four	34
Chapter Five	49
Conclusion	54
Reference List	56

## ABSTRACT

*The problem we are going to investigate is that of mental causation.*

*Human beings often explain to each other why they do what they do by citing their reasons, that is, mental states like beliefs and desires. Explanations of behaviour which invoke mental causes are known as psychological explanations. Mechanistic explanations, on the other hand, make sole reference to the neurophysical going-on in our bodies causally to explain behaviour. This kind of explanation has become entrenched in a naturalistic attitude towards any phenomenon which requires causal explanation. This naturalistic attitude, furthermore, holds explanation by reference to natural causal law in especially high esteem.*

*Our problem is going to be how, given the availability of mechanistic explanation, psychological explanations can still apply. That is, given the availability of explanation of human behaviour, by reference to physical ( or neurophysiological) laws, can mental states still constitute causes of behaviour. This problem has become very well defined since a series of papers authored by Donald Davidson who showed, very clearly, how the problem arises given the following three principles : (i) the Principle of Causal Interaction which holds that mental states interact causally with physical states. (2) the Principle of the Nomological Character of Causality which states that events related as cause and effect fall under strict laws and (3) the Principle of the Anomalism of the Mental which states that there are no strict laws, known or unknown, on the basis of which the mental can be predicted and explained.*

*A central feature of Davidson's philosophy is (3) and one of the most significant sections of this thesis will concern why Davidson formulates (3) as a principle. We need to understand Davidson's arguments for (3) if we are to gain a proper understanding of how he derives the token-identity thesis from (1), (2) and (3). This is important for our purpose here since the latter part of this essay will be an examination of the token-identity thesis as a proposed solution to the problem of mental causation.*

The central problem in the philosophy of mind has always been to give a clear account of what a mind is. It is in pursuing such an account, as we shall see in chapter one, that the problem of mental causation arises. Here we will survey two theories of mind, namely, Dualism and Behaviourism in order to see how the problem of mental causation arises. Although rejected as false we will see that this rejection entails that we accept certain principles which any theory of mental causation must subscribe to. The latter of these is that of causal closure and we will examine the broad impact this principle has had on the philosophy of mind. We will see that accepting it leads to a search for a naturalised solution to the problem of mental causation and we will also see how the mechanistic form of behaviourism is an example of a naturalised explanation of behaviour.

Chapter two deals with a very explicit attempt to provide a naturalised solution to the problem of mental causation, namely, the type-identity theory. Here we examine the problems facing the type-identity theory in order to see how they have shaped proposed solutions to the problem of mental causation. One of these, that of extrinsic reduction, will, in the course of this essay, come to be seen as some kind of a principle which any proposed solution must respect. In chapter two we also look at why some philosophers think that consciousness cannot ever be successfully naturalised.

## INTRODUCTION

Psychological explanations of behaviour contend that our reasons, that is, mental states like beliefs and desires, can be, and often are, causes of behaviour. Mechanistic explanations, on the other hand, make sole reference to the neurophysiological goings-on in our brains when causally explaining behaviour. The problem we are going to investigate in this essay is that of showing how, given the availability of mechanistic explanations, psychological explanations of human behaviour can still apply. This is going to be our problem in this essay: is mental causation possible given the availability of mechanistic explanations?

The central problem in the philosophy of mind has always been to give a clear account of what a mind is. It is in providing such an account, as we shall see in chapter one, that the problem of mental causation arises. Here we will review two theories of mind, namely, Dualism and Behaviourism in order to see how the problem of mental causation arises. Although rejected as false we will see that this rejection entails that we accept certain principles which any theory of mental causation must subscribe to. One such principle is that of causal closure and we will examine the broad impact this principle has had on the philosophy of mind. We will see that accepting it leads to a search for a naturalised solution to the problem of mental causation and we will also see how the mechanistic form of explanation is an example of a naturalised explanation of behaviour.

Chapter two deals with a very explicit attempt to provide a naturalised solution to the problem of mental causation, namely, the type-identity theory. Here we examine the problems facing the type-identity theory in order to see how they have shaped proposed solutions to the problem of mental causation. One of these, that of variable realisation, will, in the course of this essay, come to be seen as some kind of a principle which any proposed solution must respect. In chapter two we also look at why some philosophers think that consciousness cannot ever be successfully naturalised.

We begin our exposition of psychological explanations of behaviour with chapter three. Our aim here is to see exactly what it is that psychological explanations have in common and, more importantly, to give an exposition of what many philosophers feel is a naturalised solution to the problem of mental causation, namely, Donald Davidson's "Anomalous Monism." Given Davidson's conception of the nature of mental states and the conception of mechanistic explanations developed in chapter one we will thus have a better appreciation of the problem of mental causation facing us today.

Chapter four deals with the main problem facing anomalous monism and, that is, that under anomalous monism, mental states are causally inefficacious and causally irrelevant. The charge is that under anomalous monism, the mental is epiphenomenal. Some replies will also be considered and it will briefly be argued that these replies are unsuccessful.

In chapter five we deal with Davidson's (1993) defence of anomalous monism against the epiphenomenalist charge. Davidson claims that anomalous monism coupled with the supervenience relation can deflect the epiphenomenalist charge. We briefly examine this claim.

My conclusion will be that though the problem of mental causation remains, anomalous monism has afforded us new insights into the nature of our mentality.

Descartes' epistemic starting point was that, whatever else may be doubted, what is absolutely certain, from the fact that he can doubt, is that he exists. Of the nature of his body, Descartes understands "whatever has determinate shape and a definable location and can occupy a space in such a way as to exclude any other body." By the evil demon hypothesis the existence of his body is doubtful and the only attribute which Descartes finds "really inseparable from who he is... is thought. In answer to the question, "but what, then, are I?" Descartes answers: "A thing that thinks." What is a thing that thinks? That is to say, a thing that doubts, perceives, affirms, denies, wills, does not will, that imagines also, and which feels." (p117). In the Sixth Meditation Descartes asserts that "my essence consists solely in the fact that I am a thinking thing... a thinking, non-extended thing" while his body "is an extended, non-thinking thing." The body is, furthermore, by its nature divisible whereas the mind is completely indivisible. This argument on its own persuades Descartes that thought is completely distinct from the body. According to Cartesian Dualism, then, man is composed of two parts, (1) a material body and (2) an immaterial or spiritual mind or soul. These immaterial mental states are fully capable of interacting with our bodies. They cause bodily behaviour and our bodies are said to cause certain mental states. For instance, burning your finger gives rise to the mental state pain. The problem confronting the dualist has always been how to explain this mental causation. How does an immaterial mental state interact causally with physical states of the body? If these two entities are so entirely different how is causal interaction between the two possible? This problem can be clarified by considering "the principle of causal closure of the physical

## CHAPTER 1

world." (p117). In the Sixth Meditation Descartes asserts that "my essence consists solely in the fact that I am a thinking thing... a thinking, non-extended thing" while his body "is an extended, non-thinking thing." The body is, furthermore, by its nature divisible whereas the mind is completely indivisible. This argument on its own persuades Descartes that thought is completely distinct from the body. According to Cartesian Dualism, then, man is composed of two parts, (1) a material body and (2) an immaterial or spiritual mind or soul. These immaterial mental states are fully capable of interacting with our bodies. They cause bodily behaviour and our bodies are said to cause certain mental states. For instance, burning your finger gives rise to the mental state pain. The problem confronting the dualist has always been how to explain this mental causation. How does an immaterial mental state interact causally with physical states of the body? If these two entities are so entirely different how is causal interaction between the two possible? This problem can be clarified by considering "the principle of causal closure of the physical

See Jaegwon Kim's (1985) 'The Myth of Nonreductive Physicalism' for a further discussion of this principle.



## DUALISM, BEHAVIOURISM AND THE PROBLEM OF MENTAL CAUSATION

Descartes epistemic starting point was that, whatever else may be doubted, what is absolutely certain, from the fact that he can doubt, is that he exists. Of the nature of his body Descartes understands "whatever has determinable shape and a definable location and can occupy a space in such a way as to exclude any other body." By the evil demon hypothesis the existence of his body is dubitable and the only attribute which Descartes finds totally inseparable from who he is, is thought. In answer to the question, "but what, then, am I?" Descartes answers: "A thing that thinks. What is a thing that thinks? That is to say, a thing that doubts, perceives, affirms, denies, will, does not will, that imagines also, and which feels." (p107). In the Sixth Meditation Descartes asserts that "my essence consists solely in the fact that I am a thinking thing ..., a thinking, non-extended thing" while his body "is an extended, non thinking thing." The body is, furthermore, by its nature divisible whereas the mind is completely indivisible. This argument on its own persuades Descartes that the mind is completely different from the body. According to Cartesian Dualism, then, man is composed of two parts, (1) a material body and (2) an immaterial or spiritual mind or soul. These immaterial mental states are furthermore said to interact with our bodies. They cause bodily behaviour and our bodies are said to cause certain mental states. For instance, burning your finger gives rise to the mental state pain. The problem confronting the dualist has always been how to explain this mental causation. How does an immaterial mental state interact causally with physical states of the body? If these two entities are so entirely different how is causal interaction between the two possible? This problem can be clarified by considering "the principle of causal closure of the physical domain."<sup>1</sup>

---

<sup>1</sup> See Jaegwon Kim's (1989) "The Myth of Nonreductive Monism" for a fuller discussion of this principle.

According to this principle the physical world is “closed” in that all causes of physical events will be found to be other physical events. We don’t need to go outside the physical world to look for causes. The same applies to human behaviour. All scientific evidence available to us suggests that bodily movements are controlled and caused by neural events in the brain. If we are to have immaterial mental causation then it must presumably take place in the brain. Judging by the neurological evidence the human body is a causally closed system in that all causes of bodily movement will be neurophysiological events. There is no space available that only an immaterial mental state can fill.

Accepting the principle of causal closure has certain consequences for how a theory of mind can be conceived and places limits on the theoretical entities these theories posit.

Firstly, accepting the principle of causal closure is to accept that only physical events can cause physical events. There cannot thus be immaterial mental causes. Secondly, and more importantly, if we accept this principle we implicitly accept that some naturalistic view holds true. Thus, Gilbert Ryle (1949) argues for a behaviourist solution to the problem of mental causation. Behaviourists deny that the mind is a thing and consequently assert that there is no mind-body problem to be solved.

According to the behaviourist, when we ascribe mental states to others we are doing no more than describing their behaviour and their dispositions to behave. When we, for example, say “he is in pain”, all we mean is that they have the disposition to display pain behaviour. On this account, to have a disposition is for certain “iffy” statements about one’s behaviour to be true (Smith and Jones; 1986; p143 - 51). Consider the behaviourist analysis of Smith and Jones’s example:

Jack’s belief that it is about to rain is simply a matter of it being true that: if conditions A obtain, Jack will get in the washing; if conditions B obtain, Jack will take the umbrella; if conditions C obtain, Jack won’t water the garden; and so on.

The problem with this account of mental states is that it does not leave sufficient difference between the mental state and the behaviour for the mental state causally to explain the behaviour. If we look at the first conditional of the example then we see that citing the behavioural disposition has no real causal explanatory value. Consider:

It is true that, if conditions A obtain, Jack would get in the washing because Jack believes it is about to rain.

If we now substitute back into the behaviourist analysis we get:

It is true that, if conditions A obtain, Jack would get in the washing because if circumstances A were to obtain, Jack would get in the washing; if conditions B...

What we want to explain is Jack's behavioural tendency to get in the washing when it rains and we want to explain this behaviour by appealing to his belief. But, on the behaviourist account, beliefs just are dispositions or tendencies to behave. All we are doing here when we cite the belief to causally explain the behaviour is citing the tendency to perform that behaviour. The problem with behaviourism is that mental causation has disappeared. "Iffy" statements about behaviour do not amount to causes.

