IS THERE FREE WILL?

Quantum Indeterminism and the Reality of Choice

Stephen Priest

Stephen Priest is Senior Research Fellow in Philosophy at the University of Oxford.

Stephen Priest
Faculty of Philosophy
University of Oxford
Woodstock Road
OXFORD OX
ENGLAND UK
01865 426625
stephenpriest99@hotmail.com

Also by Stephen Priest:

The British Empiricists
Theories of the Mind
Merleau-Ponty
The Subject in Question

Hegel's Critique of Kant (ed.)
Jean-Paul Sartre: Basic Writings (ed.)
A Dictionary of Philosophy (ed. with Antony Flew)

Preface

Despite a large literature, there is no plausible theory of free will consistent with the findings of science. There are three requirements on any plausible theory of free will: There has to be free will. The theory has to explain what free will is, and why it exists. The theory has to be logically consistent with the best science we have. If successful, a further feature will accrue: We will be able to do better science using the theory than previously. In particular, social science will be science. In order to solve the problem of free will and determinism we have to pay far more attention to physics than is usual in philosophy or theology. Quantum physics presupposes indeterminism and, I shall argue in this book, free will.

Stephen Priest

Oxford
Michaelmas 2021

Acknowledgements

I am grateful to the following for conversations I have had over the years: Sir A. J. Ayer, Renford Bambrough, Hazel Barnes, Graham Bird, Daniel V. Came, Edward Craig, Antony Flew, Lazar Fotev, Benedikt Göcke, P. M. S. Hacker, R.M. Hare, Michael Inwood, Sir Anthony Kenny, Fergus Kerr, Saul Kripke, John Leslie, David Lewis, Grahame Lock, Bryan Magee, A. W. Moore, Thomas Nagel, Alexander Norman, G. E. L. Owen, Robert Nozick, Terence Penelhum, Sir Karl Popper, Howard Robinson, Roger Scruton, Shogo Shimizu, Kerry Smallman, Richard Swinburne, W. H. Walsh, Bernard Williams, and Timothy Williamson. Some of these people disagree with me profoundly. The world is much poorer for the fact that some of them are dead.

Daniel Came organised a one day conference on some of the ideas in this book in the Faculty of Philosophy of the University of Oxford. I thank him, Benedikt Göcke, Michael Inwood and Alexander Norman for reading papers about my work that day, and Stephen Mulhall for his comments. Michio Andho organised a one day conference on the manuscript at Keio University, Japan. I thank Kenji Murota and Michio Andoh for their papers and the audience for their comments. I thank Masataka Murama of Hokkaido University and Hitoshi Nagai of Nihon University, Tokyo, and Grahame Lock and the University of Nijmegen for organising seminars on my work. I thank Lazar Fotev and the Faculty of Philosophy at Skopje University, Republic of Macedonia, for their hospitality and for useful discussions. I thank Klaus Müller and Benedikt Göcke for a one day meeting on the ideas in this book at the Wilhelms Universität, Münster, Rheinland-Westfalen.

I am grateful to the members of the Faculty of Philosophy of the University of Oxford, the Master of Blackfriars Hall, Oxford, and the Fellows and Dominican brethren, and the Principal and Fellows of Wolfson College, Oxford, for environments conducive to thought.

Contents

I A Proof of Free Will 7

II Quantum Indeterminism 25

III The Neuroscience of Free Will 114-122

IV Free Will and Empiricism

V Kant's Concept of Freedom 122-185

VI Existentialist Freedom

VII Moral Cosmology

Appendix I: Freedom and Prediction

Appendix II: Aristotle and the Sea Battle: The Problem of Future Contingents 185

Notes

Bibliography

I

A Proof of Free Will

What is the problem of free will and determinism? If you do something freely, *you could have not done it*. You could have refrained, and done something else, or perhaps nothing. The doctrine that there is free will is libertarianism. However, if determinism is true, the event called 'your action' is the result of causes sufficient for its occurrence. It could not fail to occur, if its causes occur, so *you could not have not done it*, and so did not do it freely. The problem is, given that libertarianism and determinism are incompatible, which, if either, is correct?

Are you both free and determined? Libertarianism and determinism are mutually inconsistent because their conjunction entails a contradiction: 'It is true and it is false that you could refrain from acting'. It follows that at least one of libertarianism and determinism is false. They are mutually exclusive: If we have free will, determinism is false. If determinism is true, we have no free will. Yet, *prima facie*, there no empirical difference between a world in which one theory is true and the other false.[1]

Does the problem matter? Newtonian and Einsteinian science is deterministic: Events are necessitated by initial conditions and laws which facilitate their reliable prediction. Ethics and law are libertarian: I can only be rightly blamed or praised for what I do if I am responsible for what I do. If I am responsible for what I do, then I could refrain from what I do. (Or, in cases of negligence, I could have done what I failed to do.) But this possibility of refraining is what freedom consists in.

The problem of free will and determinism matters because, given their incompatibility, if there is free will, Newtonian and Einsteinian science is false but, if there is determinism, there are no sound ethical or legal judgements.[2]

---

[1] '[…] where a Determinist would reason "I shall take my usual glass of brandy tonight, so there is no use resolving not to take it", the Libertarian's reasoning would be "I mean to leave off that brandy, but it will be just as easy to leave it off tomorrow as today; I will therefore have one more glass, and leave it off tomorrow.' Henry Sidgwick *Methods of Ethics* (London) p. 67

[2] These presuppositions are not incontestable. Elizabeth Anscombe used to tell a joke, roughly, like this: Judge: 'Is there anything you wish to say in your defence?' Defendant: 'My hand reaching out for the goods

How are free will and determinism thinkable? Free-will and determinism presuppose a shared modal notion: ‘could be otherwise’. Libertarianism entails actions *could be otherwise*, and so occur contingently. Determinism entails that putative actions *could not be otherwise*, and so occur necessarily. Libertarians and determinists rely on a positive or negative version of one and the same concept of possibility, expressed by ‘can’ or ‘could’.

What is this notion of possibility? Suppose there is only what is the case: Suppose, no matter how thoroughly we search the universe, or universes, we only ever find objects that exist, events going on. The libertarian says about some, or all, of the universe: 'This not only is, but, *might not have been*' or, 'It could have been otherwise'. The determinist says: 'This not only is but *could not have not been*', or, 'It could not have been otherwise'. Libertarianism entails contingency. Determinism entails necessity.

The difficulty faced by both free will and determinism is answering: *What information does either doctrine add to existence?* Given that there is only existence, only what there is, what extra information has the libertarian provided by saying it could be otherwise? What extra information has the determinist provided by claiming it could not be otherwise? Both doctrines seem equally obscure at root. The notions are interderivable: 'possibly *p* means 'not necessarily not *p*', and 'necessarily *p*' means 'not possibly not *p*', but each resists analysis in more primitive terms. Each denotes a positive or a negative version of a property that existence seems to have no room for. Advocates of free will or determinism typically do not notice that that their philosophies entail the application of this root notion common to ‘□’ and ‘◊’.

If we find it difficult to analyse a concept (or understand what something is) we should not conclude it denotes nothing or denotes nothing real. There is no valid inference from human failures of comprehension to non-existence. We do know what ‘could be otherwise’ and ‘could not be otherwise’ mean because we know how to use them.

Necessity and contingency are not natural or empirical properties. From the fact that some object, or event exists, it does not follow that it exists contingently or necessarily and necessity and contingency cannot be observed. Every property is either metaphysical or empirical so they are metaphysical properties of the world.

---

was an inevitable result of my poor upbringing, and the neurology of my brain immediately prior to the incident. I could not have refrained from acting. That shews I am not responsible, so should not be sentenced'. Judge: 'Six months. This judgement is the inevitable result of my privileged upbringing, my legal training, and the state of my brain immediately prior to this judgement.'

Are free will and determinism concrete or abstract? Freedom is concrete, real, and directly experienced by you in everyday life. Determinism is an abstract theory we are taught to believe.

Argument from action:

(1) The reality of free will is a necessary condition for action.
(2) There is action.
(3) Therefore, there is free will.

Premise (2) is true because there is a distinction between what you *do* and what *merely happens* to you. Premise (2) is true because what you do depends upon the possibility of *not* doing it, that is, upon your having free will. If your actions were automatic, like your heartbeat or the operations of your liver, we would not say that you *do* them (except in the minimal sense that they belong to you rather than anyone else). Of course, many thoughts and actions do seem automatic. If you do not choose to think this rather than that, thinking typically goes on anyway. Many actions are performed unconsciously. It is possible to wonder 'Did I just do that?'. But I is a necessary condition for any of these being your actions that they are not all like that. Unless some things are done deliberately, consciously, freely, then nothing is done.

Argument from belief:

(1) The reality of free will is a necessary condition for belief.
(2) There is belief.
(3) Therefore, there is free will.

Admittedly, we are caused to acquire most of our beliefs in an involuntary way. For example, your belief in the existence of the objects around you now is caused by your conscious or unconscious perception of them. Perceptions are belief causers. However, a necessary condition for these being *beliefs* is that they are not all acquired involuntarily. Premise (1) is true because there could not be any *beliefs* unless some beliefs are acquired by weighing of evidence, choosing to accept or reject, allocating a probability of being true or false. A mechanical disposition to allocate the truth value 'true' to a proposition is not sufficient for belief, even though believing entails allocating the truth value 'true' to a proposition.

Argument from desire:

(1) The reality of free will is a necessary condition for desire.
(2) There is desire.
(3) Therefore, there is free will.

Admittedly, we are caused to acquire most of our desires in an involuntary way. For example, the desire for a doughnut might be cause by the sight or the smell of a doughnut or watching a commercial for doughnuts. However, a necessary condition for these being *desires* is that they are not all acquired involuntarily. Premise (1) is true because there could not be any *desires* unless some desires are acquired by weighing the consequences of action, choosing to accept or reject, exerting will, or resisting temptation. A mechanical disposition to think, feel, or move is not sufficient for a desire. Belief and desire presuppose free will.

An empirical argument for the reality of free will:

(1) There is experience of the moment of choice.
(2) If there is experience of the moment of choice then there is the moment of choice.
(3) If there is the moment of choice then there is free will.
(3) Therefore there is free will.

Premise (2) is true because the experience of the moment of choice is only possible if it is veridical. The apparent opening of the possibility of doing something or refraining from doing it entails that possibility. For example, you are riding a bicycle carrying a letter which you intend to post. If you do not choose to peddle, peddling just carries on automatically. Similarly, if you do not choose to think about anything in particular, thinking just carries on automatically, switching of its own accord between the traffic, the letter, the location of the mailbox, etc. However, if you become *conscious of* peddling, at that moment, the possibility opens up of peddling faster or slower, or even stopping. If you become conscious of thinking you can decide to think about something else. We seem to have the power to *intervene* in the course of our own actions. There is an intimate relationship between self-consciousness and freedom. Only a self-conscious being can be free. Only a free being can be self-conscious.

Argument from moral dilemmas:

(1) There are moral dilemmas.
(2) Moral dilemmas presuppose the reality of freedom.
(3) Therefore freedom is real.

In the twentieth century, when philosophy was still in the grip of out of date science, the view used to be put forward that free will is an illusion. The view is defensible in a philosophy seminar but anyone who has ever faced an agonising moral dilemma knows that it is false. Suppose my mother is in a coma in hospital, with no hope of recovery, and a doctor asks me, as next of kin, to decide whether her life support should be cut off. In the most drastic sense, *the decision rests with me*. I am deliberating. If I choose one way, my mother's life will end. If I choose another way, her life will continue. The switch is within me as much as in the hospital machinery. If this choice is not real then what is real? Suddenly freedom of choice looks like reality and determinism just looks like a theory, a theory held by philosophers who think it scientific. Even if determinism is true it is of no use to me in my dilemma, which is as good as saying determinism is false.

Is free will an illusion?

Argument from illusion to the reality of free will:

(1) Only a being with free will can be subject to illusion.
(2) There is illusion.
(3) Therefore there is free will.

Argument against illusion to the reality of free will:

(1) Only a being with free will can be subject to illusion.
(2) Free will appears to be an illusion.
(3) If free will appears to be an illusion free will is not an illusion.
(4) Therefore, free will is not an illusion.

In these two arguments, premise (1) is true because being subject to illusion entails acquiring a false belief and only a being with free will can hold beliefs.

Are Free Will and Determinism real or only theoretical? Freedom is concrete, real, and directly experienced by you in your everyday life. Determinism is abstract; a pseudo-scientific theory we are taught to believe.

Human beings are deceived if they think themselves causally determined. The illusion of determinism is caused by believing theory to the point of being blind to the reality of experienced choice. Freedom is real.

Could you do otherwise?

(1) There are moral dilemmas.
(2) Moral dilemmas presuppose the reality of freedom.
(3) Therefore, freedom is real.

Is freedom power? Although I always have some choice or other, much is beyond my power. The scope of freedom and the scope of power coincide only in consciousness, not in human beings. Freedom is not power. (Sartre) Freedom is being able to do as one chooses.

Does predictability entail determinism? From the fact that you can predict my behaviour it does not follow that I do not behave freely. Perhaps I exercise my freedom in regular patterns with which you have become aquainted and you inductively infer my next action. It does not follow that I could not have not done it. From the fact that my actions may be predicted it does not follow that I am not free. You might be able to predict my actions because you have become acquainted with the ways in which I exercise my freedom. In that case, my actions may be free yet predictable. Knowing the causes of events is not the only way of predicting events.

Does unpredictablity entail freedom? From the fact that events are unpredictable it does not follow that they are freely brought about. Events might be unpredictable if they are random or chaotic, and if they are random or chaotic it does not follow that they are freely brought about. Events may be unpredictable but determined if there is insufficient information about them to predict them. That an event is brought about freely is neither necessary nor sufficient for that event's unpredictability.

Can the future be predicted? The future often is predicted, and accurately, so the future can be predicted. If the future was unknown to us we could not funtion from moment to moment. We do function from moment to moment, so the future is known to us.

Are You free? Your are a cause of your own actions. You cause yourself to do what you do.

Does freedom admit of degrees? My freedom admits of degrees because I am amongst the causes of my own actions.

Am I determined? I am determined because I am not the only cause of my own actions.

How can I be both free and determined? I can be both free and determined because I am both a cause and an effect. I am the cause of what I do and the effect of what is done to me. I am a link in a causal chain. Every link in a causal chain, except the first link and the last link, is both a cause and an effect: the cause of subsequent links and the effect of previous links. I am the effect of previous links and the cause of subsequent links.

Sometimes I am free, sometimes I am determined. I am free when I am a cause, determined when an effect.

What is modality? It seems impossible to analyze modal concepts in non modal terms because from the fact that something never happens it does not follow that it is impossible, and from the fact that something always happens it does not follow that it is necessary. The concepts of unrealized or perpetually realized contingencies are not contradictory. From the fact that something only happens sometimes it does not follow that it does not necessarily happen when it does happen. The concept of an intermittently realized necessity is not incoherent.

