Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Rosemary Susan Atkins

M.Phil

2011

Keele University

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Acknowledgements

This thesis would not have been possible without the help and support of my tutors James Tartaglia and Giuseppina D'Oro. I would like to express my thanks to them, for their patience, knowledge and untiring assistance.

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Abstract

This thesis concerns the problem of mental causation and four contemporary approaches committed to its solution or dissolution. The problem of mental causation is the question of how, given the causal closure of the physical, the mental can have causal efficacy in physical events. The first approach I discuss is Davidson's Anomalous Monism. This non-reductivist approach claims that token mental events are identical with token physical events, but that there are no strict bridging laws between mental and physical generalisations. Davidson appeals to supervenience and causal extensionalism to answer the objection that his account renders mental properties epiphenomenal, but I argue that neither of these tactics are ultimately successful. The second approach is proposed by Dray, who claims that the problem of mental causation dissolves if the debate is considered as one of different methodological practices. This approach argues that by realising that different disciplines employ different methodologies that are not in competition with each other, there is no problem of mental causation. Against this, I argue that Dray has exaggerated the dissimilarities between these practices. The third approach is proposed by Fodor, Baker and Dretske, who all stress the importance of not emphasising micro-causation to the detriment of macro-causation. Against this, I argue that Fodor's account undermines the autonomy of the mental, Baker's reliance on common-sense explanation has unintended and unacceptable

consequences, and that Dretske's distinction between two types of causes cannot provide an explanation of one unified event. The fourth approach I discuss, and ultimately defend, is proposed by Rorty, who attempts to dissolve the problem of mental causation by arguing that it is based on historical confusions which need to be recognised and then rejected.

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

CONTENTS

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Chapter 1: The Problem of Mental Causation

Consider the world around you. The trees, the sky, animals, people, the solar system are all included in our ordinary objective conception of the world. However, what we cannot place in this conception are our minds as there seems no obvious place in this world for them. One problem that prevents us from conceiving of the mind in the physical world is that many of the most characteristic predicates that are used to describe the physical cannot be applied to describe the mental, for example size and shape. Another problem is that of privacy, as the mind is the only thing that is essentially private. For example, we could, in principle, delve into every part of a person's brain and dissect it into the smallest molecules. This is not a possibility with the mind, as you cannot perceive another mind, at least not *as* a mind, since if the mind is the brain and you can see brains, you can see minds too, but not as conceptualised as a mind, rather as conceptualised as a brain, so normal third-person investigation is impossible. You can only know a mind directly through introspection or else through third-person testimony. If the mind is truly not of the physical world then, as these reflections seem to suggest, what is the nature of the mind, and how can we ever come to know it? Many philosophers are convinced that the mind-body problem has already been solved (Dennett 1991), whereas some philosophers, such as McGinn, believe that we are naïve in thinking

that we are cognitively open to the problem and argue that we are not capable of ever discovering the solution (McGinn 1989: 1-22).

However, arguably even more troubling than this is the problem of mental causation; how, if the mind is indeed not a physical entity, it can have any causal role in the physical world. If the mind and the body have entirely different natures then it seems impossible that they could interact or have a causal relation with each other. Given the 'Completeness of Physics', (Spurrett & Papineau 1999: 25-29) which states that every physical effect has a sufficient physical cause and therefore anything that causes a physical effect must therefore be physical itself, it might seem that the mind can surely have no causal role in the physical world. This is the problem this thesis will address.

To find answers, we must first begin with considering the nature of the mind. When we contemplate the mind and mental phenomena we find ourselves with conflicting ideas; on the one hand we recognise that the mind is apparently very different from the physical and on the other hand we recognise that the mental must play a role in the physical world. The claim that the mind is distinguishable from the body is acceptable as phenomena, such as subjectivity, consciousness and rationality, pertain to it and none of these are properties that can apparently be found in the physical world. This leads to the conclusion that the mind is different from the physical. The claim that the mental must play a role in the physical world can also be accepted as the mental cannot be completely outside of the physical world as mental phenomena clearly have a relation to the physical, for example hurting your body gives you the feeling of pain, or having a belief of something makes you act physically in a certain way. This is where the mind-body problem arises, as these two differing ideas are clearly in contention; one of them leads us to believe that the mind could not be physical,

while the other claims that it has to be. If the mind is indeed physical then how can we explain privacy, and if the mind is immaterial then how can the mind and the body interact with each other?

Most of the theories attempting to solve the mind-body problem choose to side to a greater or lesser extent with one of these standpoints. On the one hand we have the forms of dualism which claim that the mental is so different from the physical in nature that it is of a different essence and the mind and body are only causally connected. On the other hand we have the positions of the various types of physicalism which claim that the mind is material and mental phenomena are only an arrangement of matter.

Cartesian, or substance Dualism, developed by Descartes (Descartes 1641: 1-62) proposes, at one level at least, a very simple account of the mind and a solution to the mind-body problem. Descartes argues that the mind can fit into reality, even though it is indivisible and indescribable in physical terms, because it does not belong in the physical world; it belongs in the mental realm. If the world consists of physical objects and the mind cannot belong to this world, then it must belong to the mental realm which is not described by physical predicates. However, Descartes' theory is neutralised by the Interaction problem, which states that if the mind and the body are indeed different substances and belong to different worlds then how can they possibly interact with each other and find common ground? Surely non-physical thoughts cannot control physical objects when there are no particles that can collide with the physical particles, just as an injury to the body cannot transform into a mental pain in the mind. This suggests that the mental must be a material substance. As Dennett claims,

> How can mind stuff both elude all physical measurement and control the body? (Dennett 1991: 35)

This problem is linked to the thesis of the Completeness of Physics (Spurrett & Papineau 1999: 25-29). As we have already seen, this thesis states that the domain of physics is complete, in that all physical events are determined according to physical laws and in accordance with prior physical events. We need never look beyond the realm of the physical in order to identify the antecedents of the subsequent event. Not all sciences are as closed as physics, for example psychology is not complete given that numerous mental events can arise from non-mental events, such as a physical accident and the feeling of pain. However, physics is special as it is completely closed in the respect that if we take a physical event and look at the preceding conditions that gave rise to the event then all that is needed to give a full explanation of the event are the prior physical factors. Acceptance of the Completeness of Physics, however, generates a problem of overdetermination, as it would lead to a competition between distinct physical causes for an event. An event is overdetermined if there are two or more distinct and sufficient causes of it. For example, a soldier facing a firing squad who is simultaneously hit with two bullets that individually are enough to cause death at the same instant. This event is overdetermined because there are two distinct and sufficient causes for the soldier's particular death, although one of the causes would have been a sufficient cause of it. The relevance of overdetermination to the problem of mental causation is that even if mental causes cannot be identified with physical causes, the completeness of physics does not automatically rule out the possibility of mental causation, since mental causes might overdetermine their physical effects; this would be compatible with also claiming that there is always a sufficient physical cause for any physical effect. However, overdetermination seems incredibly unlikely: the incident with the soldier was

clearly a fluke, and we would not want to admit that all instantiations of mental causation are similarly a fluke. So if we rule out massive and implausible overdetermination, and accept that the physical world is causally closed, then it does seem that we must also rule out the possibility of any non-physical thing causally affecting the physical world. Therefore, it would be impossible for Descartes' non-physical mental substances to cause physical events.

In this thesis I shall be assuming that this argument for the causal closure of the physical is true based on the overwhelming evidence for it. The argument for the Completeness of Physics is based on the premise that physics is special as it is the only discourse that is causally closed (Papineau 1993: 16-17). For example, meteorology is not causally closed, as there are some weather phenomena that are not caused by other weather phenomena. More obviously, psychology is not causally closed as there are some mental events that are not caused by other mental events, such as stubbing your toe and feeling pain. Thus arises the notion of the specialness of physics, as if we take any physical result and analyse its causes, these can always be traced back to one or more physical factors that give the full and comprehensive explanation of the event. While others, such as Putnam (Putnam 1978: 42) may criticise the use of explanation in the previous statement, as the use of explanation here has been used in a different sense than how it is used in psychological and biological terminology etc, they have merely confused the notions of explanation and causation, an argument that will be reprised later in this thesis. While these physical antecedents may not be ‘explanatory’ in the way that biological or psychological explanations are, in that they will be ‘illuminating’ (Papineau 1993: 16) to humans, all that is required is that they cause the physical outcome, not explain it. In this case, it does seem highly unlikely that psychological terms are needed when describing what caused a physical event. This does not mean that they

cannot be useful when describing or explaining a physical event, as when I say that my arm rose because I desired it to, this is an explanation of a physical event. What the causal closure of the physical does, however, is rule out any inclusion of psychological terms when providing a full and sufficient cause of a physical event. Given the history of science in investigating and determining the kind of causes that are responsible for such physical events as the motion of objects, for example, (ibid.: 31) and given that in this history we have no evidence to suggest that mental categories are among those needed to cause these events, it seems conclusive that mental occurrences, described as such, do not play any part in giving the full and sufficient cause for physical events.

So, since the Completeness of Physics shows that Descartes' solution to the mind-body problem is unsuccessful, where next to turn? If it is not possible for the non-physical and the physical to interact with each other then the mind must surely be made up of the same kind of material as the body and must therefore be physical. The doctrine of physicalism claims that everything is physical and is underpinned by the Completeness of Physics. If you have a physical effect then it will be possible to find a full and sufficient cause within the physical realm to explain the effect. So, if the completeness of physics is right, then anything that has an effect in the physical realm must itself be physical, including the mind. This premise was not always as widely accepted as it is today, but once evidence started emerging to support it, the completeness of physics was available to be used as a crucial argument for physicalism.

The first influential attempt to provide a physicalist account of mind was the Identity Theory proposed by U.T. Place (Place 1956) and J.J.C. Smart (Smart 1959). This account claims that the mind is physical in nature and is identical to the brain, so that mental states are identical to physical states in the brain. This is a reductivist account of the mind as it is reducing the

mental properties of mental states to physical properties. So, when a person experiences a sensation or another type of mental phenomena then this mental state is nothing more than the relevant portions of the brain that could hypothetically be observed. There is a difference between the extrinsic conception of something and the intrinsic conception of something, for example, the difference between the ordinary conception of water, being the extrinsic conception, and the scientific conception of H_2O, being the intrinsic conception. When a mental state occurs in a being's mind there is an identical physical state occurring in the brain. This type-type identity theory argues that just as H_2O and water are identical to each other, but have different appearances depending on how we are experiencing them, e.g. through seeing a lake or seeing the H_2O molecules through a microscope, so the mind and brain are identical. The conceptions of H_2O and water may differ, yet they are one and the same property, so just because the concept of the mind and the brain differ does not mean that they must therefore refer to different properties. For example, if we are hurt, the extrinsic conception of this would be based on the causal role of the pain, according to Place and Smart, whereas the intrinsic conception of it would be the concept of C-Fibres firing in our brain.

However, just as dualism suffered from the problem of how the mind and the body can interact if they are different substances, the identity theory crumbles under the problem discovered by Putnam (Putnam 1967: 37-48) of variable realisation. This thesis claims that the same mental property can be realised by different physical states. This is a problem for the Identity Theory as at its roots is the claim that each individual mental property is identical to an individual physical property and therefore these mental properties cannot be realised by any other physical states. This also means that as the Identity Theory claims that the mental

properties are identical to the properties of the human brain only, then these mental properties cannot be realised in any organism that has a different physiological makeup to humans.

Dissatisfaction with reductionist accounts, due to variable realisation, led to non-reductive physicalist theories. First proposed in his 1970's paper 'Mental Events', Donald Davidson's Anomalous Monism is one of the most famous accounts of mind from a non-reductivist standpoint. This theory makes two bold claims; the first is that mental events are identical to physical events, and the second is that the mind is anomalous, i.e. that mental events under their mental descriptions are not underpinned by strict laws. By claiming this, Davidson is proposing a kind of Identity Theory without the strict reductive bridging laws between mental and physical states. Davidson achieves this by maintaining that the identity holds between mental and physical tokens, i.e. particular mental and physical occurrences, rather than types which are general kinds of events. So, for Davidson, two people could be experiencing pain that is identical to some particular physical state in each of them but these physical states might not be type identical. This non-reductive physicalism is trying to combat the problem that both dualist and reductive physicalist theories encountered by attempting to adhere to the completeness of physics while also granting the mind autonomy. However, even Davidson's Anomalous Monism seems to be in peril as the problem of mental causation rears its head in a new and distinctive form, and Davidson must face criticism from philosophers both within the debate of philosophy of mind and from philosophers who have an entirely new take on the problem. This problem of mental causation which blights this, arguably the strongest form of contemporary physicalism, is the focus of this thesis.

In this thesis I intend to explain how Davidson's Anomalous Monism is affected by the problem of mental causation and why it is ultimately unsuccessful in solving it. Then I will

be considering the methodological approach proposed by Dray that claims that once you view problems such as these as problems of methodological practice rather than metaphysically real problems, then the question of mental causation dissolves. However, Dray also has to face criticism from none other than Davidson on his mistake of confusing causation with causal explanation. I will then be discussing the work of three philosophers; Fodor, Baker and Dretske, all of whom stress the importance of macro-causation in making the mind matter in the physical world. The position that they hold is that while there is micro-causation in the world, it is not the only causation that occurs and should not be held as the primary bearer of causal explanations in events. Instead we should accept our common sense views on the importance of the mental in physical events, despite what believers of a strict, physical nature of causation may argue. Finally I will look at the critiques made by Rorty of all of those philosophers mentioned above and decide whether Rorty's argument to try and undercut this whole philosophical debate is successful or not.

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Chapter 2: Davidson on Mental Causation

Not wishing to accept the idea of epiphenomenalism, Davidson (Davidson 1970) devised a theory that would attempt to give the mind a causal role in events but would not reduce mental properties to physical properties. Davidson's avoidance of a reductionist claim about mental properties has distinct advantages. His theory would simultaneously give the mind a causal role in physical causation and preserve the completeness of physics while also accommodating variable realisation. This would mean that we would not have to abandon our common sense views that what we think affects what we do, and we would also not have to abandon the completeness of physics which has garnered much support over the years. However, Davidson's non-reductivist approach has been considered by critics to continually generate a new and distinctive non-reductionist version of the problem of mental causation, one which is not altogether unlike the problem of mental causation which blighted Cartesian Dualism for hundreds of years.

Davidson's Anomalous Monism is derived from three premises; the principle of causal interaction, the nomological character of causation and the principle of mental anomalism.

The principle of causal interaction assumes that 'mental events are causally related to physical events' (Davidson 1993: 3) Davidson claims that this principle is obvious**,** (ibid.:3) as causal interaction between the mental and the physical is based on common sense. Every time I act I believe that this is due to some mental stimulus, such as a desire or a belief. For example, some physical events can be caused by mental events, as can be demonstrated by the desire to raise my arm and the act of my arm raising. The interaction also occurs the other way; that mental events can be caused by physical events, for example I may believe that a volcano has erupted because I have seen it on television. Thus, Davidson claims that events that have a mental description or instantiate a mental property are caused by and can cause physical events that either have a physical description or instantiate a physical property.

The second premise is the nomological character of causation which states that 'singular causal relations are backed by strict laws' (ibid.: 3). This premise claims that all causation must be subsumed under strict laws, as otherwise there would be no difference between causation and accidents. This premise is generally accepted to be true, though there have been critics, such as Anscombe (Anscombe 1983: 174-190) who claims that causation should not be a form of necessary condition and that the act of 'being caused' is not an instantiation of some exceptionless generalisation. Anscombe argues against the view that causation is necessarily deterministic and think that there is no general causal connection between cause and effect. She claims that something not being determined does not imply that it is not caused and that if something is not determined it does not imply that it occurred accidentally. Her wish is to shake, what she considers to be, the dogmatic conviction that determinism is a presupposition, or even a conclusion of scientific knowledge through showing that not all physical effects are necessitated by their causes. If this view of Anscombe's were indeed true, then the problem of mental causation would never arise. If the nature of causation was not

necessarily deterministic, then it would be possible for the mental to have a causal role in physical events. The idea that physical effects may not be necessitated by their causes means that there would not have to be strict and unbreakable laws governing the physical world and therefore the mental could remain anomalous, while also being able to interact with the physical world. However, in physics, at the most basic level, there are strict laws which govern the most basic components of the world: this is what the Completeness of Physics tells us. Given the Completeness of Physics, then, positions such as Anscombe's are simply not tenable, and Davidson's premise of the nomological character of causation is unexceptional.

The third premise is that of the anomalous nature of the mental and that 'there are no strict psycho-physical laws' (Davidson 1993: 3). In the physical world all science and practical reasoning is premised on the empirically well-confirmed assumption that all laws governing the physical are strict and unbreakable. In the mental realm however, the same rules do not apply, as all generalisations governing mental states with intentional contents, such as beliefs and desires, are normative. For example, if we find an exception to a rule in physics we would abandon the rule. If an exception occurred with mental phenomena however, we would put the blame on the anomaly and uphold the rule, as generalisations concerning mental states can only be normative, i.e. governed by a norm of behaviour which means that this is how a person ought to act given that the person is rational. Davidson uses this to show that there can be no psycho-physical laws, as the two different types of generalisations cannot be compatible. Davidson believes this premise is true because the principles which govern the physical and the mental are so different in kind that there cannot possibly be any psycho-physical laws that connect them when discussing the connection between intentional states and physical conditions. There are obviously psycho-physical laws in existence, such as with

sensations and pain. For example, if I have a headache and take a paracetamol tablet and my headache recedes, this is an instantiation of a psycho-physical law. Davidson is only limiting his denial of psycho-physical laws to the connection between intentional states, such as beliefs, and physical conditions.

In the case of the intentional-physical relation however, Davidson's argument seems strong. This is because the principles of the physical are descriptive, and so their application is strict, and the principles of the mental are normative, so their application is non-strict. This, he believes, means that there is no chance of connecting the two together by strict laws. For example, there is a difference between the physical act of a snooker ball colliding with another snooker ball and a person who desires a drink. In the first case, if the first ball collides with the second ball then it will cause it to move unless there are other mitigating factors. The principles governing the act are strict and, all things being equal, there is no other way that the event could have occurred. In the second case however, the principles are normative and so the person's desire for a drink may or may not cause the person to act in a way that would satisfy that desire. Therefore there could hardly be a strict law governing the relation between the desire and the snooker ball, because the person might not act on the desire, so whatever rule is said to strictly govern the snooker ball simply might not happen.

Taking into account all of these premises Davidson's argument is thus: if we accept the validity of the nomological character of causation and that there can be no psycho-physical laws, if we have to avoid epiphenomenalism, and if we accept that for two events to have causal interaction they must have a true physical description, then mental events must have a true physical description. Therefore mental events are physical events. However, unlike the reductive Identity Theory of Place and Smart (Smart 2004), Davidson claims only that every

particular mental state is identical to a particular physical state with no generalisations or strict laws connecting them, and does not identify mental properties with physical properties.

The crucial problem for Davidson is that his theory seems to leave the mental epiphenomenal. Just as the dualist had trouble explaining how non-physical properties could interact with the physical world, so, it seems non-reductive physicalists inherit a parallel problem about how irreducible mental properties can exert causal efficacy in the physical world. The similarity can be seen most clearly when considering Cartesian Dualism for which the problem of mental causation was insoluble. If we believe that the mental is a non-physical substance and that the physical domain is causally closed, then it is impossible in principle for the mental to affect the physical or vice versa. Something similar is true of Davidson, as if there is no relation between mental properties and physical properties, only one of which can be the true cause of the event and given that only physical properties instantiate strict laws, then how can the mental affect the physical? Surely without a relation of dependency between the mental and the physical, the mental would just be considered as floating free from the physical world. As Jaegwon Kim, in his article 'Concepts of Supervenience' (Kim 1984a: 53-78) has argued: if there are no psycho-physical laws of some sort then the mental will be shown to be irrelevant as there would be nothing connecting the mental and the physical and so the mental would have no causal role in events. As only physical properties instantiate strict laws and causation is nomological, if Davidson cannot find some relation between the mental and the physical, then the mental will remain causally inert. Kim claims that Davidson cannot solve this problem, that he terms explanatory exclusion (Kim 1989a: 250-264), because it is impossible to have two complete and independent explanations of the same event. If we accept Davidson's theory then surely we would have to admit that there are two explanations for an event; one of them mental and one of them physical. Take the

example that I want to raise my arm, and it rises and take another example of a brick being thrown and breaking a window. Critics of Davidson, such as Kim, say that in the case of the brick there are only certain properties of the brick that are causally relevant, such as the weight. Yet the brick possesses other properties, one of which being colour, that are not causally relevant to the fact that the brick caused the window to smash. Taking this into account, which properties can we consider to be causally relevant in the example of my arm being raised? Given the Completeness of Physics, (Spurrett & Papineau 1999) how is it possible that mental properties can have causal effects if they are not reducible to physical properties? Surely it would be only the physical properties, such as the neurones firing in the brain, that would be causally efficacious to my arm being raised, leaving mental properties being likened to the colour of the brick, merely an irrelevant detail in the causal chain. This exclusion problem is a general problem in the area of mental causation and one Davidson will have to counter in order to give the mental a causal role in events.

Davidson's main response to the problem of mental causation is his theory of supervenience (Davidson 1993: 4-16) which is used to account for the relation between mental and physical properties. Davidson argues that mental properties are supervenient on physical properties, thus giving them a causal relevance to physical events. The relationship between the mental and the physical properties is thus; the physical properties are the subvenient properties which the mental properties supervene on. The supervenience of the mental properties on the physical properties makes it the case that the event in question is only *that* particular event due to its having *those* particular mental properties. The supervenience relationship claims that if there is a change in the mental properties then there must be a change in the physical properties. However, it is possible for there to be a change in the physical properties without there being a change in the mental properties.

> a predicate *p* is supervenient on a set of predicates *S* if and only if *p* does not distinguish any entities that cannot be distinguished by *S*. (ibid.: 4)

The crucial point about supervenience for Davidson is that it provides a dependency relationship between mental and physical properties which stops short of a commitment to psycho-physical laws. The dependency relationship between the mental and physical properties is crucial, as without this the event could only be seen to have been caused by the physical properties, as the mental properties would have no link to the physical.

Without an answer to the problem of mental causation, Davidson's theory would just be a similar version of Dualism, with the mental properties being separate and distinct from the physical properties and predictably concluding once again with the problem of interaction, or else an overdetermination of explanations in contravention of Kim's principle that there can only be one complete and independent explanation of an event. Supervenience is Davidson's attempt to provide an account of the relation between the mental and physical properties, which stops short of a psycho-physical law, since this would undermine his entire argument. The presence of such psycho-physical laws would commit Davidson to a type-type identity theory and then the problems of variable realisation, which a type identity reductionist view cannot solve, would be raised once again.