The failure of behaviourism as a solution to the mind-body problem does not, though, entail that all naturalistic theories will be false, nor does it entail a return to Dualism. As yet we have not developed what we mean by a naturalistic view. To this task we now turn.

## **NATURALISM AND THE CAUSAL-MECHANISTIC VIEW EXPLAINED**

Naturalism is the view that all things physical will have physical explanations and "in the philosophy of mind is the thesis that every property of mind can be explained in broadly physical terms" (McGinn; 1988; p23).

Naturalism is not though peculiar to any particular science even though physics is viewed as the basic science. Rather, it is an attitude that is pervasive across all sciences in that the principle of causal closure is included as a guiding principle. All sciences are naturalistic, or rather materialistic, in that only physical causes are admitted. The further physicalist claim that the special sciences are reducible to physics must be distinguished from the sense in which all sciences are materialistic. This physicalist claim we will turn to shortly.

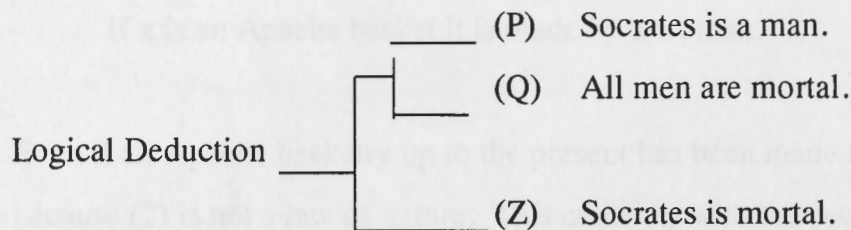
Physics, though, is taken to be complete in that "all physical events are determined, or have their chances determined by prior physical events according to physical laws" (Papineau; 1993; p16). All physical phenomena have a perfectly acceptable causal explanation in terms of natural physical laws. This is the crucial claim of the naturalist which we have to bear in mind.

Callebaut (1995), for instance, regards naturalism "as the view that whatever exists or happens in the world is susceptible to explanation by natural scientific methods, it denies that there is or could be anything which lies in principle beyond the scope of scientific explanation" (pxv). It also implies that what is discoverable in science (matters of fact) are as relevant for philosophy as for science. Thus certain philosophical views may be ruled out because of scientific discoveries. Before returning to the philosophy of mind we will first spend some time in the philosophy of science in order to arrive at a general idea of what scientific explanation involves.

The aim of scientific explanation is often characterised as answering to why-questions. For instance, why did this or that event occur? And more often than not the answers to these questions are framed in terms of what caused the event. But the scientist does not only want to find out what caused the event; insofar as he wants to predict similar events occurring he concerns himself with formulating laws which covers those events and by which he will be able to predict those events given the causes. Specifically he wants to formulate and discover laws of nature. Laws of nature are regularities that hold throughout the natural

world and a law statement is a statement that such a regularity holds. Hempel and Oppenheim (1948) formulated what is known as the deductive-nomological model of scientific explanation. Explanation is here divided into two constituents : (1) the explanandum, which is the statement describing the phenomenon to be explained, and (2) the explanans, which is the statement thought to account for the phenomenon. In short, this kind of explanation explains by subsuming the explanandum under a general law. Explanation is furthermore regarded as a deductive argument, the premises containing at least one natural law.

Consider the following example:



Given (P) and (Q), a natural law, we can logically deduce (Z). According to Hempel and Oppenheim, this is the basic pattern of scientific explanation and deductive-nomological explanation is causal explanation.

Accepting Nelson Goodman's account of what is meant for a lawlike sentence genuinely to rank as a law, Hempel and Oppenheim accept that (1) lawlike sentences have a general form, that is , "All x's are y's", and (2) that they also have a conditional form which, is to say, they support counterfactuals as in "If Socrates is a man then Socrates is mortal."

Salmon ( 1992 ) refers to the deductive-nomological model as the covering law model.

Consider two of his examples:

- (1) No signal travels faster than the speed of light

(2) All Apache basketry is made by women.

Laws of nature, according to Salmon, following Goodman again, have two capabilities. As stated, they support counterfactuals, contrary-to-fact conditionals.

Thus, we can assert

(3) : If x is a signal it cannot travel faster than the speed of light.

We can assert (3) because (1) is a law and laws are general in that they hold without exception. Now consider:

If x is an Apache basket it is made by a woman.

Even if all Apache basketry up to the present has been made by women we cannot assert (4) because (2) is not a law of nature. This connects with the second capability of laws of nature, which is that they support modal statements of physical necessity and impossibility.

It is not impossible for an Apache basket to be made by a man whereas it is impossible for a signal to travel faster than the speed of light. Rescher (1970) states that scientific explanation derives its modal aspect of necessity directly through its use of laws. When we explain by subsuming under laws we assert that things must happen the way they do and could not have happened otherwise. It is this subsumption under laws which endows scientific explanation, according to Rescher, with this modal aspect. What this modal aspect of necessity means is that things necessarily had to turn out this way rather than that way. Explanation by this kind of scientific law reveals a very deterministic picture of the world. Salmon identifies this as the causal mechanistic tradition because it attempts to discover the causal mechanisms which account for the events we want to explain. Scientific causal laws must also support counterfactuals. They must apply, and this is crucial to understanding Davidson and McGinn's arguments, "to all possible cases, specifically including unexamined ones, future ones and even hypothesized ones" (Rescher; 1970; p13). It is this

modal aspect that qualifies a generalization as lawlike. Generalizations which are accidental, such as (2), do not have this aspect of nomic necessity.

A causal law will then have the following form: If conditions x obtain, then y will obtain.

As Rescher asserts, philosophers and scientists operate with a Principle of Causality - which states that every event has a cause. The principle of the Nomological Character of Causality, espoused by philosophers like Davidson, states that events related as cause and effect fall under strict deterministic laws which, in view of what has been said above, means that, when applied to human behaviour, human behaviour is seen as completely determined by antecedent physical conditions and couldn't have been otherwise than what it is.

It should be noted that the Nomological principle as well as the principle of Causality are guiding principles of science in that it is by such nomological principles that science predicts and explains. Physics, being the study of the fundamental elements in our universe, holds a special position amongst the sciences. We must now look at the reductionist claims that have been advanced in regard to physics as the "bedrock" of the sciences.

## **PHYSICAL LAWS AND REDUCTIONISM**

As indicated earlier physics is the basic science in that it investigates and attempts to predict the behaviour of the fundamental elements, that is, atoms, neutrons, and other subatomic particles, out of which the world is constructed. The reductionalist claims that all objects, processes, events or properties are nothing more than the physical ingredients that compose them and according to Trout (1992) "[our] best scientific theories would seem to support such proclamations."

As Trout points out, the reductionist program can be understood in various ways. One version, for instance, claims that the vocabulary of the special sciences will reduce to or be

eliminated in favour of the more basic vocabulary of physics. Another version, which Trout identifies as law reductionism, argues that the laws of special sciences can be explained in terms of the laws of physics. It is this version which will be under consideration here. The reductionist programme is understood here in the following way: psychological causal statements (which we will define later) are pseudo-causal statements and should be eliminated in favour of physical causal statements.

Fodor (1974) conceives a successful reduction in the following way: Suppose the following law to be a law of the special science S :

- (1)  $S_{1x} \Rightarrow S_{2y}$  : reads : all events which consist of X's being  $S_1$  , bring about events which consist of y's being  $S_2$ .

Fodor individuates a science by reference to it's typical predicates. Thus if  $S_1$  and  $S_2$  are mental predicates they are not predicates of basic physics. For S to be successfully reduced to physics, argues Fodor, formulae 2 and 3 should be laws.

(2a)  $S_{1x} \Rightarrow P_{1x}$

(2b)  $S_{2y} \Rightarrow P_{2y}$

(3)  $P_{1x} \Rightarrow P_{2y}$

$P_1$  and  $P_2$  are predicates of physics and (3) is a physical law. Formulae (2a) and (2b) are 'bridge' laws and their characteristics feature is that they contain predicates of both the special science and physics. Bridge laws link predicates of the special science to predicates of basic physics. Of the connective " $\Rightarrow$ ", which is read as 'causes', Fodor asserts that it must be transitive, meaning that the relation carries over from one science to another because the reduction of psychology to physics proceeds via bridge laws that connect the predicates of



psychology to the predicates of neuroscience; neuroscience, in turn, being reducible to physics.

Suppose then that S represents psychology. Let S\* represent neuroscience. Then, argues Fodor, psychophysical bridge laws will obtain which connect psychological predicates with neuroscientific predicates. There will then also be bridge laws which connect neuroscientific predicates to physical predicates.

Our first observation here is that neuroscience, being the intervening science, will contain causal mechanistic laws. Construed generally, neuroscience is the study of the structures which make up the central nervous system and the brain. Being physical systems, they will obey the causal mechanistic laws which obtain here. The causal mechanist's claim is that states of the central nervous system, including the brain, are firstly the effects of other neurophysiological states. Secondly, these states are accompanied, as a matter of scientific law, by particular mental processes. Thirdly, all bodily movements are sufficiently caused by physical and neurophysiological states. The crucial claim, to borrow from Papineau, is that the completeness of physics (and neuroscience) "leaves no room for mental differences, or any other differences, to make a difference to physical consequences, once physical - antecedents are given" (p18).

The "worry" about this naturalist view being true is that it leads to a mechanistic view of human behaviour according to which the behaviour of human beings is as determined as the behaviour of billiard balls on a billiard table. Consider the following two examples: (A) a doctor taps me below the knee to check my reflexes, and my leg rises. (B) I have a desire to exercise my leg and a belief that raising it is a form of exercise, thus I raise my leg.

The proponent of the mechanistic explanation will cite as causes, in both cases, those neurophysiological changes which are necessary for my leg rising. These neurophysiological changes are furthermore said to be the effects of other neurophysiological changes. And so on. Mechanistic explanations of behaviour seem to

imply a hard deterministic view of human behaviour. In Malcolm's words: "The human body is assumed to be as complete a causal system as a gasoline engine" (1968; p127). It is thus postulated that if we know the exact neurophysiological condition of a person we will be able to predict exactly what that person's bodily movements will be regardless of what that person believes, desires or intends at that moment. Neurophysiological explanation makes no distinction between "action" and bodily movements; here there is no action, only bodily movements. This kind of explanation will thus employ strict deterministic laws that connect neurophysical states or processes with movements (Malcolm; 1968; p128). We can formulate the mechanistic claim as follows:

MC: Human behaviour can be adequately predicted and explained using strict deterministic causal laws derived from a complete neurophysiological science.

There is of course the widely held view that mechanistic and teleological or psychological explanation is compatible. Compatibilists generally attempt to show that the fact that behaviour is determined does not mean that it is not free. Though this is an important issue in philosophy we will not deal with it here as it is not central to our purposes. Anomalous monism, the view to be examined here, is though an attempt to show such compatibility

Our second observation is that if mental states are real states which play a real causal role, then an identification between mental states and physical states is inevitable. Psychological explanation, if it is to remain a viable causal explanation of behaviour must be naturalized. Specifically, consciousness must be naturalized.

We turn now to an attempt to naturalize consciousness and an examination of the problems involved in such attempts.

## CHAPTER 2

In the foregoing chapter we saw that a need exists for mental phenomena as well as consciousness to be naturalized if mental phenomena are to play a real causal role in producing behaviour. As an example of an attempt to naturalize mental phenomena and consciousness we will review the type-identity thesis as well as the problems attending it. We will also look at McGinn's arguments which purport to show why, though mental phenomena have a real causal role, we won't be able ever successfully to naturalize consciousness.

We will first, though, offer a common characterization of mental phenomena.