Where do modal properties come from? We know what 'could be otherwise' and 'could not be otherwise' mean in the minimal sense that we know how to use them. The experience of being a cause and being an effect allows us to understand modality. Hume right that there is no perception of the causal relation, only of constantly conjoined events. However, Hume misses the point that there is experience of causation, experience of being a cause and experience of being an effect. The concept of necessity is derived from the experience of being under pressure. The concept of contingency is derived from the experience of the release of such pressure. These empirical facts go some way towards explaining where the concepts of necessity and contingency come from but do not explain where necessity and contingency come from. There is the experience of being caused, the feeling of being subject to pressure or constraint against your will. You are aware of being blown by a strong wind. You are physically pushed by an aggressive human being. You are confined by the pressure of a crowd. There is also the experience of being thus affected and the noting of it inwardly. Therefore determinism is experienced. The experience of determinism allows us to

understand modality. In the experiences of being a cause and being caused there is direct experience of causation. Here, the first person singular case is the model for the third person case rather than vice versa. In one's own case, it is obvious that the idea of causation is not just derived from experience of constantly conjoined events. One feels the pressure. One feels the exhilaration: the pressure of being an effect, the exhilaration of being a cause.

Necessity and contingency are not products of language. How could language have the power to render something *not possibly not* the case or allow something to *possibly not* be the case? Language has semantic, syntactic, phenomenological, physical and psychological properties.

The necessity of some fact or thesis is reluctance on our part to 'give up' that fact or thesis. (Alvin Plantinga said years ago that one might be unwilling to give up the view that one is a fine fellow but that does not make it a necessary truth.) Nor is the contingency of some fact or thesis our willingness to give it up. Necessity is not 'ungiveupability'. Contingency is not 'giveupability'. These are epistemic attitudes that we ought to adopt to necessities and contingencies, respectively. If we do adopt them, that is posterior to the modal facts. We can only rightly adopt the attitudes if the facts are already necessary or contingent. Nor is it impossible for us to adopt the wrong attitudes. One might refuse to give up a contingent truth or willingly give up a necessary truth. Aristotle was wrong in Metaphysics Book Gamma to say that someone cannot believe a contradiction. He is even wrong if he means no-one can believe what they know to be a contradiction.

Contrary to a widespread misconception, necessity is not entailed by being analytic. Suppose:

p is analytic if and only if p is true by definition

'true by definition' means 'made true by the way human beings define words'. If anything is a paradigm of contingency, it is the ways human beings define words. For any word, that word could have been defined differently. For any true analytic sentence, that sentence could have been false. "But surely, analytic sentences express identities about meanings and these are necessary truths." It is right that any meaning, like anything else, is self-identical and propositions expressing these necessary facts are necessary truths. The point of an analytic proposition, however, is two say that two distinct expressions share one meaning. That is a contingent fact because they might not have done.

Argument that necessity and contingency are bestowed by consciousness:

(1) Only consciousness has the power to make what exists *possibly otherwise* (in the sense mentioned above).

(2) Only consciousness has the power to make what exists *necessarily so*.
(3) There is contingency.
(4) There is necessity.
(5) Therefore, contingency and necessity are bestowed on what exists by consciousness.

The conclusion, (5), implies that if an event obtains, or does not obtain, consciousness *adds ontologically* that it could not-happen, or could not not-happen. Twentieth century philosophy bizarrely ascribed some of the powers of consciousness to language. It is therefore no exaggeration to say that language was your own consciousness of twentieth century phiosophy.

The reason why it is difficult to give a non-modal analysis of modal concepts is that the modal is irreducible to the non-modal. The modal *qua* modal is primitive. The logically primitive is the sign of the quantum facilitating. Necessity and contingency come from *out of this world*. When we are dealing with necessity and contingency, knowingly or not, we are doing consciousness .

Does the past affect the present? It is a false prejudice that the present is determined by the past.

Argument that the past cannot determine the present:

(1) Necessarily, the past is over.
(2) The past is over. (from (1))
(3) Whatever affects the present is present.
(4) 'The past is present' is contradictory.
(5) The past is not present. (from (4))
(6) Therefore, the past does not affect the present.

Premises (4) and (5) might be denied by saying that, say, a Roman villa persists from the past into the present so, in a way, what was past is present. Something that did exist does exist. Or, some process, such as the rise and fall of civilisations might be held to be continuing from the past into the present. What was going on is still going on.

These remarks are plausible but the premises mean something different. 'The past is not present' means that the stretch of time before now is not now. Existing in the stretch of time before now is not existing now, not being simultaneous with us. To suppose otherwise is as incoherent as claiming 'Wednesday of last week is the same day as today' or 'Caesar's crossing of the Rubicon, although over, is going on this morning'.

That the past is present is incoherent in this sense is fully consistent with the past being present in other senses. We might maintain that the past has some ontological status, for example, that it exists as fully as the present. It is just that we are not simultaneous with it.

Is the present determined by the past?

Argument that the present is not determined by the past:

(1) The past does not exist.
(2) The present does exist.
(3) What does not exist cannot affect what does exist.
(4) The past cannot affect the present.
(5) If anything determines anything it affects it.
(6) Therefore, the present is not determined by the past.

All causes and effects are present causes and effects, so the present causes itself to be what it is. It follows that nothing outside the present causes the present to be what it is.

Argument from spontaneity:

(1) The present is a free spontaneity.
(2) No free spontaneity is causally determined.
(3) The present is not determined. (from (1) and (2))
(4) Therefore, the present is not determined by the past.

The past can narrow down, or determine, which present out of infinitely many possible presents becomes actually present. If the Germans had won World War II then the actual present would have been different. Different propositions would have been true now. Different states of affairs would have obtained, even though it would still have been now.

Is anything contingent? In nature, everything is contingent. There is no contradiction in the supposition that anything empirical is not the case.

Is anything self-determining? The universe makes itself what it is. Each part of the universe makes the universe what it is. You are part of the universe. You make this particular human being what you are. You help make the universe what it is.

Am I the cause of my own actions? If I am the cause of my own actions it does not follow that I am not free. Indeed, this is part of what my freedom consists in.

Can the future be predicted?

Argument that the future can be predicted (I):

(1) If the future could not be predicted, science would be impossible. (A)
(2) There is science. (A)
(3) Science is possible. (2)
(4) Therefore, the future can be predicted. (1), (2), (3) MT

Argument that the future can be predicted (II):

(1) If the future could not be predicted, we could not function from moment to moment. (A)
(2) We do function from moment to moment. (A)
(3) Therefore, the future can be predicted. (1), (2) MT

The future is habitually predicted, and accurately. If you open a door to the next room, you do not expect to step into an alternative reality, fall as the floor opens beneath your feet, or enter a vast empty space. You do not expect people around you to suddenly speak in an ancient language or turn into aliens. By and large, we predict the future so successfully that we manage the transition from present to future seamlessly. To see the plausibility of this, imagine that the future is a complete blank, that we can know *nothing whatever* of the future, not even the next microsecond. Unpredictability depends on extremely widespread predictability.

Is freedom experienced? There *experience of being a cause*, the feeling of agency, the observation that one's actions have effects. There is the experience of choice, or control, over which option to select.

Is determinism experienced? There is *experience of being caused*, the feeling of being subject to pressure or constraint, the observation that one is affected. The experience of not being able to do otherwise.[3]

What is the relationship between political freedom and freedom of the will? Political freedom is the absence of constraints by the state on which actions I am the cause of. Free will is a necessary condition for political freedom because only a being capable of exercising free choice may have that choice politically legitimated or constrained.

What is the relationship between political repression and causal determinism? Political repression is being caused or constrained , against one's desires, by the power of the state. Causal determinism is a necessary condition for political repression because only a being capable of being affected may be constrained or prohibited from acting.

'political freedom' is the absence of constraints by the state on choice.

Argument that free will is a necessary condition for political freedom:

(1) Only a being capable of political freedom is capable of having their choice politically legitimized or constrained.
(2) Only a being capable of exercising free choice may have that choice politically legitimated or constrained.
(3) Only a being capable of exercising free choice is capable of political freedom.
(4) Therefore, political freedom presupposes freedom of the will.

Argument that free will is a necessary condition for political freedom:

(1) Only a being capable of exercising free choice may have that choice politically legitimated or constrained.
(2) Only a being capable of political freedom is capable of having their choice politically legitimized or constrained.
(3) Only a being capable of exercising free choice is capable of political freedom.
(4) Therefore, political freedom presupposes freedom of the will.

---

[3] In the experience of being a cause, and the experience of being caused, we find what Hume was looking for when he failed to detect any direct experience of the causal relation. There is both the experience of compelling and the experience of being compelled: a natural necessity which is empirical not 'logical'.

What is choice? Choice is the possibility of choosing or not choosing, the possibility of choosing between alternatives. A choice is either a possible course of action, or the choosing of a possible course of action. 'Choice' is therefore ambiguous between 'what is chosen' and 'the choosing of what is chosen'. There may, therefore, be several choices in the first sense but only one in the second sense when an action is chosen.

Is the present determined by the past? The past does not exist, but the present does exist. What does not exist cannot affect what does exist, so the past cannot affect the present. All causes and effects are present causes and effects so the present causes itself to be what it is. It follows that nothing outside the present causes the present to be what it is. The present is a free spontaneity and no free spontaneity is causally determined. It follows that the present is not determined, so the present is not determined by the past. It is a false prejudice that the present is determined by the past. Your actions are a pure spontaneity issuing from the now.

How is weakness of will possible? Weakness of will is my not performing an action that I know I can, and wish to, perform.

How can weakness of will be overcome? Weakness of will is overcome by performing an action without will-power. Weakness of will can rarely be overcome by an act of will. Will is part of the problem. A stance towards the action is required that does not require will. For example, the action can be judged to have extra benefits, or may be seen under new aspects, for example interesting rather than boring aspects. Then the action may be performed without will-power, that is, performed without reluctance to perform it.

Does predictability entail determinism?

Argument that from the fact that you can predict my behaviour it does not follow that I do not behave freely:

(1) It is known at an earlier time that I will act at a later time.
(2) If it is known at an earlier time that I will act at a later time I will act at the later time.
(3) I will act at the later time (from (1) and (2).
(4) Had I not acted at the later time it would not have been known at the earlier time that I would have acted at the later time.

(5) I could have not acted at the later time.
(6) Therefore, I act freely at the later time.

From the fact that you can predict my behaviour it does not follow that I do not behave freely. Suppose, of my own free will, I go to The Royal Oak at 5.00 pm every Friday afternoon during Full Term. Suppose you witness these visits on successive Fridays and, on that basis, predict that I will go to The Royal Oak next Friday. If I go, you have predicted successfully. It does not follow that I did not go of my own free will. Perhaps I exercise my freedom in regular patterns with which you have become acquainted, and you inductively infer my next action. In that case, my actions may be free yet predictable. Knowing the determining causes of events is not the only way of predicting them.

It is often assumed that:

If an event is predictable it is causally determined.

and

If an event is causally determined it is predictable.

Indeed, some philosophers and scientists use ‘predictable’ and ‘determined’ synonymously. This is not right because predictability is this:

‘predictability’ (def.) ‘It is knowable at an earlier time what will happen at a later time.’

which is clearly not synonymous with the determinist thesis. Determinism is not predictability. Nor is it right that predictability and determinism are mutually dependent. It is not follow from an event’s being predictable that it is causally determined. If we see someone do something every Friday over a period of weeks, we might accurately predict their behaviour on the next Friday but it does not follow that their action was determined. Perhaps they acted of their own free will and could have not done what they did. If someone acts freely but in regular patterns we can be acquainted with those patterns and know now what they will do later.

It does not follow from an event’s being determined that it is predictable. An event might, in fact, be determined but we might not be in a position to know all its causes. Then we would not be in a position to know

whether the event will happen or not, even though those causes make the event inevitable. One might object that, in this situation, the event is predictable *in principle*. *Were* we to know all the causes, we could predict the effect. This is plausible so long as we already know what *sorts* of causes produce what *sorts* of effects. (The first time we see a new kind of event we might have very little idea of its effects.) Knowledge of natural laws is essentially knowledge of which kinds of cause have which kinds of effect. Knowledge of the laws facilitates prediction. Prediction does not give us logical certainty because it is a contingent fact that any given law holds. It does however give us empirical certainty.

Does unpredictablity entail freedom? Actions might be unpredictable if they are random or chaotic, and if they are random or chaotic they are not free. Actions might be unpredictable but determined if there is insufficient information to predict them. That an action is free is neither necessary nor sufficient for that event's unpredictability. From the fact that actions are unpredictable it does not follow that they are free.

If consciousness knows what you will do, are you free? The problem is: If consciousness knows that I shall act then I cannot not act. I can only act freely if I am able to not act. consciousness knows everything, so I have no free will. But consciousness gave me free will so I have free will. So, paradoxically, it is true and false that I have free will.

The problem does not have as much to do with consciousness or time or modality as appears. It arises because knowledge is truth entailing. If anyone knows anyone is going to do anything then it is true that they are going to do it.If anyone knows at any time, or timelessly, that I act there is a *prima facie* problem about my not acting because knowledge is truth entailing. Ironically, because consciousness is timeless, the case of consciousness might be the only one in which the problem of foreknowledge does *not* arise. Nevertheless I shall take the case of consciousness.

consciousness's knowledge is sufficient for my action but my action is necessary for consciousness's knowledge. If consciousness allows me free will, the way I exercise it is necessary for consciousness's knowledge of my acts. This is what it consists in for consciousness to give me free will. The exercise of free will consists in the truth of certain counterfactuals.

Suppose consciousness has 'foreknowledge' of all my actions. It does not follow that they are not freely performed by me. consciousness's knowledge that I will act is sufficient for my acting, because knowledge is truth entailing. However, my acting in the way that I do is necessary for consciousness's knowledge of my action. Had I not chosen to act as I did,

consciousness would not have known that I would perform that action, not because consciousness cannot know some things, but because there is no such thing as false knowledge.[4]

Where do modal properties come from? It is extraordinarily difficult to analyse modal concepts in non modal terms. Wherever analysis comes to an end, this is a sign of the presence of consciousness.

Argument that necessity and contingency are bestowed by consciousness:

(1) Only consciousness has the power to make what exists *possibly otherwise* (in the sense mentioned above).
(2) Only consciousness has the power to make what exists *necessarily so*.
(3) There is contingency.
(4) There is necessity.
(5) Therefore, contingency and necessity are bestowed on what exists by consciousness.

The conclusion, (5), implies that if an event obtains, or does not obtain, consciousness *adds ontologically* that it could not-happen, or could not not-happen. Twentieth century philosophy bizarrely ascribed some of the powers of consciousness to language. It is therefore no exaggeration to say that language was your own consciousness of twentieth century phiosophy.

The reason why it is difficult to give a non-modal analysis of modal concepts is that the modal is irreducible to the non-modal. The modal *qua* modal is primitive. The logically primitive is the sign of the quantum facilitating. Necessity and contingency come from *out of this world*. When we are dealing with necessity and contingency, knowingly or not, we are doing consciousness .