In his attempt to counter the problem of mental causation Davidson also appeals to extensionalism (Davidson 1993: 6-7) in order to give the mental a causal role in the physical world. A sentence is extensional if, and only if, its truth value is not affected by substituting co-referring terms within it. For example, the sentence 'Hesperus is a planet' is extensional

because if we substitute a co-referring term, which are terms that refer to the same object as Hesperus, then the sentence maintains its truth value. Therefore if the sentence 'Hesperus is a planet' is true and 'Hesperus', 'Phosphorus' and 'Venus' are all co-referring terms for the same object then it is impossible for the sentence to be false if you substitute any of these terms in the sentence. Therefore, the sentence 'Phosphorus is a planet' is true because it is extensional and Phosphorus is a co-referring term for Venus. Non-extensional, or intensional, sentences by comparison, do not allow you to interchange the co-referring terms in the sentence. For example, say that Joe is uneducated and while he knows that Venus is a planet he does not know that Phosphorus is the same planet, but believes it to be something else. Then, the sentence 'Joe believes that Venus is a planet' is true. However, despite the fact that 'Venus' and 'Phosphorus' are co-referring terms, the sentence 'Joe believes that Phosphorus is a planet' is a false claim. Using this distinction between extensional and intensional sentences, Davidson argues that causation is an extensional relation which means that the truth of it cannot be affected by interchanging the co-referring terms. If Davidson is correct in his view then if the sentence 'Brain state X caused Joe to have a drink' is true and the terms 'Brain state X' and 'the desire to have a drink' have the same referent, then the sentence 'The desire to have a drink caused Joe to have a drink' must also be true. By comparing the examples of Hesperus and Venus and the examples of Joe and his desire to drink we can see what Davidson's argument is here. As 'Hesperus is a planet' is true and extensional and as Venus and Hesperus are identical to each other, the sentence 'Venus is a planet' is true. Similarly 'Mental state M caused Joe's arm to raise' is true and extensional and, according to Anomalous Monism and Supervenience, Mental state M and Physical state P are identical, therefore the sentence 'Physical state P caused John's arm to raise' is true. If the claim that 'Mental state M caused Joe's arm to raise' were non-extensional then we could not infer that 'Physical state P caused Joe's arm to raise' as this claim might be false, just as 'Joe believes

that Phosphorus is a planet' is a false claim. Davidson's point here is that since causation is extensional it makes no difference whether we use mental or physical properties to pick out events. If the sentence 'Physical state P caused Joe's arm to raise' is true, then the sentence 'Mental state M caused Joe's arm to raise' must also be true. It cannot be the case that by picking out the causation using a mental property the causal relation will cease to hold.

Thus, an event may be described physically, psychologically, etc, but this does not change the efficacy of the event in question. It is the events alone that have the power to causally affect things, not the different ways in which we describe them. If causal relations pertain to particular events then however we decide to define, describe or pick out the properties that characterise them, it cannot change what those events cause. Naming a specific event one way or another cannot change what caused it to happen, as Davidson points out,

> Naming the American invasion of Panama 'Operation Just Cause' does not alter the consequences of the event. (Davidson 1993: 8)

Therefore, if an event is a mental event, i.e. it can be described in psychological terms, this does not entail that this way of describing it can change how the causes and effects of the event occur. It surely is the event, not the vocabulary used to describe it that has the power. As Davidson puts it;

> An event, mental or physical, by any other name smells just as strong. (ibid.: 12)

He claims that his critics have made the mistake of confusing causation with causal explanation, as causation is extensional, but causal explanations are non-extensional.

For example, suppose that the sentence ‘the fact that Joe desired a drink causally explains why he got up’ is true, and also suppose that this act of Joe desiring a drink is a certain brain state X. It could not be concluded that the sentence ‘the fact that Joe had brain state X causally explains why he got up’ is true because causal explanations are used relative to what we know or want to explain. The explanation of Joe’s getting up being caused by his having brain state X will mean nothing to someone if they know nothing about neuroscience. Whereas the explanation for Joe’s getting up being his desire for a drink makes sense to everyone when trying to understand why he got up.

However, McLaughlin arrives at a different conclusion to Davidson as he claims that this appeal to causation as an extensional relation will not save mental properties from being epiphenomenal (McLaughlin 1993: 27-40). To demonstrate his argument he asks us to consider the statement that ‘Tom weighs more than Mary’ (ibid: 32). McLaughlin claims that the statement ‘weighs more than’ is an extensional relation, just as Davidson claims that ‘causes’ is an extensional relation. However, McLaughlin makes the point that we can still ask which properties it is in virtue of that make it the case that ‘Tom weighs more than Mary’. Just as with the example of the brick breaking the window that was mentioned earlier, there is only one property that makes the statement true, in this case it is the property of weight, rather than any other properties, such as colour. Likewise, we can legitimately ask in virtue of which properties a mental state M causes Joe’s arm to raise. If this statement is true in virtue of the physical properties rather than the mental properties, this would make the mental properties epiphenomenal as they would be irrelevant to the causal relation, just as the property of colour was irrelevant to the weight relation in McLaughlin’s example. As Mental state M and Physical state P are identical, Mental state M does cause Joe’s arm to raise but only in virtue of the physical description of the event. If this is indeed true, then Davidson’s

appeal to the extensionality of causality to save the mental from becoming epiphenomenal, has failed.

Moreover, the conclusion that Davidson's critics draw from the distinction between causation and causal explanation is that it is Davidson who has misunderstood the problem. Despite being able to describe an event in many different ways, it is still only physical descriptions that capture strict laws, and therefore, causation. Only the physical properties have causal power and no matter how we choose to describe the event, whether it was Mental state M or Physical state P, it is only because the event falls under the description of Physical state P that the causal statement is true. As Kim (Kim 1993: 19-26) claims

> The issue has always been *the causal efficacy of properties of events – no matter how they, the events or the properties, are described.* (ibid: 21)

Therefore, unfortunately for Davidson, appealing to extensionalism to counter the problem of mental causation cannot work. If there are mental and physical explanations of an event, then the physical explanations must take priority as it is the physical properties which are causally efficacious over the mental properties because only physical properties instantiate strict laws.

Davidson's attempt to solve the problem of mental causation through his theory of supervenience is next in the line of fire. In his paper 'Concepts of Supervenience' Kim (Kim 1984a: 53-78) notes that the supervenience that Davidson ascribes to is a weak, or within world, form of supervenience. This supervenience claims that within this actual world there is no possibility of mental properties existing without physical properties. In this case the supervenience is contingent as there is the possible scenario that in another possible world

there could be a case where mental properties do not rely on physical properties for their existence and therefore there could be free-floating mental properties in this possible world. The other type of supervenience is strong, or between worlds, supervenience. This claims that there is no possible scenario, either within this world or in any other world that mental properties could exist without physical properties. The crux of the issue is best understood in terms of Kim's distinction between 'causal relevance' and 'causal efficacy' (Kim 1993: 23). As Kim claims, an epiphenomenalist may agree with Davidson in arguing that mental properties do have causal relevance as, according to his theory, the mental properties of an event makes a difference to the physical properties of the event, which of course are causally efficacious. For example, the colour of the brick will make a difference to the event as it is this property that makes it this particular event of brick throwing. However, this does not mean that the epiphenomenalist would be contradicting himself by refusing mental properties the right to have causal efficacy. Thus, at best, supervenience can allow mental properties the right to have causal relevance, and yet cannot grant them causal efficacy. Therefore Kim claims that Anomalous Monism, its premises and Davidson's supervenience do not provide mental properties with causal efficacy satisfactorily. The problem for Davidson, that Kim (ibid.: 23-24) describes, is that by choosing to subscribe to weak supervenience to maintain his claim that there are no psycho-physical laws, he has left himself open to a major criticism. If, with weak supervenience, it is entirely possible for a causal relation to take place in the absence of the mental properties instantiated in the actual world, or indeed, in the absence of any mental properties at all, then those mental properties cannot be causally efficacious. It cannot be that mental properties are what make the causation take place, since they might have been absent. At most it would be possible for them to be causally relevant. Obviously, here Davidson could save his theory by changing his allegiance and supporting strong supervenience. However, in doing so he would have to commit to the existence of psycho-

physical laws, as in order for strong supervenience to work the mental and the physical must have a law bonding them together, and therefore a crucial premise of Anomalous Monism would be null and void.

To conclude this discussion of Davidson's attempt to avoid the problem of mental causation, it seems that the only decision can be that it is impossible for him to solve the problem of mental causation. In trying to bring his non-reductivist stance, in his theory of Anomalous Monism, and his theory of supervenience together, Davidson is continually causing himself problem after problem as together they can only entail epiphenomenalism about mental properties. If Davidson is not willing to give up either his theory of Anomalous Monism, or his theory of Supervenience, then unfortunately the problem of mental causation and overdetermination will continue to circle him like a committee of vultures waiting for the kill. So where next to turn to find a kind of non-reductivism that is not vulnerable to the problems which brought down Davidson's Anomalous Monism? Maybe a methodological approach is needed and instead of embroiling ourselves in the complicated technical discussions that seem to suggest that the mind cannot be causally efficacious, we should instead take a step back and see the problem from a different angle.

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Chapter 3: Methodological Responses

In the previous chapter we saw how Davidson's appeals to Supervenience and extensionalism, which are metaphysical responses to the problem of mental causation, ultimately seem to have failed. In this chapter therefore, we will consider a completely different kind of approach to the problem of mental causation. William Dray in his work *Laws and Explanation in History* (Dray 1957) poses the question from a methodological standpoint. He tries to show that the way in which we are approaching the subject matter needs to be reappraised, and that a different methodology may help to solve the problem. The different approach that Dray suggests consists in claiming that the debate is purely methodological; so therefore the questions we ought to be addressing are not 'how can mind matter?', 'how can the mind fit in the physical world?', or 'how can desires have causal efficacy?', but 'what is the logical structure of action explanations?' and 'do action explanations have the same logical structure as event explanations?'. If Dray is correct in his belief that since the causal mind-body problem was merely a problem about the relation between differing explanatory practices then, by using the correct methods of investigation there will be no metaphysical issues or questions to be solved about the problem of mental causation.

Dray, who owes much of his thinking to Collingwood (Collingwood 1940), draws the distinction between explanations for actions and events in the social sciences, such as history or psychology, and explanations in the natural sciences, such as physics. His belief is that there should be no ontological priority between these explanations of events because the different disciplines have differing domains of enquiry, i.e. the methods used to discover answers in the area of enquiry that is under investigation. Two disciplines have different domains of enquiry, not because they study different things, but if their goals, objectives or methods are different. A domain of enquiry is defined by the objectives of a discipline; the goals of the natural sciences are the control and prediction of the natural world, whereas the social sciences' objectives are to achieve understanding of human behaviour and why agents act the way they do. The natural sciences are concerned with putting forward explanatory hypotheses that have predictive powers, whereas the social sciences are concerned with explaining the actions of agents. The argument that is put forward is that the social sciences are not concerned with discerning behavioural patterns for the sake of making predictions and deducing generalisations for events as they are concerned rather with interpreting the actions of the agent involved. What is different about the social sciences is that they try to understand and interpret behavior. If we accept Dray's view then what Kim has called 'the principle of explanatory exclusion' (Kim 1989a: 250-254) that was problematic for Davidson to solve, will be undercut. The principle of explanatory exclusion claims that there cannot be more than one single complete and independent explanation of any one event. The principle is most persuasive when discussing causal explanations of individual events. Unless an explanation can be reduced to another explanation, such as with reductive physicalism where mental properties were reduced to physical properties and thus there was only a physical explanation for an event, then two complete and distinct theories cannot co-exist. If two complete and

independent explanations of an event do occur then this is a case of overdetermination. This overdetermination, brought about by two explanations for a single event, causes an epistemic and metaphysical tension which can only be relieved when the true causal explanation for the event is found. If this principle of explanatory exclusion is true, then we can see where the problem for Davidson arose, as mental and physical properties would always be in contention in explanatory terms. However, if we believe Dray's view then this theory is undercut because the explanations of the natural sciences and the explanations of the social sciences are not in competition with each other as they operate in different domains of enquiry with different explanatory aims and so there is no problem of explanatory exclusion.

The idea of methodological unity (Hempel 1942: 344-356), which is the argument that the natural sciences and the social sciences use the same methods of enquiry to find explanations, is a contentious issue and at the heart of this debate. If the social sciences and the natural sciences can be proven to use the same methodology then the problem of explanatory exclusion would reappear as the two different explanations of the event would be in competition and this would ruin Dray's attempts to show that the problem of mental causation can be avoided by using a different methodology. Thus, to bring new light to the problem and to be able to dismiss the problem of mental causation, Dray must argue that the social sciences require different methodologies to those of the natural sciences and prove that methodological unity is unattainable.

To show that the human sciences and natural sciences need to use different methodologies in order to properly explain events in their different domains of enquiry Dray, (Dray 1964: 4-8) discusses the theory of explanation and appeals to Hempel's 'covering law theory' (Hempel 1965: 345-46). Hempel's deductive-nomological model or 'covering law theory' model

claims that the explanation of an event is only achieved if what is to be explained is subsumed or 'covered' by a general law. This covering law model claims that a scientific explanation of an event consists of several statements and laws that must be met. Firstly it consists of a statement of the determining conditions for an event, which are a set of statements that assert the occurrence of certain events at certain times and places. Secondly the model consists of general laws on which an explanation is based, such that if events of a certain kind occur then an event of the type that is to be explained will take place. As Dray himself claims; the central statement of the 'covering law theory' is that the occurrence or events to be explained are accounted for by being;

> logically deducible from statements setting forth certain antecedent, together with certain empirically verified general laws. (Dray 1964: 5)

Without both the antecedent or simultaneous conditions and the universal laws being present, the alleged explanation would be incomplete. The meaning of explanation in this case is to show the deductibility of what is being explained from something else in accordance with universal 'covering' laws. Despite the fact that this type of explanation uses deduction it does not imply that we can deduce the law from the initial conditions *a priori*. Rather we have to prove the law from the observation of such events; *a posteriori*. The theory is just stating that if such laws are established and in occurrence with a certain set of initial conditions that generally cause a certain event, then we can deduce that this event will occur.

Hempel (Hempel 1965: 335-38) asks us to consider the phenomenon observed by John Dewey (Dewey 1910: 70-71) while he was washing dishes. During this task he removed some glass tumblers from the hot, soapy water and placed them upside down. He noticed that

soap bubbles emerged from under the glass tumbler's rims that grew for a while before receding back into the tumbler. The explanation for this event is that by transferring the hot glasses he had trapped cool air in them which the heat of the glass slowly warmed. This led to the volume of the air inside the glass increasing and causing the expansion of the soap film which had formed between the tumbler's rims and the plate. However, as the glass cooled off, so did the air trapped inside the glass leading to the soap bubbles receding. This explanation uses the covering law theory, as it has both the antecedent conditions for the event; the presence of the hot water and the air being cool etc, and the presupposition of empirically testable universal laws, such that under constant pressure the volume of a gas will increase as the temperature rises.

Dray (Dray 1957: 122-137) uses this example of a scientific method of explanation and contrasts it with the methods of explanation used by the social sciences. What this appeal to the covering law theory model shows is that the ways in which scientists explain events is very different to the way in which the special sciences explain events. When we offer an explanation of a human action it does not conceptually coincide with showing an action's deducibility from other conditions in accordance with empirical laws. In considering how historians attempt to explain an event and discern the reason that an agent had for causing the event we can see the difference between this type of method and a scientific method of discovering explanations. To reach understanding of an event the historian seeks information and background on the agent's motives, knowledge of the situation that the agent would have, the likely results that would be gained from different courses of action undertaken by the agent and what the agent would want to accomplish; simply put the historian would examine the agent's goals, motives and purposes. When the historian can see why the agent did what he did for good reasons in respect to his previously mentioned beliefs and purposes, then the

action can be seen to be appropriate and understanding of the action has been gained. What Dray is emphasising by considering the way in which historians explain events in comparison to how scientists investigate them is that considerations such as these bring about a conceptual connection between understanding an agent's actions and discerning the rationale that underlies it. This kind of explanation, ones that attempt to establish a conceptual connection between beliefs, motives and the actions, Dray coins 'rational explanations'. (ibid: 125).

Rational explanations, unlike scientific explanations, are not concerned with the truth or the falsity of the premises for the argument that people use to justify their actions. This is because, unlike in the natural sciences, chains of reasoning in humans do not necessarily have to begin with true statements or sound beliefs. For example, a person may believe that there are devils in the mountains so will not venture there for fear of death. The fact of whether there may or may not be devils in the mountain is not in question; if the person believes that there are devils and thus makes the rational argument that as he does not want to die, he must not go into the mountains, that is the explanation for the agent's actions. To understand an agent's action, the pre-requisite is to discover the premises of the agent's thought process, however questionable those premises may be. It must be stressed that the establishment of a deductive-logical connection between the explanans and the explanandum, based on the empirical laws of the former, is neither a necessary nor a sufficient condition of explanation (Dray 1961: 109). It is not a necessary condition because the aim of these types of explanation is not to show that the agent is the sort of person who would always act in this way in the sorts of circumstances he thought he was in. The aim of this type of explanation is only to show that what the agent did in this circumstances was perfectly rational from his own point of view on this occasion; on another occasion he may act irrationally or be guided

by different reasoning. This type of explanation, unlike explanation of the physical sciences, is not about finding a rule that can predict behaviour; rather it is to show the relation between the agent's actions and their consequent behaviour. It is also not a sufficient condition because it would not *in itself* represent the relation between the agent's beliefs or purposes and his action, as there may have been many reasonable or unreasonable actions the agent could have taken. Here, Dray stresses the difference between the studies of conceptual connections for rational explanations in this case, and the covering law model of explanations that some philosophers believe to be necessary for metaphysical problems relating to objective conditions of the natural environment. The difference between the two discourses and their methods is that the objective or causal types of explanation are required to have necessary and sufficient conditions between the explanans and the explanandum. This is because in the physical world, given the strict causal laws and a certain set of physical conditions, the outcome would have to be the same every time otherwise the completeness of physics could not be upheld. However, the conceptual connections of rational explanations are non-strict.

One way to understand the difference between Dray and Davidson then is that Dray believes that this is a problem between differing methodologies, which in principle, have equal status, whereas Davidson, as previously mentioned, believes that this is a metaphysical problem that is precipitated by the confusion between causation and causal explanation. However, both Dray and Davidson have the same goal; to grant the special sciences or the mental, respectively, as much relevance in explanations as the physical sciences are given. They also both agree that actions have a normative character; they describe what 'ought' to be. This is the difference between the natural sciences and the special sciences as you cannot attribute deterministic causes to the social sciences and you cannot attribute rational causes to the

natural sciences as it would be false to say that, for example, water ‘ought’ to freeze at 0% but did not because it made a rational choice as a free agent. This difference between the way the social sciences explain events and the way that the natural sciences explain events seems to show that the two methodologies cannot be combined. Instead of attempting to bring these explanations together ontologically, as Davidson believed his anomalous monism would accomplish, Dray believes that pointing out the differing explanatory practices used in the natural and the social sciences is sufficient to undermine the idea that there is a problem of mental causation that needs to be solved.

Dray’s view, that the debate needs to be focused on the differing types of explanations used by different domains of enquiry, is supported by R.G. Collingwood. In his work *An Essay on Metaphysics* (Collingwood 1940: 285-338) he argues against considering the problem as exclusively ontological. The contrast between Collingwood and Davidson is that for Davidson causation is an extensional relation that is metaphysical and that holds between events irrespectively of how they are described. However, for Collingwood causation is an intensional relation that is relative to the goals of a given form of enquiry and that varies in accordance with explanatory context and with the explananda of different forms of investigation. This is similar to the view of Dray as Collingwood is arguing from a methodological standpoint that it is the domain of enquiry that determines the type of explanation that is most appropriate for explaining an event. On this view then, the difference between reasons and causes is one of different explanatory practices rather than one of explanation versus causation. For Collingwood there is no ontological primacy for one particular kind of explanation as all explanations, including causal explanations, are dependent on the subject matter they are investigating. There is no neutral ground for the debate to be held on as all explanations are relative to the discipline within which they are

constructed. Therefore, all types of explanation are equal from within their own domains of enquiry.

Collingwood makes the distinction between three types of meaning of the word ‘cause’ (ibid:285). He claims that there is the first meaning of the word ‘cause’ which is defined by the freedom and rationality of humans when acting. This is a meaning of the word ‘cause’ that historians use when explaining an event. So, when explaining what caused the outbreak of the Second World War historians will focus on the rational explanations of those people involved. The second meaning of the word ‘cause’ is used in the type of case where natural events are investigated from a human point of view. According to Collingwood, this is the meaning used by the practical sciences of nature, such as engineering and medicine, whose primary aim is to enable humans to enlarge their control of nature rather than gain a theoretical understanding of nature. So, if a car breaks down the engineer will use the word ‘cause’ in a practical sense of the word, for example, that x caused the engine to overheat and thus the car broke down. The third meaning of the word ‘cause’ is used by those who do not attempt to consider events as those that can be controlled by man, but by those who wish to obtain objective theoretical knowledge of the world as it really is. This is the sense of the word that is used in the theoretical sciences of nature, such as physics and chemistry. It is this use of the word that is causing the controversy with philosophical problems, such as the problem of mental causation, according to Collingwood. By showing that there is this difference between the use of the word ‘cause’ in these different explanatory practices, Collingwood claims that this is the crux of the problem for philosophers who have a metaphysics based on ontology and who are trying to find out what there really is in the world independently of human reasoning. For Collingwood, however, we cannot practice metaphysics in isolation to what we know and cannot claim knowledge of what is in the

world without referring to how we know it. What philosophers such as Davidson are trying to accomplish is to make the mind matter by using the latter sense of the word ‘cause’ and then trying to apply human reasoning to it which is a different sense of the word ‘cause’. They are taking the sense of the word ‘cause’ that is independent of the human will and then are confused as to why this seems to rule out the possibility of making the mind have causal efficacy in the physical world. This view of Collingwood’s is claiming that the arguments used by Davidson and Kim are confused, but that the problem they address can be solved by admitting that the different sciences use different senses of the word ‘cause’ and therefore need different explanations to be suitable, and therefore useful, to their domains of enquiry. If this is the case then there should be no hierarchy or competition between the differing explanations as they are trying to explain events using different explanatory practices and uses of the word ‘cause’. Therefore, as there is no ontological priority between the different explanations and uses of the word ‘cause’ then there is no problem of explanatory exclusion or mental causation.