Philosophers have generally divided mental phenomena into two categories: (1) sensations, and (2) propositional attitudes. With regard to category (1) we find bodily feelings like pains, aches, tickles. We also find perceptual experiences like seeing a reddish after-image, hearing a ringing noise, etc. As McGinn (1982) states, there is an importance difference between these two kinds of sensation. With perceptual experiences we distinguish between the experience and what it is an experience of. Unlike bodily feelings, perceptual experiences, like the propositional attitudes, have an intentional object. We do though classify these sensations together because they both have a phenomenological qualitative character in that there is something it is like for the subject who is experiencing the sensation.

Frank Jackson (1982), amongst others, identifies himself as a 'qualia freak' and argues that a naturalized psychology would leave out the phenomenological qualitative character of sensations. For instance, it wouldn't include the particular stabbing hurtfulness of a particular headache, the smell of a TexMex pizza from St. Elmo's, etc. Jackson's point is that no matter how far we advance with our knowledge of physics or neuroscience it would still leave out the qualitative character of sensory experience.

The class of mental phenomena categorized as sensations won't concern us here, but whatever conclusions we arrive at in regard to the propositional attitudes will have similar implications for sensations.

Turning to propositional attitudes - category (2) - these mental states have a propositional content in that they involve a 'that-clause' such as : Jack believes that it is going to rain. They include cognitive states like beliefs and affective attitudes - "for example, desiring or intending that you get an apple, and fearing that you will be run over" (McGinn; 1982; p8).

Propositional attitudes involve two factors : the type of attitude and the proposition onto which it is directed. Mental states like belief and desire, insofar as they have a propositional content, are directed towards the world. They are not just beliefs and desires but beliefs about this and desires for that. They are intentional states in that they are directed onto objects or states of affairs in the world. Furthermore - here I am following Fodor (1987) - they are semantically evaluable, that is, we evaluate them in terms of their relations to the world. For instance Jack's belief that it is raining outside is true or false depending on a certain state of affairs, that is, whether or not it is indeed raining. One may, on the other hand, desire that a certain state of affairs come about, for instance, that it rains.

We are now in a position to review an attempt to naturalize consciousness which arose out of the failure of dualism and behaviourism to solve the problem of mental causation.

## **THE TYPE-IDENTITY THESIS**

U. T. Place (1956) argues that the question whether or not consciousness is a brain process is an empirical question and a reasonable scientific hypothesis. In this essay Place laid the groundwork for what came to be known as the type-identity thesis. Smart (1959), taking up Place's cause, identified mental states with brain states. The identity, states Smart, must be taken in the sense of strict identity, so that whenever a person is in mental state x that person is in brain state y. Types of mental states are identified with types of brain states.

Three compelling objections followed which showed such a type-identity thesis to be untenable.

1) According to Saul Kripke (1972), a rigid designator is a term which picks out the same object in all possible worlds and the terms "pain x" and "brain state y" function in this way. They pick out their object essentially because they pick them out by essential properties. Suppose now that pain is identified with the firing of c-fibres, a neural event in the brain. If such an identity is proposed, then according to Kripke, the identity must be one of necessity so that necessarily, that is, in all possible worlds, every time one is in pain one's c-fibres are firing. But one could surely imagine the pain to be present without the firing of c-fibres as well as the firing of c-fibres without pain. There is a strong intuition to say that such an identity is contingent. The type-identity theorist is in the uncomfortable position of having to explain away these powerful intuitions. If the identity is not necessarily true (such that every time one is in pain one's c-fibres are firing) then the type identity theory is false because then we obviously don't have an identity between types of mental states and types of brain states. If true, Kripke is saying, it is necessarily true. It is not. Therefore the identity claim is false.

2) Hilary Putnam (1975) argues that mental states cannot be identified with physical states. If we consider the type-identity theorist's claim, with Putnam, then we see that it is highly unlikely to be true for the following reason : The type identity theorist must specify a brain state such that any organism is in pain if and only if (1) it possesses a brain of the proper physical-chemical structure and (2) it's brain is in that state. Furthermore it must not be a possible state of an organism that cannot feel pain. It is, for instance, highly unlikely that only creatures who have c-fibres feel pain and that all creatures who feel pain must have c-fibres. Intuitively, we would not deny that snakes and snails feel pain just because they do not have c-fibres. Putnam's point is that it is more likely to be true that the same mental states will be differently physically realised in different creates.

3) Nagel (1974) in "What Is It Like to be a Bat?" sets out the following problem. Consciousness, states Nagel, "is what makes the mind-body problem really intractable" (p 159). Using bats as his example Nagel argues that to be conscious means "that there is something it is like to be that organism" and refers to this as the subjective character of experience, which

includes the phenomenological features of experience, for example, the characteristic feels of pains. Because subjective conscious experience is “essentially connected with a single point of view”, Nagel thinks an objective theory of the causation of human behaviour will inevitably abandon that point of view, and is thus inadequate. For instance, knowing all the objective facts about what causes pain and how it is manifested physically in the brain will never tell us what the pain feels like. If we pursue an identification between mental states and physical states, then there is still something it is like to undergo that physical process which is left out. This is the subjective character of experience, what it means to be conscious. Nagel argues furthermore that there are facts which humans will never understand because “they will never possess the requisite concepts to represent or comprehend “ (p162).

Colin McGinn (1991) has an acute perception of the problems consciousness poses for the naturalist's view and finds common cause with Nagel here.

According to McGinn, consciousness connects to the world along two major axes. The vertical axis connects consciousness to the body and brain and is the axis of embodiment. “Given that we have conscious mental experience and that we accept that some naturalist position holds true, the question is how does physical matter give rise to consciousness?” Presumably somewhere along the evolutionary line consciousness comes into existence. Evidence from psychosurgery strongly suggests that the brain gives rise to consciousness. We even have evidence which shows that brain damage can lead to grotesque changes in personality. A famous example is that of Phineas Gage, a construction worker in America, who in 1848 survived “an accident in which an iron bar, over 3 ½ feet long and 1 ¼ inches thick was blown through the front of his head, entering at the lower cheek and exiting from the upper forehead”(Beaumont; 1983; p61). Gage, first being a capable and efficient worker, became impulsive, wilful, inconsiderate and obstinate. Other examples would be the psychosurgery performed on schizophrenics in the 1940's and 50's.

The problem, in McGinn's words, is "How could the aggregation of millions of individual insentient neurons generate subjective awareness"(1989; p1). As McGinn states, naturalism about consciousness is not an option, it is a condition of understanding and existence. To suppose that consciousness cannot be accounted for naturally within the physical realm would be to reject the principle of causal closure and to do so would be to accept the possibility of some kind of immaterial mental causation. As we saw earlier we have enough reason to suppose that this kind of mental causation just doesn't occur. Since we know that all things have physical causal explanation, it must presumably be in virtue of some property P (or collection of properties) of the brain that consciousness is caused (or realized) by the brain. This property P would be the psychophysical link connecting mental states to physical states. P must, as it were, face both the mental realm and the physical realm. McGinn thinks we can never find this property P and introduces the concept of cognitive closure in this regard. A mind is cognitively closed with respect to certain properties or theories if those properties or theories cannot ever be grasped by that mind. For instance, dogs are cognitively closed to evolutionary theory while we are not. We, according to McGinn, are cognitively closed to that property of the brain, as well as that theory which, taking account of that property, explains how brains cause minds. Thus, though there must "exist objective natural laws that somehow account for the upsurge of consciousness" McGinn thinks we will never discover those laws.

While Nagel doubts that consciousness can be accounted for within the physical realm, McGinn doubts that we can know how the physical gives rise to the mental. There is an important difference to note here: Nagel's position allows for (and seems to require) Cartesian Dualism because the physical realm is deemed to be inadequate as an account of our mental lives. According to Nagel, physical facts cannot ever account for the subjective character of our experience. McGinn, on the other hand, is pointing out that there are limits to the extent to which we can know ourselves and that, though the psychophysical connection between consciousness and brain will systematically elude us, we must assume that there is such a connection if we are to avoid reverting to Dualism. It is important to see why McGinn thinks the psychophysical connection between consciousness and brain will elude us because he thinks



that we can and have made progress in our investigations of mental causation and he thinks that mental states can cause behaviour and cannot be reduced to physical states.

McGinn identifies two ways through which we could investigate that property P in virtue of which brains give rise to consciousness: (1) we could investigate consciousness directly by introspection or (2) we could look to the study of the brain for P.

Introspection will not reveal to us how P is the psychophysical link because it does not reveal consciousness as depending on the brain. Firstly, if introspection could reveal P to us then it seems highly likely that it would have done so by now. Centuries of introspection have not helped us in identifying P. Since introspection doesn't help us in identifying P and P must be a property of both consciousness and the brain (it must face both directions), introspection doesn't reveal to us the full nature of consciousness. The fact that it doesn't deliver us P is reason to think that there is more to consciousness than introspection shows and that P may reside in the "hidden structure of consciousness". Secondly, introspection is severely limited in what it can tell us about consciousness. It can only be directed onto the subject's mental states. It tells us nothing about the mental states of others. If I, for example, want to know anything about your mental states, I must stop introspecting and start perceiving. Also, it cannot tell us anything about mental states we haven't experienced. The mental concepts we form are determined by the kinds of mental states we experience. Introspection doesn't give us access to concepts that lie beyond our own form of consciousness. Our concepts of consciousness are inherently constrained by our own form of consciousness.

Similar attempts to identify P by studying the brain must, according to McGinn, fail. Here McGinn argues that P is perceptually closed because of the role of perception in our studies of the brain. Firstly, nothing we can imagine discovering about the brain would convince us that we had found the psychophysical link. We would always be perplexed as to how neural states give rise to consciousness. Secondly, inferences from observing the brain to the unobservable won't do the trick because we always expect physical phenomena to have purely physical

explanations. Our data, to begin with, doesn't show which property of the brain gives rise to consciousness. Any inference to the best explanation that something unobservable is taking place is thus disqualified. That is, any inference that something unobservable, (maybe a Cartesian immaterial entity), best explains consciousness, is disqualified.

McGinn thus concludes that though the psychophysical link has a natural scientific explanation, this explanation will be inaccessible to us as a matter of principle. If McGinn is right, then not only does the type-identity theorist's attempt to naturalize consciousness fail but any such attempt is doomed to failure from the start. Consciousness will never be fully naturalized and certain features of mental phenomena will never be incorporated into a naturalistic scientific explanation.

It is at the horizontal axis that consciousness is connected to the objects and properties represented by conscious states. This is the axis of intentionality and it is here that the propositional attitudes connect to the world. They have a representational content and, as stated earlier, are about worldly states of affairs. Though the problem of embodiment (the problem of how physical matter gives rise to consciousness) is closed, McGinn thinks that the question of individuation (the question of what makes a content a content of this rather than that) is open. It is here that McGinn and Davidson argue that our reasons (mental states like belief and desire) can be causes of behaviour, and that we find Anomalous Monism as a proposed solution to the problem of mental causation.

## **OBSERVATIONS AND CONCLUSIONS**

We have now looked at the type-identity thesis as a solution to the problem of mental causation and, though rejected, an identification between mental states and physical states is still inevitable if we are to maintain a naturalistic attitude and not revert to Dualism. The problem of mental causation is then one of identifying a causal role for mental states given that all causes are physical events. The problems identified by Kripke and Putnam must, furthermore, be

answered by any theory of mental causation which purports to be a naturalized solution. Finally, given that we accept the principle of causal closure I think it is safe to reject Nagel's position, for that principle requires that all events be given a physical causal explanation while Nagel's position requires that we allow for the possibility that only a Cartesian mind can explain the subjective character of experience. Thus, though it seems that we cannot explain the psychophysical connection between consciousness and the brain (for reasons given by McGinn), we must accept that brains somehow give rise to consciousness.

We turn now to our exposition of psychological explanations and how Anomalous Monism, which invokes a token-identity, is offered as a solution to the problem of mental causation.