---

[4] (1) $(\forall x)\,(\forall p)\,\mathrm{K}xp \rightarrow p$
(2) $(\forall x)\,(\forall p)\,\mathrm{K}xp(\mathrm{t1}) \rightarrow p(\mathrm{t2})$
(3) $(\forall x)\,(\forall p)\,\mathrm{K}xp \rightarrow \in p$
(4) $\mathrm{KG}p \rightarrow p$
(5) $-p \rightarrow \text{-}\mathrm{KG}p$
(5*) $-p \in\!\!\rightarrow \text{-}\mathrm{KG}p$ [1]
(6) $\text{-}p(\mathrm{t2}) \rightarrow \text{-}\mathrm{K}xp(\mathrm{t1})$
(7) $(\forall p)\,\mathrm{KG}p \rightarrow \in p$
(8) $\text{-}\in p \rightarrow \text{-}\mathrm{K}xp$
(9) $\Diamond\text{-}p$
(10) $\text{-}\in p$
I use ' $\in\!\!\rightarrow$ ' in the sense introduced by David Lewis in *Counterfactuals* (Blackwell, Oxford, 1973).

Does the past affect the present? It is a false prejudice that the present is determined by the past.

They are bestowed by consciousness.

What is free? Is there absurdity in the conjecture that quantum events are themselves free? The subject of freedom is consciousness. Choice is the efficient cause, that is, the sufficient condition, of action. So, if choices are mental and actions are physical, there is mental causation. Nevertheless, Aristotle provides no account of how mental causation is possible.

The Libertarian Argument for the existence of consciousness:

(1) You have free will.
(2) There is no empirical explanation of free will.
(3) Any explanation is empirical, logical or metaphysical.
(4) There is an explanation of free will.
(4) There is a metaphysical explanation of free will.
(6) If there is a metaphysical explanation of free will then consciousness exists.
(7) Therefore, consciousness exists.

In actualising the universe, consciousness actualises a free, or self-determining, system. All events are in the Now. Your actions are a pure spontaneity issuing from No-thing-ness.

What is an explanation? Empirical explanation entails *subsuming the unfamiliar under the familiar*. Metaphysical explanation entails *subsuming the familiar under the unfamiliar*. Empirical explanation facilitates predictions. Metaphysical explanation discloses presuppositions. Empirical explanation requires evidence. Metaphysical explanation requires proof.

Are there limits to explanation? Wherever analysis comes to an end, this is a sign of the presence of consciousness. Explanations end in consciousness. However, the gaps in explanation can ultimately only be filled by metaphysical explanations. In particular, if we find it difficult to analyse a concept (or understand what something is) we should not conclude it denotes nothing or denotes nothing real. There is no valid inference from human failures of comprehension to non-existence. Explaining is not explaining away.

How do we know that an explanation has come to and end? We know that an explanation has come to an end when it terminates in something that is its own explanation, that is, not in a necessary being but in necessary Being: in consciousness.Yes: explanations have to come to an end somewhere. They come to an end in ultimate reality. They come to an end in consciousness and the soul. Explanations terminate in *the kind of reality capable of terminating explanations*.

Ultimate explanation is explanation that ends in something that is its own explanation.

Explanations are truly explanatory when they end in consciousness. Otherwise, for any chain of scientific explanation, we are still left asking; Why? Without consciousness, any explanation is incomplete.

Argument that there is inference to the ultimate explanation:

(1) Explanations have to come to an end somewhere.
(2) Anything that stands in need of explanation cannot end an explanation.
(3) Anything can be explained.
(4) Something is its own explanation.
(5) Explanation ends in something that is its own explanation.
(6) Therefore there is inference to the ultimate explanation.

For example, any empirical explanation depends on the existence of the empirical world. The empirical world is not its own explanation so it is ultimately made possible by something that is its own explanation. Any mathematical proof presupposes the existence of numbers. Numbers are not their own explanation, so numbers are ultimately made possible by something that is its own explanation.

Scientific explanation is grounded in consciousness. consciousness is necessary for science. Therefore, the truth of science is, knowingly or not, sufficient for consciousness.

How could you do otherwise? You could do otherwise because consciousness entails free will. Cosmic questions: meta-physical answers.

How is free-will possible? In nature everything is contingent. When consciousness created the universe he created a free of self-determining system. The universe make itself what it is. Each part of the universe makes itself what it is. You are created as part of the universe so you make yourself what you are. Human beings are deceived if they think themselves causally

determined. The illusion of determinism is caused by believing theory to the point of being blind to the reality of experienced choice.

The problem of freedom and determinism in philosophy may be formulated in different ways but one of them is this:

If we are free then what we do is the outcome of our choices and we could refrain from doing what we do and do something else. If we are determined then all our actions are caused and, if caused events are inevitable, then we cannot do other than what we do and our freedom is illusionary. Prima facie, freedom and determinism are mutually inconsistent theories, so at most one of them is true and at least one of them is false. The problem of freedom and determinism is stating correctly whether human actions are free or determined.

Do we have free will or not? The problem of freedom and determinism is whether our actions really result from our choices, or whether our choices are illusions and our actions are inevitable because they have other causes or because what we choose is inevitable. Determinism is the doctrine that every event has a cause, and in its strong version it includes the view that caused events are inevitable. Libertarianism is the doctrine that thought and action are produced by choices which are free. By an action's being 'free' is meant, that if I choose to perform some action and perform it, I could have nonetheless chosen otherwise and not performed that action, and if I choose not to perform some action and do not perform it, then I nonetheless could have chosen otherwise and actually carried out that action. It seems, then, if I am determined I am not free and if I am free I am not determined.

II

Quantum Indeterminism

Free Will

Penrose says

‘It is quite extraordinary that quantum mechanics enables you to test whether something might have happened but didn’t happen. It tests what philosophers call counterfactuals. It is remarkable that quantum mechanics allows real effects to result from counterfactuals!’ (Penrose, 1997: 67)

This is ‘extraordinary’ compared with what we find in classical physics (or, for that matter, chemistry and biology). A *counterfactual* is a state of affairs which might have obtained but does not obtain. The notion has been made reasonably precise by the Princeton philosopher David Lewis. (Lewis) Three aspects of counterfactual states of affairs are relevant here.

Firstly, the existence of counterfactuals entails the reality of *contingency*. A contingent state of affairs is one which, if it exists, might not have existed and, if it does not exist, might have existed. The reality of contingency rules out the truth of determinism because determinism implies that what is the case could not have been otherwise, at least given its causes.

Secondly, quantum mechanics allows us to test for counterfactuals. As Penrose says, this is quite remarkable because, as we normally think of them, counterfactuals do not exist. They are *ways things could have been*, *but are not*. In other words; we can detect things that do not exist but might have existed. Philosophically, the only test for counterfactuals we can offer is logical and *a priori*. We search for contradictions in descriptions of the putative counterfactual. If a contradiction is found, the putative counterfactual is eliminated. If none is found, then the counterfactual is tentatively allowed as genuine. Quantum mechanics enables us to go much further. On the quantum level, at least, we can tell whether the counterfactual is really there.

Thirdly, and perhaps most remarkably, quantum mechanics exhibits counterfactuals as having causal efficacy: *The way things could have been affects the way things are*. Again, we normally think that only something that exists can bring about something that exists. We assume only what exists can make a difference to what exists.

To accommodate these three features of conunterfactuals in the light of quantum physics, we have to make fundamental revisions in what we think.

Determinism, no matter how ‘scientific’ it was previously thought to be, has to go. Determinism is false, at least at the quantum level, and almost certainly at many others.

Possibilities exist: either in the sense David Lewis recognised, as an infinite number of possible worlds, or in the sense increasingly accepted by modern

physics, as an infinite number of parallel universes. Ours is not the only reality. If this seems outrageous, or ontologically extravagant, we may look at it this way: There is a difference between something's being possible, and its not being possible. If something is possible, then there is a sense in which its possibility exists. Quantum physics forces us to take this thought seriously. All the possibilities are really 'out there'. (Popper)

These possibilities make a difference. Universes which admit of different possibilities are different universes. What happens, what causes what in a universe, depends on what is possible or impossible in that universe.

From a logical point of view, there exists an infinite number of possible worlds. In some of these, the laws of nature of the actual universe do not hold. In order to do quantum mechanics we have to assume that such possible universes affect our own:

'in these funny superpositions of space-times, you have to consider at least the possibility that some of these potential universes might have closed time-like lines […]. In these, causality has gone all crazy, the future and the past get mixed up and causal influences go round in loops. Now, although these need only play a role as counterfactuals […] they still have an influence on what actually happens. I would not say that this is a clear argument, but it is at least an indication that there could easily be something of a non-computational nature in the correct theory, if we ever find it.' (Penrose, 1997: 123)

We need to make sense of this. In a superposition of space-time, a particular spatio-temporal position is not realised. In a most literal sense, a particle is 'neither here nor there', and yet, it exists. A possible universe has a 'closed time-like line' if the particular laws governing it do not govern other universes. They form a set. In Penrose's examples, 'the future and the past get mixed up' and 'causal influences go round in loops'. The future and the past get mixed up if, for example; an event was past but *then* becomes future. Here 'then' has broken with the meaning we usually give it. This, of course, looks impossible to us, but would be the case if, for example, time travel occurred and if 'then' consequently acquired a meta-sense. Causal influences go round in loops if, for example; An earlier event causes a later event, and that later event causes that earlier event. We do not usually think in this way, because we assume (by and large) that causes predate their effects, but if a later event causes an earlier event, a cause postdates its effect. The putative effect is 'over' by the time the putative cause happens, so it is hard to see how it could be brought about by it. The putative cause 'has not happened yet' at the time that the putative effect happens, so it is hard to see how it could bring it about. Nevertheless, these relations are not logically or theoretically impossible. There is nothing

incoherent about backwards causation, which is to say, there are possible worlds in which there is backwards causation. By 'computable' Penrose means provable algorithmically, that is, in a series of logical steps which could in principle be carried out by a machine.

At the quantum level, a counterfactual state of affairs can have an effect. Although only a possibility, it makes a difference to actuality. How is this possible? Quantum events which occur (not the ones which are merely possible) behave *just as though they were taking decisions*: just as though they were consciously 'taking account of' the merely possible states of affairs, the counterfactuals. There is an important lesson here for free will.

The way things could have been affects the way things are *at the human level*. Human ideas and beliefs about the possible and the impossible affect human decisions, and so actions.

The quantum world provides us with a model for the human individual, the human group, and the human world. The behaviour of sub-atomic particles and the behaviour of any human individual, or anything made out of human individuals, have common quantum features:

(1) Resistance to strict predictability.
(2) Amenability to probabilistic or statistical predictability.
(3) Indeterminacy caused by observation.
(4) The blurring of the distinction between subject and object.
(5) Behaviour 'in the light of' unactualised possibilities.
(6) Contingency, randomness. Absence of determinism.

Counterfactual situations in other possible universes affect what happens in our universe. Where the existing natural laws break down in quantum mechanics, counterfactual laws governing other universes intervene, take over. At the quantum level, there is *seepage* between possible universes. We are used to the idea of an object, or a person, existing in more than one possible world. Philosophers and logicians have however thought of them as *causally sealed*: one possible world does not *affect* what happens in another. Quantum mechanics forces us to give up this assumption. Quantum mechanics makes us recognise that one universe *spills over* into another.

Crucially, counterfactuals are causally efficacious in mental causation and the exercise of free will. If we wish to understand human action, we have to understand motives, intentions, interests, perceived interests, decision-making, emotion. All these are intelligible only against a background of *possibilities*, perceived or real. It seems obvious to us that we inhabit a world of physical objects. This is not false, but superficial. At a profound level *we inhabit a world*

of possibilities. This is true of us at the level of the individual, the social and the historic. In other words, it is true of you, true of classes, nations and governments and true of human reality as it moves through time. Human action is taken 'in the light of' possibility.

Penrose says 'determinism and computability are different things' (Penrose, 1997: 124) This is right because computability is a property of *theory*. A theory is computable if it is decidable (as to truth or falsity) algorithmically. Determinism, on the other hand, is an alleged property of *events* (or whatever is made of events). If one event determines another then, if the first happens, the second cannot fail to happen. If human action is at least partly caused by counterfactuals, even if essentially mediated by human belief or knowledge, then the prospects for determinism look hopeless. We need a rigorous theory of free will.

What can we learn from Penrose about free will? He says about '[…] the issue of *free will*':

'In philosophical discussions, free will has always been talked about in terms of determinism. In other words "Is our future determined by our past?" and issues of that nature. It seems to me that there are lots of other questions which might be asked. For example, "Is the future determined *computably* by the past" – that is a different question.' (Penrose, 1997: 124)

'Always' is too strong here. Free will is sometimes talked about in contrast with *fatalism* rather than determinism. Determinism is a thesis about causation, fatalism is not. Fatalism is the doctrine that the future is inevitable. One particular future is inevitable. The thought is that the future will only happen in one way. It is just our ignorance of that one future that misleads is into thinking the future is 'open' or could be implemented in different ways. Free will is also sometimes talked about in contrast with *foreknowledge*, including divine foreknowledge. If consciousness, or anyone else, knows you are going to do something, are you free not to do it?

Penrose asks 'Is the future determined *computably* by the past?' If the future is determined by the past then, given the way the past has happened, the future can only happen one particular way. If that outcome is computable then it could, theoretically at least, be calculated logically using a machine. If determinism is true, then there seems little theoretical obstacle to computability. If determinism is false, and human beings freely fashion their futures by implementing a subset of a range of possibilities, then computability looks impossible. Probabilistic prediction looks like the only way forward.

Penrose says

‘There are always great arguments about the extent to which our actions are determined by *our heredity* and *our environment*. Something which strangely is not often mentioned is the role of *chance elements*. In some sense, all of these are beyond our control.’ (Penrose, 1997: 124)

Chance elements are ‘not often’ mentioned but they are sometimes. For example, Bernard Williams and Thomas Nagel discuss ‘moral luck’. (Williams, Nagel) We should distinguish *external* from *internal* chance elements. External chance elements subject the agent to pressure in decision making. Internal chance elements constitute a degree of randomness in the decision making process itself. Penrose suggests:

‘The normal kind of non-determinism just involves random elements, but that does not help you very much. These chance elements are still beyond our control. You might have *non-computability* instead. You might have *higher-order types of non-computability*.’ (Penrose, 1997: 124)

It is right that non-determinism, or indeterminism, implies the existence of random elements where ‘random’ means ‘not exactly predictable’ and ‘contingent’. However, Penrose sees clearly that we have to distinguish freedom of action from pure randomness. If one’s thoughts and bodily movements were utterly arbitrary, occurred at random, one would hardly say that one had free will. (Imagine being *surprised* by the suddenly flailing of your arm.) Penrose is right that something else is necessary; not being ‘beyond our control’. If I have free will, I am the cause of my own actions. Or, at least, I am *a* cause of my own actions. I am *amongst* the causes of my own actions. In order for this to count as a kind of control, I need to be self-conscious (whatever I am) or in some way monitor, or intervene, in the course of what I do. As Penrose says, there might not only be non-computability but higher order non-computability: meta-computability.