The view that Dray and Collingwood propose is to try and highlight the difference between the explanatory practices in order to dissolve the illusion that there is a metaphysical problem. This is closely related to an argument that Wittgenstein was making around the same time. In his famous work *Philosophical Investigations* (Wittgenstein 1953), Wittgenstein claims that there are many differing discourses and depending on what problem you are attempting to explain or solve depends on the methods of enquiry. Take these two examples of questions that are to be solved: ‘Could A have had the car that B has now’ and ‘Could A have had the sensation that B is now having?’ (Pears 1971: 99). Wittgenstein claims that it is a mistake to treat these two questions the same as they belong in different types of discourse. By trying to ascribe the same meaning to them we would be transgressing

the boundary between the differing discourses and by doing this philosophical problems are created. Wittgenstein bases his argument on his idea of 'language games' (Wittgenstein 1953: 5). Language can be used in many different ways depending on which language game is being used, so language changes meaning depending on the use to which it is put. For example, the word 'water' can be used as an exclamation, a demand, a question or an answer. The meaning of the word depends on the language game in which it is being used. Wittgenstein claims that this is how philosophical problems arise as the meaning of a word is dependent on the language game being used, and if philosophers are using different language games when debating philosophical problems then this is bound to cause confusion and controversy. By using this example we can see how it applies to Dray's and Collingwood's argument, as two people could be using the same word, but if they are using it in terms of different language games then the word would have a totally different meaning for each of them. If the natural sciences and the special sciences are investigating the same thing but bringing a different meaning to it depending on the domains of enquiry that they are investigating it within, then this may be how the problem of mental causation has arisen. If this is true, then the special sciences and the natural sciences will never be able to resolve their differences on the subject as the meaning they derive from the explanations they use will always differ.

This difference between methodologies suggests that Dray may have found a solution to the problem of mental causation that seemed so problematic for Davidson. By showing the difference between the methodologies of the natural and social sciences in their attempts to explain events, Dray has apparently avoided both the problem of explanatory exclusion and shown that the problem of mental causation is not a problem at all. The main differences dividing Dray and Davidson are their aims and goals when arguing these points. Davidson

wishes to solve a metaphysical problem that has plagued philosophy for a great deal of time. On the other hand we have Dray, who believes that it is a mistake to try and solve the problem of mental causation at a metaphysical level. If this is the case then Davidson has misconceived the problem and this is the reason why he cannot solve it. When the correct methodology for solving the problem of mental causation is understood, Dray thinks there really is no problem to be solved.

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Chapter 4: Against Methodological Responses

Following the attempts of Davidson (Davidson 1970) to solve the problem of mental causation, we looked at a different approach by Dray (Dray 1957), who claimed that by viewing the problem as a methodological one, the problem of mental causation can be bypassed altogether. This approach by Dray seems to have done what Davidson could not, by showing that the problem of mental causation is not a problem at all. However, which is the best approach to solve this problem of mental causation; a methodological approach or a metaphysical approach? As trying to solve the problem of mental causation from a metaphysical standpoint appears to cause more problems than it solves, maybe this different methodological approach is needed.

Both Dray and Davidson have followed a very similar line of argument so far. Davidson believes that there is a difference between the special and the natural sciences; that the special sciences are normative and the natural sciences are descriptive. Similarly, Dray believes that there is a difference between the methods used in the social and the natural sciences for the same reason. However, here the similarity ends as Davidson takes a metaphysical approach to

the problem of mental causation and Dray takes a methodological approach to the problem of mental causation. A metaphysical approach to the problem involves seeking to understand the nature of reality, by determining what things exist and what it means to exist in the first place. By taking a metaphysical approach to the problem, Davidson is trying to find the answer to a question about reality that will enhance our understanding of the world around us. A methodological approach to the problem on the other hand merely involves pointing out and describing the different explanatory practices implicated by cases of mental causation, such as the practices of the natural and the social sciences, and the methods used to investigate them. These differing methods are very important to the fields they are investigating as they define the areas of study and the domains of enquiry. Therefore, even though both Davidson and Dray recognise the different methodologies of the natural and the social sciences, their different approaches to these philosophical problems mean that they interpret this significance differently and thus come to different conclusions concerning the problem of mental causation. For Davidson, it means that reductionism is not an option, as he wants to grant the mental autonomy from the physical. Davidson also believes that there are many different types of causal explanations used when describing events, whether these are mental or not and he himself agrees with Dray when discussing the importance of psychological explanations for the social sciences. However, he also believes that the nature of causation still needs to be explained, which he did so by providing his extensionalist view of the nature of causation. This extensionalist view means that events remain the same no matter how they are described and that it makes no difference whether we use mental or physical properties to pick out events, as the causal relations between those events will still hold. Dray on the other hand, believes that there is no problem of mental causation to be solved as it is merely a problem of different methodological practices; all that is needed is to point to the different methodological practices used to investigate problems and the idea of a

problem of mental causation dissolves. However, Dray's success in proving that this problem is not really a problem at all lies in his argument that there is no need for methodological unity as the difference between the methodologies used in the natural sciences are very different from those used in the social sciences. Here, Dray will face attacks from not only those philosophers who work from within a metaphysical arena and so believe that the question of mental causation is purely a metaphysical one, but also from those philosophers who work from within his own methodological arena. If Dray's arguments of the difference between the methodologies used in the natural and the social sciences can be proven wrong then his whole argument is undermined, as a problem similar to that of explanatory exclusion will again be raised and the problem of causation will return.

In response to the claim put forward by Dray to attempt to dissolve the problem of mental causation, I believe that Davidson and Kim would claim that Dray had misunderstood the problem by trying to make the solution about methodology rather than metaphysics. It is perfectly plausible to begin this problem from a methodological standpoint, but it seems unacceptable to leave the debate at a purely methodological level without solving anything real. Philosophy may begin with conceptual analysis but it certainly should not end there as the point of philosophy is to move on to investigate what there really is, if indeed we can say what there really is. Conceptual analysis, which is the analysis of different explanatory practices, should be a mere preamble to this more serious job. If this 'real' kind of metaphysics shows that any mention of cause that is not purely physical is merely loose talk, then we must either, as Kim would say, be prepared to bite the reductionist bullet, or, as Davidson would claim, find new theories to disprove this conclusion. Even by remaining in the methodological camp, Dray is not immune from problems that blighted the metaphysical approaches. By introducing the concept of differing methodologies for the natural and social

sciences Dray is re-introducing a problem that is similar to that of the problem that Davidson encountered; explanatory exclusion. When Kim (Kim 1989a: 250-254) first introduced the problem of explanatory exclusion in philosophy of mind his argument ended by undercutting Davidson's theory of Anomalous Monism and showing that non-reductivism could never be a viable position to hold when discussing the matter. He claimed that Davidson, and other such non-reductivists, could not have their cake and eat it by defending the autonomy of the mental while still remaining faithful to the completeness of physics. The outcome of Kim's argument was to revert to the reductive physicalism which saves the mental only in virtue of the physical, and gives the mental no autonomy or independent causal efficacy in any events. The principle of explanatory exclusion maintains that if the causal explanations of the physical are complete then any other type of explanation is surplus and has no causal role in explaining the same event, otherwise the event would be overdetermined. The problem,

> seems to arise from the fact that a cause, or causal explanation, of an event, when it is regarded as a full, sufficient cause or explanation, appears to *exclude* other *independent* purported causes or causal explanations of it. (Kim 1989b: 281)

Kim could claim that this problem of explanatory exclusion can also be leveled against Dray's view of differing methodologies, as if there are numerous ways of explaining an event or an action, then surely one of the methodologies must take precedence as regards causal efficacy, thereby leaving the other methodologies as surplus in this regard. The claim that Davidson and Kim make is that only the methodology of the natural sciences is concerned with strict physical laws and causation, therefore it will only be the explanations of the natural sciences that will give true causal explanations, leaving the explanations of the social sciences irrelevant to causation.

Dray argues that because there is more than one way of explaining actions and events we need to work out the different types of methods that we could use to explain them and then decide which one is the correct method to do the explanatory work. It is not that one methodology takes precedence while another is irrelevant, it is that depending on what you are trying to explain and your goals in explaining it, one methodology will be of more use than another. For example, if we are trying to explain why it is raining, then the method to use would be that of the natural sciences as we would be trying to explain a physical occurrence. It would make no sense to try and explain why it is raining using a rational explanation. However, if we are trying to explain why Hitler invaded Poland, then we would need the methods of the social sciences as it would only be a rational explanation that would make sense in this case. Explaining the biological or physical occurrences surrounding the event would be useless in actually explaining why Hitler invaded Poland. Instead, what we need to explain the event are historical and rational explanations.

In response to this argument of Dray's however, Davidson would claim that he had not fully appreciated the nomological character of causation and its restriction to the physical domain. The nomological character of causation, the second premise that Davidson uses in his theory of Anomalous Monism, claims that 'singular causal relations are backed by strict laws'. (Davidson 1993: 3) Thus all causation must be subsumed by strict laws, as otherwise we would be unable to tell the difference between causation and accidents. This is a major bone of contention for Davidson, Kim and other philosophers, who take for granted that causation is extensional and debate the problem of mental causation from the same standpoint despite the fact that they will disagree with each other from within this standpoint, and Dray who believes that in order to solve this problem we must think outside the metaphysical box and

find other methodologies to explain events. The difference between Dray and Davidson is that Dray believes that the debate is conceptual, whereas Davidson believes that the debate is ontological and that Dray is ignoring the essence of the problem. For example, in his argument for the difference between the methodologies of the natural and the social sciences, Dray gives an example of a person who believes that there are devils in the mountains and therefore will not venture into the mountains. Dray claims that it does not matter whether there are devils in the mountains or not, as long as the person believes there are devils: that is all that is needed to make a rational argument for the agent and to explain the agent's action. As causation should not be viewed as ontological and there should be no hierarchy of explanations or explanatory exclusion issues, then, according to Dray, the explanations of the special sciences are just as relevant and necessary as those of the physical sciences.

Davidson disagrees with this view, as he claims that while these rational explanations are necessary when it comes to understanding the actions that people make, they can have no role in physical causation as it is only physical descriptions that truly capture causal laws as strict laws are only captured by physical descriptions. Only the physical domain has strict laws and is therefore causal. Events may be able to be explained in many different ways and this will allow such rational explanations that Dray argues for to have relevance when explaining the event. However, because causation is extensional, rational explanations are only relevant when explaining an event and are not causally relevant, as it is under strict laws that causation occurs. Therefore when it comes to discerning the cause of the action, it has to be physical as the action must fall under a physical law and, consequently, have a true physical description**.** There may be other ways of using the word 'cause', but just like his mentor Quine, Davidson would regard these alternative uses of 'cause' as simply 'loose talk' (Quine 1960: 216-221). People discuss causation in numerous ways, of course, whether it is physical,

rational, or emotional, yet any mention of cause that does not ultimately lead back to a physical description is merely loose talk. It is the same with explanations of events. There are many different ways to explain events, yet it is only the physical description of the event that captures the causation as only the physical domain has strict laws. All the other explanations of the events may be useful as explanations, but do not capture the causation of the event. Thus for Davidson, Dray's attempt to just view the problem as a methodological one is avoiding the core issue of the problem of mental causation. Whether we explain an event in a physical or a rational way it does not change the event in question, therefore whatever methodologies we use to explain an event, the causation in question remains unchanged.

Not only does Dray face criticism from philosophers such as Davidson and Kim, who argue against his whole approach to the problem of mental causation, he also faces criticism from those whose interest in aspects of the debate is also purely methodological. Hempel disagrees with Dray about the differing natures of scientific and historical explanations and therefore believes that causation should be a matter of hierarchy (Hempel 1962: 95-126). He claims that despite the view that the social sciences search for meaning and descriptions of particular events, rather than searching for the laws that may govern those events as the natural sciences do, the social sciences, such as history, actually do use general laws when seeking to predict and explain events and these laws are taken from the scientific methodologies:

> Those universal hypotheses to which historians explicitly or tacitly refer in offering explanations, predictions, interpretations, judgments of relevance, etc. are taken from *various* fields of scientific research……In addition, historical research has frequently to resort to general laws established in physics, chemistry, and biology. (Hempel 1942: 355)

This claim suggests that the scientific explanation of an event would be the only explanation that has relevance, while the social explanation of the event is a secondary consideration. Dray claims that scientific explanations are subsumed by general laws whereas historical explanations cannot be subsumed by such laws as they have a normative character. Hempel (ibid: 344-355) believes that there is ultimately no difference between the two types of explanations and that historical explanations are subsumed by general laws as well as scientific explanations. Hempel claims that Dray's views on the difference between the natural sciences and the social sciences are confused, as both types of explanations are subsumed by general laws. The only difference between the two is the extent to which the different explanations are expanded, as the natural sciences are searching for general laws to subsume their explanations and so expand their explanations fully, whereas the social sciences are not searching for general laws and so when an explanation is found for an individual event they take the explanation no further. So, in order to find full explanations for the social sciences they must adopt the same methodologies as the natural sciences; this is the argument of methodological unity (ibid: 356).

If this argument for methodological unity is correct then Dray cannot argue that there is a fundamental difference between the methods of the natural sciences and the methods of the social sciences. After all, it does seems strange that Dray would deny history's reliance on general laws and try to separate this subject from that of the natural sciences as the two disciplines have so much in common, the difference between them not being about different methodologies, but, as Hempel plausibly suggests, about how far the explanations are taken. The explanatory analysis used in history to explain historical events is not so much an explanation in itself, but what Hempel terms an 'explanation sketch'. (ibid: 351). An

explanation sketch indicates the laws and initial conditions relevant to explaining the event, but needs filling out to be able to turn it into a full and complete explanation. Such sketches are common in explanations in the social sciences, such as history and psychoanalysis. To complete an explanation sketch and make it a full explanation what is needed is empirical investigation. However, as the aim of historians is not to empirically research general laws to complete their explanations, they can produce nothing more than explanation sketches, whereas scientific investigation can produce complete explanations. Therefore, scientific and historical explanations are no different from one another, in that they both need to be covered by general laws. The only difference between them is that historical explanations are weaker and less definitive than scientific explanations. This is because they will never be completed as full explanations due to the goals of the social sciences, which do not aim to discover general laws, so there is no need to take the explanation further than what is required for the explanation of individual events. This differs from the natural sciences, where the goal is to discover general laws that enable them to explain generalities rather than specific events.

This view that methodological unity is essential and that the explanations of the natural sciences and the social sciences are the same is criticised by Alan Donagan (Donagan 1987: 113-136). In his book *Choice: The Essential Element in Human Action,* Donagan claims that Hempel is wrong in his view on historical explanations being subsumed by general laws and his views on explanation sketches. Even with the possibility that there is no strict connection visible between cause and effect in historical explanations due to the law being sketchy, there is no evidence that historical explanations show any law-like tendencies. This suggests that, while the explanations in the natural sciences are backed by causal laws, the explanations used in the social sciences are not, and therefore methodological unity is unattainable. However, for Hempel, scientific explanations fall under two categories; the deductive-

nomological model and the inductive- probabilistic model. Both of these models are covering-law explanations as they subsume events under general laws and they both explain an event by showing that in view of certain circumstances and general laws, this occurrence was to be expected, either with deductive certainty or inductive probability. Hempel defines general laws as

> a statement of universal conditional form which is capable of being confirmed or disconfirmed by suitable empirical findings. (Hempel 1942: 345)

In the deductive-nomological model a general law must be strictly universal, so it must be true at any time and in any place. General laws must also be capable of empirical verification; they cannot be verified or falsified *a priori* otherwise they would not be able to be empirically falsifiable. In this model the explanations take the form of deductive arguments and are the type of explanations used in the natural sciences. Deductive-nomological explanations subsume particular events under general laws and are easily falsifiable; if an instance is found which does not fit the law then the law is abandoned. This is in contrast with the social sciences where if an instance is found that does not fit the law, the law may not be abandoned and the instance may be ruled an anomaly, thus deductive-nomological explanations cannot be used for the social sciences. The inductive- probabilistic model, in contrast with the deductive-nomological model, is inductive in nature and does not have a strictly universal form; thus these types of explanation can be used in the social sciences. An explanation has a probabilistic form if it states that if

> certain specified conditions are realised, then an occurrence of such and such a kind will come about with such and such a statistical probability.(Hempel 1962: 100)

The premises of a probabilistic explanation do not necessarily entail the conclusions, as they are not deterministic in nature; they only render the conclusion probable. Both of these types of explanation are law-like and so there is no fundamental difference between the explanations of the natural and the social sciences. The only difference is that deductive-nomological explanations are stronger and more certain than the probabilistic explanations, but this difference does not rule out the argument for methodological unity. Thus arises the main difference between Dray and Hempel, as Hempel disagrees with Dray and claims that there is no difference between the historical and scientific explanations as they are all subsumed by covering laws, whether these laws are general or probabilistic. While Hempel accepts that inductive- probabilistic explanations are genuinely scientific he also claims that they are second best to deductive-nomological explanations. (ibid: 106-107).This is in direct contrast to Dray, who not only believes that there should be a fundamental methodological difference in the explanations used in the different sciences, but also believes that there is no priority in the explanations. However, if we accept that there are inductive-probabilistic explanations that are subsumed by covering laws and that these are the type of explanations that are used in the social sciences then a problem of explanatory exclusion re-appears. By agreeing with the idea of methodological unity, that the two types of explanation are of the same kind, then yet again a hierarchy of explanations arises for an event. This means that the more scientific and deterministic explanations will have priority over the humanistic explanations and thus we are back to the beginning of the debate. By claiming this, Hempel is not arguing that this will conclude with the problem of epiphenomenalism. Both Hempel and Dray, despite their differences, are trying to decide how explanations for actions and events should be described and used. They are not trying to solve the problem of mental causation from the standpoint that Davidson and Kim argue; that the problem of mental causation is a

real problem that can be solved by a metaphysical approach, as they look no further than the methodological argument in this debate. For Hempel, the main aim is to criticise the kind of view held by Dray, namely that the explanations of the natural sciences and the social sciences are different in kind. Thus, in arguing for methodological unity Hempel is comparing the differing methodologies to discover which is the most useful for explanatory practices. However, by following the argument that Hempel has described we can see that methodological unity entails that a hierarchy of explanations will occur and, in that case, it will be the explanations of the natural sciences that take precedent leaving the explanations of the social sciences as second order. It seems that however Dray defends his case, whether against Davidson and Kim who believe that the problem of mental causation is a real problem that can be solved metaphysically, or against Hempel in the debate on methodological practices, he cannot avoid the problem of explanatory exclusion.

A final criticism of Dray to be discussed in this chapter comes from another philosopher who believes that this problem is really methodological. Amie Thomasson, in her work *Ordinary Objects,* (Thomasson 2007) claims that science and common sense views on such things as ordinary objects are thought to be in competition, thereby leading people to say ridiculous eliminativist things such as ‘tables do not exist, only molecules do’. Thomasson claims that this problem is due to the analytic entailments between scientific and common sense ontology being overlooked. Once we understand this we can see that the statement ‘the baseball broke the window’ entails that ‘the molecules broke the window’ because there are analytic entailments between the two. This point of view could be used to apply to the problem of mental causation, because we would not think physical causation would rule out mental causation if we understood the analytic entailments between the mental and the physical. This view, that there are analytic entailments between the mind and the body, was most famously

developed by Gilbert Ryle in his work *The Concept of Mind* (Ryle 1949). Ryle claimed that statements that strictly separate the mental from the physical are nonsensical because of the analytic entailments between the two. Just as with the example above concerning the baseball and the molecules, it makes little sense to say that it was both the baseball *and* the molecules that broke the window due to the analytic entailments between the baseball and the molecules that the baseball is composed of. Take another example that there are two cows in a field. You could either claim that there are two cows in the field or that there is a pair of cows in the field, but it would make little sense to claim that there is both two cows in a field *and* a pair of cows in the field because of the analytic entailments between the words 'two' and 'a pair'; that if you have two of anything you therefore have a pair of something. Taking all this into account, if there are analytic entailments between the mind and the body then Ryle argues that statements that claim that there is both a mental state and a physical behaviour make little sense as the analytic entailments between the two means that statements such as these are similar to that of the baseball and the molecules, or the cows in the field. So, to claim that a mental state and a distinct physical behaviour occurred may be nonsensical if the one necessarily entails the other. So, for example, the physical behaviour of John walking towards the beer fridge entails that John desires a beer. By comparing statements about the mind and body to statements that consider objects such as the cow and the pair of cows or the baseball and the molecules we can see how this applies to the problem of mental causation. If the mind and the body do have analytic entailments between them then it might mean that there is no competition between the mental and the physical and the problem of mental causation would not arise.

However, this view of Ryle's, that there are analytic entailments between the mind and the body, has been proven to be erroneous. The main problem with this analytical behaviourism

is that it seems obvious to us that when we have a mental state there is something more to this than merely the observable behaviour exhibited. For example, it is possible for you to feel pain without outwardly showing that you are feeling it. However, would you then deny that you had been in pain just because there had been no behaviour? Ryle would claim that there is merely a disposition, but it seems obvious that a pain is a current reality, not just a disposition to complain of pain, one that we might manage to suppress. The behaviourist view also denies that beliefs are internal mental states that cause you to act in certain ways but are in fact only dispositions that you sometimes exhibit externally. For example, say that I have the belief that London is in England. When questioned on where I believe London is I say that it is in England. According to the behaviourist view my belief is not an internal mental state that caused me to answer this question, as it happens, correctly, but it is merely the disposition to give that answer. Yet, clearly we can see that this is counter-intuitive as I can believe that London is in England whether or not I am asked the question. Therefore, there must be something, some sort of current reality, such as a brain state, occurring that is more than just a behavioural disposition. However, if we take it that physicalism is true, then the only current reality that could be happening would be a brain state, and it is obvious that there are no analytic entailments between mental states and brain states, just as there are no analytic entailments between water and H_2O. Just as if there were analytic entailments between water and H_2O, if there were analytic entailments between mental states and brain states then we would know something about brain states just by understanding the concept of mental states. Therefore, although there may be some analytic entailments between mental states and outward bodily behaviour, there are none between mental states and brain states of the sort that would be required to account for mental causation involving conscious states. Thus, the arguments from both Ryle and Thomasson are ultimately irrelevant to this debate on mental causation.

While Dray defends his argument about the nature of the debate being conceptual, rather than ontological, and about the differing explanations that can be used to describe events, I feel that he cannot defend himself sufficiently against the criticism leveled against him. From Davidson and Kim he faces the criticism that he has misunderstood the nature of the problem and that the problem of mental causation is a real problem that can only be solved by a metaphysical approach. Even when discussing the debate from a purely methodological viewpoint, Dray still comes across a problem similar to that of explanatory exclusion which leaves the explanations of the social sciences second-order to the explanations of the natural sciences. Even from within his own methodological field Dray comes across criticism that leads us to believe that his approach cannot eliminate the problems that Davidson faced and prove the problem of mental causation to be irrelevant. Even if it is accepted that mental and physical explanations employ different methodologies, this has nothing to do with the physical causation of an event. The fact remains that if it is only physical descriptions that truly capture causal laws then mental events must have physical descriptions to be causally efficacious in physical events. In the next chapter we shall investigate other methods of trying to solve this problem of mental causation that appears to be so troublesome.

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Chapter 5: Macro-causal approaches to the problem of mental causation.