### CHAPTER 3

Davidson's token identity thesis is an attempt to neutralize psychological explanation. It is thus a proffered solution to the problem of mental causation. The token-identity thesis is materialistic without being physicalist, that is, it views all causes as physical events but rejects the further claim that the mental is reducible to or eliminable in favour of physical explanations. Davidson rejects this claim, because, though he thinks causal-psychological explanations are available, he also thinks there are no purely psychological laws or psychophysical bridge laws on the basis of which mental events can be reduced to physical events. The mental is claimed to be anomalous, that is, not like *Q* or *R* in *Q* → *R*. Our aim in this chapter is to gain an understanding of why Davidson puts forward this claim and why he formulates it as a principle. We will also look at McGinn's argument to see why he thinks the mental is anomalous. First, though, we need to be more specific about what we mean by a "purely psychological law" and a "psychophysical bridge law".

### CHAPTER 3

We saw earlier that the most central aspect of scientific explanation is that it explains by formulating laws on the basis of which physically occurring phenomena can be predicted. Thus, if we have a law formulated in the conditional "if-then" form (if *x*, then *y*), then once we identify the antecedent conditions of the "if-clause", we will be able to predict the consequent given by the "then-clause" with increasing accuracy. This follows from our conception of what it means to be a law. This means that we will be able to predict to unobserved future cases, over hypothetical ones, where we know the antecedent conditions hold. The occurrence of the consequent is determined by the occurrence of the antecedent. By this conception we can formulate one of the relevant kinds of laws in the following way:

- (1) Physical law: If physical state *x* occurs, then physical state *y* follows.  
 Example: When neural state *x* occurs, then neural state *y* follows.

This is the kind of law we find in physics and some of the special sciences.

Davidson's token-identity thesis is an attempt to naturalize psychological explanation. It is thus a proffered solution to the problem of mental causation. The token-identity thesis is materialistic without being physicalist, that is, it views all causes as physical events but rejects the further claim that the mental is reducible to or eliminable in favour of physical explanations. Davidson rejects this claim because, though he thinks causal-mechanistic explanations are available, he also thinks there are no purely psychological laws or psychophysical bridge laws on the basis of which mental events can be reduced to physical events. The mental is claimed to be anomalous, that is, unlawfullike. Our aim in this chapter is to gain an understanding of why Davidson puts forward this claim and why he formulates it as a principle. We will also look at McGinn's argument to see why he thinks the mental is anomalous. First, though, we need to be more specific about what we mean by a "purely psychological law" and a "psychophysical bridge law".

We saw earlier that the most central aspect of scientific explanation is that it explains by formulating laws on the basis of which physically occurring phenomena can be predicted. Thus, if we have a law formulated in the conditional "if-then" form (if x, then y), then once we identify the antecedent conditions of the "if-clause", we will be able to predict the consequent given by the "then-clause" with unerring accuracy. This follows from our conception of what it means to be a law. This means that we will be able to predict to unobserved future cases, even hypothesized ones, where we know the antecedent conditions hold. The occurrence of the consequent is determined by the occurrence of the antecedent. By this conception we can formulate our three relevant kinds of laws in the following way:

- 1) Physical law : If physical state x occurs, then physical state y follows.  
Example : When neural state x obtains then neural state y follows.

This is the kind of law we find in physics and some of the special sciences.

2) Psychological law : If mental state x obtains in conjunction with certain other specified mental states, then mental state y occurs (or action Z occurs).

Example : If Smith believes it 'is raining' and has a desire to get wet, then he won't take an umbrella.

Here we find certain mental states or actions being predicted from the occurrence and interaction of certain other mental states.

3) Psychophysical bridge laws : If neurophysiological state x obtains, then mental state y obtains.

Example : If Smith is in neurophysiological state x then Smith believes it is going to rain.

These kinds of laws are those required by Fodor for a successful reduction of psychological explanation to physical explanation. As stated, these kinds of laws connect typically used predicates from the reduced science with typically used predicates from the reducing science. Furthermore, there will be further bridge laws connecting the intervening special science with physics.

Though he recognizes the availability of physical laws such as (1), Davidson denies that there can ever be laws such as (2) and (3). This is because the mental is claimed to be anomalous, which is the claim that there cannot be any laws whatsoever on the basis of which mental events can be predicted. If indeed the mental is anomalous, then, whatever the status of psychological explanation, psychological explanation cannot be expected to develop with the same kind of precision we find in scientific explanation:

The nomological irreducibility of the psychological means, if I am right, that the social sciences cannot be expected to develop in ways exactly parallel to the

physical sciences, nor can we expect ever to be able to explain and predict human behaviour with the kind of precision that is possible in principle for physical phenomena. This does not mean there are any events that are in themselves undetermined or unpredictable; it is only events as described in the vocabulary of thought and action that resist incorporation into a closed deterministic system. These same events, described in appropriate physical terms, are as amenable to prediction and explanation as any (1974; p230).

With these words Davidson expresses his central thought on the problem of mental causation. As can be seen he clearly believes that human behaviour does in fact obey physical laws. We should thus have a clear understanding why Davidson (and McGinn) think that the anomalousness of the mental is known a priori. Davidson attempts to show that, by the principles of normativity and holism which guide the interpretation and translation of behaviour, the anomalousness of the mental is known a priori. I will discuss these principles in turn.

## **THE NORMATIVITY OF THE MENTAL**

According to Davidson, the only way we can intelligibly interpret and translate a person's verbal behaviour is by establishing his attitude "towards his sentences, such as holding, wishing or wanting them to be true" (1970; p222). What Davidson means is that if we are even to begin to understand what a person means by his verbal behaviour then we must attribute to that person beliefs and desires. We must attribute a content to his utterances and an attitude towards those utterances, that is, a propositional attitude. But to interpret verbal behaviour, we must also be able to tell when a speaker holds a sentence he speaks to be true. Sentences are held to be true partly because of what is believed and partly because of what the speaker means by his words.... "the basic strategy must be to assume that by and large a speaker we do not yet understand is consistent and correct in his beliefs" (p238). In ascribing mental states like beliefs and desires we, according to Davidson, are committed to finding a large degree of rationality and consistency. It is not only that we are committed to finding such a degree of rationality but, for

Davidson, belief-desire ascriptions “necessarily impose conditions of coherence, rationality and consistency” (p231). This is a much stronger claim for it implies that it is a part of what it means to be a belief or desire that it conforms to the demands of rationality, coherence and consistency. These are the normative principles that specify which kinds of belief an agent should or should not hold and which actions should be performed on the basis of which beliefs and desires. In interpreting we have to assume that the believer is rational in thinking his beliefs true and that the whole set of his beliefs and desires cohere and are consistent with each other to a significant degree. We do not, for instance, expect a person to believe p and to believe not p at the same time.

Dennett (1987) offers similar considerations in arguing for his version of the belief-desire thesis. There are many strategies, according to Dennett, which we can use in attempting to predict the behaviour of a system. One of these, the intentional strategy, treats people as true believers who are intentional systems. By this strategy, if you want to predict the behaviour of a person, you first treat the person as a rational agent, then you figure out what beliefs and desires that agent ought to have given its place in the world, and given, especially, the states of affairs it is disposed to bring about in the world. Finally, you predict that the agent will act in a certain manner given its beliefs and desires.

Suppose, for instance, that you observe Smith exercising his legs in the gym. You do not now need to be told that Smith desires to build up the muscles in his legs, nor do you need to be told that he believes exercising them is the way to build the muscles in his legs. The explanation derives its authority from the fact that it makes sense of what we can observe, and what we have observed about the agent in the past, and what we already know independently about the agent. This is what Davidson and McGinn mean in asserting that reasons justify and rationalize actions. They are the means to intelligible interpretation of behaviour. During our everyday personal transactions with each other we often explain to each other why we do the things we do by citing our reasons, that is, pairings of beliefs and desires. When asked, for example, why I lifted my leg, I answer: I wanted to exercise it. When we answer in this way the people who



are questioning us do not usually press for further answers. They gain an understanding of what we did and our behaviour which once seemed strange becomes understandable. Our reasons supply a justification for our behaviour and in this way, according to Davidson (1963) and McGinn (1979), our reasons rationalise our behaviour. When we learn the reasons why someone did something we have, according to Davidson, "a new description of what he did which fits into a familiar picture. The picture certainly includes some of the agents beliefs and attitudes, perhaps also some goals, ends, principles, general character traits, virtues or vices" (1963; p10).

Reasons rationalise behaviour, in the sense that when we learn an agent's reasons for acting "we learn what it was about the action, given his beliefs and desires, that made it appeal to him, and we learn that he acted because it thus appealed to him. This type of explanation may be called rationalisation" (McGinn; 1972; p25)

According to the belief-desire theorist, then, unlike what the mechanist thinks, people do sometimes act. They act when they do something for a reason.

For Davidson the pro-attitude (desires, urges, etc.) toward an action of a certain kind and the belief that one's contemplated action is of that kind is the primary reason why an agent performed an action, and primary reasons are causes. It is important to see that reasons don't simply cause us to act but we act because of them. There is an important difference to note here. If beliefs and desires were just causes of our behaviour they could be causes in a very mechanistic way coupled, of course, with the right intention. When the belief-desire theorist says that we act because of our beliefs and desires he implies that we exercise a certain kind of autonomy in acting - as stated, we use our beliefs and desires in our deliberations and act because of them. Thus Drestske (1989) asserts: "I am a materialist who thinks that we sometimes do things because of what we believe and want, we need beliefs and desires because our wanting this and believing that, besides being our reasons for doing what we do, are sometimes at least the reasons why we do it" (p1). It is thus the content of our beliefs and

desires, in Dretske's words "the semantic aspect of reasons, the what it is we believe and desire" that the rationalist thinks plays a causal role in explaining behaviour. This is what proponents of psychological explanations have in common: thinking that the content of our beliefs and desires plays a causal role in our behaviour. It is on this point, as we shall see later, that Davidson bases his view that mechanistic and psychological explanations are compatible, for it is when a person acts for a reason that Davidson regards an action as free. We return now to the irreducibility of the mental.

These conditions of coherence, rationality and consistency will "have no echo in physical theory which is why we can look for no more than rough correlations between psychological and physical phenomena", states Davidson (1974; p231). This is one of the reasons why Davidson thinks psychological explanation cannot be eliminated in favour of or reduced to neurophysiological explanation. There will be no neurophysiological correlates for the conditions of coherence, rationality and consistency. Regardless of how far we advance with our knowledge of the neurophysiological workings of the central nervous system the predicates and concepts used in that science will never allow for the kind of interpretation the conditions of coherence, rationality and coherence demand. The predicates used there will never allow for the interpretation of meaning required by the normative principles. The concepts of a complete neuroscience will never get us to the semantic aspect of beliefs and desires, the attitudes a speaker holds towards his utterances; they won't allow us to interpret a speaker's attitude toward his utterances. It must be remembered that it is this attitude towards the content of one's utterances that the proponent of psychological explanations thinks are causes of behaviour.

Psychological explanation will not thus be eliminated in favour of neurophysiological explanation because neurophysiological explanation won't enable us to interpret and translate a speaker's meaning in the way that belief-desire ascription does.

The normative principles necessary for successful interpretation of behaviour thus shows, for Davidson, the indispensability of psychological explanations of behaviour. We end this discussion with a suitable quote from Davidson (1974):

Psychological concepts, I have been arguing, cannot be reduced, even nomologically, to others. But they are essential to our understanding of the rest.

We cannot conceive a language without psychological terms of expressions - there would be no way to translate it into our own language (p243 - 4).

We turn now to Davidson's argument for the a priori known anomalousness of the mental.