In a stroke of brilliance, Penrose applies the idea of non-computability to yet further levels:

‘Indeed, it is a curious thing that the Gödel-type arguments which I have given can actually be applied at different levels. They can be applied at the level of what Turing called *oracle machines* […]. So one has to consider the question of whether or not there might be some kind of higher-order type of non-computability involved in the way that the actual universe evolves. Perhaps our feelings of free will have something to do with this.’ (Penrose, 1997: 124-5)

Now we have non-computability at at least four levels: quantum, human, oracle machine, and the universe as a whole. Exact predictability gives way to probabilistic predictability in the behaviour of sub-atomic particles, the behaviour of human beings and things made of human beings, computers running inductive or probabilistic programs, and, finally, the gaining and losing of properties by space-time as a whole and its contents. Penrose is right that this is 'a curious thing'. Non-computability seems *to be there*, just as he says. What is curious is that the four levels look otherwise unrelated in any informative way. In our current state of knowledge, it seems as though non-computability 'breaks out' at these four levels. We lack an overarching theory of non-computability, its origin and its applications, which would make sense of the existence of non-computability as a whole, and in its various manifestations.

Whatever metaphysical view we take, free will as we know it seems to have empirical prerequisites, and it is the job of the scientist to discover these. Penrose says:

'I claim that we must search for structures in the brain with some very clear cut physical properties. They should be such as to enable well-shielded spatially extended quantum states to exist, persisting for at least the general order of a second, where the entanglements involved in this state give it a spread over fairly large areas of the brain, probably involving many thousands of neurons all at once.' (Penrose, 1997: 179)

Whatever makes free will empirically possible has to be 'well shielded' because being determined from the outside is incompatible with freely, randomly, or indeterministically, realising one of a range of possibilities. The states constitutive of free will are 'quantum states' because they entail the superposition of choice: being *poised* between one option and another. The 'entanglements' involved in a quantum state are its relations to other quantum states. Penrose thinks these extend over quite large areas of the brain because the exercise of free will causes human action. The empirical prerequisites for free will therefore have to link into the neurological and other biological prerequisites for human action.

Penrose is almost certainly right to argue that the comparatively macroscopic neurology available to us at the moment lacks the properties necessary to ground free will empirically:

'To support such a state, we need biological structures with very precise internal construction, probably with a crystal-like structure, and able to have an

important influence on synapse strengths. I do not see that ordinary nerve transmission can be sufficient on its own because there is no real chance of obtaining the needed isolation.' (Penrose, 1997: 179)

Two features are currently missing from a possible neurology of free will: 'isolation' from causally sufficient conditions, and an explanation of the existence of the capacity to alter synaptic strengths. Penrose explains this:

'We can refer to the reliability of a synapse in passing the message from one neuron to the next as the *strength* of the synapse. If the synapses all had fixed strengths, the brain would be very much like a computer. However, it is certainly the case that these synaptic strengths can change and there are various theories about how they change.' (Penrose, 1997: 125)

Variation in synaptic strengths is empirically necessary for free will in two ways. It allows decision making to be probabilistic. It implies that the brain is not a purely algorithmic machine. Decision making is probabilistic, because if it is *inevitable* that a person chooses one way, any unchosen option was not really an option, the choice no choice at all. It was false that they could have chosen a different course, so no different course was an option. Something is only an option if it could be chosen. The outcomes of the exercise of free will cannot be algorithmic because then they would be exactly computable *a priori*. We cannot predict human choices with anything like strict accuracy, so whatever constitutes free will is not algorithmic.

Because of the parallels between quantum phenomena and free will, Penrose is almost certainly right that 'It might be that quantum mechanics is important in understanding these processes. (Penrose, 1997: 131) cf: 127-131 He offers us this hypothesis for the isolator: 'Microtubules seem to be a good candidate for the structures within which this large-scale quantum coherent activity might take place.' (Penrose, 1997: 133) His reason is:

'One of the things that excites me most about microtubules is that they are *tubes*. Being tubes there is a plausible possibility that they might be able to isolate what is going on their interiors from the random activity in the environment.' (Penrose, 1997: 131)

He says:

'we need some new form of […] physics and, if it is going to be relevant, there must be quantum superimposed mass movements which are well isolated from

the environment. It may well be that, within the tubes, there is some kind of large scale, quantum coherent activity, somewhat like a superconductor. (Penrose, 1997: 131-2)

and

'Significant mass movement would be involved only when this activity begins to get coupled to the (Hameroff-type) tubulin conformations, where now the "cellular automaton" behaviour would itself be subject to quantum superposition.' (Penrose, 1997: 132)

Penrose says

'As part of this picture, there would have to be some type of coherent quantum oscillation taking place within the tubes which would need to extend over very large areas of the brain.' (Penrose, 1997: 132-3)

Penrose says

'There were some suggestions of this general type put forward by Herbert Frölich many years ago, making it somewhat plausible that there might be things of this nature in biological systems.' (Penrose, 1997: 133)

Penrose says '[…] the effects of quantum non-locality, [...] show that effects which are widely separated cannot be regarded as separate from one another.' (Penrose, 1997: 133) The notion of 'separated' but 'not separate' can be applied to free will and mental causation. This is *separated* from that if it is not it, if this not that. This is *not separate* from that if this affects that (or *vice versa* or both). Now, in mental causation, mental events cause physical events. In terms of ordinary physics, we have no idea how this is done. The mental events are not the physical events and, even if they were, we would be left with the problem of how those events *as mental* have causal efficacy. (Davidson, Honderich) Nevertheless, mental events make physical events happen and physical events make mental events happen. Mental events and physical events are, in Penrose terminology, separated but not separate. In free will, one of a range of possibilities is opted for or decided upon. Possibilities are not decisions, but the decision selects the possibility. Motives and intentions are not physical actions but they cause them. Not only are these phenomena separated, not one another, they do not in any intelligible sense 'contact' one another. We need a concept of causation which does not involve touching. Penrose says:

‘Non local effects […] occur in quantum mechanics and they cannot be understood in terms of one thing being separate from another – some sort of global activity is taking place.’ (Penrose, 1997: 133) cf pp. 133-4

An effect is ‘non local’ if it is remote from its cause in time or space, or both. Being remote rules out being juxtaposed or contiguous, so we are being invited to entertain ‘action at a distance’, something we normally think completely impossible.

It is a mistake to reject a view only because it is unfamiliar. We seem to be exposed to causation through physical contact many times every day of our lives. This conditions us to assume that action at a distance is impossible. However, there are surprisingly familiar cases of it. For example, although gravity is a force known to Newton, Galileo, and, in a way, the ancient Greeks, we have next to no idea of what the existence of gravity consists in. We know that large objects, such as the earth ‘attract’ small objects, such as cannon balls or apples, when they are released. But what is this attraction? What is this force? It seems to be nothing over and above *the fact that* the smaller objects move towards the large object. This is all there is to the existence of the ‘pull’. However, if this is what gravity is, *gravity is action at a distance*. Similarly, school pupils are familiar with the attraction of iron filings by a magnet altering their distribution on a sheet of white paper into a curve. The force we call ‘magnetic attraction’ just is the obtaining of the totality or facts like the filings moving towards the magnet (or perhaps the potential, or the disposition, for this to happen). Magnetic attraction is action at a distance. Despite the sophistication of the physics of forces and fields, especially since Einstein, fundamentally, we have no idea how these cases of action at a distance are possible. Nevertheless, they exist.

The case for quantum non-locality is far more acute. Penrose says ‘It is very difficult to understand quantum non-locality within the framework of General Relativity.’ (Penrose, 1997: 137) and

‘My own view is that, to understand quantum non-locality, we shall require a radically new theory. This new theory will not just be a slight modification of quantum mechanics but something as different from standard quantum mechanics as General Relativity is different from Newtonian gravity. It would have to be something which has a completely different conceptual framework. In this picture, quantum non-locality would be built into the theory.’ (Penrose, 1997: 137)

What could this new theory look like? Conspicuous in the difficulties in explaining quantum behaviour is the 'selection' of a particle from a range of possible positions or trajectories. Immense caution is necessary here. Ernst Cassirer is right to warn us against anthropomorphism:

'To identify the 'selection' that an electron is able to make from the set of different quantum orbits – in accordance with Bohr's theory – with 'choice' in the ethical sense of that concept would be to succumb to a purely linguistic confusion. For a choice exists only where there are not only different possibilities, but where also a conscious differentiation and a conscious decision are made. To credit the electron with such acts would constitute a severe relapse into anthropomorphism.' Ernst Cassirer *Determinism and Indeterminism in Modern Physics* trans. O.T. Benfey p. 204

Of course, this does not mean that quantum mechanical models do not apply to free will. They do. Cassirer is cautioning us about doing the reverse: attributing free will to quantum events.

An adequate quantum theory would have to combine *quantum non-locality*, one quantum event affecting another remotely, with an explanation of the *actualisation of possibilities*. As things stand, we have no idea what it consists in for a possibility to become actualised, to be realised. It is obvious that we can in many cases *detect the occurrence of empirical events* but much more is needed here. We need a theory of the transition from possibility to actuality and this in turn requires an explanation of what possibility and actuality are. We need to know how 'mere' possibilities can affect actualities. As Penrose says, 'non-locality […] although it is very puzzling, can still be described mathematically.' (Penrose, 1997: 133) but we need an explanation of what it consists in for non-locality to exist.

Free Will

Along with the immortality of the soul, and the existence of consciousness, a central tenet of the great monotheistic religions is that we have free will. According to Stapp, modern physics shows that we have free will. On this issue, science and religion now coincide. Stapp says

'[Conscious] choices are not fixed by the laws of the new physics,
yet these choices are asserted by those laws to have important causal

effects in the physical world. Thus contemporary physical theory annuls the claim of mechanical determinism.' (Stapp, 2007: vii)

If there is free will, anyone could, at least in principle, not do what they do, but do something else instead (or, perhaps, nothing at all). Free will entails the ability to refrain. If there is mental causation, then mental events, such as motives, intentions, decisions, bring about physical events. Stapp is claiming that physics confirms both free will and mental causation. Stapp says

'it is the revised understanding of the nature of human beings, and of the causal role of human consciousness in the unfolding of reality, that is, I believe, the most exciting thing about the new physics, and probably, in the final analysis, also the most important contribution of science to the well-being of our species.' (Stapp, 2007: 6)

All three facts follow from Stapp's position. If consciousness has a *causal* role in the unfolding of physical reality, then mind body dualism, not idealism, is true. Idealism entails that the physical world is really mental, or logically depends upon consciousness. Stapp, however, thinks physical events are distinct from mental events but caused by them. This is dualism. If we have free will then ethics has a metaphysical foundation: Ethics implies responsibility. Responsibility implies the ability to refrain, and this is implies free will. As we shall see, Stapp spells out the implications of this. If we have free will there is now room for the social sciences to be sciences, not, out of date, Newtonian determinist, or physicalist, sciences, but sciences which model human beings, and human organisations, as quantum systems.

Stapp draws a crucial distinction. Stapp says 'The aspects of nature represented by the [quantum] theory are converted from elements of being to elements of doing.' (Stapp, 2007: 20) Stapp distinguishes *things* from *actions* and claims that the universe is constituted by actions not things. There are in fact two distinctions which are relevant here. Firstly, that between

(1) The universe is made up of things.

and

(2) The universe is made up of events.

secondly, that between

(1) events in general.

and

(2) actions in particular.

By a 'thing' is meant a physical object, or anything like a physical object in its macroscopic properties. A physical object essentially has shape and size, occupies space-time, and is made of a substance called 'matter'. Physical objects interact causally, in manner classically held to be deterministic. In the case of physical objects, if we have enough knowledge of their initial conditions, we can predict their behavior. Einstein, in physics, and Russell and Whitehead in philosophy, argued that physical objects should really be understood as events. In relativity theory, a physical object is a process because, instead of taking about a physical object being 'in' space or time, we talk about a physical object having spatio-temporal properties. Space and time are, so to speak, 'absorbed' into the physical objects, as the set of spatio-temporal relations between them. In philosophical terms, Einsteinian space-time is a kind of Leibnizian space-time. The effect of relativity theory is to construe physical objects as processes, or events. They have not only the dimensions of height, width and length but a beginning, a duration, and an end. During their existence, they change, perceptibly and imperceptibly. This is enough for physical objects to really be events, even if slow events, or processes (where a process is a set of closely related events). The difference between a physical object and, say, an explosion, is one of speed.

Planck, de Broglie and Schrödinger found out that physical objects are have an event constitution. (Russell, Priest) The existence of event-like quanta is a necessary condition for the existence of physical objects, and, in some way that still needs to be spelled out, sufficient. (If the world is really quantum, a good question is: *Why does the world look Newtonian?* Decoherence theories are our best answer at present.) If we conjoin relativity theory and quantum physics we obtain an event ontology: The universe is a universe of events, not physical objects.

Stapp goes much further. Having absorbed the lessons of the Einstein-de Broglie-Schrödinger universe, Stapp argues that events are *actions*. In our normal thinking, all actions are events but not all events are actions. Consciously or not, we work with a distinction between:

(1) what happens.

and

(2) what is done.

We might think of what happens as being the deterministic outcome of what happened earlier. Either way, what happens, of itself, carries no obvious implication about the existence of consciousness or free will. Actions, on the other hand, entail the existence of both freedom and awareness. To say that someone acted is to say they could have done otherwise. (Of course, they might have acted under duress but they are still performing in a way they normally could have refrained from.) We think of our heartbeat as happening rather than something we do. We think of going for a walk as an action we perform, rather than something that happens to us. Breathing and thinking straddle both categories. We can breath more or less deeply, or rapidly, but, if we do not intervene in this way, breathing goes on anyway. We can choose to think about this or that but if we make no choice, thinking goes on anyway. Actions are doings. They presuppose an agent. Stapp says

‘The effect of this change is profound: it replaces the world of material substances by a world populated by actions, and by potentialities for the occurrence of the various possible observed feedbacks from these actions.’ (Stapp, 2007: 20)

Stapp says

‘Thus this switch from 'being' to 'action' allows – and according to orthodox quantum theory demands - a draconian shift in the very subject matter of physical theory, from an imagined universe consisting of causally self-sufficient mindless matter, to a universe populated by allowed possible physical actions and possible experienced feedbacks from such actions.’ (Stapp, 2007: 20)

Stapp says

‘The most salient philosophical feature of quantum theory is that the mathematics has a causal gap that, by virtue of its intrinsic form, provides a perfect place for Homo sapiens as we know and experience ourselves.’ (Stapp, 2007: 143)

In the out of date Newtonian physics, it was thought that the physical universe is a closed system which leaves no room for mental causation. The Third Law

of Thermodynamics entails that the quantity of energy (or matter) in the universe is constant. If every physical event is an inevitable consequence of prior physical events, then non-physical causation is either excluded or superfluous. Suppose causation is understood in terms of necessary condtions or sufficient conditions. Either:

(1) Mental events are necessary for physical events.

or

(2) Mental events are sufficient for physical events.

or both. In the old materialist physics, physical events cannot have mental prerequisites or the universe would not be wholly physical. On the other hand, physical events cannot be necessitated by mental events or there would be overdetermination. There is overdetermination if and only if there is more than one sufficient condition for one and the same event. In the old determinist physics, every physical event is necessitated by previous physical events. If physical events are necessitated by mental events they are necessitated twice over. It follows on the old physics that mental causation is either impossible or redundant, depending what we mean by 'causation'. Either way, the reality of mental causation was left wholly unexplained. (Davidson, Honderich)

If two views are logically inconsistent with one another then at least one of them is false. For a long time, Newtonian physicists and materialist philosophers assumed the non existence of mental causation and the truth of the old physics. We now know, not just from common sense and social science, but from quantum physics, that the reverse is the case. Consciousness has real causal efficacy and it is false that physics implies determinism and materialism. As Stapp puts it 'the mathematics has a causal gap'. This gap allows self-consciousness and free-will. Stapp says

'Man's free will is no illusion. It constitutes his essence. And it rests upon the law of necessity. Any play of chance would falsify the idea that I, from the ground of my essential nature, make a true choice.' (Stapp, 1993: 91)

Having free-will is not an accidental 'add-on' to being a human being. Free will is not something we could lack yet still be human. (Sartre, Priest) If we have an essence, it is part of our 'essence': Free will is one of those properties singularly necessary but jointly sufficient for our being what we are. If we had no free will we would perform no actions. Everything you thought was an action would just

'happen' to you. If you perform no actions you 'do' nothing. It might seem as though you would be reduced to a passive witness of what you previously considered your behaviour. You observe yourself get up from the chair and leave the room but there is nothing you can 'do' to stop this happening. You are 'taken along' by your body. But even this seemingly passive witnessing presupposes free will; an order of priorities in which you pay attention to this rather than that. Free will is ineliminable. Without free will, you would not exist.