Introduction

As we have seen in the last few chapters, there have been many approaches to the problem of mental causation that have all, ultimately, failed in solving or dissolving this most troubling of problems in philosophy. In this chapter I shall be discussing three different views that, given what has gone before, are all still attempting to make the mind matter and to give mental states the causal relevance that common sense tells us they have. These three philosophers all believe that they can give the mental causal relevance, while avoiding the problem of mental causation, by claiming that there are other kinds of causation that we can use to describe an event instead of just micro-causation. These philosophers have a contrasting view to that of philosophers such as Kim (Kim 1984b: 92-109), who we discussed in the second chapter, who held that all causation is not only physical, but that due to the upward nature of causation, it must be dependent on the most basic physical components. The fundamental claim at the heart of this argument is that there is only one kind of real

causation, and that is causation at the level of microphysical events. Thus, causation extends upward all the way from the physical domain to the higher-level domains, such as the special sciences. However, if this is to be believed then macro-level explanations are epiphenomenal, which includes the explanations of the special sciences, and leaves the mind causally irrelevant. The three philosophers who I shall be discussing take a different approach on the problem of mental causation to the approaches that we have previously discussed. Each of these philosophers is attempting to solve this problem from a metaphysical standpoint, rather than a methodological one, but, like Dray (Dray 1957), they are not arguing from the assumption that we should take science as the primary giver of causal explanations. Not that they are taking Dray's side by agreeing that the problem is methodological in nature. The stance that these philosophers take is against the view of micro-causation as the only causal explanation and in support of macro-causal explanations. So, instead of the bottom-upwards causation that scientists, and a great deal of philosophers, assume, where causation is dictated by the basest levels of physics, there can also be causation that runs top-down so that we take our common sense views and the explanations of the special sciences as providing equally, or perhaps even more, valid causal explanations. By doing so they are disputing the view that the only kind of causation is micro-causation and thus are disputing the conclusion that the mind and the explanations of the social sciences are irrelevant.

The first philosopher I shall consider is Jerry Fodor. He proposes a theory whereby intentional mental states have causal relevance as long as there can be intentional laws. Fodor wishes to make the special sciences similar to the natural sciences in their empirical and evaluable nature, but wishes them to only concern intentional mental states. In doing so, he aspires to make intentional mental states causally relevant without succumbing to the problem of mental causation. The second philosopher I shall discuss is Lynne Baker, who has

a radical theory of how to give the mental causal relevance, so that what we think does affect what we do. Her proposal is that we abandon certain parts of our metaphysical background that concern causation and that are generally taken for granted. In doing so, we will give the mental causal relevance as it will not be constrained by the strict physical nature of causation. However, she does not only argue for this conclusion from the point of granting the mind causal relevance. She claims that if we do not abandon this part of the metaphysical background, then it will not only be the mind that is eliminated from having a causal role in explanations; it will also be the explanations of every discipline, except the most strict of physical explanations. If we take our current conception of causation as true then it is only through micro-causation that we can talk of causation, and any macro-causal explanations, including such disciplines as chemistry, geology and meteorology will not be valid. The third philosopher I am discussing is Fred Dretske, who believes that he can give mental states a causal role in explanations by making the distinction between triggering causes for events and structuring causes for events. The triggering causes for events are the token, physical causes, while the structuring causes are the mental causes that are used to explain an event in the context of its surroundings. Dretske thereby aims to grant the mental causal relevance in explaining events by showing that without the explanation of the mental, the event would not be fully causally explained. In this chapter we shall investigate all three claims of these positions.

Jerry Fodor

One of Jerry Fodor's main aims is to defend a theory of mind whereby intentional states, such as beliefs or desires, can have causal implications to our behaviour. Fodor has a different

conception of laws that subsume events. Unlike Davidson and Kim, Fodor believes that not all causal laws have to be strict, we can have laws that are hedged by *ceteris paribus* conditions, but that these are no less causal than strict laws. By accepting that only strict laws have roles in causal events, this not only denies causation for mental states but also seems to restrict causation to fundamental physics. In doing so, we are not only claiming that mental properties do not have any causal implications but we are also claiming that any causal description of an event that is not at the micro-level of physics is epiphenomenal. This would leave not only psychology as epiphenomenal, but also disciplines such as geology, psychology, chemistry and economics as all of these use macro-level explanations of events. Instead of accepting this conclusion, Fodor (1990:137-159) argues that, just as there is causation according to the strict laws in basic science, i.e. physics, there can be causation according to the *ceteris paribus* laws of the special sciences. In Fodor's view all that is required for causation is subsumption under a law, whether it is a strict law or a non-strict law, otherwise we would have to concede to statements such as 'it is not the case that it is because Mt. Everest is a mountain that is the reason for it having glaciers on top of it'. This is because the description of 'being a mountain' is not a statement that is covered by a strict physical law. The crux of the argument is that there are *ceteris paribus* laws in the special sciences, including psychology, and that these laws are sufficient to back up causal claims. However, they are not basic laws like those of physics, which do not need to be hedged by *ceteris paribus* laws. The difference between strict laws and *ceteris paribus* laws is that for *ceteris paribus* laws there is a further story to be told about the underlying mechanisms of causation. If Fodor is correct in this view then there can be causation without reduction to physical law and without encountering the problem of mental causation.

What Fodor wishes to argue is that there are two philosophical mistakes that have led to this problem of mental causation and that by realising these mistakes the problem of mental causation just melts away. The first mistake that philosophers have made in Fodor's opinion is that they have got the wrong idea about what it is for a property to be causally responsible. The second mistake that is made involves the relations between events and the special science laws that subsume them. By showing these mistakes for what they are and by putting an end to the confusion, Fodor claims he can prove that there are intentional causal laws which mean that intentional properties can be causally responsible without having to be reduced to physical properties and the problem of mental causation vanishes.

Despite a subscription to physicalism, Fodor has a conception of the special sciences according to which higher-level laws and the properties that figure in them are not reducible to lower-level laws and properties. Fodor (Fodor 1987) claims that his motivations for attempting to grant intentional states causal relevance arose from common sense. The predictive powers of intentional states, such as beliefs and desires, are so great that they cannot be denied. People do what they believe and desire to do, and so to ignore intentional states or claim they have nothing to do with our behaviour is nonsensical:

> The theory from which we get this extraordinary predictive power is just good old common sense belief/desire psychology. … If we could do that well with predicting the weather, no one would ever get his feet wet; and yet the etiology of the weather must surely be child's play compared with the causes of behavior. (1987: 3-4)

If a theory that details what it is for a property to be causally responsible leads to the epiphenomenalism of, for example, mountainhood as we saw in the earlier example, or of

such intentional states as desirehood and beliefhood, then this is not a theory that should be endorsed. Indeed, it is not only intentional states that would succumb to epiphenomenalism if the microphysical view of what makes a property causally responsible were held. This is where the first mistake that philosophers make arises as Fodor believes that what it means for a property to be causally relevant is different from what the generally held conception is. Fodor believes that as a minimal condition a theory should not be endorsed if what it holds for a property to be causally responsible is that which entails the epiphenomenality of mountainhood, desirehood and other aspects that are considered in disciplines that use macro-level explanations. Instead, Fodor proposes a theory that he claims can meet this condition of granting intentional states causal relevance while avoiding the problem of mental causation. What is considered for a property to be causally responsible is just that there are laws that subsume those properties. So, for Fodor, intentional properties are causally responsible in the case that there are intentional causal laws; geological properties are causally responsible in the case that there are geological causal laws, etc. As long as there are causal laws about the property in question this is all that is needed to claim that the property is causally responsible.

So, on the present view, the question of whether intentional properties are causally responsible reduces to the question of whether there are intentional laws concerning said properties. However, what if there are no intentional laws, or if there are, what if they cannot cover individual causes in the way that causal laws are supposed to cover the events that they subsume? If this is the case, then it is obvious that this attempt by Fodor to give intentional properties causal responsibility has failed. To give a response to this criticism Fodor highlights the three premises of Anomalous Monism, which together cause tension and lead to the problem of mental causation. To discuss these three premises we must go back to the beginning of this thesis to Davidson's Anomalous Monism. This non-reductivist proposal

was based on three premises all of which Davidson believed to be correct. The first of these is that of the nomological character of causation. This premise claims that all causation must be subsumed with strict laws, as otherwise we would be unable to tell the difference between causation and accidents. The second premise to be highlighted is that of the anomalous nature of the mental. This premise compares the strict laws of nature with the normative laws of psychology and shows that there can be no strict psycho-physical laws due to the difference between the nature of the mental and the nature of the physical. The third premise highlighted is that of the causal responsibility of the mental. This premise states what we all believe to be obvious; that our mental states do affect our behaviour and that there is a causal interaction between mental and physical states. As we saw earlier, when faced with these conflicting premises, Davidson concluded that the only solution was Anomalous Monism which would grant all three premises while adhering to physicalism. However, as we saw his theory faced severe criticism and ultimately failed in solving the problem of mental causation.

Fodor now faces the exact same problem that Davidson did. If there must be exceptionless laws covering causal transactions, and if the laws of physics do differ from the laws of the special sciences, as special science laws are typically hedged with *ceteris paribus* clauses, then how can we have causally responsible intentional properties when the laws covering events must be strict laws and, therefore, physical laws? However, instead of concluding that these three premises are undoubtedly true and thus siding with Davidson, Fodor approaches the problem by analysing whether these three premises are as watertight as Davidson would have us believe. Now, Fodor has been arguing all the way through for intentional causal laws to grant intentional properties causal responsibility so it will come as no surprise to find that he argues that the premise of the anomalous nature of the mental is an unavoidable principle. He claims that if even geological laws have to be hedged with *ceteris paribus* clauses then the

'all else equal' proviso in psychological laws is certainly not eliminable. So, what does Fodor suggest? He argues that if the choice is between choosing to adhere to either the nomological character of causation or the idea that there can be causal responsibility for the mental then it is obvious where we should lay our allegiance. Our intuitions will not be satisfied with an account of mental causal responsibility that is only in virtue of supervenience upon the physical. Given the choice we should abandon our traditional conception of causation in favour of solving the problem of mental causation to the satisfaction of our common sense views. What we need is what Fodor claims he can deliver; a robust theory of causal responsibility for intentional properties that can also be compatible with the premise of the anomalous nature of the mental. For this robust construal of mental causal responsibility, what is needed is to reconcile the idea that mental properties are nomologically sufficient for behavioural properties with the fact that psychological laws are hedged with *ceteris paribus* clauses. Surely it cannot be the case that special laws can only cover particulars *ceteris paribus, and* that it must be the case that whenever we get mental properties we *must* get behavioural properties? This is where the second mistake that Fodor thinks philosophers have made arises as he believes that they have mistaken the nature of what it is for a law in the special sciences to subsume an event. He believes that all it is for a special science law to subsume an event is that the law covers the event whenever the *ceteris paribus* conditions are satisfied. Though laws are needed when covering events, Fodor does not see that *ceteris paribus* laws are less explanatory than strict covering laws, as if the conditions are satisfied then the outcome is guaranteed. This is no different from strict covering laws, except that with special science laws, the *ceteris paribus* conditions are not spelled out, and in the strict covering laws they are. The special sciences rely on 'mediating mechanisms' (Fodor 1990: 155) which are not articulated because sometimes they are not known and sometimes a law can be implemented in many different ways. *Ceteris paribus* clauses cover the explanations

for these mechanisms, so that when we use the term *ceteris paribus* all we are saying is that there is some mechanism such that when it is in place A's cause B's. This is what Fodor claims the use of *ceteris paribus* laws in the special sciences are and argues that this makes them no less efficient and useful than strict covering laws. Fodor argues that we can get an account of sufficient causality that is just as good as the ones used in physics. If it is a law that mental states cause behavioural outcomes *ceteris paribus*, then it follows that whenever you get mental states and the *ceteris paribus* conditions are satisfied, you get behavioural outcomes. So, intentional laws do necessitate their consequents sufficiently when their *ceteris paribus* clauses are satisfied.

Therefore, Fodor claims that there is now no convincing argument for accepting that the only causal laws are strict, and therefore physical. Even if causes needed to be covered by laws that necessitate their consequences this does not mean that they need to be strict laws, as hedged laws have been shown to necessitate their consequents when the *ceteris paribus* conditions are satisfied. So, if intentional properties are already covered by these hedged laws why should they need further covering by strict physical laws to have causal responsibility? Strict laws and hedged laws, where the ceteris paribus conditions are satisfied, operate similarly concerning their roles in covering law explanations and their roles in covering causal relations:

> Strict laws are just the special case of hedged laws where the *ceteris paribus* clauses are discharged *vacuously;* they're the hedged laws for which 'all else' is *always* equal. (Fodor 1990:154)

If Fodor is correct in his arguments then we can see how tempting his conclusions are. By claiming that strict physical laws and so called non-strict laws of the special sciences operate in the same way this means that the laws of the special sciences can ground mental properties in a causally responsible way without denying physicalism. This means that Fodor has apparently found a way of reconciling the opposing notions that make the problem of mental causation such a troubling one; the notions of micro-causation versus our common sense views of macro-causation. In doing so, we can still adhere to physicalism, but we can also accept our views that what we think does affect what we do, and that the special sciences are not irrelevant in causal explanations. By letting go of the notion that only strict laws can be involved in physical causation, Fodor can reconcile the anomalous nature of the mental with the claim that intentional properties have causal responsibility and the problem of mental causation vanishes. Therefore, Fodor believes he has proposed a theory that can grant intentional states causal relevance while avoiding the problem of mental causation. In doing so, he has denied a view of strict physical causation in which only micro-causation occurs, in favour of a more lenient account of causation in which macro-causation as well as micro-causation occurs. This is a contentious issue, but he has done what Davidson could not do; grant the mind autonomy while also being able to explain how mental states can have an effect on physical events without resorting to physical reduction.

Lynne Baker

While Dray viewed the debate from a purely methodological standpoint, Baker attempts to solve the problem of mental causation by undermining our generally accepted assumptions of causation. She suggests that by abandoning certain suppositions of the traditional

metaphysical background to the problem, we can preclude the problem of mental causation and therefore dissolve this problem that has been the downfall of so many philosophical theories. The claim that Baker (Baker 1993:75-95) puts forward is that much of contemporary philosophy has taken for granted the validity of this traditional metaphysical background to the detriment of our common sense beliefs and explanations, which are then deemed to be untrustworthy if they do not fit into this conception of reality. Instead, it should be argued that we need to begin with a range of tried and tested explanations, both common sense and scientific, and work downwards from these, rather than work from the basis that everything at its core has to be based on, or reduced to, micro-causation. This is a view that is closely connected to that of Fodor's which we have just discussed. Both these philosophers believe that we should take our common sense views on the validity of the mind in a causal role in events as being just as relevant, and in Baker's case more relevant, than strict scientific explanations. They both also believe that they are not just rescuing the mental from being causally relevant in explanations, but that we should also be saving the special sciences, such as geology and meteorology, from becoming second order explanations. Not only that, but they believe that if we accept that there can be both top-down and bottom-up causation, then other disciplines, as well as psychology will have just as much causal relevance in explanations as the strict physical sciences. If Baker's arguments are to be believed, we would be left with a theory of mind and body whereby the mental can have causal relevance and the 'problem' of mental causation would not be a problem at all.

Baker begins by discerning that the problem of mental causation arises because of two metaphysical theses; the thesis of materialism and the thesis of 'the causal closure of the physical' (Baker 1993:78). Materialism, as Baker describes it, is the 'thesis that every property-instantiation supervenes on physical property-instantiations' (ibid: 78) and the

causal closure of the physical, as Baker describes it, is the 'thesis that every physical property-instantiation that has a cause at *t* has a complete physical cause at *t*.' (ibid: 78). Both of these theses we have encountered before in previous chapters and Baker's attribution of blame on them for the problem of mental causation is unsurprising. As has been shown in previous examples, if a person has a desire to do something and follows through that desire with the appropriate physical actions then according to the causal closure of the physical there is a complete micro-physical causal explanation of the event. There is also a sufficient cause for the event as there would have been some neurological event that was the cause of the body moving that the mental property would have been supervening on. Given that there is a complete physical causal explanation for such events, is there any room for a mental property to have a causal role? In this metaphysical background, as we have seen in previous chapters, it seems that, barring reductionism, there is no room for the mental to have causal efficacy, and yet it seems perfectly obvious that our desires and thoughts do have a role in our daily lives. As, Baker claims, all significant materialist theories are committed to both the causal closure of the physical and some type of supervenience and, as adherence to these theories means that mental properties can never have true causal efficacy, then we appear to have reached a crossroad. Do we give up a part of the metaphysical background, and if so which part, or do we follow the path that takes the metaphysical background for granted which would leave the problem of mental causation unsolved?

Before deciding on the path ahead of us, Baker first wants to make us aware of what is actually at stake if we decide to take the route that philosophers such as Davidson and Kim have taken which results, in her view, in leaving the problem of mental causation unsolved. Just like Fodor, Baker argues that the problem of mental causation can be generalised to a problem of macro-causation. The underlying assumptions of the causal closure of the

physical and supervenience are that everything is dependent on the micro-level; that every instantiation of every property is supervenient on an instantiation of a micro-physical property and that instantiations of micro-physical properties have a complete physical cause. As we have seen in previous chapters, this leads us to the problem of overdetermination. How can macro-level properties have any causal relevance at all if they are only supervenient on or reducible to micro-properties that have a complete physical cause and so need no other causal explanation? As with attempting to prove the autonomy of the mental, we could claim that the advances and successes in the special sciences prove that macro-level properties have causal efficacy. However, given the metaphysical background that we are working against, explanations of the special sciences are second order and not as relevant as the purely physical explanations. This conclusion would also extend so far as to lead to scepticism about causation not only in the special sciences, but in many other scientific disciplines as well. Baker argues that claims, such as 'the acidity of the liquid caused the litmus paper to turn pink' (Baker 1993:90) in chemistry, would not be needed as it is a macro-level description and therefore is not causal in this case. The actual cause in this event occurs at a micro-level, as a real relation in nature, not an explanatory one. If we take this metaphysical background for granted then it will not only be philosophical questions concerning the relation between mental and physical that will be affected. What we will be left with will be statements of causation that do not count as proper explanations because the nature of causation would only occur at the micro-physical level. One solution would be to find a metaphysical solution to the problem of mental causation, in order to give genuine credence to not only the mental, but to macro-properties and any explanations that mention macro-properties. However, if this is not a possibility, then we will need to abandon part of the metaphysical picture in favour of maintaining our common sense views and special science explanations.

Baker argues that we should choose to abandon a part of our metaphysical background and believes that what should be abandoned is the part that concerns our conceptions of reality and causation, as these by themselves give rise to the problems surrounding macro-causation. Baker criticises this aspect of our traditional metaphysical background as she denies the idea that we can do metaphysics in isolation from epistemology; that we can claim knowledge of what is in the world without referring to how we may know this. Her argument is that the best way to explain and understand events is by referring to what we already know and believe, and that nothing should take us away from our subjective experiences. So, in short, we should listen to our common sense intuition that the mind does have causal relevance in what we do. She also criticises the metaphysical claim that causal closure of the physical has at its foundation the idea that causation is an objective relation between events. Baker claims that this is a very narrow conception of causation as it would rule out intuitive causal connections. There are many examples of intuitive causal connections, for example, a student's failing of Maths making them ineligible to play football in the team would not be given as the explanation of the event if we were using this strict and narrow conception of causation. On this basis, so many obviously true claims of causal connection would have to be false, and so Baker abandons this narrow conception of causation as inadequate. Once this conception is abandoned, however, how are we to understand causation in a way that can help solve the problem of mental causation?

Baker proposes her own theory of causation that takes at its core a range of explanations, both scientific and common sense, that have earned their acceptance, instead of a metaphysical doctrine of reality that has not. She suggests a conception of cause as follows;

c caused *e* (i) if *c* had not occurred then, all things being equal, *e* would not have

> occurred, and (ii) given that *c* did occur then, all things being equal, *e* happening was inevitable (ibid.: 93)

This conception of cause is one where for one thing to have caused another depends on what other things get held constant and what sorts of things get held constant is dependent on the different explanatory contexts that the causation is being explained within. This is a view of causation that is similar to the one that Fodor proposed about the nature of *ceteris paribus* clauses. The *ceteris paribus* conditions that are in question depend on the context or discipline from which an event is being explained. For example, when explaining why a person died, depending on the context or discipline that you are explaining it from, it could either be explained by a detailed bystander statement, a biological statement or a psychological one. This is in contrast to the metaphysical picture which has the flow of causal power running from bottom upwards;

> from subatomic particles through atoms and molecules through simple organisms through intermediate organisations up to persons with beliefs and other intentional states. (ibid.: 93)

This traditional conception of causation views causation as an objective relation in nature with strict physical causation as the only true causation. If we reverse this priority of causation and explanation and abandon the causation aspect of the metaphysical background then the problem of mental causation just simply is not an issue anymore. We began by wondering if what we thought had any effect on our actions, but this question was taken and changed by the metaphysical background to become one of finding out how mental states can be causally relevant to producing behavioural events. If we reverse the priority of causation

and explanation however, this question does not arise and the original question becomes one that is easy to answer. What we think does affect what we do, since it evidently *explains* what we do. If someone is under the belief that they have left their wallet in their car and then they go back to the car to collect it, they have acted in virtue of the explanatory fact that if they hadn't thought they had left their wallet then, all things being equal, they would not have returned to the car. Given that they did think that they had left their wallet, then all things being equal, returning to the car was inevitable. Baker admits that on this premise we will end up with a mixture of statements concerning many different explanatory practices in areas such as politics, psychology, law, economics and science, but this is preferable to the alternative where we cannot allow the mind and our thoughts to affect how and what we do. As Baker claims concerning the reduction of explanations; 'Unity is merely desirable, not inevitable.' (ibid.: 94).

Therefore, Baker claims she has found a solution to the problem of mental causation by abandoning the causation aspect of the metaphysical background picture that the debate has based itself on for many years. In doing so, we are not left with any strict physical causal doctrine to ascribe to or any metaphysical ideas about objective reality, but we are left with explanations that have been proven to work and without which we would not have the interactions that we do. While Fodor and Baker reach different conclusions on whether we should abandon our traditional conception of causation or just change it to incorporate the special sciences, their arguments on why we should give due credence to our common sense views on the causal relevance of the mental and the special sciences in events are incredibly similar. By accepting Baker's argument and claiming that we have no better access to reality than what is required for cognitive success, and by beginning with successful explanatory

practices while letting the metaphysics go, Baker would claim that at least we can avoid the problem of mental causation.