## **THE HOLISM OF THE MENTAL**

When we ascribe beliefs and desires we never ascribe them one by one; whatever beliefs and desires we ascribe to an agent, by the normative principles, depends on other beliefs and desires we may ascribe. "There is no assigning beliefs and desires to a person one by one on the basis of his verbal behaviour, his choices or other local signs no matter how plain and evident, for we make sense of particular beliefs only as they cohere with other beliefs, with preferences, with intentions, hopes, fears, expectations and the rest" (1970; p221). Secondly, when an agent acts, he generally acts by first consulting his background beliefs and desires. This holism of the mental shows, for Davidson, that if we want to predict an agent's behaviour on the basis of his beliefs and desires we must incorporate all the agent's relevant beliefs and desires.

Suppose Smith informed us that he had no desire to build up the muscles in his legs. Ought we to infer that Smith is suffering from a peculiar mental disease and that he should book himself into the nearest mental health clinic? Surely not, for on further enquiry, Smith informs us that he has had a particularly nasty injury in which the muscles in his legs were damaged. Though he has no particular desire to build the muscles in his legs he does have a desire to walk

properly and he believes that exercising his legs will strengthen his muscles sufficiently to enable him to walk again.

What this kind of example shows, for Davidson, is that we can never determine the necessary and sufficient conditions for acting on a reason. We can never specify precisely or completely enough the antecedent conditions of a psychological or psychophysical law. This is because each "antecedent condition" or reason is related to further reasons, forming an "open" system and "[b]eliefs and desires issue in behaviour only as modified and mediated by further beliefs and desires, attitudes and attendings, without limit. Clearly this holism of the mental realm is a clue both to the autonomy and to the anomalous character of the mental" (1970; p217). In order to formulate psychological or psychophysical laws we need to determine in advance those belief's and desires an agent will act on. We saw earlier that by our conception of what it means for a statement genuinely to rank as lawlike we will be able to predict to future unobserved cases once the antecedent conditions are given. Because we can never specify these antecedent conditions we cannot formulate psychological or psychophysical laws. Davidson thus concludes that the mental is anomalous.

McGinn (1978) offers a very different argument to show that there are no strict laws under which mental events can be subsumed. McGinn utilises Kripke's contingency argument and Putnam's variable realization argument in order to show that mental states enjoy no real essence whatsoever and that "it is precisely because mental predicates can be shown not to denote natural kinds but rather to express concepts of a fundamentally different character that authentic psychological laws can be ruled out in advance." (1978; p127). First, says McGinn, for a mental state M to have a physically real essence P a person must not be able to instantiate M without P or P without M. Kripke's intuition, that any proposed identity be necessary rather than contingent, which it seems to be, proves for McGinn that the relation cannot one be one strict identity. Any proposed type identity is false.

Secondly, following Putnam, the argument for variable realization shows that various physical states can realize a given mental state at different times, not only in different subjects, but in the very same subject.

McGinn offers the following example in support of variable realization. Imagine that a person comes to believe that "Hollywood is seedy" while standing on a Hollywood street corner. The belief is formed on the basis of sensory stimuli received by the person. Another person forms the same belief on the basis of overhearing a conversation while in an aeroplane flying over Hollywood. Yet a third forms the same belief while reading some travel books. Now, says McGinn, it would be extremely implausible to hold that the internal physical states of these believers coincided. The properties of the internal physical states of these believers will reflect "the particular modes of presentation and manifestation surrounding the formation of belief" (1980; p175). This shows, for McGinn, that the external environment plays a role in fixing the content of belief. It is thus small wonder for McGinn that the same belief can be realised by different physical states.

McGinn thus also concludes that we know a priori that there cannot be psychophysical laws because for such laws to be generalisable they must be projectible to future unobserved cases. Such a generalization requires a physical state that can regularly instantiate a mental state and we have enough reason to suppose that such a physical state cannot be found.

Davidson now faces a serious problem. Firstly, he wants beliefs and desires to be causes of physical events. Secondly, the principle of causal closure demands that only physical events can be causes of physical events. Thus, if mental events are causes of physical events, they must be physical events. Ontological dualism such as Cartesian dualism is ruled out, and ontological monism is demanded. This means that there can only be one sort of entity, that is, physical entities. Mental events cannot constitute a distinct ontological category; an ontological reduction appears to be inevitable. Thirdly, if mental events are causes of physical events, then the nomological principle requires types of mental states to be identified with types of physical

states. This is because laws are generalizable and must support counterfactuals. We must be able to predict to future unobserved cases. This suggests that we must be able to say: if A is in mental state x, he is in physical state y and if B is in physical state e, he is in mental state f. It suggests we be able to formulate psychophysical laws of the kind identified earlier. Now, we know that such a type identity is probably false. This is not merely because of variable realization but because the mental is, as Davidson claims, anomalous. An ontological reduction which requires a type identity is untenable because the mental realm is governed by normative principles. These principles show, for Davidson, the irreducibility of the mental while the holism of the mental shows that there cannot be any laws whatsoever on the basis of which mental events can be predicted.

Since an identification is inevitable in spite of the failure of the type identity theory, Davidson's problem is how to reconcile the anomalism of the mental with the nomological principle of causality, how to achieve mental causation while respecting the anomalous character of mental states.

It is here that Davidson has attempted to provide a solution in the form of Anomalous Monism, which invokes a token identity between mental states and physical states.

### The Token Identity Thesis

Davidson views the problem described above as arising from the apparent contradiction entailed by the acceptance of the following three principles (with abbreviations offered by Macdonald and Macdonald - 1995) : (1) PCI - The Principle of Causal Interaction - mental events interact causally with physical events; (2) PNCC - The Principle of Nomological Character of Causality - events related as cause and effect fall under strict deterministic laws; (3) PAM - The Principle of the Anomalism of the Mental - there are no strict laws on the basis of which mental events can be predicted and explained.

Davidson takes it as unproblematic that mental events interact causally with physical events, perception and action providing the most obvious cases where mental and physical events

interact causally. PNCC is accepted as an assumption, but we have seen that it is a guiding principle of scientific explanation. PAM is accepted because it is reasonably entailed by the irreducibility of the mental and the "open" character of psychological explanations. Since Davidson accepts the principle of causal closure, he accepts ontological monism (because it accepts that all events are physical) but rejects the claim that there are laws on the basis of which mental events can be predicted and explained.

In order to show how the three principles lead to Anomalous Monism I need to say a bit about Davidson's views on events as particulars.

We need events as an ontologically distinct category, according to Davidson, because they allow for a better account of actions under which the same action can be differently described.

Let's consider Davidson's example from "Actions, Reasons and Causes" (1963). I flip the switch because I want to turn on the light. In doing so I alert a prowler, unwittingly, to the fact that I am home. Here, argues Davidson, I haven't done four things, only one, of which four different descriptions can be given, namely; I flipped the switch, turned on the light, illuminated the room and alerted the prowler. Reasons rationalize under those descriptions under which the action is intentional, in this case "I turned on the light", "I illuminated the room", "I moved my finger", etc. Actions can be given different descriptions because actions are events and events are unrepeatable dated individuals such as the birth or death of a person and the start of World War Two. Furthermore, statements about individual events can easily be framed as identity statements. For example:

The death of Clinton = the death of the President of the United States.

The birth of Jack = the birth of the son of Skunk and Ansie.

Now, for Davidson, an event is mental if describable in purely mental terms and physical if describable in purely physical terms. Because events are dated unrepeatable individuals, a single event can be given both a mental and physical description. Furthermore, Davidson views

causality and identity as extensional, that is, as a relation between events no matter how they are described. Also, because laws are linguistic, "events can instantiate laws, and hence be explained and described in one way or another." (p215). It is because Davidson subscribes to such a view of events, that is, as particulars, that he seems able to solve the problem of mental causation while avoiding the reduction of mental events to physical events.

Suppose then that *m*, a mental event, caused *p*, a physical event, under PCI. Then, under PNCC there must be some law which related *m* and *p*. According to PAM there are no psychological or psychophysical laws. So the relevant law must be a physical law. But if *m* falls under a physical law it must have a physical description and hence, must be a physical event. But, and this is why the anomalousness of the mental is respected, it is identical with a token physical event, which is to say, a dated unrepeatable physical event. Every mental event that is then causally related to a physical event is a physical event. But though it is necessarily identical to the physical event it instantiates, the identity doesn't generalize to other physical events. The identity is not that of a type-type identity relation and thus allows for variable realization because we can now have a mental event realized by different physical states, not only in different organisms at different times, but in the same organism at different times. The important point is that a mental event is said to be identical with a dated unrepeatable individual physical event. The time specification shows that the identity doesn't generalize.

Davidson can avoid law reductionism, which requires that there be bridge laws linking the mental to the physical, because it is only under their physical descriptions that mental events instantiate laws; and here we can only predict to other physical events. These very same events, described mentally, do not fall under such laws. We will never be able to predict which mental descriptions can be applied to which physical events and this is why the anomalism of the mental is respected. It thus seems as if Davidson can have mental causal efficacy that does not require reduction of the mental to the physical.



We can now reconstruct Davidson in the following way: All causally efficacious events are physical events. Events are dated unrepeatable individuals. When events interact causally they can be given some description under which they fall under laws. Some causal events also satisfy mental descriptions, that is, they are mental events. Thus every causal mental event is a physical event. It is in terms of these mental descriptions that we primarily understand and interpret people. This interpretation is guided by principles (normativity and holism) which show that the mental is anomalous, that is, cannot be predicted and explained on the basis of physical laws. They also show that there cannot be psychophysical bridge laws or purely psychological laws, known or unknown.

It is the anomalousness of the mental coupled with the identity thesis that allows Davidson to show how the mental can play a causal role in human behaviour while remaining irreducibly mental.

Despite holding that the mental is irreducible to the physical Davidson still thinks that the mental is in some sense dependent on the physical by holding that mental is supervenient on the physical. As we will examine the supervenience relation in chapter five, we will set it aside and turn to a critical examination of Anomalous Monism as a proposed solution to the problem of mental causation.

In this chapter we are going to examine a frequently voiced objection against Anomalous Monism and, that is, that its antecedent causal claims are epiphenomenal. The version of the objection to be considered here is given under a reductionist interpretation of anomalous monism which has been referred to as the *reductionist* or *theory of events or properties*. (Macdonald and Macdonalds 1975). We will, very briefly, first take a look at what this interpretation involves.

On the co-instantiation view an event can instantiate both a mental property and a physical property. Thus the very same event can be an action if it possesses mental properties, a neurologically indexed behavior if it possesses the appropriate physiological properties and a biological event if it possesses biological properties, depending on what you are looking at. Davidson is denying that one can predict the occurrence of mental events, because the only laws governing the domain are physical laws. For instance, if a certain neural state is a sufficient condition for the occurrence of a certain mental property, then it is also a sufficient condition for the occurrence of a certain neurophysiological property. If, for instance, some neural state *N* instantiates some mental property *M* at *t*, then at *t* the same state *N* is also instantiated by neurophysiological property *A*. If *A* is a sufficient condition for *M* at any other time it occurs, or just now, after *t*, then *A* will be instantiated by *N*. Events are said to be *caused* if they are caused by other events.

## CHAPTER 4

We can now look at the objection that epiphenomenal mental properties have no causal efficacy and are causally irrelevant.

### The Epiphenomenalist Charge

Consider the following objection (also known as the *reductionist* or *theory of events or properties*) as formulated by Malcolm (1968):

In this chapter we are going to examine a frequently voiced objection against Anomalous Monism and, that is, that on anomalous mental events are epiphenomenal. The version of the objection to be considered here argues against a particular interpretation of anomalous monism which has been referred to as the co-instantiation theory of events or properties. (Macdonald and Macdonald; 1995). We will, very briefly, first take a look at what this interpretation involves.