Stapp says free will rests on 'necessity' not 'chance'. He means we have to distinguish genuine free will from mere chance. If your bodily movements or your thoughts suddenly occurred in a random or chance way, we would not count that as free will. Far from it. You would have been overcome by some chaotic tendency incompatible with freedom. Being free in your actions is quite different. It involves being the cause of your own actions; acting with the possibility of acting otherwise, in other words, choice.

Stapp says

'In 1935 Albert Einstein, together with two young colleagues, Boris Podolsky and Nathan Rosen, published a paper that focused attention on a paradoxical feature of quantum theory. The theory appears to require this: What is experienced by one person must depend, in certain situations, upon what a faraway and seemingly disconnected person freely decides to do. An intense scrutiny of this puzzling situation by physicists has made clear the fact that the structure of quantum mechanics is profoundly compatible with the idea that [the Process 4] choices can be consistently regarded as free choices. But this element of freedom entails a deep level of interconnectedness of the conscious experiences of persons situated in far-apart regions. This non-local connectedness has been endlessly discussed by physicists and philosophers and is known to be strictly incompatible with any ordinary---that is, local mechanical---idea of how the world operates.' (Stapp, 2005: 5)

In other words, there is action at a distance. One person acts 'in the light of' what another person is doing, or even might do. Free will is exercised in ways informed by absent, remote, or even purely hypothetical situations.
This kind of causation is far removed from the model of two distinct physical objects colliding, and the force of one transmitted to the other. Indeed, Newton emphatically denied that there was action at a distance.

Stapp says

‘The subtle connectivity---revealed by these purely secular scientific studies---between the experiences associated with physically separated persons seems to demand the existence of a reality that can provides the needed connections. But these connections go far beyond anything that classical materialism can accommodate.’ (Stapp, 2005: 5)

Free will, and the action at a distance it implies, needs an ontology. I mean, we need to aks the question: *What has to exist if free will and action at a distance exist?* As Stapp says, the old world of physical objects, or even discrete physical events, will not do. Stapp says

‘What seems to be called for is a pervading immaterial global reality that is informed by our thoughts, and that can subtly act back upon far-away other persons. (Stapp, 2005: 5)

Nothing less than consciousness will do. In order to ‘ground’ free will, and acting-in-the-light-of, there has to exist ‘a pervading immaterial global reality’. This is what religious people call ‘consciousness’. It is immaterial, that is, not physical. It is pervading, that is, it is present in the world but is not the world. It is global, that is: There is not anywhere it is not. It is ‘informed by our thoughts’, that is: It knows what we are thinking.

Stapp’s spirituality is two-way. The ‘pervading immaterial global reality’ acts back on us in turn, not deterministically but ‘subtly’. In theological terms, not only does consciousness know our thoughts but consciousness guides us, in a way that allows our free will.

Wolf says there are mysteries. Wolf says ‘mysteries. They are awesome and marvellous, because no one knows why the world is made this way.’ (Wolf, 1996: 29) It is a contingent fact that ther world is the way it is. We already think of free will as entailing contingency: If you act freely, you could have not acted or acted otherwise. Now, it looks as though, in order to explain why the world is the way it is, we have to admit that contingency presupposes free will. The world is the way it is either because consciousness, as the ultimate source of energy in the quantum field, wants it that way, or else because consciousness has made a self-determining universe; a universe that makes itself be how it is.

Free Will

Walker says 'the mind, through the agency of the will, affects matter and transcends the limits of space and time.' (Walker, 2000: 330) Walker says 'our individual wills, each of which can intervene in the material events of the world, can and must be constrained at some common source.' (Walker, 2000: 330)

Walker says 'We have seen matter and space as the natural consequence of nothing more than the fact that conscious observers exist.' (Walker, 2000: 331) Berkeley and, on most readings, Kant, hold this view. Space and time depend upon consciousness. Matter depends upon space and time, so matter depends upon consciousness. Although shocking to the materialist mind (the materialist mind), as Berkeley sees, this view is strictly empiricist. Logically, no one has ever perceived the physical world unperceived. On pain of incoherence, no scientist has ever observed the physical universe nor any part of it unobserved. Idealism turns out to be the most strictly scientific of philosophies.

Walker says

'science can no longer ignore the fact that our conscious observation affects the quantum potentialities of matter'(Walker, 2000: 331)

Science can no longer ignore consciousness. Why did science ignore consciousness in the nineteeth and twentieth centuries? Clearly, science was successful in explaining physical phenomena but explanation halted abruptly if the subject matter was consciousness. There was an overwhelming human temptation to conclude from 'science explains physical phenomena' that 'only physical phenomena exist', even though the failure of the inference is obvious.

If you are an autonomous agent then you have the capacity of freely choose which actions to perform and which actions not to perform. If you freely choose to worship you have not in any way detracted from that capacity. Of course, whenever you choose *any* course of action over others, those others are not chosen. This does not show you are not free. It shows you are free, (or 'autonomous'). If you are a moral agent you are capable of freely doing good and doing evil. You do not lose this disjunctive moral capacity if you choose to worship.

It is terribly difficult to escape one's own freedom, to cease to be a chooser, even though we may deny our freedom, as the (atheist) philosophy Jean-Paul Sartre clearly sees. Free will, being a chooser, is not an extravagant postulate but an ineliminable presupposition.

Freedom

At the quantum level, free will is scientific and determinism is unscientific. Goswami says 'if we measure an electron possibility wave (technically called a wave-packet) we sometimes find an electron here, sometimes over there.' (Goswami, 2006: 20 There is both unpredictability and indeterminism. The behaviour of the electrons is probabilistic. Goswami asks

'What does this mean, to say that the electron's behaviour is probabilistic, that we can predict only the probability of the electron being here or there, but nothing definite? If we cannot predict things definitely, where is determinism? Determinism is out of the window.' (Goswami, 2006: 20)

Determinism is false, but not because of some failure in our predictive powers which we could put right. The problem is much deeper. The probability, the uncertainty, the randomness, is *in the electrons themselves*. We are not just mapping our own ignorance: If we want to know the determinate truth about the electrons' location, *there is nothing to know*. If the probabilities were not objective, but only degrees or our own belief and ignorance, then there would be hope that we could patch our theories up, and make them deterministic. The problem is, there is no discovery to be made, no experiment to be conducted, which could exhibit the behaviour of the electron as deterministic.

As Goswami says, according to the uncertainty principle: 'we cannot measure both the position and the velocity of a quantum object simultaneously with the utmost accuracy.' (Goswami, 2006: 20) Goswamis argues as follows:

[1] 'we cannot determine both position and momentum (or velocity) simultaneously with utmost accuracy.'

[2] 'Therefore, we can never determine those coveted initial values of Newtonian determinism (initial position and velocity).'

[3] 'Without accurate initial values, it is impossible to predict the trajectories of movement of objects'(Goswami, 2006: 21)

The 'uncertainty' is not just of the form 'We cannot be sure'. There is nothing to be sure of.

As Goswami says 'If material forces, material causes, cannot determine all the future, room is made in the world for creativity and for free will.'

(Goswami, 2006: 21) Free will is not just made possible by indeterminism in quantum mechanics, it is entailed by the possibility of quantum mechanics. If we start to ask *What collapses the probability wave?* or *What actualizes the possible location of the particle?* then the answer is *the conscious observer* or *consciousness*. The point is that *we can choose* whether to measure position or momentum. Freedom of choice is at the basis of quantum physics and, *a fortiori*, at the basis of all physics.

Uncertainty and probabiliy extend to the Newtonian and Einsteinian world, but, at this level, they are far less apparent:

'quantum effects, uncertainty and the possibility wave nature of most macro-objects are camouflaged. The possibility wave of a macro-object spreads between measurements, but it spreads very slowly […]

Nevertheless […] even macro-objects are quantum probability waves, they do spread.' (Goswami, 2006: 21)

There is covert indeterminism, uncertainty, in the behaviour of physical objects. We return to overt and radical indeterminism as soon as we try to predict human behaviour, and the behaviour of human institutions such as governments, armies. Goswami says

'If consciousness is the ground of all beings, then matter exists as possibilities within consciousness. And consciousness chooses among the available possibilities by recognizing a particular one for a particular event.' (Goswami, 2006: 30)

For explanations of human realities, we will need Quantum Social Science, Quantum History. Much of the explanatory lag of social science behind natural science is due to a failure to go quantum: more than a century after Planck published his seminal papers. Goswami rightly points out that quantum events presuppose free will. Goswami says 'the basis of all this self-action is conscious choice.' (Goswami, 2006: 55) If we ask the question 'Why does anything happen?' the answer cannot with any plausibility be that the human individual chooses everything. Our powers are known to be limited. Goswami says 'Consciousness is the chooser.' (Goswami, 2006: 55) Why does anything happen? There is Gaia. Consciousness infinitely exceeds human consciousness.

Goswami says

'the choosing consciousness is unitive, and this one consciousness always recognises and chooses and collapses the quantum probability wave. Behind all

our apparent individuality and separateness, there is one transcendent unity that consciousness is.' (Goswami, 2006: 35)

We can only be choosers because consciousness is a chooser. The soul only has free will because consciousness have given it free will. As soon as we start to ask 'How is the observer possible?', whether we like it or not, we have to start to do theology.

Free Will

Hawking and Mlodinow say about quantum indeterminacy:

'Since this behaviour is not observed on a macroscopic scale, scientists have long wondered just how large and complex something could be and still exhibit such wavelike properties. It would cause quite a stir if the effect could be demonstrated using people.' (Hawking & Mlodinow, 2010: 68)

but they are pessimistic about the prospects:

'but as we've said, in general, the larger the object the less apparent and robust are the quantum effects. So it is unlikely that any zoo animals will be passing wavelike through the bars of their cages.' (Hawking & Mlodinow, 2010: 68)

We do not know that quantum mechanics do not apply to human beings. In history and the social sciences, the models currently deployed are Newtonian and so hopelessly out of date. They were never appropriate to the subject matter. Models taken from quantum physics can and should be tried out in history and the human sciences. They are likely to prove explanatory.

Hawking and Mlodinow miss the application of quantum physics to human reality because they are looking at the wrong level. Their biological counter example expresses pessimism that large organisms will behave more like waves than macroscopic physical objects. For quantum explanation of human behaviour, it does not much matter if they do or do not. Nevertheless, it is already a consequence of relativity theory that any object (including, therefore, any biological organism) can rightly be regarded as a process. Instead of thinking of three things: the physical object, the space, and the time, in which it is located, it is legitimate, and more accurate according to the Einstein theory, to think of one process. The object is not 'in' space and time but the object has spatio-temporal properties. This means it is best viewed as a four dimensional

space-time process. The object is 'strung out' over space-time. It is a space-time worm.

There is no reason why a human being should not be regarded as a space-time worm rather than a physical object. The psycho-physical totality which is the individual human being has a beginning, a duration and an end. It changes its spatial location, which is to say, it changes its spatial relations to other beings. The human being is therefore a space-time process. The human being is 'strung out' over space-time. The human being is a space-time continuant: a psycho-physical process.

Human beings behave like quantum phenomena. Not just human individuals, but human groups are quantum like. Events, processes and institutions made out of human beings such as wars, revolutions, reforms, acts of diplomacy, economic organisations, states and governments exhibit quantum indeterminacy. Human reality is quantum reality in at least these respects:

(1) Human actions are not determined.
(2) Human actions are performed freely.
(3) Human actions are not precisely predictable.
(4) Human actions are probabilistically predictable.
(5) Human beings, and beings made out of human beings, are the cause of their own actions.

Hawking and Mlodinow appreciate that there are experimental attempts to apply quantum ideas to larger phenomena and to the biological. For example:

'Still, experimental physicists have observed the wave phenomenon with particles of ever increasing size. Scientists hope to replicate the buckyball experiment someday using a virus, which is not only far bigger but also considered by some to be a living thing.' (Hawking & Mlodinow, 2010: 68)

While we await the results of virus experiments, two points are relevant.

Firstly: A virus may be already regarded as a process, not just a physical object or group of physical objects. A virus has a genesis, a duration, and an end. It has measurable spatial dimensions and relations. It is strung out in space time. It is a process.

Secondly: Whether a virus is a 'living thing' is more a matter of conceptual stipulation than empirical discovery. We need to decide, if possible, upon the necessary and sufficient conditions for being alive. 'Life' is a family resemblance concept. There is no set of properties common to all and only the entities which we count as alive. Rather, there is a cluster of properties

associated with life. Any given candidate for life might have some percentage of these. They include: Having an organic chemistry, having a genetic make-up, reproduction, nutrition, and excretion, self-locomotion, growth, a capacity for death. (Death might itself be the cessation of some majority of the other properties.) Nothing that is alive need possess all of the properties, nor any one rather than another, but anything that is alive has to possess some cluster of the properties, perhaps a majority. How large that cluster has to be is a matter of choice; conceptual choice on our part. It does not follow that being alive admits of degrees (although that is, in any case, likely). This might just be how concepts work. (Wittgenstein 19 . Bambrough, 19 .)

Experimental attempts to apply the principles of quantum mechanics to ever larger phenomena are *bottom-up*. There is nothing to stop historians and social scientists starting to deploy quantum concepts, straight away, but *top-down*. Their aim should be to meet the physicists coming the other way.