Fred Dretske

A stern critic of Davidson's theory of Anomalous Monism, Dretske (Dretske 1993: 121-136) formulated his own theory to overcome the problem of mental causation that he believed Davidson had been unable to solve. Just as Fodor and Baker argued, Dretske claims that causal explanations are context-sensitive; that what we use to decide what the cause of an event is depends on our interests and our purposes. Moreover, such events are always dependent on a variety of preceding events, meaning that there are a great number of choices, given the right context, to select as a cause for a causal explanation. Dretske wants to describe the difference between two types of causes, what he calls triggering causes and structuring causes, to aid in the understanding of two different types of causal explanation. He also aims to prove that the difference between psychological and physical explanations of behaviour is thus; that psychological explanations give us structuring causes of behaviour, while physical explanations give us triggering causes of behaviour. Take the example that Dretske uses of a terrorist planting a bomb in a politician's car (Dretske 1993: 122-23). The bomb remains inactive until the politician enters his car and turns the key in the ignition and the bomb is detonated. To explain this event we would claim that the triggering cause was the turning of the key in the car that set off the bomb beneath the car. However, this causal explanation alone would not be enough to explain or understand the situation of why this event occurred. In this situation it was not the politician who killed himself; it was the terrorist who planted the bomb in the car. So, although the politician's own action was the

triggering cause, it was the terrorist's action that was the structuring cause of the event that needs be singled out as the explanation for the explosion. Using this example, Dretske argues that we should use the same theory to deal with the problem of mental causation. Instead of attempting to provide mental properties with a causal role in physical events, we should see that the role of mental properties is to explain the context of a physical event. This theory, unlike Davidson's, has no conflict with the completeness of physics and neither does it succumb to the problem of explanatory exclusion.

Dretske begins by making the distinction between the relations that structuring causes have to their effects and the relations that triggering causes have to their effects. Triggering causes are sufficient for a certain event and give rise to causal regularities such as follows; that

> whenever *T* occurs in *these* conditions (the conditions existing at the time of *T*'s occurrence) *E* occurs. (Dretske 1993: 123)

Unlike triggering causes, structuring causes occur in conditions that are not sufficient to explain a certain event, as later events that are independent of the structuring cause, for example, the key turning in the ignition, must occur for the structuring cause to be an explanation for the event. Unlike the triggering cause, therefore, there are no such causal regularities; for example, the terrorist plants the bomb, but no-one ever turns the key. This would mean that what would have been the structuring cause in this example would never be activated as there is no explosion that warrants the description 'structuring cause'. The structuring causes are not necessitated even though the intent is still there; the structuring cause is only a cause in potential. Yet if someone does turn the key then the structuring cause of the resulting explosion is the terrorist's action. Another difference between triggering and

structuring causes is that the structuring causal relationship is a 'one-many' relation, while the triggering causal relationship is a 'one-one' relation. Let's consider another example that Dretske uses to emphasise his argument. Imagine that Tom, a computer operator, moves a cursor on a screen by pressing a key on a keyboard (Dretske 1993: 122-123). The triggering cause in this example is the pressure on the key that is the cause of the movement, and the structuring cause is the fact that the previous day Tom rewired the computer so that pressure on that certain key will make the cursor move. For every token movement the cursor makes there is a distinct pressing of the key that is the cause of it. Every distinct effect is produced by a distinct cause and every distinct cause produces a distinct effect. However, as Tom's action the previous day, when he rewired the computer, is the structuring cause of that token event, then it is also the cause of each movement of the cursor. Tom does not have to repeat his previous activities, of rewiring the computer, over and over to make the cursor move; the one action can cause the numerous movements of the cursor in the future. His token action will become the cause of the many token movements that are caused by the key continually being pressed. The previous example of the terrorist is unsuitable to demonstrate this as this event is caused by both triggering and structuring causes and is a one-off; the condition the terrorist created by planting the bomb is destroyed as soon as the bomb explodes and would require another structuring cause to repeat the event. However, say, for example, that one day I decide to rewire my light switch and then proceeded to use this light switch regularly for the rest of the week. The structuring cause for the light going on on subsequent days is my action of rewiring the light switch earlier on in the week and it is the same individual cause. However, the triggering causes are different every time, as each token, distinct time that I flip the switch to put the light on or off is different on each occasion. One structuring cause can be the cause of many distinct events, whereas a triggering cause is an individual, token cause. Dretske claims that just as actions can operate as triggering causes and structuring causes,

similarly different states of the same object, for example, an object possessing different properties, can function in the same way. For example, a property of an object might be the triggering cause of a certain event while a different property of the same object can be the structuring cause of the same event.

This is the main point of Dretske's argument as he attempts to show that the difference between biology and psychology is that biological explanations provide the triggering causes of an event, while psychological explanations provide the structuring causes of an event. The conclusion that Dretske is aiming for is that the mental provides the structuring causes for physical events, while the physical provides the triggering causes for the events. Dretske attempts to provide some independent specification of the difference between a biological and psychological explanation of behaviour (Dretske 1993: 131-132). He begins by differentiating between an object's intrinsic and extrinsic properties. Extrinsic properties are those which are relational, as they consist of the object's relation to other things. For example, a £10 note has intrinsic properties such as shape, size etc, and extrinsic properties, such as monetary value, which are dependent on the place of origin where it was printed. Extrinsic values are potentially changeable depending on the context. For example, historically a £10 note could have had a completely different monetary value if the legal system had developed differently. Similarly, printed words have physical properties by which we identify them, as well as a meaning which is an extrinsic property of the word. Dretske believes that mental states, such as desires and beliefs are the same as words, in that their properties are extrinsic and relational. Therefore, Dretske argues that if materialism is true, and if we do not accept epiphenomenalism, then it is plausible to believe that these material states with content and meaning, beliefs, desires, purposes and intensions, etc.., derive their

meaning from their relation to other things, just like words. The content of the belief, desire, or intension, is no more in the head than the meaning of the words that we use.

> If there are beliefs, internal states that not only represent, but can also misrepresent, the state of the external world, and if these states are material states of the believer, then they derive their representational content – and, hence, their identity as beliefs – from their extrinsic relations to other things. (ibid:130)

If this is true, then mental states will gain their causal explanatory role in behaviour because it is both extrinsic and intrinsic explanations that are needed to explain behaviour. Dretske assumes that this distinction between extrinsic and intrinsic is a fundamental component in what is judged as mental and what is judged as physical. A thought is relational: it concerns how the brain in which a thought occurs is related to the representation of the thought. He therefore decides that he has gone some way in demonstrating the causal relevance of the mind, whereby it does not get disqualified from having a causal role in the explanation of behaviour, and therefore avoiding the problem of mental causation. It is the brain's intrinsic properties that are influential in triggering behaviour of the body, and it is the brain's extrinsic properties that are influential to structuring the same behaviour. By showing that the extrinsic properties, as well as the intrinsic properties, are necessary when it comes to the explanation of behaviour, Dretske believes he has shown how the mental can have a causal role in explanation of behaviour, while also avoiding the problems of explanatory exclusion and mental causation.

Conclusion

In this chapter we have considered three different proposals that have attempted to give the mind causal relevance while also avoiding the problem of mental causation. These proposals all took a stance to defend macro-causal explanations against the view that only micro-causal explanations are relevant in events, and to save the mental as well as the special sciences from becoming irrelevant in causal explanations. Fodor attempted to give intentional states causal powers by claiming that the special sciences do not have to be abandoned or ignored in favour of the natural sciences as macro-causal explanations are just as causally valid as micro-causal explanations. Baker proposed that instead of concluding that the mind cannot have any causal role in our behaviour, or that it can only have causal power in virtue of the physical, we should abandon the causal aspect of our metaphysical background and accept top-down causal explanations. In doing so, we are not only saving the mental but we are also giving credence to explanations other than micro-causation explanations. Dretske argued for the mind to have a causal role in events in virtue of its being the structuring cause of an event, without which we would not have a full explanation of the event in terms of the context in which it occurred. By contextualising explanations, he is showing how micro-causal explanations are not always as relevant as we have been led to believe and that macro-causal explanations should be taken into account and sometimes given as much relevance as strict physical explanations. In the next chapter we shall discuss the criticism that all three of these theories come under and see whether they can stand up to the criticism or have to fold beneath it.

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Chapter 6: Against macro-causal approaches

Introduction

The three theories we discussed in the last chapter all attempted to make the mind matter by granting the mind autonomy from the physical, while also giving mental states causal relevance. Jerry Fodor (1990) argued for this by proposing a theory which claimed that mental states have causal relevance as long as there are intentional laws that cover intentional properties. In doing so, he not only tried to give intentional properties causal relevance, but also prove that the laws in the special sciences are good enough to back up causal claims while also not being reducible to the laws of physics. Lynne Baker (1993) proposed that instead of going against our common-sense views in favour of the metaphysical background against which this debate is conducted, we should instead embrace these views and abandon certain parts of our metaphysical background that concern the nature of causation. By doing this we will give the mental causal relevance as it will no longer be constrained by strict physical causation. Not only this, we will also be saving explanations in other disciplines that would otherwise be eliminated from having a causal role in explanations if we do not

abandon this conception of causation that we are currently taking for granted. Fred Dretske (1993) believed that he could grant mental states a causal role in explanations by making the distinction between triggering causes, which are the token physical causes, and structuring causes, which are the mental causes, used to explain the context of the event. He claimed that without mental explanations an event would not have a full causal explanation.

However, while these explanations are intriguing and satisfy our common sense beliefs, can they really hold up against the tide of philosophers and scientists who take for granted that the nature of causation is ultimately strict and physical? Also, when considering the previous chapters that discussed Dray (1957) and the idea of methodology versus metaphysics we found that while Dray put forward a good argument for the difference between explanations for the physical sciences and explanations for the social sciences, his methodological approach did not solve the problem of mental causation as it was merely confusing causation with causal explanation. Maybe these three new theories that attempt to make the mind matter more are making the same mistake and are involved in what Quine (1960:216-221) might describe as 'loose talk'. In this chapter we will be discussing these criticisms of Fodor, Baker and Dretske and discovering whether any or all of them can still put forward a theory that gives the mental causal relevance while also managing to avoid or solve the problem of mental causation. We shall begin by evaluating each individual theory and the problems they face before moving on to evaluate these theories' whole approach to the problem of mental causation. If we find that they can manage this task then we can consider our work done. However, if we find that these techniques for solving the problem of mental causation either fall prey to the problem or have made the mistake of confusing causation with causal explanation and so are not successful, then we have not only ruled out a micro-causal metaphysical solution, we will have ruled out a macro-causal metaphysical solution as well.

If this is indeed the case, and methodological solutions also cannot help, as we have already seen, then what other options do we have to consider to find the answer to this problem?

Jerry Fodor

As we saw in the last chapter, Fodor (1990:137-159) argued that the view of causation that Davidson (1970) and Kim (1984b) hold is too strict and that our common sense views about causation concerning sciences other than physics should not be overruled. This conclusion was based on his argument that all that is needed for there to be causation is that the event be subsumed by a law, and that these laws can be provided by the special sciences *ceteris paribus* laws. The only difference between strict physical laws and *ceteris paribus* laws is that with the latter there is a further story to be told about the underlying mechanisms of causation. If it can be proven that there are such intentional causal laws in the special sciences then, according to Fodor, this is all that is needed for causation: mental properties could have causal efficacy without resorting to supervenience on the physical.

However, here Fodor seems to have made a crucial mistake in his argument. If he is basing his argument for the autonomy of the mental and the special sciences on there being intentional laws that differ from the strict laws of the physical sciences, then surely it is a mistake to claim that there is a great similarity between the two laws. This leaves him open to criticism from philosophers such as Hempel (1942). Hempel we discussed earlier in his debate with Dray over the idea of methodological unity and the similarities between the explanations in the special sciences and the physical sciences. During this discussion we found that Dray could not argue against his claim that the explanations in the special sciences

and the physical sciences had no major differences between them; the only difference being that with the special science explanations there was something more to be told about the causal story. This is a very similar argument to that of Fodor, but unfortunately this seems to be his undoing. For if the two types of law that Fodor mentions are so similar, then there seems to be nothing, in principle, preventing a reduction of intentional laws to physical laws.

It is interesting to compare Fodor with Quine. Quine claims that if we wish to discuss the nature of reality then the only language we should rely on is that of the physical sciences. When we discuss intentional statements that concern meanings and mental states this is only because these are the most practical terms in which to discuss such matters. However, these statements are not ontologically significant. For Quine the idea of intentional laws which govern intentional properties and that are autonomous of the physical sciences merely convinces him of,

> the baselessness of intentional idioms and the emptiness of a science of intention. (ibid:221)

Fodor takes almost entirely the opposite view, by attempting to grant the laws of the social sciences the same status as the laws of the physical sciences. But in doing so, Fodor makes a huge mistake. Claiming that there is little difference between the two types of laws does give intentional laws a certain standing, but it also begs the question of why, if there is such little difference between the two, the laws of the mental cannot be incorporated into physical laws.

So, it seems clear that Fodor has succumbed to the same criticism that was levelled against Dray. By attempting to grant the mind and the special sciences autonomy from the physical, he has fallen into a trap from which he cannot return. His theory hinged on the fact that the special sciences had their own laws which were autonomous from those of the physical

sciences, but as we have seen they were not immune from unity with physical laws. In which case, Fodor's whole argument collapses as it is clear that the only laws that can truly capture causation are those of the strict physical sciences.

Lynne Baker

In the previous chapter, Baker (1993:75-95) proposed a theory that would grant the mind causal relevance by arguing that we should abandon certain parts of our metaphysical background that concern causation and that are generally taken for granted. In doing so, we would be granting the mental causal relevance as it would no longer be constrained by the strict physical nature of causation. It would also be saving explanations in all but the most basic of physical sciences, which would otherwise end up becoming either irrelevant or reducible. The explanations in such disciplines as geology, meteorology and chemistry would become invalid or second-order when compared to the explanations of strict physical science if we do not abandon the micro-causation prejudice within our traditional metaphysical background and accept a macro-causal explanation for events which coincide with our common sense views of explanations.

This argument has at heart the view that we should believe in our common sense views and not reduce everything to basic physics. However, how does Baker determine which common sense explanations we should embrace instead? We may not be left with any metaphysical doctrine to ascribe to or any ideas about objective reality, but how will we fill the void that these explanations leave? What Baker claims is that we are left with explanations that are tried and tested and have been proven to work effectively, a sort of pragmatic view of

explanations. We use the explanations that are useful to us at the time we are using them. Yet, common-sense views generally, including those about causation, cannot always be trusted. Baker has advocated a U-turn to abandon a part of our metaphysical knowledge that has crucially underpinned philosophical theories for years and all of this just so that our common sense views concerning the mind can be upheld. However, when our common sense views that the Earth was at the centre of the universe and that the Sun revolved around it were overturned by Galileo's scientific discovery we had to let those common sense views go rather than abandon proven scientific fact. So why should this instance of common sense versus what is held to be scientific fact be any different? There is much evidence to suggest that the causation which Baker wishes to abandon is indeed at the basis of all causation, namely the evidence for the Completeness of Physics, and ignoring that fact will not make the problem go away or make the mind matter more. As we saw with Dray, common sense explanations are just that: explanations. If we are to believe in a true, physical causation, then we may well require something more substantial than a common sense explanation.

According to Baker then, what we are left with is a hypothetical range of tried and tested explanations, both common sense and scientific, which we will then work downwards from, rather than work from the basis that everything at its core has to be based on, or reduced to, micro-causation. Baker claimed that if a common sense view, such as that my mental states affect what I do, does not fit in with the strict nature of causation that is in our metaphysical background, then we should abandon the strict nature of causation rather than our common sense view. However, this argument for top down causation is not without its problems. If we are left with top down causation then what does Baker suggest we do if we have a top down explanation of causation that does not fit in with well-confirmed microphysical theories? Must we abandon such theories to accommodate common sense views? Why not instead abandon the idea of causal efficacious mental states in favour of strict physical explanation?

For example, it was once believed that burying a potato relieved a sufferer of warts. Yet, if science were to discover proof that there is no causal mechanism between burying a potato and curing warts, should we cling to our common sense views, or abandon science? In short, without micro-causal explanation, we may have no way to know whether to trust common sense.

It seems that while Baker has an intriguing theory that does justice to and appeals to our common sense views, there would be nothing suitable left to replace the strict nature of causation and ground causal explanations. Her theory also begs the question of where this would end in terms of abandoning scientific explanations in favour of our common sense notions. While we all believe that what we think affects what we do, we must find a way to incorporate this view into the tried and tested physical nature of causation rather than start to abandon aspects of this metaphysical background to save our beliefs at any cost.

Fred Dretske

Dretske (1993:121-136) believed that he could give mental states a causal role in explanations by making the distinction between triggering causes and structuring causes. Triggering causes are the token, physical causes, while structuring causes are the mental causes that are used to explain an event in the context of its surroundings. Dretske then went on to claim that physical objects have intrinsic properties, whereas mental states have extrinsic properties, but that both are necessary when explaining an event. Dretske aimed to grant the mental causal relevance in explaining events by showing that without the explanation of the mental, the event would not be fully causally explained. So, for Dretske, it

is this distinction between the triggering and the structuring causes of an event that grants the mental a causal role in explanations, as it distinguishes the role that the mental plays in an event from the role that the physical plays in an event. The structuring cause is the background condition that the triggering cause occurs in. So, for example, the structuring cause of a bomb exploding is that the terrorist planted the bomb, whereas the triggering cause of the event is the key turning in the car that the bomb had been wired to.

Dretske claims that it is right to describe both triggering and structuring as causes, as which cause we will need to explain an event depends on what we wish to explain. However, we have to consider whether these types of causes are both necessary when explaining an event. Of course, it could be claimed that the structuring cause of an event is necessary otherwise the triggering cause could not occur. Yet, when describing the true, physical cause of an event we would have to claim that the bomb went off because someone turned the key. As Dretske himself notes, the bomber could have planted the bomb, but with no-one to turn the key, the bomb would not have gone off.

It is also suggested in Dretske's argument that what we are looking to explain is not just one event, but two (Dretske 1988: 42-45):

> The first type of cause, the triggering cause, causes the process to occur *now*. The second type of cause, the structuring cause, is responsible for its being *this process*, one having *M* as its product, that occurs now. (Dretske 1988:42)

Later in his argument, Dretske also claims that a;

> s*tructuring* cause…helps explain, not why *D* or *M* is occurring now, but why now, *D* is causing *M* (rather than something else). (ibid.:114)

So here, it seems that we do not have two different *explanans* to explain one *explanandum*, but we have two different *explanans* to explain two different *explananda*. Instead of having, as Dretske seemed to claim, one event to explain, we have two. We have to explain why the bomb exploded when it did, and we have to explain why it *was* a bomb exploding rather than anything else instead. This means there are two things in need of explanation for the bomb explosion. This is a problem for Dretske because it suggests that the distinction between the triggering and structuring causes corresponds to the distinction between what the mental cause is explaining and what the physical cause is explaining. Instead of having a structuring mental cause that is invaluable to a physical event, then, what we seem to be left with is the dilemma that Davison faced; how to make the mental causally relevant in physical events given the problem of explanatory exclusion. As Dretske concedes that what the structuring and triggering causes are explaining is not the same thing, then how can he claim that the structuring cause can be contributing to the explanation of the physical event and therefore the overall explanation? What Dretske has done is to show that the structuring cause is not indispensable to the physical event, because all it was doing, in essence, was explaining the cause of the mental event; the structuring cause explains how the bomb came to be wired up to the car, and not why it exploded. By giving us two things in need of explanation, it seems obvious that the structuring cause is explaining the 'mental' event and the triggering cause is explaining the 'physical' event. The triggering cause is the real explanation of why the bomb went off, leaving the structuring cause to just provide the context of why the bomb happened to be there. This means that the structuring cause has no causal explanatory relevance to the bomb exploding as the triggering, or physical, cause can explain that with no competition. All that the structuring cause is useful for is non-causal explanatory purposes, i.e. explaining how it came about that the world contained a certain cause that consequently brought about a

certain effect. So, we yet again have the split between the mental and the physical with no way to bridge the gap and make the mental causally relevant.

Unfortunately for Dretske, bringing this problem to light means that he has failed in granting the mental causal relevance in physical events. His theory attempted to do this by trying to show that the mental explanations had invaluable significance to physical events as without them we could not explain a part of the event that the triggering cause did not cover, and therefore we could not understand the event fully. However, by mentioning the difference between what the structuring cause and the triggering cause explained, it brought to light the problem that the structuring cause was not explaining the same thing as the triggering cause and therefore could not be as relevant to the overall explanation as Dretske had hoped.

Fodor and Baker

All of these criticisms, however, pale into insignificance when considered next to what appears to be a reoccurring problem. Just as we saw with the arguments against Dray, these approaches by Fodor and Baker appear to have mistaken the difference between causation and causal explanation. If we take for granted, like so many scientists and philosophers do, that causation is a real relation in nature and not just an explanatory tool, then Fodor and Baker are not answering the problem of mental causation, they are merely putting forward theories on how the mind can have relevance in causal explanation. This is by no means solving the really contentious issue as it is already clear that mental explanations are necessary when explaining both physical and mental events. Yet what they have not put

forward is a theory of how mental properties can be causally efficacious in physical events, which is the problem in hand.

So, in conclusion it appears that once again the criticisms against the theories that have been proposed by Fodor, Baker and Dretske are more persuasive and damning than the advantages of accepting them. While Fodor and Baker have indeed put forward arguments that seem to make the mind matter more in the physical world as well as in explanations for physical events, they have confused causal explanation with causation in nature and have not managed to solve the problem of mental causation. Dretske, likewise, put forward a theory that tried to give mental explanations explanatory relevance in physical events, but ended up negating his whole argument by claiming that what the two types of causes were explaining was not the same thing at all and thereby limiting the amount of relevance that the structuring causes had to the causation of physical events.

Now, the question is; where do we go from here? It might seem to the casual observer that we have already covered all the options. A metaphysical approach manages to be compatible with the idea of the physical nature of causation but could not discover a way in which mental properties could be causally relevant without resorting to reductionism. The methodological approach decided to tackle the problem differently and instead concentrated on the difference between the physical sciences and the social sciences and how they have different ways of explaining events. By highlighting the two differing methodologies, advocates of the methodological approach believed that they had given the mental causal relevance as both methodologies have equal importance and should not be compared to each other. The problem of mental causation arises, it is argued, when these methodologies are thought to be in competition with each other. By acknowledging the importance and equality of the two methodologies, this approach claimed to have given the mental causal relevance

and shown the problem of mental causation to not be a problem at all. However, while this theory focused on the importance of mental explanation it faced several criticisms. The first criticism concerned the alleged differences between the explanations of the physical sciences and the social sciences and we discovered that actually both the sciences used similar methods to deduce their explanations putting them in competition with each other. The second criticism levelled against this approach is that it was merely confusing causation, as a real relation in nature, with causal explanation. In doing so, the approach does not actually manage to give the mental causal relevance and thereby avoid the problem of mental causation. The third kind of approach we have just discussed, and while it attempted to make the mind matter in the physical world, it too could not avoid the criticism that it was confusing causation with causal explanation. Thus, we seem to have reached a problematic standstill since all of the conventional approaches we have discussed have been unsuccessful. Therefore, maybe we should be considering less conventional approaches to this problem. In the next few chapters we shall be doing just this by discussing the ideas of Richard Rorty, and considering whether his arguments could be applied to the problem of mental causation.