On the co-instantiation view an event can instantiate both a mental property and a physical property. Thus the very same event can be an action if it possesses mental properties, a neurologically induced behavioural movement if it possesses neurophysiological properties and a biological event if it possesses biological properties. In asserting mental anomalism Davidson is denying that one can predict the occurrences of mental properties, because the only laws governing the domain are physical (or neurophysiological) ones. Variable realization means that mental properties can be co-instantiated with a number of different neurophysiological properties. If, for instance, some neurophysiological property B instantiates some mental property M at  $t_1$ , then at  $t_2$  the same mental property M could be instantiated by neurophysiological property A. If B instantiates M at  $t_1$ , it doesn't mean it instantiates M at any other time it occurs, or that any other occurrence of M will be instantiated by B. Events are still taken to be dated unrepeatable individuals.

We can now look at the objection that, on anomalous monism, mental properties have no causal efficacy and are causally irrelevant.

### The Epiphenomenalist Charge

Consider the following argument (that includes a neurophysiological law) as formulated by Malcolm (1968):

Whenever an organism of structure *s*, is in neurophysiological state *q*, it will emit movement *m*. Organism *o* of structure *s* was in neurological state *q*. Therefore, *o* emitted *m*.

The charge is that a mental difference makes no causal difference to what we do with our bodies once neurophysiological antecedents are given.

If psychological explanations are necessary for explaining behaviour, and they are not reducible to physical (or neurophysiological) explanations, then we must be given some argument which shows neurophysiological explanation to be an inadequate causal explanation of bodily behaviour, or some argument which shows neurological explanation to be a bad empirical theory as to what constitutes causes of behaviour. If, following Kim (1989), we accept something like the Explanatory Exclusion Principle, which states that no event can be given more than one complete and independent causal explanation, then we should reject mental causation as a causal explanation of bodily behaviour.<sup>2</sup> Mental states are causally inefficacious in explaining bodily movements and hence causally irrelevant.

We will now look at Honderich (1982) and Kim's (1989) formulation of the objection.

Honderich (1982) formulates the following Principle of the Nomological Character of Causally Relevant properties - namely, that events enter into causal relations in virtue of certain of their properties. He now asks: "If a mental event causes a physical event, what is the causally - relevant property, or what are the causally - relevant properties of the mental event?" (p62). If we follow Davidson's token - identity thesis, argues Honderich, then we have either to give up PCI or PAM. Firstly, if, as Davidson states, it is only as a physical event that a mental event

---

<sup>2</sup> Kim (1989a) gives good argument as to why we should accept his principle. I won't reproduce his arguments here but will accept the principle as intuitively plausible.

can be a cause, then we must give up PCI. Davidson states by PNCC that the relevant law will be a physical law. But now it seems that only physical properties can enter into causal relations. Thus the mental properties are doing no work. We must therefore give up PCI. Secondly, if we accept that mental properties enter into causal relations as mental properties then we must give up PAM, because then there must be laws linking mental properties to physical properties. Since PAM is what distinguishes Davidson's view, it is hard to see him giving it up. We will thus concentrate on the first claim, that is, that anomalous monism leaves no causal work for mental properties: the mental is epiphenomenal.

Kim (1989) states the objection in the following way:

On anomalous monism, events are causes or effects only as they instantiate physical laws, and this means that an event's mental properties make no causal difference. And to suppose that altering an event's mental properties would also alter its physical properties and thereby affect its causal relations is to suppose that Psychophysical Anomalism, a cardinal tenet of anomalous monism, is false (p269 - 70).

In a recent paper Macdonald and Macdonald (1995) have attempted a comprehensive defence of anomalous monism against this objection. I will therefore concentrate on their defence though I will mention one or two other commentators as well as Davidson's (1993) defence.

Macdonald and Macdonald identify three distinct claims that can be grouped under the epiphenomenalist charge:

- (1) Mental events are causally inefficacious.
- (2) Mental properties are causally irrelevant.
- (3) Mental properties are explanatorily redundant.

The Macdonald's argument against (1) is based on Davidson's view of causality as an extensional relation. According to Davidson (1993) it does not matter to the causal efficacy of the event whether it is described in mental or physical vocabulary. Thus, "it makes no sense to suppose that describing it in the physical vocabulary might deprive the event of its potency. An event, mental or physical, by any other name, smells just as strong" (p12). It is events, for Davidson, as for the Macdonald's, that have causes and effects "... it makes no literal sense to speak of an event causing something as mental, or by virtue of its mental properties, or as described in one way or another" (p13). As stated, causality is extensional. The Macdonald's formulate this argument further. Causation is viewed as a relation between events, an event being an instance of properties. We must remember that on the co-instantiation view one and the same event can instantiate both mental and physical properties. The question, for the Macdonald's, then becomes: is the instance of event A, which instantiates mental property M and physical property P, causally efficacious? If so, then the instance of the mental property has causal efficacy. Neither is PNCC violated. Events, claim the Macdonald's, instantiate laws only in instantiating certain properties. As Davidson states "... laws are linguistic; and so events can instantiate laws, only as those events described in one or another way... 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 descriptions that satisfy a law" (1970; p215). Though the mental property does not fall under a law, it is co-instantiated with a physical property- "[s]o if a mental property is causally efficacious, then its instantiations are covered by a causal law, even if that law mentions only physical (or neurophysiological) properties" (Macdonald and Macdonald; 1995; p17).

The second charge, that mental properties are causally irrelevant, Macdonald and Macdonald reply, results from a confusion between the dual role that nomologicality plays with regard to causation and explanation.

On the one hand PNCC requires that, where events are related as cause and effect, there be a law covering them. On the other hand, as we saw in chapter one, explanation by law is

intimately connected to the deductive–nomological account of explanation as formulated by Hempel and Oppenheim.

Macdonald and Macdonald accept PNCC but reject the assumption that this form of explanation is the only acceptable kind and the assumption that only nomological properties are causal properties. The fact that a property fails to figure in a causal law and consequently in explanation pertaining to that law does not mean, for the Macdonalds', that it is causally irrelevant. Non-nomological properties can be explanatory, it's just that they don't figure in causal laws. Although causation entails nomologicality it does not entail that every property co-instantiated with a nomological property must itself be nomological. There need not be a "general relation between the non-nomological property and the nomological property of the effect. When this happens the property instance will be causally efficacious, but the property will not be causally relevant to the nomological type of effect." ( p69 - 70). That is, the mental property will not be relevant to the kind of effect that figures in causal laws.

McLaughlin (1989) offers a similar argument and reconstructs the epiphenomenalist charge in the following way:

- 1 : (P1) PNCC implies that only strict nomic properties are causal.
- (P2) PAM implies that no mental property is a strictly nomic property.
- (C) Thus PNCC and PAM imply that no mental property is a causal property (i.e. that mental properties are epiphenomenal (p121)).

According to McLaughlin, proponents of I take (P1) to imply an Exclusion Principle, namely, that events are causally related only in virtue of falling under strict laws.

McLaughlin claims that (P1) does not imply the Exclusion Principle. Very briefly, he argues in the following way: Suppose that physical property P, which falls under a strict law, causes some event A. According to McLaughlin this does not mean that P is the only property that can cause A. (P1) is false because it leaves open the possibility that there can be properties which

do not fall under strict laws but which can fall under non-strict laws of the kind Davidson recognises "...[c]ausality does not even imply that only nomic properties are causal ...For it is compatible with the Principle of Causality that events can participate in causal relations in virtue of instantiating non-nomic properties" (p122). McLaughlin's claim is that I does not imply the Exclusion Principle because (P1) is unsound. PNCC may imply that when an event participates in a causal relation it does so in virtue of some nomic property but it leaves open "whether an event can participate in a causal relation in virtue of having a non-nomic property" (p127).

### Macdonalds

Both the Macdonald's and McLaughlin then accuse the proponent of the epiphenomenalist charge of making unsound assumptions, namely, that (i) PNCC implies that only nomic properties are causal properties and (ii) that explanation by causal law is the only kind of acceptable explanation.

Consider that some event E is an instance of both an action property A and a behavioural property B. Suppose further that A is caused by some mental property MP which is identical with an instance of a physical property PP. On a very intuitive test of causal relevance, state the Macdonald's, it could be argued that the mental property is causally irrelevant on condition that

- (a) if MP hadn't been instantiated E would still have occurred, and
- (b) if PP hadn't been instantiated, then E would not have occurred.

But according to Macdonald and Macdonald, (a) is false because E, by hypothesis is an instance of both the action property A and behavioural property B. If MP hadn't been instanced, A would not have been instanced and E would not have occurred.

Davidson's (1993) offers a similar reply to an example credited to Ernest Sosa (1984).



Suppose a gun goes off, a shot is fired and it kills someone. The loud noise is the shot. Sosa argues that had the gun been equipped with a silencer, the shot would still have killed the victim. The loudness of the shot thus has no causal relevance to the person's death. Likewise for mental properties.

Note Davidson's response: "The crucial counterfactual is fatally (sorry) ambiguous ..." (p17).

If the gun had been equipped with a silencer it would no doubt have caused a death but it would not have caused the same death nor would it have been the same shot.

The proponent of the epiphenomenalist objection is thus accused of confusing types for particulars.

According to Macdonald and Macdonald, in the above kind of counter-example, the failure of the mental property to be generally efficacious is offered as a reason for thinking it to be inefficacious in a particular instance. If an event instance is causally efficacious, this does not mean that all properties instantiated by that event will be generally linked such that a mental property co-instantiated with a physical property will always be co-instantiated with that physical property. Macdonald and Macdonald can here be seen to be invoking the variable realization of the mental.

This links up closely with Macdonald and Macdonald's reply to the third epiphenomenalist claim, namely, that mental properties are explanatorily redundant. Firstly, according to the Macdonalds" the causal relevance of properties are type relevant so that some properties will be relevant to some kind of effects and not others. Secondly, the generality captured by causal laws is not the only kind of generality found in nature. Functional biological explanations, especially of the kind endorsed by Dretske and McGinn, show a different pattern. To explain the action aspect of the effect we need to invoke mental causes. Mental properties are causally and explanatorily relevant to the action aspect. The action pattern found in nature is that of rationality and this distinctive pattern secures the generality required for causal relevance. An instance of a cause-effect relation can then be an instance of more than one pattern. Mental

properties, according to Macdonald and Macdonald, may not causally explain physical (or neurophysiological) effects but they do (and are needed) to explain actions. This is exactly what was demanded earlier in this chapter when it was said that psychological explanations must, if they are necessary to explaining behaviour, do something that neurophysiological explanations are incapable of doing and, that is, causally explain actions.

This point can be better seen if we consider McGinn's argument as to why we need content based psychological explanations.<sup>3</sup> I will briefly set out this argument.

According to McGinn, the biological function of mental states can only be revealed if we look at their content. Mental states have an environmentally directed function – relational proper function – which enables the organism to cope with the contingencies of the environment: “to locate food, evade predators, protect itself from heat or cold, and so on” (p145). Furthermore, though each function can be subserved by different causal mechanisms likewise, “the same causal mechanism can subserve distinct functions in different environments in which the organism has evolved” (p146). Variable realization is thus preserved.

This teleological taxonomy thus has information not contained in a description of how the causal mechanism works and this functional understanding, argues McGinn, enable us to make sense of the causal transactions we observe.

Relational proper function coincides with extrinsically related content and can only be revealed if we look at the content of mental states. Content bearing states are thus explanatorily relevant and “looking at the causal mechanisms will never reveal the functions such mechanisms serve and this is why content cannot be recovered from causal basis” (p149).

---

<sup>3</sup> McGinn is one of a number of philosophers, for example, Papineau (1993), Dretske (1988), Sterelny (1990), who have recently attempted to naturalise psychological explanations by opting for some kind of combined teleological, functional, biological classification of mental states.