Hawking and Mlodinow presuppose the truth of determinism when they write 'Since people live in the universe and interact with other objects in it, scientific determinism must hold for people as well.' (Hawking & Mlodinow, 2010: 30) From the fact that people live in the universe does not follow that determinism holds for people, even if determinism holds for the universe (which may be doubted). Suppose every event in the universe is inevitable, given its causes. It is perfectly possible for human beings to live in such a universe and their actions *cause* some of those events but not be *causally necessitated* by any of them. Our actions are then very much like quantum events (and not much like, say, billiard balls in motion).

Hawking and Mlodinow tendentiously construe their opponents as arguing 'in order to preserve the idea of free will'(Hawking & Mlodinow, 2010: 30) Nobody does this. (Sartre, for example, regards the reality of freedom as a terrible burden.) Free will is not the object of some sentimental attachment. Free will is a fundamental presupposition which is terribly difficult to eliminate. Unless you presuppose free will, it is impossible to make sense of guilt, blame, responsibility, morality, immorality, achievement, failure. Without free will it is impossible to make sense of the human world. Determinism is an out of date and abstract theory. Whether we like it our not, our lived reality is a world of choice. (Sartre, Priest)

Hawking and Mlodinow ask the rhetorical question: 'If we have free will, where in the evolutionary tree did it develop?' (Hawking & Mlodinow, 2010: 31) They expect the answer 'At no point'. The answer to this question is: At that stage when the brain became large enough to develop a neurological feedback mechanism in the cerebral cortex which allowed human beings to affect their own behaviour: at that stage when there emerged a with-holding and

implementing system regulated by the brain itself. The powers of decision-making presupposed by hunter gather societies suggest this evolutionary emergence is Pre-Palaeolithic. If the theory of evolution is true, it confirms the existence of consciousness rather than refutes it. Creation evolves.

Hawking and Mlodinow make several mistakes when they write:

'Recent experiments in neuroscience support the view that it is our physical brain, following the known laws of science, that determines our actions, and not some agency that exists outside those laws.' (Hawking & Mlodinow, 2010: 32)

and

'For example, a study of patients undergoing awake brain surgery found that by electrically stimulating the appropriate regions of the brain, one could create in the patient the desire to move the hand, arm, or foot, or to move the lips and talk.' (Hawking & Mlodinow, 2010: 32)

There are several logical confusions here. Firstly, from the commonsense platitude that we use our brains to think with, it does not follow that determinism is true. Secondly, from the fact that using the brain is a necessary condition for human action, it does not follow that it is sufficient. Something else might be necessary (for example: consciousness). Thirdly, even if determinism holds true in this experimental situation (which may be doubted) it does not follow that it holds true when our brains are not being experimentally manipulated. Fourthly, we do not know that stimulating the brain causes *desire*. Desire is hidden and inward. It can be *felt* or undergone but not *observed* experimentally or otherwise. Fifthly, if the patient does feel desire, and if that desire does have a role in their speech and action in the experiment then this mental causation, which lies outside the physical causal chain, is left wholly unexplained. *How does consciousness push atoms around?*

What has gone wrong here? How could Hawking and Mlodinow, and possibly the scientists conducting the experiments, *jump to conclusions*? Hopefully, the most likely explanation is not that they are doing exactly what atheists accuse religious people of doing. Do they *want* a certain view to be true? Do they *want* determinism to be true? Do they therefore conclude that it is true? Even in some of our top scientists, the old Newtonian picture is still there as a hang-over. They expect a Newtonian picture so they see one. Perception is pre-conception. The observer affects the observed. (We need a quantum theory of scientific progress and the lack of it.)

Hawking and Mlodinow say that in a quantum system: '[...] given the initial state of a system, nature determines its future state through a process that is fundamentally uncertain.' (Hawking & Mlodinow, 2010: 72) They do not realise that, in that case, we have to give up the concept expressed by 'determines'. What is uncertainty?

The term 'uncertain' is ambiguous between a subjective, or psychological, sense and an objective sense. In the psychological sense, if I am 'uncertain' then *I am unsure about the facts*. There might well, nevertheless, be facts, one way or another. In the objective sense, if there is uncertainty, *there is no fact of the matter*. Reality is indeterminate: neither one way nor another. In the objective sense of 'uncertain' there is uncertainty irrespective of the state of anyone's mind (indeed, irrespective of whether there are any minds). The two senses of 'uncertain' are compatible to the extent to which I can be uncertain, and there might be real indeterminacy 'out there'. Hawking and Mlodinow are talking about 'fundamental' uncertainty. They mean uncertainty at least in the objective sense, and probably in the subjective sense also.

Now, uncertainty in the objective sense is incompatible with determinism. In a determinist system, a state of a system at an earlier time entirely determines the state of the system at a later time. Here something 'determines' something if and only if it makes it what it is. 'Determines' means 'causes' in the sense of 'is a sufficient condition for'. Given the earlier state, the probability of the later state obtaining is 1. In the quantum case, given the earlier state of a system, the later state of the system is indeterminate or, as Hawking and Mlodinow put it, the later state comes about through 'a process that is fundamentally uncertain'. If something is indeterminate there is no one way it is, so it is not determined. It is not caused, in the sense that there are no sufficient conditions for its obtaining. The probability of its coming about is between 1 and 0, so not 1 (and not 0). On the other hand, Hawking and Mlodinow are right to conclude: 'In other words, nature does not dictate the outcome of any process or experiment, even in the simplest of situations.' (Hawking & Mlodinow, 2010: 72 Here they realise that determinism has to go. If nature 'does not dictate the outcome' nature does not determine the outcome. Nature does not make the outcome inevitable. It is, exactly as Hawking and Mlodinow say: 'Rather, it allows a number of different eventualities, each with a certain likelihood of being realized.' (Hawking & Mlodinow, 2010: 72)

The essential point for free will is this. If, as historians or social scientists, we are trying to predict the outbreak of a revolution, or a war, or the result of some piece of diplomacy, *we cannot do this with strict scientific accuracy*. It is often said that history and the human sciences are not sciences. This is right. They are not *Newtonian* sciences. As such they will always fail. They can

however be *quantum* sciences. The historian needs to reach for the quantum physics textbook.

Suppose we identify, in a scholarly way, what we consider the preconditions for a revolution. We proceed empirically, looking at, at least; the 1640s in Europe, the 1770s in America and Poland, 1789 in France, 1848 in Europe, 1917 in Russia, 1949 in China, and so on, finding perhaps a cluster of preconditions forming a 'family resemblance', rather that strict prerequisites: mass poverty followed by a sudden amelioration of economic conditions, an emergent middle class, a need for industrialisation or urbanisation, an intelligencia, the dissemination of new ideas, a dissatisfied group willing to use violence, a j-curve of expectations, etc. Suppose these conditions, or some majority of this cluster, are met. We can then allocate with some probability the likelihood of a revolution taking place. Nevertheless, it might not take place. To use Hawking and Mlodinow's words: The initial conditions 'do not dictate the outcome'. They 'allow[...] a number of different eventualities, each with a certain likelihood of being realised'. Revolutions are quantum events.

Physicists often describe the behaviour of particles, metaphorically, as though the particles themselves have free will. Hawking and Mlodinow, themselves, say the particle 'had to "decide" whether to travel through just one slit' and a particle had 'The choice whether to take one or both paths' and they speak about a particle having a 'choice'. They say 'with our observation in the laboratory we will be affecting that choice' (Hawking & Mlodinow, 2010: 83) Of course, Hawking and Mlodinow do not think quantum particles have free will, or really make any choices. Nevertheless, the reverse application does hold. The exercise of free will exhibits quantum features. The animism and anthropomorphism of Hawking and Mlodinow raise a question about scientific explanation. We need to know: *What would be left of scientific explanation if we dropped all metaphor?* If we stopped talking about genes carrying information, nerves carrying messages, sub-atomic particles choosing this path rather than that etc.how much would remain? In science, perhaps only the mathematical is the literal.

Hawking and Mlodinow say 'It is, to paraphrase Einstein, as if consciousness throws the dice before deciding the result of every physical process.' (Hawking & Mlodinow, 2010: 72) In the human case, it is as though the human being throws the dice before deciding on every physical action.

Perhaps Hawking and Mlodinow are right. Perhaps consciousness does throw dice before deciding the result of every physical process. With the physics we have, we have no idea why quantum events come into and out of existence. If we take seriously my suggestion that consciousness causes quantum events then, in a very real sense, consciousness does play dice.

consciousness creates an indeterministic universe. consciousness has not created a universe in which each event is an inevitable consequence of a prior event, or even the whole set of prior events. The indeterministic universe emerges with a degree of randomness from *the here and now*. Quantum events come in and out of existence. They come in and out of Being, and Being is the being of consciousness. They come in and out of Presence and Presence is the presence of consciousness. It follows that quantum events come from consciousness, and return to consciousness.

Be that as it may, in the human case, choices are quantum events. Whatever the initial conditions, whether specified neurologically or in terms of a Freudian unconscious, or the sociology of the individual, any choice could indeterministically go one way or another.

Hawking and Mlodinow say 'Quantum physics might seem to undermine the idea that nature is governed by laws but this is not the case.' (Hawking & Mlodinow, 2010: 72) Yet the deterministic natural laws are now like Hawking and Mlodinow's view of free will: an 'effective theory'. They map our ignorance. They *work* macroscopically but the fundamental truth is quantum and indeterministic.

Hawking and Mlodinow say about quantum mechanics:

'[...] it leads us to accept a new form of determinism: given the state of a system at some time, the laws of nature determine the *probabilities* of various futures and pasts rather than determining the future and past with certainty.' (Hawking & Mlodinow, 2010: 72)

This is the very last stand of determinism. The 'new form of determinism' is in fact *indeterminism*. This is like saying we need to accept a new form of democracy: 'dictatorship', or a new form of freezing: 'boiling', or a new form of darkness: 'light'. Determinism is dead. We might as well drop the word.

Nevertheless, Hawking and Mlodinow are clearly right to say we can provide probabilistic explanations in quantum mechanics. Indeed, their conception of multiple futures and pasts could be given wide application in the humanities. To begin to explain political, or economic, or military, events we not only need possible futures and possible pasts but possible *envisaged* futures and possible *envisaged* pasts: envisaged by the historical agents. Identifying options, and possible trajectories issuing in those options, is necessary for allocating probabilities to outcomes. In each case, the human *nexus* is a point of indeterminism: a switch which switches itself, in many possible directions. The human *nexus* is a human being or anything made out of human beings.

Hawking and Mlodinow think that in the *non*-quantum case:

‘Our use of probabilistic terms to describe the outcome of events is […] a reflection not of the intrinsic nature of the process but only of our ignorance of certain aspects of it.’ (Hawking & Mlodinow, 2010: 74)

One can see the plausibility in this. You throw a dice. We say there is a 1 in 6 chance of a six coming up. On the other had, we are tempted to say *if only* we knew enough about the velocity, and angle, with which the dice left your hand, if only we knew how the dice was very slightly weighted in its materials, or even its number of pips, if only we knew about the surface of the table where it will land, then, surely, we could predict with certainty whether a six would come up. There is ‘subjective uncertainty’ but no ‘objective uncertainty’.

Nevertheless, Hawking and Mlodinow say ‘Probabilities in quantum theories are different. They reflect a fundamental randomness in nature.’ (Hawking & Mlodinow, 2010: 74) In quantum mechanics, the randomness is really ‘out there’. Nature, in itself, is indeterminate. In other words, in quantum reality there is ‘objective uncertainty’ in the sense defined. Finally, Hawking and Mlodilow admit it:

‘According to quantum physics, no matter how much information we obtain or how powerful our computing abilities, the outcomes of physical processes cannot be predicted with certainty because they are not *determined* with certainty.’ (Hawking & Mlodinow, 2010: 72)

Determinism has gone. The science of the human world will be the science of free will. The human world is a quantum world.

The same applies to the human world. It is as crude a mistake to apply materialism and determinism in explaining human action as it would be to apply Newtonian mechanics to quantum phenomena.

Hawking and Mlodilow are right to say about indeterminism: ‘Though that is distasteful to some, scientists must accept theories that agree with experiment, not their own preconceived notions.’ (Hawking & Mlodinow, 2010: 72) There is not only a delay in science reaching philosophers and theologians. There is a delay in the latest science reaching the scientists. Although Hawking and Mlodilow are right to be critical of philosophy: ‘Philosophy has not kept up with modern developments in science, particularly physics.’ (Hawking & Mlodinow, 2010: 5) there is no room for complacency:

Science has not kept up with modern developments in physics, particularly quantum mechanics.

Davies on Free Will

'And you can decide what to look for. Look for its location and you get an atom at a place. Look for its motion and you get an atom with a speed. But you can't have both. The reality that the observation sharpens into focus cannot be separated from the observer and his choice of measurement strategy.' (Davies, 1983: 103)

Is there choice? Kant argues that we have free will. Each of us is a 'spontaneity' (or, perhaps as we would say in the light of physics, a 'singularity') so the observer has the sort of freedom of scrutiny, or freedom of observation, that Davies describes. Kant, however, assumes that how the world appears to us is inevitably dictated by the categories. We have no freedom over whether to apply or withhold them: If there is observation the categories are applied. Davies, on the other hand, is saying that observing location determines location and observing motion determines speed. Freely creating the world to this degree is simply not envisaged by Kant. Indeed, he thinks appearances are a deterministic system.

Davies fully appreciates that these idealist and indeterminist conclusions are counter-intuitive, at least to the 'common sense' of the Newtonian mind. He says:

'If this all seems too mind-boggling or paradoxical to accept, Einstein would have agreed with you. Surely the world out there really exists whether we observe it or not? Surely everything that happens does so for its own reasons and not because it is being watched? Our observations might uncover the atomic reality, but how can they *create* it?' (Davies, 1983: 103)

Somehow, the mind-dependence of quantum phenomena has to be reconciled with the apparent mind-independence of the Newtonian world. How is this possible? Davies thinks the assumption that the physical objects are really 'out there' has to go.

Tipler on Free Will

‘The very fact that humans are machines of a very special sort allows us to prove that we humans probably have free will, that we shall have life after death in an abode that closely resembles the Heaven of the great world religions, and that life, far from being insignificant, can be regarded as the ultimate cause of the very existence of the universe itself.’ (Tipler, 1994: xi)

Tipler is not alone in thinking that materialism is compatible with free will. For example, Epicurus advocates the materialist thesis that existence is only atoms moving in the void. However, some of the atoms move in random, unpredictable ways. Atoms collide, and cause one another to change direction, but it is a contingent fact that an atom moves one way rather than another. This leaves room for what human beings call ‘free will’. (Epicurus)

Normally, in Kant’s philosophy, we have no knowledge of things as they are in themselves, only as they can appear to us. However, free will is an exception. Each of us really is free according to Kant, in our own intrinsic nature or as a thing-in-itself. Tipler is right that science does not show this, but Kant thinks it is demonstrated by his own ‘critical’ philosophy.