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Chapter 7: Rorty, Davidson and Mental Causation

Introduction

In the previous chapters we have discovered that many different approaches to the problem of mental causation have fallen short of the goal of making the mind matter in the physical world without having to reduce the mental to the physical. In the last section of this paper we shall be discussing the work of Richard Rorty, whose controversial goal of trying to undermine the whole of philosophy, thus bringing the subject to a halt altogether, is at the heart of all of his arguments in philosophy. Over the next few chapters we shall be discussing his controversial claims, from his attempts to undermine the reasons for believing that the mind-body problem and the problem of mental causation are problems at all, to his ideas on the nature of truth and language. Along the way we shall be investigating where he truly stands on the problem of mental causation by comparing his views with those of Davidson. In

this first chapter on Rorty we shall be introducing him in contrast to Davidson's views. Rorty is very well known for taking philosophers' words and twisting them to suit his argument in a way that they claim they did not mean. Such a philosopher is Davidson who has claimed that Rorty had committed this misdemeanour, yet this is the philosopher of whom he argues he has the greatest similarities with. So, has Rorty just manipulated Davidson's views to fit into his overall agenda, or has he seen something in his work that does agree with his own pragmatic views? The goal in this chapter is to discover whether Rorty and Davidson are more alike than the latter has previously admitted to.

First of all, I think I should give a general background of Rorty's most controversial work that is at the heart of the next few chapters; *Philosophy and the Mirror of Nature* (1979). In this book, Rorty, sometimes just for the show of it, tackles many of the biggest philosophical questions that are around today in his attempt to bring about his overall goal: the end of philosophy as it has been traditionally conceived. Not that he thinks that people will ever stop discussing philosophical problems or reading philosophical texts, but what Rorty is aiming for is that when people are doing so they do not actually think that they are making any progress or solving anything 'real'. He wishes to undermine the;

> reader's confidence in 'the mind' as something about which one should have a 'philosophical' view, in 'knowledge' as something about which there ought to be a 'theory' and which has 'foundations,' and in 'philosophy' as it has been conceived since Kant. (Rorty 1979: 7)

He does not wish, as Kant did, that philosophy be put on the 'secure path of science' (Kant: 1933: 17) and thereby become an autonomous subject with the same authority as geology, biology and history, for example. The main objection against philosophy as he sees it is its

reliance on and perpetuation of the concept of an objective truth that can be discovered by relating our language to the world. The reason for this objection to philosophy's conception of an objective truth is that Rorty sees this as transferring our past dependence on the church to provide privileged access to the truth, to our dependence on the objective world to provide the truth. Instead of a blind faith in God, we now have a blind faith in science and believe that there is an objective world out there that can provide all the answers that we desire, and that there is one true description that can capture the entire nature of reality. Rorty, as a pragmatist, will happily use scientific explanations if they are the most useful at the time, and science undoubtedly has been very useful to us. However, there is no reason why we should abandon looking for other ways to describe the world and settle instead on scientific descriptions only, as it is only down to historical contingencies that we rely on science as much as we do. Because of this, the only true way to give up on the old religious ties and become 'consistent' atheists (Rorty 1998: 62) is to abandon any notions of realism that hold that the nature of reality is intrinsic and independent of our experiences of it or how we choose to describe it. If we did embrace Rorty's view and abandon realism he claims that we would realise that,

> what matters is our loyalty to other human beings clinging together against the dark, not our hope of getting things right. (Rorty:1982:166)

The consequences of this view would be something similar to the argument put forward by Dray (1957) or Baker (1993) earlier in this thesis, that there are many forms of description that are not absolutely ranked above or below each other, albeit with the pragmatic overtone that they can be ranked on their usefulness in certain situations, which would change depending on the context. With no ontological beliefs thus forced upon us, science would no longer be the provider of the one and only truth.

Given this view, it is no wonder that Rorty places a lot of effort into his arguments for language and truth. What follows in the remainder of this thesis are Rorty's arguments against this objective truth and science as the true description of reality as well as his historical deconstructionist argument on the issue of the mind-body problem. We will see how he uses Davidson's arguments, as well as his own, to ward off the criticism put forward by Quine against this pragmatic view. Rorty only claims to enter into the metaphysical ring for therapeutic reasons, namely to attract the attention of fellow philosophers, rather than actually believing that he is solving a metaphysical puzzle.

> Although I discuss 'solutions to the mind body problem' this is not in order to propose one but to illustrate why I do not think there is a problem. (1979:7)

Yet, he has made some rather interesting claims on Anomalous Monism, which as we saw earlier, was proposed by Davidson. The superficial differences between them are great; Davidson, unlike Rorty, does believe that the mind-body problem and the problem of mental causation are real problems that can have a metaphysical, real answer, and it seems clear, as we shall see, that for Davidson, the conclusions drawn by Rorty that spell the end not only for an objective reality and truth, but also for the demise of philosophy as a whole subject, are not the conclusions that he would endorse.

To begin discussing the similarities between Rorty and Davidson, we shall start with the topic this thesis is dedicated to: the problem of mental causation. Davidson proposed his non-reductive theory of Anomalous Monism, thinking that it could solve the problem of mental causation without resorting to reductionism or epiphenomenalism. It would be assumed by

anyone familiar with the famous arguments of *Philosophy and the Mirror of Nature*, that Rorty's view on this topic would be that this problem between the mind and the body is just a historical mistake that dates back to the Ancient Greeks and that it should be abandoned accordingly. You certainly would not entertain the notion of him getting his fingers dirty by joining in the debate and actually taking up a metaphysical stance. Yet, in one of his papers (Rorty 1991: 113-125) this is exactly what he seems to do. He argues that Davidson's non-reductive physicalism is the answer to the problem of what relation the human self, and more specifically intentional states such as beliefs and desires, have to the world. He also claims that Davidson's views on language and philosophy of mind bring about his own pragmatic conclusions that mean that there is no hierarchy between disciplines, so that physics and arts such as poetry can enjoy the same status within culture. However, is the view that Rorty puts forward the same as Davidson's non-reductive physicalism, or is Rorty manipulating Davidson's work to turn it into a pragmatic stance on the problem of mental causation?

To determine this, it is useful to step back for a moment to examine the disagreements between Davidson and Quine. While both Davidson and Quine are generally considered to be physicalists, despite Rorty's preference of casting them as pragmatists, there is a great difference in their views with regards to psychological descriptions. Quine is ontologically hostile to psychological descriptions as he believes that it is only the language of physics that has ontological significance and should be thought of as the ultimate vocabulary for 'limning the true and ultimate structure of reality.' (1960: 221). The language of physics is the only language that can capture the true nature of the world, according to Quine, and psychological descriptions are merely 'loose talk'. Quine believes that if we are to discuss the true structure of reality we should not use a conceptual scheme that refers to propositional attitudes, such as beliefs and desires, instead we should employ a conceptual scheme that refers to the

constitution and behaviour of physical things. This is because Quine believe that physics is the only area of culture whose job it is to provide 'full coverage': to formulate strict and exceptionless laws governing everything that happens in the world,

> nothing happens in the world, not the flutter of the eyelid, not the flicker of a thought, without some redistribution of microphysical states…If the physicist suspected that there was any event that did not consist in the redistribution of the elementary states allowed for in his physical theory, he would seek a way of supplementing his theory. Full coverage in this sense is the very business of physics, and only of physics. (Quine 1981: 98)

However, this does not mean that he would claim that intentional idioms, which admit of exceptions and do not refer to anything not already 'covered' by physics, should be renounced in everyday life. Instead, we should consider them as expressions of how we unreflectively perceive and respond to things, or as a useful technique for carrying out our everyday activities, rather than as true descriptions of reality. Only the physical discourse is properly descriptive of reality, while other discourses are only an inaccurate 'rule of thumb'.

For Quine, then, physics is central to our conceptual scheme. It is the area of culture with responsibility for drawing up ontological distinctions and formulating exceptionless causal laws. Psychological descriptions cannot be reduced to physical descriptions, in Quine's view, since the former are intensional while the latter are extensional. This means that the reference and truth-value of sentences employing the former are variable according to context, rather than invariably capturing the state of the actual world and therefore there is a clear sense in

which Quine does not think mental causation is possible. The only causation is physical causation, because the only causal laws are physical. Davidson, by contrast, does want to grant mental events causal efficacy. To argue against Quine's conclusion, that only the vocabulary of physics can be of relevance to causal descriptions in the world, Davidson attacks the scheme-content distinction in the hope that this can undermine the hostility of Quine's feelings to psychological descriptions; this was not the stated aim of Davidson's attack on the scheme-content distinction, but it does seem like an obvious motivation for him.

Davidson (2001:183-198) rejects the scheme-content distinction as he does not believe that within knowledge or experience it is possible to distinguish between concepts and content: holism makes it impossible to separate truth from meaning. He claims that if we want to undertake a scientific study of meaning, the most objective perspective possible is that of the 'field linguist' (1986: 314-15) trying to make sense of an entirely unknown language where all they have to go on is behaviour in response to environmental conditions. As there is nothing to lift us out of the confines of language to objectively see how it relates to the world itself, the best approach is to project ourselves into the language and see how it relates to the people's environmental conditions until we can eventually interpret it. The only way to distinguish the truth in a foreign sentence, then, is by translating it into our own language and evaluating it in our own terms. Thus, there is no prospect of a sentence being true but untranslatable; the only way we could ever be in a position to say whether it was true or not would be by translating it. An objection to this would be that the translation may just be the best available approximation within our conceptual scheme, even though the sentence says something in the foreigner's conceptual scheme which is true, but nevertheless untranslatable by us. However Davidson argues that to be able to even begin to interpret a foreign language, we must presuppose that most of what they believe is true. This supposition is known as the

'principle of charity' (2001:197). This is due to the idea that we could never find a coherent pattern in the noises other beings make unless we assume they believe roughly what we do and are talking about the world as we understand it. This is because the field linguist only knows the conditions in which foreign sentences are uttered, but does not know what they mean or if they are used to express true or false beliefs. Only by assuming that the sentences reflect true beliefs can the noises be correlated with environmental conditions, and finding such correlations is the only way to translate the language. However, if we cannot translate their language then we are in no position to make claims about what they mean or believe, and hence we could never be in the position to assert that they are saying something true, but untranslatable. Even God could never be in such a position, Davidson claims.

This means we could never be in the position to assert that there are alternative conceptual schemes, and so Davidson concludes that the scheme-content distinction is an illusion or the 'third dogma of empiricism' (ibid: 198), coming after the two dogmas Quine exposed, namely reductionism and analyticity. In fact, the only way of understanding the world is the one we have always employed, says Davidson, since most of our beliefs about the world must have always been right, and it is not really a 'conceptual scheme' at all, given that we cannot coherently talk about a 'content' it is supposed to 'organise'. While Quine's case for the supremacy of physics might initially seem convincing, Davidson's argument against the scheme-content distinction undermines it by showing that we cannot separate our conception of the world from the world itself in order to claim that the conceptual scheme of physics is superior to any other, or that physics is the ontological centre of our conceptual scheme. Davidson's innovation is to realise that the content, or the world-as-it-is-itself, is not doing any work in making our beliefs true and is simply a dogma. He thus undermined the distinction between the world itself and our conception of the world; hence there is no

scheme-content distinction. Rorty sees this view, in very grand terms, as representing Western culture's final realisation that we do not need external guidance (Rorty 1979: 257-311): we do not need to think of our way of seeing things as just a scheme guided by the world itself or as a scheme organising the content so that we can interpret it correctly. Given Davidson's argument, we can see that without the world making our beliefs true, we can use whatever methods we wish to describe the world, whether these are from the physical sciences or disciplines such as psychology and history. While physics is certainly more useful when describing the microphysical nature of the world, there are other vocabularies that are equally useful when describing other aspects of the world, and these should not be ignored.

Assuming Davidson's defence of psychological descriptions has been sufficiently validated, Rorty can now defend his non-reductive physicalism to bring an end to the problem of mental causation. However, the question still arises as to why Rorty is putting forward a defence of this view as it is a metaphysical stance on a puzzle that Rorty would apparently not believe we should even be bothering with. We can begin to guess his motives when he aligns Davidson's views with his own and claims that Davidson is committed to a fully pragmatic account, whether or not Davidson admits this to himself. By aligning himself with one of the world's most famous philosophers, he grants his views a credence that he would not get without the alliance. So we shall examine the similarities and differences between the two and see whether Rorty is right in claiming that Davidson should hold to the pragmatic view, or whether he was overplaying his hand. As we saw in chapter 2, Davidson's Anomalous Monism was not without its problems, which despite Davidson's best efforts could not be overcome. Rather surprisingly, Rorty took it upon himself to defend Davidson's arguments and also combine physicalism with a pragmatic view on the problem of mental causation. In doing so, Rorty claims that Davidson's Anomalous Monism can solve the problem of mental

causation without having to resort to reductionism, as well as fulfilling his own pragmatic ends. However, do Davidson's arguments for non-reductive physicalism really have pragmatic undertones that lead to a fully pragmatic account of the problem, or has Rorty misrepresented Davidson's views?

To see how Rorty comes to the conclusion that Davidson's non-reductive physicalism can simultaneously solve the problem of mental causation and contain pragmatic overtones, we shall see how Rorty uses the similarities between his own pragmatic arguments and Davidson's arguments to reconstruct this non-reductive physicalist view. The first position that both Rorty and Davidson share, according to Rorty, is the one that we have just discussed; the abandonment of the scheme-content distinction. In doing so, we can see that if there is no distinction between the world and how we conceptualise it, then there is no distinction between beliefs that are made true by other beliefs and beliefs that are made true by the world. Therefore, there is no point in ranking vocabularies, in the way that Quine ranks physics above psychology.

The second position that causes agreement between Rorty and Davidson, according to Rorty, is the abandonment of a representationalist view of beliefs. Davidson does not construe belief as a 'mixture' of subjective belief and objective content, since these elements cannot be separated. Therefore, we cannot investigate knowledge or aim for an accurate representation of reality, as we can never accurately make a distinction between the parts of the knowledge that come from us and the parts that come from the 'world outside'; this all follows from the rejection of the scheme-content distinction. For Rorty, then, there is no sense in researching,

> the idea of knowledge of, or successful linguistic reference to, a reality underlying the appearances that nature presents. (Sorell 1990: 11)

So, according to Rorty, there is no legitimate distinction between objects as they appear and as they are in themselves. Rather, knowledge is a continuous interaction in which we develop ways of coping with reality. Davidson also believes that we should not ground knowledge 'on something that counts as an ultimate source of evidence' (Davidson 1986: 313). Given this agreement we can see how, if there are no representations of the world then we will cease to think of physics corresponding to anything which it accurately represents and so we will just use the best tools, or vocabularies, for the job, rather than thinking that physics is the ultimate provider of knowledge and truth.

The third position is the abandonment of the notion of the world 'making true' any of our beliefs. This view entails that if we have causal relations that hold between the world and the self and have internal networks of justifications for holding those beliefs, for example the world giving us reasonable justification for holding a belief such as seeing rain and the acquisition of the belief that it is raining, we need no more relations to add to the picture. What we have is a continual reweaving of intentional systems which occurs when there are acquisitions of new beliefs and desires. We need not worry about finding things in the world that make our beliefs on algebra, aesthetics or morals true. While there are causes for acquiring beliefs, and reasons for retaining these beliefs in our internal network of belief systems, there is no correspondence or representational relation to the world through which we make contact with a 'truth-maker' that is the cause for the truth of beliefs. As Davidson claims,

> Nothing…no *thing*, makes sentences and theories true: not experience, not surface irritations, not the world can make a sentence true. (Davidson 2001:194)

So, given all of these arguments, according to Rorty, Davidson has provided us with the best of both worlds with his non-reductive physicalism. What we are left with is a physicalism whereby our reasons for acting, such as beliefs and desires, can be classed as causes, and yet we also have no reason to deny the 'full coverage' of physical causation. There is no more difference between talking about the mind and the body as between the micro-structural and macro-structural descriptions of a table. The physicalist can give the praise to the physical sciences which they duly deserve as they describe strict physical laws which apply to everything in the universe, but he cannot claim that these sciences grasp the true nature of reality or are above any other discourses; none of them correspond or fail to correspond to language-independent 'truth-makers'. So, on this account, rather surprisingly, Rorty has taken Davidson's theories and fashioned himself a non-reductionist physicalist account that he agrees with.

However, is Rorty justified in claiming that Davidson's non-reductive physicalism can solve the problem of mental causation and fit in with Rorty's own pragmatic views? As we saw earlier in chapter 2, Davidson wished to grant causal relevance to mental states in physical events, but he also maintained that the vocabulary of physics is the only language that can capture causal laws. There is no hint in Davidson's views that physics is simply a useful vocabulary, as Rorty thinks. Rather, Davidson's thesis of the nomological character of causation holds that causation can only occur under strict laws, and that physics is the only vocabulary that captures these strict laws. So, we can see that Davidson does not seem to hold the pragmatic views that Rorty ascribes to him as this claim means that he does still privilege

the physical vocabulary. Even more damning for Rorty is that, despite wanting to align himself to Davidson to abolish the problem of mental causation once and for all, Rorty has no other reason for agreeing with this view. Davidson's underlying reasons for developing Anomalous Monism are to balance the nomological character of causation and the principle of mental anomalism, but Rorty rejects both of these notions. For him, there is no 'true' causation in the world, any more than there is anything 'special' about the nature of the mind. The fact that Davidson does not wish to give up on the anomalous nature of the mental, nor on the notion that there is strict causation in the world which only the laws of physics can grasp, shows how mistaken Rorty is in believing that Davidson is a true pragmatist. To see how and why Davidson holds on to these notions while Rorty is happy to leave them by the wayside, we shall compare their views on the notion of the specialness of the intentional and the notion of the physical world having an effect on our beliefs.

For Rorty, the idea of philosophical questions and topics is something that needs to be abandoned. Therefore, he does not understand why Davidson continues to insist on there being a philosophically interesting difference between the intentional and the non-intentional, the mental and the physical (Rorty 1996: 575-594). He criticises Davidson for not coming to the realisation that his previous views on language and truth undermine the Quinean distinction between the intentional and the non-intentional and claims that Davidson should not fall into the trap of taking this distinction seriously (ibid.: 575-576). By accommodating this Quinean distinction, according to Rorty, Davidson is not following his argument to the pragmatic conclusions he should accept. Since Davidson denies that there are any sentences that are made truer than others in virtue of their having a relation to reality, there is no reason to consider the notion of whether we can or should be able to reduce one vocabulary to

another, as the only distinction between them is how well vocabularies fulfil a purpose. To draw a distinction between the intentional and the non-intentional is just as pointless:

> All our idioms are tools for coping with the world. This means that there can be no philosophical interest in reducing one idiom to another, nor in asking whether and how a non-extensional language might be replaced with an extensional one. (Rorty ibid: 576)

Once we have given up the notion of 'facts of the matter' and 'objective' truths, there is no reason to think of different vocabularies or disciplines as falling into different classes of 'philosophical interest' and so there is no reason to draw distinctions between intentional and non-intentional statements.

So Rorty sees no sense to the claim that Davidson and Quine make, that the gap between the disciplines of psychology and biology is greater and more insurmountable than the gap between such disciplines as biology and chemistry. Quine makes this claim because he believes that science should be given greater preference and because he distrusts 'intentional idioms'. Davidson makes this claim because he believes that there is something interesting about the irreducibility of the mental that makes it different from other vocabularies. So, while Davidson and Quine make the same claim, they have very different objectives in mind. Rorty believes that they are both wrong and claims that Quine's distinction,

> between the 'baselessness of intentional idioms' and the better 'based' idiom of the physical science strikes pragmatists like me as a residue of the unfortunate positivist idea that we can divide culture into the part in which there is an attempt to correspond to reality and the part in which there is not. (Rorty ibid: 576)

So, Rorty believes that Davidson has not understood the consequences of his previous arguments and taken the conclusions to the pragmatic ends that he has already assumed. Davidson, on the other hand, still seems fascinated by the distinction between the mental and the physical and is unlikely to give up trying to find an answer. The 'constitutive principles' governing the mental and physical are irreducibly different, according to Davidson, and although he thinks this does not rule out mental states entering into causal relations, he never shows any desire to undermine this distinction, as Rorty does. Despite Rorty's claims that Davidson should abandon this need to make a distinction between the intentional and the non-intentional then, Davidson never did. Rorty is right that Davidson does not try to reduce vocabularies to one another or say that intentional states have any priority over non-intentional states, or vice-versa, but he does think there is a principled difference in that only physical generalisations are causal, which is a much stronger claim than just saying that physical and intentional vocabularies are useful for different purposes.

Rorty seems to have exaggerated the extent to which Davidson is committed to the pragmatist view to suit his own ends. While both Rorty and Davidson agree that it is only through communication, language and comparing beliefs with other beliefs that we can form and evaluate our beliefs, and that there cannot be any confrontation between our beliefs and the world which determines the truth or falsity of statements of belief, there is a difference in how they interpret this view. For Davidson, only communicators can have a sense of the objective world and form beliefs about it:

> Nothing can count as a reason for holding a belief except another belief (Davidson 1986: 310)

To devise a way of forming beliefs we have to form a 'triangulation' (Davidson 1982: 480); two people communicate their beliefs about the world to each other to form and justify their beliefs about the world. These three corners of the triangle for Davidson are necessary; the world and the two people communicating their beliefs to each other about the world. Rorty however denies that one corner in the Davidsonian triangle exists: the world itself. In his eyes there is no other guidance than that which we derive from other humans. Without representations, which Rorty of course rejects, he thinks there is no hope of 'getting things right' or having beliefs about the 'world in itself', and so he abandons these notions as useless. Thus, there is no objective world that we interact with to form our beliefs, just other people.

So, once we give up on any notions of our language 'hooking' on to the world, that we can ever 'get things right', or that there can ever be any confrontation between the world and our beliefs, Rorty concludes that there is no point in choosing between beliefs on the merits of how responsive they are to the causal pressure of the world. Here, we see that Rorty, by commandeering some of Davidson's arguments regarding the scheme-content distinction, has managed to withstand criticisms from realists, such as Quine, and put forward a convincing argument for the lack of an objective 'world in itself'. However, here we find that Rorty has misrepresented Davidson as agreeing with his pragmatist views, whether intentionally or not, as Davidson has no intention of abandoning the idea that our beliefs are about the world. The 'world in itself' cannot make our beliefs true, for Davidson, only because to talk about it 'in itself' is to use the scheme-content distinction. However, Davidson believes that although our beliefs cannot be paired off with parts of the world, our beliefs do holistically accommodate themselves to a real world. Because we have a holistic relationship between our language and the world, that does not rely on individual words or sentences hooking on to the world, we

can understand the features of reality by examining our ways of talking about such features. Both Davidson and Rorty believe that our beliefs must generally be true, yet for Davidson this correlation is driven by the world, and for Rorty the match is due to the world being a reflection that is cast by our generally accepted beliefs.