# Macdonalds

To return to the Macdonald's, the falsity of

- (d) if PP (the physical property) had not been instanced, A (the action) would not have occurred

is due to the difference between the generality exemplified in physical property causal relevance and mental property causal relevance. Mental properties explain particular effects under their action aspect and which renders the action intelligible. Thus if MP had not occurred, then the effect could not be described as an action, in fact, the effect would not be an action. The difference between physical and mental property causal relevance is dependent, according to Macdonald and Macdonald, "upon there being a mental aspect to the effects produced, irrespective of the physical "realizers", so that the effects are not just bodily movements, but also actions" (p273). Mental properties are relevant to explaining a different pattern found in nature, namely; actions.

If we now group together these replies against the epiphenomenalist charge we are left with the following:

1. All properties instantiated by an event has causal efficacy.
2. PNCC does not imply that there cannot be non-nomological causal properties.
3. Mental properties are causally and explanatorily relevant to construing events as actions.

I will now give argument as to why these replies are inadequate.

Reply to (1): We are led to believe (1) by Davidson and the Macdonalds because causality is extensional. But is (1) really implied by the view that causality is a relation between events no matter how they are described? I think not. Even if causation is a relation between events A and B, why shouldn't that relation hold because of some property of each of them? Consider McLaughlin's (1993) example: If Tom weighs more than Mary then they are so related however they are described. But, if they are so related, then they are so

related because Tom has some weight  $X_1$ , Mary has some weight  $X_2$  and  $X_2$  is less than  $X_1$ .

Neither Davidson nor the Macdonalds say anything which suggests that, because causality is extensional, it cannot be because of properties of events. A stronger argument is consequently needed to show that, because causality is extensional, all properties instantiated by causally efficacious events are causally efficacious.

This is then a valid question that Kim asks: "Given that the causal relation from  $m$  to  $e$  is grounded in the basic physical properties of  $m$  and  $e$  and a strict law relating them, what causal work is there for  $M$  to do?" (p25-6).

Honderich offers the following example: I put some green French pears on a scale and the pointer moves to the two kilogram mark. Honderich asks is: in virtue of which of its properties do the pears cause the pointer to move to the two kilogram mark?

According to Davidson and Macdonald and Macdonald this is not a valid question because if any of the properties were different, they would not be the same pears and would thus not be the same cause. Thus, although the greenness has no general explanatory relevance for the effect, it is causally efficacious because the instancing of greenness is the instancing of weight  $X$ .

Now, as Honderich points out, the greenness and Frenchness may have been necessary to the events being the event it was, but they are not thereby necessary to the events being the cause it was. Events are causes because some of their properties stand in certain conditional connections with the effects they are said to produce.

Firstly, the fact that causality is extensional, as stated above, doesn't mean that the causal relation cannot hold because of certain properties of events.

Secondly, the epiphenomenalist objection is that mental properties make no difference to physical effects since causal relations between physical events hold because there are laws grounding those relations and mental properties, being anomalous, do not partake in laws.

The reply above misses the point entirely. The point is that it doesn't matter to physical effects which mental properties are co-instantiated with which physical properties; as long as those physical properties remain as they are, the same physical effects will occur. This is what prompts Kim (1989) to argue that a type-identity is needed between content bearing properties and neurophysiological properties (and what Davidson doesn't see is required) if mental properties are to have causal efficacy. The same point can be seen if we consider (2).

Reply to (2) and (3): Let's suppose that (2) holds. Does this mean that mental properties have causal relevance? The answer to this question depends, as Macdonald and Macdonald point out, on what effects are being discussed.

Suppose, as before, that some event E instantiated MP and PP and E causes F which instantiated action property AP and behavioural property BP.

To briefly summarize, Macdonald and Macdonald argue that different properties are needed to explain different effects. Physical properties explain physical effects, biological properties explain biological effects and mental properties explain mental effects, namely, actions which are non-nomological effects.

Actions are thus distinct from physical effects, they are nonphysical effects. Mental effects occur because (this is an explanatory "because") of mental properties. We are thus led to (3): Mental properties are causally and explanatorily relevant for actions. PNCC only states that only nomological properties can be responsible for nomological type of effects.

PNCC holds no implications as to whether there can be non-nomological causal properties or other kinds of explanation.

This solution seems to show that there is no epiphenomenalist problem to solve in the first place, each property being responsible for its own effects. However, as I will now try to show, reconstructing the problem in terms of distinct properties and effects is no solution at all.

Given the distinction between mental and physical properties, we could, with Honderich, restate PNCC to read:

Physical events interact in virtue of their physical properties and when they do, there is a strict law covering them.

This principle excludes mental properties from any causal role whatsoever in producing physical effects. Thus if mental properties are causes, they cannot have physical effects.

Physical effects are nomological. But, now we must give up the principle of causal interaction.<sup>4</sup>

The original problem of mental causation was to explain how mental and physical events interact causally. If each kind of property is responsible for its own kind of effect and actions are distinct nonphysical effects, then there seems to be no reason to think that the mental and physical domains interact causally. We must thus give up PCI, that is, the principle that mental and physical events interact causally.

In fact, the view subscribed to by Macdonald and Macdonald is very hard to distinguish from some kind of Parallelism under which mental changes are accompanied by physical changes although no causal interaction takes place between the two domains.

---

<sup>4</sup> Lynne Rudder Baker in Heil and Mele (1993) offers a similar argument.

Campbell (1984) presents this view as follows:

Mind... → M1 → M2 → M3 → ...

Body... → B1 → B2 → B3 → ...

M1, etc. Indicates mental events and B1 etc, indicates bodily events. Thus, even if (2), PNCC does not imply that there cannot be non-nomological causal properties, leads to (3), mental properties are causally and explanatorily relevant for actions, then we must give up the principle of causal interaction.

Now, I don't think Davidson subscribes to the Macdonalds view that mental effects are distinct from physical effects. In fact, on anomalous monism, all events are physical events, as are all effects, although mental properties are distinct from physical properties. On the purely Davidsonian view, some events have mental descriptions, which is to say, they have distinct mental properties and the properties play a causal role in producing physical effects, namely, actions. Given that the arguments presented above, especially the reply to (1), go through, Davidson has two options: (1) he could argue that some events, namely, actions are overdetermined but it is highly unlikely that Davidson would argue for overdetermination as this is a notoriously difficult position to maintain.<sup>5</sup> Besides, Davidson makes it quite clear that he holds (2) mental properties are supervenient on physical properties and, in this way, make a causal difference to physical effects.

Before discussing supervenience I will first, very briefly, summarise the outcome of this chapter.

---

<sup>5</sup> See Kim's (1989) "Mechanism, purpose and explanatory exclusion" for argument.

## OBSERVATIONS AND CONCLUSIONS:

The original objection was that, by PNCC (and PCC), mental properties are epiphenomenal because, by these principles, only physical properties can enter into causal relations with physical events.

The reply was that this argument (for the epiphenomenalist objection) presupposes some unwarranted assumptions, namely, that (1) explanation by law is the only kind of explanation acceptable and (2) PNCC implies that only nomological properties can enter into causal relations. It was argued that these assumptions are unsound and that mental properties are non-nomological. Furthermore, it was argued that mental properties are causally and explanatory relevant to actions, which are nonphysical non-nomological effects.

The rejoinder here is: given that (1) and (2) are unsound, PCC and PNCC (revised) still imply that mental properties cannot be causally efficacious for physical effects and are, hence, causally irrelevant and explanatorily redundant for physical effects. This is because laws relate causes to effects and, on anomalous monism, this can only be a physical law.

The distinction between nonphysical and physical effects is thus necessary if mental properties are to have any kind of causal explanatory relevance.

But now we are again led to the conclusion that we must give up all hope of giving sense to the idea that the mental and physical domains interact causally (assuming, for the moment, that supervenience is unsuccessful). This is a highly unsatisfactory conclusion for the debate to be in (and one which, I hope, most would reject) because we are also landed in the absurd situation of having to accept the possibility of multiple causal domains with no causal interaction between any of them. We might as well accept some kind of non-interacting Cartesian Dualism.



I stress "causal" because citing an explanatory "because" (as the Macdonalds do) doesn't clarify mental causation of mental effects (for myself at least). Mental properties are supposed to be causally and explanatory relevant for actions but if they cannot be actual causes of the effects they are supposed to explain, then we have to conclude that this is a very weak explanatory "because."<sup>6</sup>

We turn now to an examination of the supervenience relation in order to see whether this frequently invoked relation can save non-reductive monism from the charge that, by it, mental causation is epiphenomenal.

There are two important aspects about supervenience that we have to note.

Firstly, psycho-physical supervenience is asymmetrical. It states that the mental is supervenient on the physical but the converse doesn't hold, that is, the physical doesn't supervene on the mental.

The physical domain is regarded as the base properties while the mental is regarded as the supervenient properties. The reason we find this idea so intuitively plausible is, I think, because, at base, the world is composed of physical properties and we expect those properties to at least affect whatever other properties there may be.

---

<sup>6</sup> I will offer a brief argument in the final section why mental properties cannot be actual causes of mental effects.

Many philosophers hold that the only way that the mental can play a causal role is if the mental is supervenient on the physical. Since "Mental Events" the literature concerning supervenience has grown abundantly and I will not therefore attempt to cover all aspects of the supervenience relation. Specifically, discussion surrounding the internalist - externalist debate and its implications for how the supervenience relation should be construed will be conspicuously absent. I will rather concentrate on Davidson's (1993) defense in which he explicitly appeals to this relation in order to deflect the epiphenomenalist charge.

The idea behind supervenience is that there cannot be "two events alike in all physical respects but differing in some mental respect, or that an object cannot alter in some mental respect without altering in some physical respect." (Davidson; 1970; p214). The idea is often expressed using the following kind of example: Suppose Smith and Jones are molecule-for-molecule identical with each other. Then, if the mental supervenes on the physical and Smith is in physical state x and mental state y, Jones being in physical state x, must be in mental state y.

There are two important aspects about supervenience that we have to note.

Firstly, psycho-physical supervenience is asymmetrical. It states that the mental is supervenient on the physical but the converse doesn't hold, that is, the physical doesn't supervene on the mental.

The physical domain is regarded as the base properties while the mental is regarded as the supervenient properties. The reason we find this idea so intuitively plausible is, I think, because, at base, the world is composed of physical properties and we expect those properties to at least affect whatever other properties there may be.

The second aspect is that the supervenience relation is one of dependence and determination.

It states that if two events are alike physically then they must instantiate the same mental

properties. Physical properties determine, to a certain extent, what mental properties an event can instantiate.

It is here that the question arises: what is the strength of the supervenience relation and does it entail the existence of psychophysical laws? Jaegwon Kim (1984) offers two definitions of supervenience.

**Strong supervenience:** A strongly supervenes on B if and only if necessarily for each X and each property F in A, if X has F then there is a property G in B such that X has G, and necessarily if any Y has G, it has F.

It is generally recognised that strong supervenience would entail the existence of psychophysical laws. The problem is that it implies in any given world, the necessary co-extension of base and supervenient properties. That is, that a change in the same physical properties entail a change in the same mental properties. As we saw earlier, this means that we will be able to formulate laws on the basis of which mental properties can be explained, that is, psychophysical laws and hence, reductionism.

## CHAPTER 5

A weaker supervenience relation is hence needed if we are to avoid reductionism. Hence,

**Weak Supervenience:** A weakly supervenes on B if and only if necessarily for any X and Y if X and Y share all properties in B then X and Y share all properties in A.

The absence of the second modal operator "necessarily" shows that a law is restricted to the given world under consideration. Kim (1990) argues that weak supervenience is consistent with each of the following:

properties. Physical properties determine, to a certain extent, what mental properties an event can instantiate.

It is here that the question arises: what is the strength of the supervenience relation and does it entail the existence of psychophysical laws? Jaegwon Kim (1984) offers two definitions of supervenience.