Tipler emphasises that physics does not give us absolute certainty (Arguably, this is obtainable in mathematics and pure logic.) He says about Kant’s questions:

‘the answers to all three are *probably* He does exist, *probably* we have free will, and *probably* He will grant us eternal life after we have died. I say “probably” because science is not in the business of giving an absolutely certain-to-be-true answer, valid for all time.’ (Tipler, 1994: 7)

Barr says

‘there is nothing in the laws of nature or the character of physics as they exist today which is logically incompatible with free will.’ (Barr, 2003: 184)

This is right. If you choose one course of action over another you do not thereby break any natural law. If you choose, your action is thereby not causally determined but there is no scientific reason why natural laws should be understood in a determinisitic manner. Determinism is a philosophical theory, not part of science. It is the thesis that causes necessitate their effects, such that if the cause happens, the effect cannot not happen. But as Hume showed, no one has ever observed this supposed necessity. There is no experiment that has ever suggested its existence. Necessity is a philosophical notion and there is no

scientific reason to think that natural laws are anything other than regularties in nature. (Hume, Priest)

Planck on Free Will

Planck presents a similar solution to the problem of free will and determinism. Plank says: 'Observed from without, the will is causally determined. Observed from within it is free.' (Planck, 1950: 75) If I act freely, I could not-do what I do, in the sense of refrain from doing it. If I am determined, I could not *not* do what I do. Being free and being determined look incompatible, because their conjunction entails a contradiction: I could, and could not, not do what I do. Refraining looks both possible and impossible. However, Planck thinks there can exist both free will and determinism. He thinks events, or actions, appear free to an internal point of view but appear determined from an external point of view. Plank says:

'The external viewpoint is the one which permits the volitional processes to be observed without being disturbed, affected, or interfered with, by the observation. (Planck, 1950: 72)

but

'The internal viewpoint is the one, from which the volitional processes cannot be observed without thereby being disturbed.' (Planck, 1950: 72)

In the internal perspective, the action is changed by the observer. When I am free, I make a difference to what I do. In the external perspective, the action is not changed by the observer. The events continue undisturbed.

Do we have Free Will?

It is a central tenet of Christianity that there is freedom of the will. Planck says 'Religion belongs to that realm that is inviolable before the law of causation and, therefore, closed to science.' (Planck, 1932, QQ: 151) Planck asks

'Is there something in the nature of man, some inner realm, that science cannot touch? Is it so that when we approach the inner springs of human action science cannot have the last word? Or, to speak more concretely, is there a point at which the causal line of thought ceases and beyond which science cannot go?' (Planck, 1932, QQ: 149)

By 'the causal line of thought' Planck means determinism. 'Determinism' is the doctrine that any event is inevitable given its causes. If the cause happens, the effect *cannot not* happen. In other words, causes *necessitate* their effects.

Determinism is usually thought to be incompatible with free will. You act freely if, when you do something, *you could have not done it*. If you act, it was within your power not to act. You could have done otherwise or (if there is such a thing) done nothing at all.

Free will and determinism are incompatible because, if you have free will, *you could have not acted* but, if determinism is true, *you could not have not acted*. The combination, or conjunction, of free will and determinism entails a contradiction: *you could and could not have not acted*. On most views, entailing a contradiction is fatal to any theory.

Determinism used to be thought a scientific view, so, to be intellectually respectable, or rigorously scientific, you had to be a determinist and show little patience with free will. However, this attitude is mistaken and out of date. Far from being scientific, it is based on ignorance of the fundamental revolutions in science which have taken place over the last hundred years or so. The Newtonian physics of the seventeenth century was indeed deterministic, and Einstein tried to rescure determinism in the last century. Quantum physics, cutting edge physics, entails the opposite of determinsim: Quantum physics is indeterministic and probabilistic in a way that not only leaves room for free will but, arguably, proves it.

Planck is certain that we have free will. Plank says:

'[…] the individual ego […] Over this realm, no outer power of fate can ever have sway, and we lay aside our own control and responsibility over ourselves only with the laying aside of life itself.

Here is the place where the freedom of the will comes in and establishes itself, without usurping the right of any rival.' (Planck, 1932, QQ: 150)

Planck's argument is:

(1) Free will is necessary for life (in the sense of someone's life, for example, your life).

(2) There is life

(3) Therefore there is free will.

This argument is valid, because everything entails its own prerequisites. It is only sound if the premises are also true.

Is Planck right that free will is necessary for life? He is certainly right in saying there can be no responsibility, no self-control, unless there is free will. If you could not do otherwise, you are not responsible for what you do. If you could not do otherwise, there is no sense in which you are in control of your actions. If your actions are the inevitable result of factors outside your control, then they are not yours. They happen but are not 'done'. If determinism were true, there would be no sense in which anyone ever 'did' anything. Everything that you 'do' would really be something that 'happens to' you. It looks, therefore, as though Planck is right. You could not live your life, or, in a sense, there would be no 'you', without the control you have over your own actions, and the responsibility this entails. Free will is necessary for self-control and responsibility, therefore free will is necessary for life. There is life. Therefore, there is free will.

When Planck says there is free will 'without usurping the right of any rival' he is leaving room for determinism to be true *at a certain level*, roughly speaking: the Newtonian, and Einsteinian, levels at which medium sized physical objects appear to us. Determinism breaks down at the limits of the Newtonian-Einsteinian world. Some realities are too small to be deterministic, some too large. Some are too subjective, too intimate, too proximal, some too vast. The cosmos as a whole falls into the second category. Your own existence falls into the first.

Planck thinks we know we have free will. In our lives we are confronted with the reality of choice. We do not *infer* this freedom from our own conscious states. It is a way in which they are directly given: '[…] the freedom of the ego here and now, and its independence of the causal chain, is a truth that comes from the immediate dictate of human consciousness.' (Planck, 1932, QQ: 150) Like Sartre, Planck thinks that if determinism were true it would be of no practical use to us in resolving the very real moral dilemmas we face. Plank says: 'In the conduct of our own lives, the causal principle is of little help.' (Planck, 1932, QQ: 150) Sartre invites us to consider the case of a young man, in Nazi occupied France, who faces the choice of joining the resistance and avenging his brother's death or looking after his bereaved mother, who would be doubly bereaved were he to be killed in the fighting. One of the lessons of Sartre's story is that it would be of no help whatsoever to be told that determinism is true, that the way you choose is inevitably dictated by your neurology, your upbringing, your unconscious. Determinism is an abstraction. Free will is a lived reality. (Sartre)

The determinist has free will. Plank says:

‘Being emancipated […] we are at liberty to construct any miraculous background that we like in the mysterious realm of our own inner being, even though we may be at the same time the strictest scientists in the world, and the strictest upholders of the principle of causal determinism.’
(Planck, 1932, QQ: 150)

It follows that determinism is existentially, or pragmatically, self-refuting. The determinist, like anyone else, has free will. The determinist is the living refutation of determinism. To put it another way: Determinism may be formulated just on condition it is false.

Planck sees free will as being exercised in the present, in the now. Although past, present and future are not scientific concepts, they are tacitly presupposed by the possibility of science. Planck says:

‘And what holds good for the present moment of our being holds good also for our own future conduct in which the influences of present ego play[s] a part. The road to the future always starts in the present. It is, here and now, part and parcel of the ego. And for that reason, the individual can never consider his own future purely and exclusively from the causal standpoint.’ (Planck, 1932, QQ: 150)

The present is an indeterministic *spontaneity* in which events occur and are extinguished. We experience this spontaneity in our own freedom. Determinism, on the other hand, presupposes a linear picture of events in causal chains. But this line does not exist.

Free Will

Bohr says ‘it is evident that the word volition is indispensable to an exhaustive description of psychical phenomena.’ (Bohr, 1958: 78) Something essential to mentality is left out if free will is left out. What?

Suppose, for a moment, you experienced determinism as true. Suppose your *body* stood up, walked across the room, opened the door, and walked outside, all completely against your will, but taking you with it. Imagine yourself protesting, from inside it: ‘Stop, go back! I want to stay where I was!’ Needless to say, you wish you had not lost control over where your body goes. The fact that this scenario is fanciful shows how fundamental free will is to our

ordinary lives. The fact is you are in control of where you go. This control is *your being able to exercise the choice* of remaining seated or crossing the room etc.

How far can we extend the scenario? Can we suppose you not only lose control over your body but also your mind? Of course, as things are now, you do not freely choose each thought before you have it. There is no inner rehearsal of the the thinking which precedes the thinking. On the other hand, you can choose to think about one thing rather than another. If you make no such choice, thinking goes on anyway. Thoughts and emotions arise and subside. Thinking is like breathing. You can choose to breath faster or slower but, if you do not do this, thinking and breathing go on anyway. Thinking and breathing are between freedom and determinism.

Suppose, however, not only does your body get up and walk against your will (even though you still view the world from it) but also trains of thought appear in your mind, completely foreign to anything you want to think. Suppose you cannot stop these thoughts, cannot 'change the subject'. Like Raskalnikov in *Crime and Punishment*, you are tyrannised by thoughts. At least Raskalnikov retains the *capability* of thinking otherwise. Suppose we lost even this. What then?

In this scenario, one is no longer the author of one's thoughts, no longer the doer of one's physical actions. They are not 'owned' by you. You are reduced to a passive witness or specatator of what your body and mind do. In this case, we have to give up talking about 'actions' and 'thoughts' and mention only 'happenings'. To have to give up talking about actions and thoughts is to have to give up talking about mentality. Free will is necessary for thought and action, therefore free will is necessary for mentality.

How is free will possible? Bohr says

'I [...] emphasize that it is just this impossibility of distinguishing, in introspection, sharply between subject and object which provides the necessary latitude for the manifestation of volition' (Bohr, 1958: 22)

Why should the dissolution of the distinction between subject and object make possible free will?

One answer is this: Introspection is a kind of self-consciousness. In self-consciousness, in a way, subject and object are one. Nobody has investigated the complexities of self-consciousness with greater thoroughness than Hegel. Hegel says that, in self-consciousness, consciousness draws a distinction which, at the same time, is for consciousness, not a distinction. (This sort of utterance, understandably, annoys some of Hegel's detractors, but it is well worth

persisting with it.) Consciousness draws a distinction, in self consciousness, between subject and object, between that which is conscious, and that which consciousness is consciousness of. This is not a distinction, both in the sense that what consciousness distinguishes from itself is only itself, and in the sense that the subject-object distinction does not appear to consciousness. In self-consciousness, the subject-object distinction is epistemological but neither ontological nor phenomenological. It is epistemological because it is a distinction between the knower and the known. It is not ontological because there is only one self-consciousness here. It is not phenomenological because only what appears directly to consciousness is phenomenological.

How does self-consciousness bear on free-will? Bohr talks of providing the necessary 'latitude' for free will. Free-will requires a space, or quasi-space, in which to operate. To appreciate this intuitively, suppose you are cycling along the High Street, in order to post a letter. If you are not particularly conscious of cycling, the cycling goes on anyway. You do not have to plan each rotation of the peddle before executing it. The cycling is automatic. If, however, you become conscious of cycling this opens up a space of choices. At that moment you can choose to cycle faster or slower, or stop the bike altogether. The power of *intervening in the course of our own actions* is available. Crucially, it was an act of self-consciousness that made possible acting otherwise.

The space of choices opened up by an act of self-consciousness, the 'latitude' Bohr talks about, is made possible by the structures Hegel describes (or something close to them).

It is not a purely accidental fact that human beings are both self-conscious and have free will. We are the beings who are *aware of themselves* and who *make themselves what they are*. Self-consciousness opens up the space for freedom. Indeed, it is our conspicuous distinguishing feature that *humanity creates a human world*. We see ourselves reflected back in our creations and manipulations.

Bohr encourages us to

'look for a direct correlation between life or free will and those features of atomic phenomena for the comprehension of which the frame of classical physics is obviously too narrow.' (Bohr, 1958: 20)

'Life' is ambiguous between 'biological life' and life in the sense of 'your life'; your life as you live it. Life in either sense cannot be understood on classical, Newtonian, models. The attempt to do so is both explanatorily futile and politically repressive. There are two ways of exceeding the determinist and classical: downwards to quanta and upwards to human reality, and the

consciousness, politics and economics it entails. Newtonian mechanism, far from being the whole truth, applies to what Austin calls 'medium sized dry goods'. (Austin)

The quantum world and the human world are probabilistic worlds. Indeed, the human world is a quantum world. It is important not to misunderstand this probability. Bohr says

'However, it is most important to realize that the recourse to probability laws under such circumstances is essentially different in aim from the familiar application of statistical considerations as practical means of accounting for the properties of mechanical systems of great structural complexity. In fact, in quantum physics we are presented not with intricacies of this kind, but with the inability of the classical frame of concepts to comprise the peculiar feature of indivisibility, or "individuality", characterizing the elementary processes.' (Bohr, 1958: 34)

The probability is ontological not merely epistemological. It is 'out there', not just in our theories. This is not the kind of probability that merely maps our ignorance or expresses the degree of our belief. The likelihoods are a feature of the world because, both in the quantum and in the human case, the subject matter affects itself.

Bohr talks about the 'individuality' of quantum phenomena in a sense analogous to human individuality: *the quanta behave as though they have free will and personality*. This is not to say the quanta *do have* free will and personality but, if they did, we would not expect them to behave much differently. Bohr is quite clear about this. Bohr talks of

'an unexpected limit for the subdivision of all physical actions extending far beyond the old doctrine of the limited indivisibility of matter [...] giving every atomic process a peculiar individual character.' (Bohr, 1958: 24)

By saying each quantum has a 'peculiar individual character' Bohr means it is as as though each has a will. Each differs from every other. Each exhibits an ineliminable unpredictability. What does 'individuality' mean? This one is not that. This one is not *like* that. Individuality is also indivisibility. An individual cannot be split.

Until 1927, determinism looked like a scientifically respectable theory which had wide application. This is why Bohr says the individual behaviour of the quantum is 'unexpected'. Who would have thought it? *The constituents of Newtonian objects are not Newtonian*. After 1927, there is little excuse for

talking about the 'mechanism' of anything outside Newtonian mechanics. 'Mechanism' is a metaphor that belongs in the 1830s.

By 'the principle of causality' Bohr means the conjunction of the claims that every event has a cause, and all caused events are inevitable. Bohr says

'the principle of causality, so far regarded as the unquestioned foundation for the interpretation of natural phenomena, has proved too narrow a frame to embrace the peculiar regularities governing individual atomic processes.' (Bohr, 1958: 25)

Quantum physics entails the falsity of both parts of 'the principle of causality'. There are phenomena which come in and out of existence, without any cause, or at least, without any physical cause in this universe. There are phenomena which occur only with a probability less than one, given the initial conditions. If there are events without causes then it is false that every event has a cause. If some causes only make their effects *likely* then it is false that all caused events are *inevitable*. Therefore 'the principle of causality' is false.