Here we can see that Davidson is not following Rorty to a pragmatic end. Given the agreement between them up until this point you can understand Rorty's mistake in thinking that they would be in agreement on the conclusion. In some ways it appears strange that Davidson would not follow through to this pragmatic end, but for Davidson his non-reductive physicalism was aimed at combining the physical world in which true causation exists with the anomalous nature of the mental without having to resort to reductionism. Thus, he will never be able to give up on the notion that there is true causation in the world and that the vocabulary of physics is the best when capturing this causation. However, this does not mean that his physicalism is as strict as Quine's as he does not believe that there is a hierarchy of vocabularies, or that other vocabularies are merely 'loose talk'. Events remain the same no matter how they are described, for Davidson, but if we want to understand their fundamental causal relations, we must employ the language of physics, whereas if we want to understand relations of justification, we must employ the language of mind. Neither has any privileged attachment to the 'way the world is' for either Davidson or Rorty, but that does not detract from the principled distinction between these vocabularies for Davidson.

Conclusion

While the works of these philosophers seem to be forever intertwined and while there are most definitely close similarities, there are also important differences. While Rorty is sympathetic to Davidson's arguments against the reductionism of vocabularies, it is only because it supports his argument against the role of the physical sciences as the provider of an objective truth. The main problem for Rorty's claim that he and Davidson have the same view, and that Davidson has secret, previously undiscovered pragmatist sympathies, is that Davidson still holds that the distinction between physical and mental vocabularies is importantly irreducible, while Rorty thinks the irreducibility is trivial. Rorty's dismissive view on irreducibility, which he first presented in his famous paper on eliminative materialism (Rorty 1965), is that irreducibility is simply a matter of convenience. Thus we cannot reduce talk of tables to talk of atoms, because it would be incredibly complicated and impractical to do so. However, Rorty never denies that this reduction is possible in principle; it might be pointless and difficult, but we could do it if we so desired, since the world will not stop us from describing it in whatever terms we wish. For Davidson, however, mental and physical vocabularies are fundamentally different and irreducible in principle; he refuses to explain this in terms of the scheme-content distinction, as Quine did, but nevertheless this belief he inherited from Quine remains central to his philosophy.

Rorty argues the finer points of metaphysical debates, such as the problem of mental causation, to prove his point and debunk other theories, but he does not really believe that these arguments have a purpose outside of his goal of attacking philosophy; ultimately he does not think they should even be discussed. Davidson, on the other hand, is a true

metaphysician. He may stray into pragmatic territory every so often, but he does not dwell there long, and while he and Rorty do tend to meet in the middle on certain topics, it is clear from the frustration and disappointment that Rorty shows (Rorty 1996: 576) that Davidson is not a pragmatist at heart and believes that there are real, objective answers that can be found, given the right arguments. Rorty believes that he has been trying to understand Davidson's views to;

> defend them against actual and possible objections, and to extend them into areas which Davidson himself has not yet explored. (Rorty 1991: 1)

This difference is obvious when we remember, for example, that Davidson believes that causation is an extensional, two-place relation between events, while Rorty would hold that we could describe causation any way we want if it serves our purposes, and could presumably describe the world without causation if we so wished. In short, Davidson's insistence on the distinctiveness of the mental, and his desire to discover the right, rather than simply the most useful answer, shows that he is not ready to join Rorty in the post-ontological camp.

Mind over matter or matter over mind?

A comparison and evaluation of four approaches to the problem of mental causation

Chapter 8: Rorty's solution to the problem of mental causation

Introduction

In the last chapter we discussed Rorty's attempt to solve the problem of mental causation by endorsing Davidson's non-reductive physicalism. Rorty argued that the pragmatic and anti-realist notions agreed on by Davidson and himself yield a pragmatic solution to the problem. However, we discovered that, while Rorty had faithfully portrayed Davidson's views in the main, he had under-estimated Davidson's metaphysical commitment to physicalism. Due to this, the form of non-reductive physicalism that Rorty endorsed turned out to not be the type of physicalism that Davidson had in mind.

In this chapter, we will delve deeper into Rorty's view, by placing it in the context of his overall, metaphilosophical position. This position is set out in his most famous work: *Philosophy and the Mirror of Nature* (1979). In this, Rorty attacks what he considers to be the foundations of philosophy as a whole, and especially the notions of objective truth and the

mind as a reflection of this truth; the 'Mirror of Nature' of the title. Rorty claims that central to the notion of philosophy as a distinct discipline is the metaphor of the mind as a mirror, it being the philosopher's job to decide which representations in the mind accurately reflect nature, and thus form the basis of our knowledge, by its correspondence to the objective world (Rorty 1979: 3-4). The aim of *Philosophy and the Mirror of Nature* was to undermine this philosophical objective and thus end philosophy as a real subject. As we discussed in the previous chapter, Rorty has already tried to undermine certain cornerstones of this aim; the scheme-content distinction, representationalism, and the notion of the world making our beliefs and knowledge true. However, it is his attempts to undermine the mind-body problem and the notion of objective truth through historical deconstruction which are most relevant here.

Rorty's overall view on the mind-body problem is simple: he believes that this problem is merely a pseudo-problem that has no answer and is not worth attempting to solve. He argues against certain philosophical views on the mind-body problem only because he wants to undermine the whole debate (ibid.:7). No one needs an answer to this problem as it has no real relevance to contemporary life. To solve this problem will not solve anything real or important, such as world poverty, the economic or oil crisis, or cure any illnesses**.** It seems safe to assume, then, that Rorty's attitude to the aspect of the mind-body problem that we are dealing with in this thesis – the problem of mental causation for non-reductive physicalism – would be much the same. So what are his motivations for endorsing Davidson's Anomalous Monism?

The answer, I believe, is that Rorty believes that what Davidson did with his Anomalous Monism was useful in that he showed how it was possible to be a physicalist while also not having to worry about the problem of freewill or having to resort to reductionism. However, other philosophers then began to poke 'scholastic', as Rorty would say, holes in this theory and discovered the 'problem' of accounting for the efficacy of mental properties: as the causation would always be in virtue of the physical properties, they claimed, Anomalous Monism was, without realising it, a form of epiphenomenalism. Rorty would not want to respond to any specific arguments and be dragged into this metaphysical debate, because his overall metaphilosophy would persuade him that the 'scholastics' would always be able to find more and more problems for him to solve, that he would consider pointless, confused, and not worth the effort, since they would not link to real-world practical problems. Therefore, to dissolve this problem of mental causation, Rorty would step back to try to persuade the critics of Anomalous Monism that they were wasting their time.

To show this, Rorty would look back to the history of the mind-body problem, to ask why the mind belonging to the physical world was ever considered to be problematic. Rorty would claim that the reason the problem of mental causation is taken seriously is because the mental-physical distinction is assumed to be intuitive. Rorty attacks this apparently obvious notion as stemming from Platonic and Cartesian dualism; if this can be debunked, then one of the crucial premises of the problem of mental causation has been eradicated and the problem dissolved (ibid: 17-22). If there is no intuitive contrast, and this problem is merely a historical confusion, there is no reason to worry about the relation between the 'mental' and the 'physical', or to try to connect the vocabularies that describe them to one another.

So why do we consider it obvious that the mental seems non-physical (ibid: 32-37)? When discussing what sets the mind apart from the physical world there are two traditional candidates; the intentional, such as beliefs and desires, and the phenomenal, such as pain. Mental causation is primarily concerned with intentional states, such as beliefs and desires, and how they can interact with the physical world without being physical in nature or without being reduced to the physical. Yet why do we assume that these intentional states are non-physical in the first place? Rorty claims that the reason that we believe that the intentional is non-physical is because intentionality is not an observable feature of physical things (ibid: 22-27). Intrinsic meaning is not inherent in observable, physical features, such as inscriptions or sounds, as they need interpretation by a person to be understood and have meaning. These inscriptions in the sand or oscillations in the air mean nothing by themselves as we would fail to interpret them if we did not know the relevant language. Yet, once we have interpreted them their meaning is apparent. This lack of intrinsic meaning in these features encourages the idea that to interpret them there has to be meaning superimposed upon them by mental states that have their own intrinsic intentionality. Therefore, intentional states are considered to be a source of intrinsic meaning. The assumption here is that a mistake in our concept of meaning has led us to believe that physical things, such as inscriptions or oscillations, must derive their intentionality from something that is intrinsically intentional and that this must be something which contrasts with the physical.

However, Rorty objects to this notion of intrinsic meaning as he believes that all meaning is holistic (Rorty 1991:126-150). Beliefs and desires have meaning in exactly the same way that writing has meaning; through their relations to other things, such as other beliefs and inscriptions. To know what inscriptions or sounds mean is merely to see them in relation to other inscriptions or sounds, such as interpreting a foreign language by comparing it to your

own. Rorty claims that intentional states acquire meaning in just the same way. If we cannot observe intentional properties in physical things all that this means is that we need to place the inscriptions or sounds in a wider context. For Rorty, the notion of being able to walk around a huge brain (Leibniz 1973), to take an example from Leibniz, and point at meaningful brain states is not impossible: it is just that to grasp the meaning, the brain states would have to be related to other things in the world. Meaning is not an isolated affair, but a relation between the bearer of meaning and the wider context that it is in. So, it is merely a misunderstanding of meaning that leads us to believe that meaning must be derived from something that is intrinsically intentional and hence, something that contrasts with the physical.

But why were intentional states, such as beliefs and desires, ever considered to be the most appropriate states to derive meaning from? According to Rorty, it is because of their connection with the phenomenal (Rorty 1979: 31) This is due to the false notion that the only things that are self-interpreting are phenomenal states, as you cannot be wrong about phenomenal states: when you are currently thinking about something, your thought does not need to be interpreted, but is rather intrinsically meaningful. As some intentional states are phenomenal, then, and phenomenal states are self-interpreting, their possession of intrinsic meaning entails that they are non-physical. Yet, according to Rorty (Rorty 1991:21-23), the only reason that we consider phenomenal states as being non-physical is because of a historical mistake originating with Plato and his notion of hypostatised universals that exist beyond the physical world (Plato 1961 Phaedrus: 247c).

Plato (Plato 1961 Republic: 514-541b) considered knowledge of universals to be the highest form of knowledge. Knowledge of particulars in the ordinary world was problematic for the Greeks, as the changing and transitory nature of the world made it difficult to talk about knowledge of the world as it was constantly in flux. At the heart of Greek philosophy was the idea that change stands in the way of knowledge. Another problem that stood in the way of knowledge for the Greeks was that there were no examples of complete or perfect attributes in the world, such as perfect beauty, complete virtue, absolute straightness, to which we could measure all examples against. Plato's account of knowledge of the universals apparently solved these problems by introducing a dualism between the physical and changing world where we live and the transcendental and changeless world where universals reside. The world of humans was the world of becoming, and the realm of the universals was the world of being. (Plato 1961: 753d) Knowledge then becomes the contact that the soul has with the changeless universals. Ordinary or real world exemplifications of the universals, such as beautiful people or lines drawn in sand, are explained as emulating these universals to a greater or lesser extent. Rorty then claims that the Greeks' choice to model our knowledge of the universals on vision sets the scene for the mind-body problem. As the universals reside in a transcendental realm and our knowledge of them is modelled on vision, then the point from which the Forms can be 'seen' must also be transcendental (Plato 1961: Republic 507-511d). This brings about the conception of the mind as doing the 'seeing' and results in the mind-body dualism that Descartes later entrenched; the mind is transcendental and non-physical - the 'eye of the soul' (Plato 1961: Republic 533d) as Plato describes it - and the body is physical and objective.

The Greeks fixated on knowledge of universals, according to Rorty, because rationality was seen as the human trait that distinguishes us from other beings, such as animals, and objects

that lack consciousness. It is only humans who can call on their conscience to decide whether they are acting morally in how they are treating other beings. Only humans can rationalise with themselves and others, formulate plans for the future, and look back and analyse the past. Consciousness is also seen as a trait of personhood as there is a clear difference between the way we feel in how we, as humans, treat each other and animals, that are considered to have a form of consciousness, compared to things that are not considered to possess consciousness, such as chairs or vegetables. However, despite this clear connection between consciousness, rationality and personhood, Rorty wishes to show that these connections are only intuitive because we have inherited a certain way of talking from the Greeks. Rorty blames the unfortunate connections between consciousness, rationality and personhood on Plato and his hypostatised universals. As reason and consciousness, according to Rorty, were connected when the Greeks decided that knowledge of universals was the distinguishing characteristic that set humans apart from animals, it is this historical mistake that is the cause of the contemporary problem between the mind and the body.

To undermine the problem of mental causation, then, Rorty's most distinctive tactic is to argue that the central premise that generates the problem of the distinction between the mental and the physical, arose from a historical confusion that has persisted this far. This he seems to have accomplished as when we compare the Greeks notions of hypostatised universals with our own mind-body problem we can see that we have inherited certain crucial concepts about the objective and subjective realms of knowledge. Just as the Greeks considered the transcendental realm to provide them with unadulterated truths, so we consider the mind to provide us with *a priori* knowledge of the world, and phenomenal states of the mind to provide us with incorrigible truths. Here we can see the similarity between Plato's theory of knowledge and our own contemporary problem of the relation between the

mind and the body. The universal forms of Plato's theory occupy the transcendental plane because the ordinary world can only contain particulars and cannot accommodate these perfect forms. This is similar to the subjectivity of consciousness which appears as transcendent because the world is conceived by us as entirely objective and so the subjective nature of the mind cannot be accommodated within the physical world. Both Plato's hypostatised universals and the phenomenal states of the mind have no place in the physical world, which Rorty attributes to the intertwining of these two concepts from the time of the Greeks. Plato's theory, in Rorty's view, has dragged the concept of the mind with it to the transcendental plane; something which Rorty believes is nothing more than a historical mistake and so should be abandoned accordingly.

Another reason Rorty claims that we believe that phenomenal states are non-physical is because people try to explain the social practice of making incorrigible reports by attempting to give them ontological importance (ibid: 88-90). We want to justify the distinction between incorrigible and corrigible reports; why certain sensations are special, such as pains which we can never be wrong about, whereas other states are not classed as special as they are fallible. However, according to Rorty, there is no difference between our phenomenal states and our physical bodily states, as when we are describing the experience of being in pain this is no different to commenting on a change in our temperature or an increase in our blood pressure (1991:121). Rorty wishes us to see phenomenal states not as something ontologically special, but merely as a matter of how we talk. Rorty claims that all there is to privileged access, the idea that we have better knowledge of our own states than anyone else can have, is that we have been brought up to talk in such a way as to allow first-person reports about phenomenal states to be taken above any third-person reports. This practice of favouring first-person reports above third-person judgements arose because taking people's genuine first-person

reports of their phenomenal states for granted meant that people were able to predict behaviour more precisely. While it is quite obvious that discussions can occur over the nature of any physical event, as all physical events are in principle third-person knowledge, phenomenal states are automatically exempt from this; no-one can legitimately argue with us that we are not feeling what we claim to feel.

> We have no criteria for setting aside as mistaken first-person contemporaneous reports of thoughts and sensations, whereas we do have criteria for setting aside all reports about everything else. (Rorty 1970: 413)

Therefore, for Rorty, the ability to report on our mental states is not proof of a 'presence to consciousness' (1991:121) but is simply a matter of how we were taught to use those words. The uses of sentences such as 'I have a belief that X' is no different in principle from the report that 'I have a temperature', it just reflects the fact that third-person tests for temperature are better than third-person tests for beliefs. We sometimes get people's beliefs wrong, so incorrigibility turned out to be a useful social practice, but that could all change as brain-reading devices improve. So according to Rorty, this means there is no reason to distinguish between 'mental' and 'physical' reports.

Given this argument it seems that there is nothing in principle that makes phenomenal states non-physical. The only difference between phenomenal states and other physical states is that we have been brought up to speak of them as if they are something different and special. If this is the case, and the whole idea of a special contrast between mental and physical properties is a historical mistake based on the failure to realise that meaning is holistic, the influence on Plato's theory of universals on our notion of the phenomenal, and the social practice of incorrigibility, then the question that plagues Davidson, namely of whether

causation is in virtue of mental or physical properties, is similarly confused. Once we know the history, Rorty argues, we should not even want to try and answer any questions about the efficacy of mental and physical properties in events.

Rorty would not wish to propose, dispute or agree with a position on the mind-body problem at all. To do that would just endorse the problem that he believes is based on ancient historical mistakes. The best position to have on the problem of mental causation, in Rorty's view, is to either ignore it or claim that it is a pseudo-problem**.** However, here Rorty would face several criticisms; the first, that by arguing that there is no reason to think that mental properties are non-physical this means that mental properties must be physical properties and therefore Rorty is committed to a reductionist position and the second being that this dismissal of the problem in favour of a theory based on the usefulness of vocabularies is no different to the views put forward by Dray (1959) or Baker (1993). Yet, Rorty would not be bowed by these criticisms. In response to the first criticism he would claim that even though there is no reason to think that mental states are not part of the physical world, this does not give cause to embark on the odd notion of trying to reduce the mental to the physical. Rorty does not claim, as Dray does, that there is a principled difference between the vocabularies, he merely points out the uselessness of a reduction as we would be limiting our explanations of the world. The only reason that people would embark on this mission of reductionism is if they believed that there is a way in which 'the world is' and that it has a preferred way of being described, i.e. through the vocabulary of physics. However, Rorty, as we have already seen, denies this notion. This leads us to the second criticism; that Rorty's pragmatic approach to the problem of mental causation has made either the mistake of confusing causation with explanation or denying that there is strict physical causation. However, Rorty can undercut this criticism by appealing to his aforementioned views on the nature of

objective truth and the world in itself to prove that there is no such thing as 'true' causation. As we discussed in the last chapter, in all of his views on language, belief, and representationalism, Rorty completely denies that the world 'as it is in itself' has any real role to play (ibid:118-121). There is no 'world in itself' that we can ever know, or that corresponds to our language to make our beliefs true.

> I urge that, rather than trying to climb out of our own minds – trying to rise above the historical contingencies that filled our minds with the words and beliefs they presently contain – we make a virtue of necessity and rest content with playing off parts of our minds against other parts. (Rorty 1991: 14)

Rorty makes this claim as he believes that it is only historical contingencies that lead to our search for an objective 'Truth' (Rorty 1991:21) which involves our attempt to step outside of ourselves in order to observe the objective world distinct from human influence. Just as we inherited the flawed distinction between the mind and the body from Plato, we are the heirs of this objectivist tradition that is built on shaky foundations. By making the distinctions between appearance and reality and knowledge and opinion, Rorty claims that Plato brought about the notion of the 'intellectual'; a person whose knowledge of such truths transcends that of the non-intellectual (Plato 1961: 514-541b). This distinction was later entrenched by the adoption of the Newtonian physical scientist who became the model of this 'intellectual'. Ever since then there has been the human desire to step outside our community to examine ourselves with a God's eye view that transcends the world as we experience it, and that it is these 'intellectuals' who are deemed to have the answers on what does or does not correspond with this outside world. However, Rorty does not believe in this realist goal that,

> The whole point of philosophical thought is to detach oneself from any particular community and look down at it from a more universal standpoint. (Rorty ibid: 30)

What Rorty claims is that these goals were brought about, in much the same way as the mind-body problem, by a historical contingency, and should be abandoned in much the same way. If we take this view, then Rorty can face the critics' claims, as he has shown that science has no privileged access to the 'outside world' and that it is only due to a historical contingency that we believe that scientists have this special connection to objective truths. This being so, as there is no notion of 'true' causation to correspond to our actions and events, then there is no principled distinction between causation and causal explanation as all useful explanations of an event could in principle be counted as causal if we found this way of talking useful. As we saw in chapter 7, Rorty has put forward, along with Davidson, many convincing arguments for the lack of an objective world that our language, beliefs and truths correspond to. Instead of holding onto this conception of an objective reality, we should realise that there is no 'skyhook' (Rorty ibid: 13) that can lift us outside of the confines of our beliefs or language and, as there is no principled supremacy of physics over any other vocabulary, given the historical deflationist position just given, there is no reason to cling on to a notion of 'true' causation, privilege any vocabulary over any other, or advocate reductionism.

However, unlike Baker, Rorty does not try and upend our notions on causation so that common sense views take priority over physical explanations as, in principle, there should be no priority in vocabularies other than according to their usefulness. Rorty also does not deny the merits of physicalism as it takes us away from past beliefs of mysticism and towards science, which has proven its worth in our society. Yet, although physical descriptions are very useful they are not always the most useful descriptions in every circumstance as different purposes require different vocabularies and there is no reason to think that one vocabulary should be the most useful in every circumstance.

It is also worth noting that Rorty would not have been concerned with any of the Quinean arguments that consider any talk of the mental as merely being 'loose talk' while the vocabulary of physics truly captures the nature of reality. This kind of talk would only be endorsed if people believed that there is a way the world is, in which case the vocabulary of physics is primary, leaving all other vocabularies substandard and second order. To be a realist and to take metaphysics and the world in itself seriously would lead you to believe that any other vocabulary than physics would have to be 'loose talk'. However, if you take the route of pragmatism then there is no way the world is and thus different vocabularies have different purposes. Therefore, we do not have to try and reduce all vocabularies to that of physics or believe in an objective world guiding our knowledge or beliefs. For Rorty, you could in principle reduce mental properties to physical properties if you developed an incredibly complex reduction, yet you could also do the same reduction for the properties of being a chair (Rorty 1991: 114-115)**.** Yet to do this would be incredibly hard and would not justify any need for such a reduction, as there is no reason to try and accomplish this unless you believe in capturing the way the world is. For Rorty, if people are not going to abandon the idea of finding a solution to the problem of mental causation then the simplest solution is to become an Anomalous Monist like Davidson, as this saves us from the pointless quest of reductionism. However, Rorty is only half-endorsing Davidson's non-reductive physicalism as he does not believe that there is a reason in principle that you could not have a reduction of the mental to the physical. Rather he just thinks there is no reason to do this. Davidson, on the other hand, did believe that ineliminable differences between mental and physical generalisations would prevent you from making such a reduction.

Therefore, Rorty would conclude that all we need talk about in relation to the problem of mental causation is the usefulness of mental descriptions in causal talk, such as a person did

something because of what they believed, or because of a pain they felt. However, there is no reason to believe that these descriptions pick out anything other than the physical states of a person unless you have fallen for the historical misunderstandings and believe that mental states are non-physical. Moreover, there is no reason to attempt to reduce these specifically 'mental' properties of these states to physical properties unless you believe that there is a way the world is and that physics is the preferred vocabulary that the world is described in. For the same reason, there is also no reason to distrust mental descriptions, as Quine did, as merely being 'loose talk'; if a vocabulary is useful, there is no need to worry about it. Therefore, the best way to express a physicalist position is with Davidson's non-reductive physicalism, as long as we do not make the mistake of believing that there is some principled reason for this lack of reduction, or get dragged into 'scholastic' disputes which are all based on taking the mental-physical contrast for granted, and hence are all based on historical ignorance. We should just think of Anomalous Monism as a nice, and useful, way of talking about the mental and physical descriptions of the world we employ in daily life.