**Strong supervenience:** A strongly supervenes on B if and only if necessarily for each X and each property F in A, if X has F then there is a property G in B such that X has G, and necessarily if any Y has G, it has F.

It is generally recognised that strong supervenience would entail the existence of psychophysical laws. The problem is that it implies, in any given world, the necessary co-extension of base and supervenient properties. That is, that a change in the same physical properties entail a change in the same mental properties. As we saw earlier, this means that we will be able to formulate laws on the basis of which mental properties can be explained, that is, psychophysical laws and hence, reductionism.

A weaker supervenience relation is hence needed if we are to avoid reduction. Hence,

**Weak Supervenience:** A weakly supervenes on B if and only if necessarily for any X and Y if X and Y share all properties in B then X and Y share all properties in A.

The absence of the second modal operator "necessarily" shows the relation is restricted to the given world under consideration. Kim (1990) argues that weak supervenience is consistent with each of the following:

1. In a world just like this one in the distribution of physical properties, no mentality is present.
2. In a world that is just like this one in all physical details, unicellular organisms are fully conscious, while no humans or other primates exhibit mentality.
3. In a world that is just like this one in all physical details, everything exhibits mentality in the same degree and kind.

Kim's point is that weak supervenience is too weak to support any dependency thesis with modal force and "modal force is arguably a necessary aspect of any significant dependency claim" (p142). Thus if strong supervenience is what is needed, then it seems we must give up our anti-reductionist claims. This is not a point I wish to argue here. I think Kim adequately demonstrates that the supervenience relation is not as straightforward as it seems, I only wish to give attention to this fact.

Davidson (1993) states that the notion of supervenience he endorses is that of weak supervenience identified by Kim. Davidson argues that anomalism monism (AM) coupled with PCI and PNCC and supervenience "hold that altering an event's mental properties would also alter it's physical properties. But supervenience does not imply the existence of psycho-physical laws" (p7). For Davidson, weak supervenience only entails that any change in a mental property of physical event will be accompanied by a change in the physical properties of e. Weak supervenience "does not entail that a change in mental property P in other events will be accompanied by an identical change in the physical properties of these other events" (p7). Davidson is here replying to Kim's (1989) argument against anomalous monism:

the very same network of causal relations would obtain in Davidson's world if you were to redistribute mental properties over its events any way you like; you would not disturb a single causal relation if you randomly and arbitrarily reassigned mental properties to events, or even removed mentality entirely from the world (p269).

Thus, Davidson states, "[if] supervenience holds, psychological properties make a difference to the causal relations of an event, for they matter to the physical properties, and the physical properties matter to causal relations" (p14).

I think there are two problems with Davidson's reply to Kim. Firstly, as pointed out, the supervenience relation is asymmetrical; the mental supervenes on the physical but the converse doesn't hold, that is, the physical doesn't supervene on the mental. This means that the supervenience relation, being a dependence relation, runs from the physical to the mental, the physical being the base determining properties and the mental being the supervenient properties. Although he does recognise that it is mental properties that supervene on physical properties, Davidson argues as if it is the physical properties that supervene on mental properties. One has to then ask, as Kim (1993) does: is Davidson perhaps not confused here and looking for the wrong kind of laws, that is, mental to physical, when he should be looking for laws going from the physical to the mental.

Secondly, if supervenience did hold that altering on events mental properties would also alter its physical properties then it would be in clear violation of PNCC as reconstructed. This principle holds that only physical properties can make a difference to which physical properties an event may instantiate. What Davidson doesn't seem to realise is that this principle is implicit throughout Kim's critique of anomalous monism. Davidson states that supervenience entails the reverse of what Kim thinks, that is, that mental properties do make a difference to how physical properties are distributed over this world. Consider, Davidson asks, two events, with the same mental property and the other with that mental property

removed. According to Davidson these cannot be the same event since one has a property the other one lacks. This mental properties distinguish two events not distinguished by physical properties. But, as argued earlier, this reply seems to rest on a confusion of identity conditions with causal conditions.

The reason why supervenience cannot entail that mental properties make a difference to physical properties is because firstly, supervenience is asymmetrical and secondly, but more importantly, under PNCC, only physical properties can make a difference to physical effects.

Davidson states:

we have the makings of a refutation of AM+P provided it can be shown that AM+P is inconsistent with the supervenience of mental properties on physical properties. The refutation would consist, not in showing AM+P inconsistent but in showing it inconsistent with supervenience, and so with the supposition that the mental properties of an event make a difference to it's causal relations. (p13-14).

This is precisely what I have maintained in the previous chapter and here: PNCC, reconstructed, allows mental properties no causal role whatsoever in producing physical effects, that is, that mental properties make no difference to what physical properties an event instantiates.

Although I haven't discussed supervenience in great detail I think enough has been said to show that it doesn't, as is often supposed, provide a clear cut solution to the problem of mental causation.

## CONCLUSION

How close are we, then, to providing a solution to the problem of mental causation?

I think the problem, since Descartes, has retained its basic form although it has been refined, especially since Davidson's early papers.

The problem with anomalous monism is that events are viewed as some kind of neutral category that can instantiate a number of different properties. The world we find ourselves in is undeniably physically biased in that, everything in it is composed of atoms, neutrons and other subatomic particles. Here, experience teaches, only physical events can be causes of physical events. Thus we reject Dualism and accept the principle of causal closure. Here we also find that laws govern the causal relations that hold between physical events. Thus we accept the principle of the nomological character of causality. This principle (revised) binds a certain kind of causal property to a certain kind of effect and states that only that kind of causal property can be responsible for that kind of effect.

The argument that PNCC holds no implications for whether or not there can be non-nomological causal properties or other kinds of explanation (besides deductive-nomological explanation) has absolutely no potential for providing a solution to the problem of mental causation. If we accept that mental properties are distinct from physical properties and we accept PNC and PCC, then mental properties, just cannot have physical effects. It seems that a type identity between properties (and hence reduction) is needed if we are going to explain mental and physical causal interaction.

Of course, PNCC doesn't imply that mental properties cannot be causal explanatory for nonphysical effects. But then it seems we might as well accept that the causal relation can hold between immaterial mental properties and effects. As stated, we might as well accept



some kind of parallelism. Stating that mental properties have causal explanatory relevance for mental effects but not causal efficacy doesn't help much either because then mental properties cannot even be a cause of that which it is supposed to explain.

The problem is that we do want to restrict the causal relation to physical objects and events and rule out mysterious kind of causal relations. If we are going to be naturalists about the causation of human behaviour then any state, event or property that is not a strictly physical state, event or property cannot be a cause of human behaviour. Physical objects and events may have other kinds of properties but, as far as causal relations are concerned, PNCC rules over these objects and events. This is exactly what makes causal mechanistic explanations so threatening to other kinds of explanation.

It seems then that we have to give up one of Davidson's three principles : (1) we give up PCI and admit that mental and physical properties do not interact causally although this seems counter-intuitive ; (2) we give up PNCC in which case we give up one of the guiding principles of science as well as any hope of having predictive power over ourselves and over our environment, or (3) we give up PAM which entails giving up a distinctive view of ourselves as rational animals and accepting that a reduction of the mental to the physical is a plausible hypothesis.

The choice here doesn't look pretty and I don't think (3) is really an option. The great virtue of Davidson's anomalous monism is that it has given us new insights into the nature of our mentality and what it means to be rational. Different principles, that of holism and normativity, rule over this domain, not PNCC.

That leaves (1), (2) and the problem of mental causation.

## REFERENCES :

1. Beaumont, J.G. (1983) Introduction to Neuropsychology  
Blackwell Scientific Publications
2. Campbell, K. (1984) Body and Mind, 2nd ed,  
Doubleday and Company
3. Callebaut, W. (1995) Taking the Naturalistic Turn  
University of Chicago Press, Chicago
4. Davidson, D. (1963) "Actions, Reasons and Causes"  
(1970) "Mental Events"  
(1974) "Philosophy as Psychology," in his Essays on  
Actions and Events (1980)  
Clarendon Press, Oxford.  
(1993) "Thinking Causes" in Heil, J and Mele, A (eds),  
1993  
Clarendon Press, Oxford
5. Dennett, D. (1987) The Intentional Stance,  
MIT Press,  
Cambridge, Massachusetts
6. Descartes, R. Discourse on Method and the Meditations tr. By  
Sutcliffe, F, E, 1968;  
Penguin Books, London

7. Dretske, F. (1989) "Reasons as Causes" in Philosophical Perspectives, 3, Philosophy of Mind and Action Theory, pp 1 - 15, Tomberlin, J. (ed)
8. Fodor, J. (1974) "Special Sciences" in The Philosophy of Science; Boyd, B; Gasper, P. and Trout, J.D. (ed)  
MIT Press  
Cambridge, Massachusetts
- (1987) Psychosemantics  
MIT Press  
Cambridge, Massachusetts
9. Hempel, C. And Oppenheim, P. (1948) "Studies in the Logic of Explanation" in Readings in the Philosophy of Science, Body, A; B; ed, 1970  
Prentice Hall, Englewood Cliffs, New Jersey
10. Heil, J. And Mele, A. (eds) (1993) Mental Causation  
Oxford University Press  
Claredon
11. Honderich, T. "The Argument for Anomalous Monism,"  
Analysis, 42, 1982, pp 59-64

12. Kim, J. (1984) "Concept of Supervenience" in his Supervenience and Mind (1991),  
Cambridge, University Press
- (1989a) "Mechanism, Purpose and Explanatory Exclusion."  
ibid
- (1989b) "The Myth of Nonreductive Monism." ibid
- (1993) In Mental Causation
13. Kripke, S. (1972) Excerpt from 'Identity and Necessity' in Readings in Philosophy of Psychology (1)  
Block, N. (ed) 1980  
Harvard University Press, Cambridge
14. Macdonald, C & Macdonald, G. (1995) Philosophy of Psychology  
Blackwell Publishers  
London
15. Malcolm, N. (1968) "The Conceivability of Mechanism" in Freewill;  
Watson, G in (ed) 1982  
Oxford University Press  
New York
16. McGinn, C. (1978) "Mental States, Natural Kinds and Psycho-physical laws" in his Problems of Consciousness (1991)  
Blackwell Publishers, Cambridge, Massachusetts.
- (1980) "Philosophical Materialism", ibid  
"Consciousness and Content", Ibid

- (1988) "Can We Solve the Mind-Body Problem", *ibid*
- (1989) "Consciousness and the Natural Order", *ibid*
- (1989) "Action and its Exploration" in Bolton, N. (ed).  
 (1979) 1982.
- (1982) The Character of Mind, Basil Blackwell
- (1989) Mental Content, Basil Blackwell
17. McLaughlin, B. (1989) "Type Epiphenomenalism, Type Dualism, and the Causal Priority of the Physical" in *Philosophical Perspectives* 3,  
*Philosophy of Mind and Action Theory*, 1989
18. Nagel, T. (1974) "What Is It Like To Be a Bat" in Readings in Philosophy of Psychology;  
 Block, N. (ed), 1980
19. Papineau, D. (1993) Philosophical Naturalism  
 Blackwell Publishers  
 Cambridge, Massachusetts
20. Place, U.T. (1956) "Is Consciousness a Brain Process" in The Mind-Brain Identity Theory  
 Borst, C.V. (ed) 1970
21. Putnam, H. (1975) "The Nature of Mental States" in Block
22. Salmon, W. (1984) Scientific Explanation and the Causal Structure of the World. Princeton University Press, New

Jersey.

(1992) "Scientific Explanation" in Introduction to the Philosophy of Science, Salmon, M.H. (ed), 1992  
Prentice-Hall, New Jersey

23. Smart, J.C.C. (1959) "Sensations and Brain Processes"  
in Borst

24. Trout, J.D. "Introductory Essay"  
in Boyd, Gasper and Trout

---