So, when engaged in a debate about whether causation is in virtue of mental or physical properties, the best answer to give is simply that mental causal descriptions can be very useful when describing a person's actions. Without these descriptions it would be practically impossible to predict or interpret the actions of people. If it is claimed that this position leads you to epiphenomenalism, then explain that there is no reason to believe that mental properties are not physical, even though this does not mean that anyone has to set out on the pointless journey of trying to reduce these mental properties to physical properties. In Rorty's opinion, the best stance to have, other than to just walk away from the whole debate, is Davidson's non-reductive physicalism. Davidson made only two mistakes with this theory; the first is to believe that there is a principled reason why mental properties cannot be

reduced to physical properties. The second mistake is to allow himself to become embroiled in the metaphysical debate of trying to answer pointless questions on the nature of mental and physical properties.

Conclusion

This conclusion by Rorty brings us to the end of our investigation into the different approaches developed to try and either solve or dissolve the problem of mental causation. As we have already seen, neither Davidson, Dray, Fodor, Baker nor Dretske has managed to put forward a fully convincing argument that can solve this problem. Rorty's unique approach offers something different by denying that there is a problem; 'problems' of this kind are pseudo-problems based on historical mistakes and should not be investigated. In the final chapter, we shall be asking whether Rorty's proposal does finally bring an end to the problem of mental causation, and whether going along with Rorty is a price worth paying to accomplish this.

Mind over matter or matter over mind?

A Comparison and evaluation of four approaches to the problem of mental causation

Chapter 9: Dissolving the problem of mental causation

Introduction

And so the curtain has been drawn on the fourth and final approach to the problem of mental causation that we shall be discussing. Other approaches are no doubt possible, but these strike me as the most distinctive approaches in the contemporary debate. In the previous chapter we discussed Rorty's arguments for dealing with the problem of mental causation. We witnessed how he would have argued, not for the solution to the problem of mental causation, but about there being a problem to be dealt with in the first place. What Rorty has done is to show us how the mind-body problem is merely a historical mistake. This construction, combined with evidence for the completeness of physics, made it seem that the mind must be reduced to the physical in order to account for mental causation, but for Rorty, there is little point in trying to reduce the vocabulary of the mental to the vocabulary of the physical as physics has no priority in principle over any other discourse. All that matters is what is useful to our society, and both mental and physical descriptions have their own uses. This line of thought is very different from the other solutions we have discussed previously in this thesis, yet is it any more credible? In this final chapter we shall evaluate this Rortian argument in comparison to the previous arguments put forth in this thesis to see whether this is indeed the answer to the problem of mental causation.

Back at the beginning of this thesis Davidson (1970) proposed his Anomalous Monism to solve the problem of mental causation by granting the mental causal relevance in physical events without having to resort to physical reductionism. To do this, Davidson argued that mental events were physical events and also that the mind is anomalous, not underpinned by strict laws. Unlike reductionist theories he argued that while mental events were physical events, there were no reductive bridging laws between mental and physical states, there was only token identity between particular mental and physical occurrences. Davidson claimed that this theory gave mental events causal relevance and could also preserve the completeness of physics. Davidson's theory came under heavy fire, most notably from Kim, and faced criticisms that his theory entailed the epiphenomenality of mental states. Davidson appealed to his proposals of supervenience and the extensionality of causation, but all these arguments merely revived the criticisms levelled against him. As we discovered, Davidson's theory, despite numerous attempts to disarm them, constantly generated problems that he was unable to solve, the most damning of all being how he can grant mental properties causal relevance in physical events given the problem of explanatory exclusion.

Dray (1957) claimed that the problem of mental causation is not metaphysical, but is rather a problem of competing methodologies. The problem is premised on there being competition between the vocabulary of physics and the vocabulary of psychology and the special sciences. Yet, Dray claimed that if we realise that there are different disciplines with differing domains of enquiry that use different vocabularies that cannot be put into competition, then the problem vanishes. By realising that these differing disciplines have different domains of enquiry and methods and are therefore so different to each other that to try and reduce them

to each other is impossible, Dray claimed that there is no problem of mental causation to be solved. However, we saw that Dray's stance was undermined by Hempel (1942) who argued for the idea of methodological unity; the claim that in reality the methodologies of physics and of the social sciences are very similar and so they are put into competition with each other. This leaves us with either the problem of explanatory exclusion**,** or reductionism.

Fodor (1990) wished to give sciences other than just physics causal relevance in explanations without having to resort to reductionism. He claimed that philosophers such as Davidson and Kim (1984) have firstly misunderstood what it is for a property to be causally responsible and secondly have misunderstood the relations between events and the special science laws that subsume them. By showing that there are intentional laws that subsume intentional properties his view entailed that intentional properties are causally responsible. He also claimed that all it is for a special science law to subsume an event is that the law covers the event whenever the *ceteris paribus* conditions are satisfied. Though laws are needed when covering events, unlike Dray, Fodor does not think that *ceteris paribus* laws are any different or less explanatory than strict covering laws. Yet Fodor claimed that this argument does not entail the problems Dray faced, as this similarity between the strict and special science laws would grant the latter as much credence for causal relevance as physical laws and therefore they would not become second-order in any way. However, as we saw, Fodor ultimately fell into the same trap as Dray, as by claiming that there is very little difference between the laws of physics and the laws of the special sciences, there seems no way to also claim that the laws of the special sciences cannot be reduced to the laws of physics. Unlike Davidson, Fodor does not claim that there is some principled reason why the laws of the intentional cannot be as strict as the laws of physics, and so there is no principled reason why they should not be reduced.

Baker (1993) claimed that we should abandon the part of our metaphysical background that concerns causation. Instead of assuming that all the roots of causation lie in micro-causation as these describe a real relation in nature, we should turn this conception on its head and begin with a range of tried and tested macro-causal explanations that give credence to our common sense views. In doing so, we can save not only the efficacy of the vocabularies of the mental, but all of the special science vocabularies that utilise macro-causal explanations which would, if we accepted that only micro-causal descriptions have causal relevance, become second-order or irrelevant. However, Baker, just like Fodor, succumbed to the criticism that she has merely confused causation with explanation. She also faces the problem of how we are to distinguish between the macro-causal explanations that are relevant and useful, and those that are irrelevant or untrue.

Dretske (1993) claimed that causal explanations are context-sensitive and that whatever cause we decide to pick out for an event depends on our interests and purposes. Dretske explained the difference between two types of causes, what he calls triggering causes and structuring causes. The triggering cause is the immediate cause of the event, such as the turning of the key in a car wired with a bomb, while the structuring cause is the cause of the situation required for the immediate cause to have its effect, such as the wiring of the bomb by a terrorist. Without the structuring cause the event could not have occurred; if the terrorist had never planted the bomb in the car then the car would never have exploded. However, Dretske does face the criticism, which he concedes, that instead of having one event with two differing and invaluable causes we have two events, one which is explained by the triggering cause and one that is explained by the structuring cause. As Dretske makes the distinction

between what the structuring cause and the triggering cause are describing, he has therefore shown that what the mental and physical causes are describing are not the same event. Dretske's argument for the causal efficacy of the mental relied on having one physical event that has two causes: the mental/structuring cause and the physical/triggering cause. However, as Dretske eventually conceded, what his theory actually entailed was the explanation of two events: the mental and the physical. This meant that there were two causes for two events, and this reiterated the problem that Davidson's Anomalous Monism brought about, of how the mental can have efficacy in physical events. Therefore, Dretske cannot claim that the structuring, or mental, cause is invaluable to the physical event, as all it was doing in essence was explaining the psychological event, not the physical event. The structuring cause explains the terrorist planting the bomb, and the triggering cause explains why the bomb exploded. This leaves the mental as epiphenomenal, as all the mental is doing is explaining the context of the event, rather than being relevant to the occurrence of the physical event. This simply resurrects the problem that plagued Davidson, of how to grant the mental causal relevance in physical events.

And so we are back to the beginning of the problem. None of the philosophers above have managed to solve the problem of mental causation without succumbing to an apparently insurmountable obstacle. Rorty however, seems immune to all of the criticisms levelled against the previous theories as I will now show. In comparison to Davidson's theory, Rorty believes that he has dissolved the problem of mental causation by showing that the whole problem is based on a historical mistake. He does not garner any criticism from those who wish to preserve the completeness of physics, as his theory does not claim that this is void, and he also rejects the notion of reductionism, thus warding off any dissenters who dislike the notion of reducing mental vocabulary to physical vocabulary. As we saw in chapter 7, Rorty

proposed that if we are to talk about the relation between mental and physical states at all, then Davidson's Anomalous Monism was the best way to do it: it is a nice story and allows physicalists to continue to believe in the completeness of physics, while also accommodating our common sense view that what we think affects what we do. However, as we also saw, Davidson did not take Rorty's pragmatic conclusions to the extent that Rorty wished him to and so could not hold off 'scholastic' criticisms, as Rorty might have put it. By refusing these pragmatic conclusions, Davidson became embroiled in the metaphysical debate and had to take seriously the problems that the 'scholastics' devised. By taking the pragmatic route, however, Rorty does not have to face these problems, as he can just claim that all we have to debate is what is useful to society and what vocabularies are useful for certain event explanations. Rorty claims that philosophers who continue to try and solve this problem of mental causation have not understood the argument, as the problem is based on a historical mistake that they should no longer be debating. The main advantage of Rorty's theory is that he takes the pragmatic conclusions that were offered to Davidson, but which were never accepted, and therefore he can counter all of the criticisms that Davidson's theory faced.

The views of Dray and Rorty in relation to this problem have a certain similarity, since neither of them wishes to take up a metaphysical stance and would rather look from the outside in at the problem; Dray prefers the methodological stance, and Rorty prefers the pragmatic stance. For Rorty, all that matters is the usefulness of a vocabulary; he would agree with Dray that we should not start reducing vocabularies to one another since all vocabularies have their uses when describing or explaining different events. However, unlike Dray, Rorty can avoid the problem of methodological unity since he concedes that there is no principled reason why we could not reduce the vocabularies of the mental to the physical, but simply argues that there is no point in doing so because none of the vocabularies relate to the 'world

as it is' and so none of them have priority over any others. As both vocabularies have their individual uses, then, there seems little point in reducing them to each other as this would be a very difficult task that would succeed only in eliminating a useful descriptive tool. Therefore, Rorty can hold off the critics that Dray could not, by appealing to the uselessness of this kind of reductionism, rather than trying to claim that reductionism is misconceived or impossible in principle.

Rorty's view is also similar in some ways to the views of Foder, Baker and Dretske, as they all aimed to undermine the idea that the vocabulary of physics is the only vocabulary that is relevant to causation. Rorty has already solved the problems with Dretske's theory when he argued against the critics of Davidson's Anomalous Monism, as both Davidson and Dretske encounter the same problem of how the mental can have relevance in physical events given the problem of explanatory exclusion. Yet, Rorty also manages to avoid falling into the same trap as Fodor or Baker, that of mistaking explanation for causation, since he denies that there are any laws or descriptions that truly capture reality: there is no 'way the world is'. He denies that micro-causation is the root of all causation and that macro-causal explanations are second order, because he does not believe in 'true' causation or that the world has one privileged vocabulary in which it chooses to be described. Thus, there is no fundamental distinction, and thus confusion, between true causation and causal explanation. However he also disagrees with Baker's claim that we should reverse the causal hierarchy and take macro-causal explanations over and above micro-causal explanations; he does not believe that *any* disciplines or vocabularies have priority or are in competition with each other. Such mistaken views, according to Rorty, are based on a belief in a world 'in itself', which allows us to judge the truth of sentences depending on the extent to which they represent this true reality.

Rorty claims that there is no principled reason why the vocabularies of the special sciences could not be reduced to the vocabularies of physics, but he does not believe that this claim entails any reductionism. Once we have realised that there is nothing special about the mental, and that mental states are just physical states, then that is the end of the story. Philosophers may as well insist that we reduce the superficial properties of a chair to their component parts for all the good that reducing mental properties to physical properties would do. The only reason we feel there is a difference, and that reductionism about the mental is somehow important, is that we are inclined to believe in a competition between the mind and the body that needs to be resolved. This, however, is merely the result of a historical confusion, and once we see this, we see that the whole problem is merely a pseudo-problem. This historical deflation of the problem allows Rorty to accommodate Fodor and Baker's wishes to avoid reducing different vocabularies to one another, without denying the reduction is possible.

So, there seem to be a number of good reasons why Rorty's view is a cut above the rest. By taking a step back instead of wading into the fray, Rorty finds a way to give credit to the physical sciences for being useful without losing the notion that mental states have causal relevance. To do this, however, Rorty has attacked our notions on objective knowledge and truth. His objection to philosophy's conception of an objective truth is that this is a transfer of our past dependence on the church to provide privileged access to the truth, to our dependence on the objective world described by science to provide the truth. This shift of a blind faith in God to a blind faith in science leads us to believe that there is an objective world out there that can provide all the answers and that there is one true description that can

capture the entire nature of reality. However, by agreeing with Rorty's assumption that it is only down to historical contingencies that we rely on science as much as we do, then there is no reason to think that there is only one privileged language that the world can be described in, and this means that science is no longer the provider of the ultimate truths.

Rorty does not believe that this is a problem. He thinks that by undermining the scheme-content distinction, Davidson showed how the world 'in itself' does not make our beliefs or knowledge about the world true. Thus once we give up on our language corresponding to or representing reality, there is no point in judging beliefs, theories or languages on the merits of how responsive they are to the causal pressure of the world. Once we can see that there is no distinction between the world and how we conceptualise it, then we can also see that there is no distinction between beliefs that are made true by the world and beliefs that are made true by other beliefs. This means that despite physics being more useful when describing the microphysical nature of the world, there are other vocabularies that are equally useful when describing other aspects of the world. So, we do not need to change our conception of causation, as Fodor or Baker suggested, just realise that this way of talking about causation is useful rather than actually corresponding to or capturing the true nature of causation. Therefore, there is no point in ranking or creating a hierarchy for different vocabularies or viewing the vocabulary of physics as above any other vocabulary.

What Rorty has done is to show how the mind-body problem can be dissolved and then, in case of any reprisals, why it is pointless to debate the problem anyway. By following up his metaphysical argument with this historical deconstruction claim, what Rorty has accomplished is to show that even if philosophers do not agree with his views, there is no

point in proposing counter-claims as the whole mind-body problem is based on a historical mistake. The penalty for following Rorty down this path, however, is abandoning the idea of science as the ultimate provider of truth and knowledge.

Conclusion

In my own mind, I think that this penalty is worth accepting. I feel that his arguments against the metaphysical physicalist belief that the vocabulary of physics is the ultimate privileged language that accurately describes reality are convincing enough to be hugely damning. Rorty's own 'physicalism', which just means acknowledging the capacity of physics for 'complete coverage', and for being exceptionally useful, is a big improvement. If there is no way the world is, then claims to knowledge of things such as 'real' causation can be abandoned. This would leave the mental not quite as redundant as physicalists have previously claimed, for although Rorty is not showing us how to capture the true nature of mental causation, or showing some principled reason why the mental cannot be reduced to the physical, he is undermining any reason to abandon the concept of the mental as relevant in causal explanations or to continually try to reduce the mental to the physical. If we agree with Rorty that the mind-body problem is merely a pseudo-problem, then we should agree that physical descriptions are very useful in our society, but that this vocabulary is not actually reflecting reality and should not be taken as the privileged language of the world. Instead, we should continue to talk about the mental, not in a condescending or metaphorical way, but as something that is incredibly useful in everyday life, just as talking about physics is useful. When we stop trying to discover the ultimate nature of reality, as it is a pointless quest, then we will question what we were actually trying to accomplish, since science no

longer needs to be defended against the church by philosophy: it can look after itself. By eradicating the mental we are not contributing any knowledge to humanity, or doing anything useful. Once we realise this, it seems we can abandon the problem of mental causation without also abandoning the crucial notion that the mind is ever-present in everyday life.

Mind over matter or matter over mind?

A comparison and evaluation of four contrasting approaches to the problem of mental causation

References

Anscombe, G.E.M. (1983) 'The Causation of Action', in C. Ginet & S. Shoemaker (eds.) (1983) *Knowledge and Mind: Philosophical Essays,* Cambridge: Cambridge University Press, pp. 174-190.

Baker, L. (1993) 'Metaphysics and Mental Causation', in J. Heil & A. Mele (eds.) (1993) *Mental Causation,* Oxford University Press, pp. 75-95.

Collingwood, R.G. (1940) *An Essay on Metaphysics,* Oxford: Clarendon Press.

Collingwood, R. G. (1965) *Essays in the Philosophy of History*, University of Texas Press, Austin.

Davidson, D. (1970) **'**Mental Events', in D. Davidson (1980) *Essays on Actions and Events,* Oxford: Clarendon Press, pp. 207-227.

Davidson, D. (1982) ‘Rational Animals’ in E. LePore & B. McLaughlin (eds.) (1985) *Actions and Events,* Oxford: Basil Blackwell, pp. 473-480.

Davidson, D., (1986)"A Coherence Theory of Truth and Knowledge", E. LePore (ed.) (1993) *Truth and Interpretation,* Oxford: Blackwell, pp. 307-319.

Davidson, D (1993) ‘Thinking Causes’, in J. Heil & A. Mele (eds.) (1993) *Mental*

Causation, Oxford University Press, pp. 3-17.

Davidson, D (2001) *Inquiries into Truth and Interpretation,* Clarendon Press: Oxford.

Dennett, D. C., (1991) *Consciousness Explained,* London: Penguin.

Descartes, R. (1641) *Meditations on first philosophy: with selections from the Objections and replies,* trans. M. Moriarty (2008), Oxford: Oxford University Press.

Dewey, J (1910) *How We Think,* London D.C. Heath & Co.

Donagan, A (1987) *Choice: The Essential Element in Human Action,* Routledge & Kegan Paul Ltd.

Dray, W. (1957) *Laws and Explanations in History*, Oxford University Press.

Dray, W (1961) 'The Historical Explanation of Actions Reconsidered', in S. Hook (ed.) (1963) *Philosophy and History: A symposium,* New York University, pp. 105-135.

Dray, W. (1964) *Philosophy of History,* Prentice-Hall Inc.

Dretske, F. (1988) *Explaining behavior: reasons in a world of causes,* Massachusetts Institute of Technology.

Dretske, F. (1993) 'Mental Events as Structuring Causes of Behaviour', in J. Heil & A. Mele (eds.) (1993) *Mental Causation,* Oxford University Press, pp. 121-136.

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

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

Hempel, C. (1942) 'The Function of General Laws in History', in P. Gardiner (ed.) (1959) *Theories of History,* The Free Press, pp. 344-356.

Hempel, C. (1962) ‘Explanation in Science and in History’, in W. H. Dray (ed.) (1966) *Philosophical Analysis and History,* New York: Harper, pp. 95-126.

Hempel, C. (1965) *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science,* The Free Press: New York.

Kant. I. (1933) *Critique of Pure Reason,* trans. N. K. Smith (2007), Basingstoke: Palgrave Macmillan.

Kim, J. (1984a) ‘Concepts of Supervenience’, in J. Kim (1993) *Supervenience and Mind,* Cambridge: Cambridge University Press, pp. 53-78.

Kim, J. (1984b) ‘Epiphenomenal and Supervenient Causation.’, In J. Kim (1993) *Supervenience and Mind,* Cambridge: Cambridge University Press, pp. 92-108.

Kim, J. (1989a) ‘Mechanism, Purpose and Explanatory Exclusion’, in J. Kim (1993) *Supervenience and Mind,* Cambridge: Cambridge University Press, pp. 237-264.

Kim, J (1989b) ‘The Myth of Nonreductive Materialism’, in J. Kim (1993) *Supervenience and Mind,* Cambridge: Cambridge University Press, pp. 265-284.

Leibniz, G. (1973) *Philosophical Writings,* trans. G. Parkinson & M. Morris, London: J.M Dent.

McGinn, C. (1989) ‘Can We Solve the Mind-Body Problem?’ in C. McGinn (1991) *The Problem of Consciousness,* Basil Blackwell Ltd.

McLaughlin, B. P. (1993) ‘On Davidson’s Response to the Charge of Epiphenomenalism’, in J. Heil & A. Mele (eds.) (1993) *Mental Causation,* Oxford University Press, pp. 27-40.

Pears, D. (1971) *Wittgenstein,* London: Fontana.

Place, U.T. (1956) ‘Is Consciousness a Brain Process?’, *British Journal of Psychology,* 47: 44-50.

Plato. (1961) *The Collected Dialogues of Plato,* eds. E. Hamilton & H. Cairns, Princeton, NJ: Princeton University Press.

Putnam, H. (1967). ‘Psychological Predicates’, in W.H. Capitan and D.D. Merrill (eds.), *Art, Mind, and Religion*. Pittsburgh: University of Pittsburgh Press, pp. 37-48.

Putnam, H. (1978). *Meaning and the Moral Sciences,* London: Routledge and Kegan Paul.

Quine, W.V.O. (1960) *Word and Object,* The Massachusetts Institute of Technology.

Quine, W.V.O. (1981) *Theories and Things,* Cambridge, Mass: Belknap.

Rorty, R. (1965) ‘Mind-Body Identity, Privacy, and Categories’, *Review of Metaphysics*, 19:24-54; reprinted in Borst (ed.) (1970), pp. 187-213.

Rorty, R. (1970) ‘Incorrigibility as the Mark of the Mental’, *Journal of Philosophy* 67: 399-429.

Rorty, R. (1979) *Philosophy and the Mirror of Nature*, Princeton, NJ: Princeton University Press.

Rorty, R. (1982) *Consequences of Pragmatism,* Minneapolis: University of Minnesota Press.

Rorty, R., (1991) *Objectivity, Relativism and Truth: Philosophical Papers, vol. 1,* Cambridge: Cambridge University Press.

Rorty, R. (1996) ‘Davidson’s Mental-Physical Distinction’, in The Philosophy of Donald Davidson: The Library of Living Philosophers, (1999) vol XXVII, pp. 575-594.

Rorty, R. (1998) *Truth and Progress: Philosophical Papers, vol. 3,* Cambridge: Cambridge University Press.

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

Smart, J.J.C. (1959) ‘Sensations and Brain Processes’, *Philosophical Review*, 68: 141-156.

Smart, J.J.C. (2004) ‘Consciousness and Awareness’, *Journal of Consciousness Studies*, 11: 41-50.

Sorell, T. (1990) ‘The World from Its Own Point of View’, in A. Malachowski (ed.) (1990) *Reading Rorty: critical responses to Philosophy and the Mirror of Nature (and beyond)*, Oxford: Basil Blackwell, pp. 11-57.

Spurrett, D & Papineau, D. (1999) ‘A Note on the Completeness of Physics’, *Analysis* 59. no.1, pp. 25-29.

Thomasson, A. L. (2007) *Ordinary Objects*, Oxford University Press.

Wittgenstein, L. (1953) *Philosophical Investigations*, trans. by G.E.M. Anscombe, Basil Blackwell: Oxford.