The refutation of determinism is thoughly scientific. Experiment after experiment falsifies determinism and confirms indeterminism at the quantum level. Understandably, Bohr says

'Certainly everyone will understand that physicists have needed very cogent reasons to renounce the ideal of causality itself; but in the study of atomic phenomena we have repeatedly been taught that questions which were believed to have received long ago their final answers had most unexpected surprises in store for us. (Bohr, 1958: 25)

There is a distinction between causation and creation. Causation is not *ex nihilo*: the materials to bring about an effect are *already there*. Creation is *ex nihilo*: the materials to bring about an effect are *not already there*. Remarkably, the coming in and out of existence of quantum phenomena looks much more like creation than causation. For example, Bohr says:

'there could be no question of attempting a causal analysis of radiative phenomena, but only, by a combined use of the contrasting pictures, to estimate probabilities for the occurrence of the individual radiation processes.' (Bohr, 1958: 34)

If determinism is false, we need to ask this question: *If an event did not have to happen why did it happen?*

Hume thinks the concept of causal necessity is not empirical. Although we observe one event followed by another in time, and can thereby establish regularities in what follows what, we observe no compulsion, inevitability, or 'necessary connexion' between one event and the next. Indeed, Hume suggests that it is only the repeated witnessing of events of one type following another type that gives rise to the idea of causal necessity in our minds. We mistake our *psychological expectation* that this event will follow that for the *objective necessity* for this event to follow that.

Depending on how we read Hume, Hume either concludes that there is no objective causal necessity, or that causal necessity really is out there but hidden from our senses, or that there is room for a great deal of scepticism about the existence of causal necessity. It is uncontroversial that Hume thinks all we directly observe are 'constantly conjoined events'. Our knowledge of what follows what is *a posteriori*, knowable only by experience. What follows what is a contingent fact: it *is* a fact but it might not have been. If one billard ball strikes another it moves along the beige. In the absence of witnessing this, for all we know, it might not move along the beige but jump up and down, glow red, and sing 'Rule Britannia'.

Hume's analysis of causation undermines determinism even in the classical case. Of course, Hume is not denying that we can make very successful predictions. He insists that we can. He is casting doubt on the idea of causal necessity, or inevitability. We can add to *If it did not have to happen, why did it happen?* a new question: *Why does anything happen?*

Bohr raises the issue of 'how to speak of the appearance of the phenomena for which only predictions of a statistical character can be made.' (Bohr, 1958: 51) He says that in 'a terminology proposed by Dirac [...] we were concerned with a choice on the part of "nature".' (Bohr, 1958: p. 51) and

'The circumstance that, in general, one and the same experimental arrangement may yield different recordings is sometimes picturesquely described as a "choice of nature" between such possibilities.' (Bohr, 1958: 73)

What is this 'choice'? Bohr does not wish to say that quanta themselves have free will: 'it is hardly reasonable to endow nature with volition in the ordinary sense' (Bohr, 1958: 51) and

'Needless to say, such a phrase implies no allusion to a personification of nature, but simply points to the impossibility of ascertaining on accustomed lines directives for the course of a closed indivisible phenomenon.' (Bohr, 1958: 73)

Understandably, Bohr resists anthropomorphism; the projection of human characteristics onto the non-human world. Yet, the point cannot be only epistemological. The ontology makes the epistemology possible.

The behaviour of quanta is *like* the exercise of free will by a self-conscious human individual in crucial respects. Both quanta and people are predictable in their behaviour but not with deterministic certainty. Quanta affect quanta, and people affect people, but in ways which cannot be tracked deterministically or physically. Perhaps most remarkable of all, both quanta and people seem to be affected by unactualized, not just actualized, possibilities. Both operate in a 'space of possibilities' or 'space of choices'. Quanta are affected by unrealized quantum situations. Human beings are affected by unrealized human situations. Both quanta and people act 'in the light of' possibilities. Under the heading of 'people' here I include anything made of people: states, governments, armies, businesses, classes, and so on.

Does consciousness play dice? No. But consciousness has created a world in which plenty of dice playing goes on, some of it at a fundamental level. consciousness has created *a world which makes itself what it is*. This is not a deterministic world, but it is a self-determining world.

Bohr draws an important distinction between an 'open process' and an 'closed process'. He says:

'the closed processes studied in quantum physics are not directly analogous to biological functions for the maintenance of which a continual exchange of matter and energy between the organism and the environment is required.' (Bohr, 1958: 76)

This is right. Biological organisms, including the living human body, constantly shed and gain matter, lose and gain molecules through, for example, nutrition and excretion, breathing, the shedding of skin cells. As Bohr says

'In fact, we cannot even tell which particular atoms really belong to a living organism, since any vital function is accompanied by an exchange of material though which atoms are constantly taken up into and expelled from the organisation which constitutes the living being.' (Bohr, 1958: 10)

Quanta, on the other hand, seem to gain or lose nothing.

The distinction between an open and a closed system has a close bearing on the possibility of free will. It could be that a system has to be in some sense, or to some significant degree, 'closed' in order for it to make choices.

Insulation from the environment is necessary for the putative agent not to be forced to act just in one way. Insulation from the environment creates what I am calling the 'space of choices'. It allows an agent to affect its own behaviour, to intervene in the course of its own actions. Insulation admits of degrees. So, therefore, does freedom. To the extent to which a system is self-determining it is insulated from the environement. A self-determining system is an insulated system.

There is an interesting question about the boundaries of self-determining phenomena. Where do states, governments, armies, businesses, classes, biological organisms end, and where does the environment begin? (In your own case, the infinite space of your own consciousness makes it clear that you do not end where the outer surface of your body ends.) There is a need for a Quantum History or a Quantum Social Science which will investigate the boundaries and the behaviour of these systems, deploying models taken from quantum physics. There is a clear case for Quantum Psychology, because: 'In the account of psychical experiences, we met conditions of observation and corresponding means of expression still further removed from the terminology of physics.' (Bohr, 1958: 92) As Bohr says:

'The question has become so acute as to whether a comparison between the organisms and highly complex and refined systems, such as modern industrial constructions or electronic calculating machines, offers the proper basis for an objective description of the self-regulating entities which living organisms present.' (Bohr, 1958: 75)

The question is even more acute now than when Bohr raised it in the 1950s. Human beings and human organisations come under the heading of what Bohr is calling 'self-regulating entities'. Something is self regulating if it (at least partly) makes itself be what it is. Anything self-regulating is self-determining and probabilistic. It is insulated from the environment to a degree which precludes its being determined in the way it determines itself. Bohr says 'the problem is how far we can speak about freedom to act according to our possibilities.' (Bohr, 1958: 78) Does free will admit of degrees?

Sartre caused controversy after World War II by saying that no people had been more free than French resistance fighters who had been captured by the Nazis and were being interrogated. How should we understand this? Sartre draws an important distinction between freedom and power, and possessing freedom is not necessarily comfortable. The prisoners had very little power but they remained hideously free: to betray their comrades, or to bear the pain for a second longer. Freedom is the capacity to choose. Power is the availability to a

person of this rather than that choice. Your power varies but your freedom is absolute. You are not, therfore, able to do anything whatsoever. You, nevertheless, remain a choosing being, a chooser, even when your power is reduced. Otherwise it would make no sense to talk about reducing or removing your power. Power presupposes freedom. Indeed, it is because we remain choosing beings that imprisonment is a punishment. If we had no free will political freedom could not matter to us. (Sartre)

By 'our possibilities' Bohr means the set of options within our power. As a sub atomic particle takes one route rather than another because of certain unactualized possibilities, so human beings adopt one course of action rather than another in the light of forseen but unrealized outcomes. As Bohr says: 'As long as unrestricted deterministic views are taken , the idea of such freedom is of course excluded.' (Bohr, 1958:. 78) Freedom and determinism are mutually exclusive. compatibilism is false. Bohr says:

'However, the general lesson of atomic physics, and in particular of the limited scope of mechanistic description of biological phenomena, suggests that the ability of organisms to adjust themselves to environment includes the power of selecting the most appropriate way to this purpose.' (Bohr, 1958: 78)

Not just the quantum world, and the human world, but biology presuppose freedom and purpose. Quantum Biology is needed because of 'the insufficiency of [the] physical approach in accounting for the characteristics of life.' (Bohr, 1958: 95)

Bohr thinks that models of explanation radically removed from Newtonian mechanics are necessary to understand the distinctively human. He advocates something very close to phenomenology, or hermeneutics:

'The decisive point is that, if we attempt to predict what another person will decide to do in a given situation, not only must we strive to know his whole background, including the story of his life in all respects which may have contributed to form his character, but we must realize that what we are ultimately aiming at is to put ourselves in his place.' (Bohr, 1958: 78)

If we could know (by some criterion of completeness) someone's entire life story up to a certain moment, it does not follow that we could predict with certainty what they will to do next. No matter what we predict, they might choose something else. People are choosers. From a classical point of view, a person is a *singularity*. If we *are* able to predict someone's behaviour (and sometimes, of course, we do) it does not follow that their action was causally

determined, and they had no freedom of choice. It could be that we have simply become acquainted with the patterns in which they exercise their freedom. In prdicting, we extrapolate inductively from these patterns. A person is not a Newtonian mechanism.

If (by some criterion of success) we succed in 'putting ourselves in another person's place', ie, accurately imagining being them, even then, we cannot predict with certaintly what they will do. Even if we knew someone's intentions up to the last moment, they might act otherwise at the last moment.

Bohr is right to imply that we cannot understand human action 'from the outside'. On one level, being human, our only option is to understand the human from a human standpoint: from within. But this is not what Bohr means. He means to understand why a person acts we have to know *what it is like to be* that person. Phenomenology, hermeneutics and good literature all involve attempts to reconstruct what it is like to be another person. (Tolstoy, it seems to me, knows what it is like to be a dog.) Nevertheless, this unpredictability even extends to *being* a person. I cannot, with certainty, predict my own behaviour. I might change my mind, or be subject to weakness of will, or be distracted, in ways I do not forsee.

Recognizing the reality of free will is essential to understanding the human world. Bohr says:

'we may both practically and logically speak of freedom of will in a way which leaves the proper latitude for the use of words like responsibility and hope, which are themselves as little definable separately as other words indispensable to human communication. (Bohr, 1958: 78)

Freedom, responsibility and hope do not exist in isolation from one another. If I am *responsible* for what I do, that carries the implication that I act freely. It can hardly be right to hold people responsible for what is inevitable or happens inspite of, or against, their will. Finding someone guilty in law, or morally reprehensible in their behaviour, implies that they could have *refrained* from what they did. In other words, it implies that they acted freely. 'They chose' entails 'they could have chosen otherwise'.

Hope and responsibility presuppose free will. Hope and responsibility are necessary for human communication. Therefore free will is necessary for human communication. There is an argument for free will here: Free will is necessary for human communication. There is human communication. Therefore there is free will. Notice that Bohr does not say that mere *belief* in freedom, hope and responsibility are necessary for communication. It is the realities which are prerequisite.

Bohr's ideas about quantum physics have yet to be assimilated by the intellectual community. When they are, Eddington's forcast might well prove apposite:

'If our expectation should prove well founded that 1927 has seen the final overthrow of strict causality by Heisenberg, Bohr, Born, and others, the year will certainly rank as one of the greatest epochs in the development of scientific philosophy.' Sir Arthur Eddington (QQ: 204)

Schrödinger on Free Will

Schrödinger calls the problem of free will and determinism an 'impasse' and says 'It is a horrible antinomy.' (Schrödinger, 1958: 43) If there is free will, then I could refrain from acting: *not* do what I do. If there is determinism then I have to act: I *cannot not* do what I do. Schrödinger insists that we have free will. He asks rhetorically: 'Are we [thus] not the doers of our deeds?' (Schrödinger, 1958: 43) He is alluding to the distinction between what is *done* and what only *happens*. Everything that is done happens but not everything that happens is done. *Doing* something presupposes the possibility of not doing it, that is, free will. Even being deliberately compelled to do something presupposes the possibility of not doing it, otherwise the compulsion would be redundant. Schrödinger says about our actions: 'we feel responsible for them, we are punished or praised for them, as the case may be.' (Schrödinger, 1958: 43)

From the fact that we *feel* responsible for our actions it does not follow that we *are* responsible for them, even though the feeling would then have to be explained. But if we *are* responsible for our actions then we have free will. No one is responsible for what they could not refrain from doing.

Schrödinger is right to think the problem is not only apparent: 'The impasse *is* an impasse.' (Schrödinger, 1958: 43) If we combine free will and determinism we obtain a contradiction: 'I could have acted otherwise' is both true and false.

Schrödinger offers us a solution. Science can drop determinism. Indeed, it was a mistake all along to think that the truth of science entails or otherwise presupposes determinism. Determinism is entailed by many philosophies but need not be assumed in doing science. The view that determinism is scientific belongs to an out of date philosophical picture of science. Quantum science is indeterministic.

Schrödinger sees the problem of free will and determinism as bound up with the problem of mental causation, and the exclusion of consciousness by the old science:

'I maintain that it [the problem of mental causation] cannot be solved on the level of present day science which is still entirely engulfed in the 'exclusion principle' – without knowing it – hence the antinomy.'
(Schrödinger, 1958: 43-4)

The problem of mental causation is: *How can mental events cause physical events?* My *decision* to fire the missile is one of the causes of the firing of the missile, as is my *wish* that it will hit the target, and my *belief* that it will. Of course, the physical state of my brain has a causal role in my finger pressing the button. But how can intentions, decisions etc cause brain states? *How does consciousness push atoms around?* If it does not, then nobody ever *does* anything. Denying mental causation presupposes mental causation. If we give up mental causation, then life becomes unintelligible.

The exclusion of consciousness by the old science follows from its construal of everything as *other*: Something only exists if it is colourless, silent matter in motion in empty space. This third-person perspective completely suppresses everything supplied to the universe by your own existence. It is absurd. Future generations will find it incredible that it was regarded as in any way 'scientific'.

The quantum solution to the problem of free will and determinism entails that *any event that happens could have not happened*. The universe contains *contingency*. Contingency is the opposite of necessity. Anything contingently the case might not have been the case. Anything necessarily the case has to be the case. Determinism presupposes necessity. In the universe, there is only contingency so no necessity. Therefore, there is room in the universe for free will but not determinism. Contingency is necessary but not sufficient for free will. We also need: Actions are caused by the agent's choices. Always and everywhere, if I choose to do *x* then I could have chosen to refrain from doing *x*.

Free Will

de Broglie says

‘We have also had to abandon the traditional idea that phenomena, even elementary ones, are rigorously determined and exactly predictable, and to substitute for the rigid determinism of classical physics a more flexible conception, admitting that there exist at each instant in the evolution of elementary phenomena verifiable by us different eventualities concerning which it is only possible to estimate the relative probabilities.’ de Broglie (QQ: 119)

Three notions need separating here: ‘elementary’, ‘determined’, and ‘predictable’. By ‘elementary’ phenomena, de Broglie means sub-atomic constituents which, so far as we can tell, have no constituents. They are fundamental because they make atoms possible, and therefore make the macroscopic physical universe possible, but, on the other hand, they have no internal prerequisites detectable by contemporary physics. By ‘determined’ de Broglie means ‘necessitated by its causes’. If anything is determined then, if its cause happens, it cannot not happen. By ‘exactly predictable’ de Broglie means: that some phenomenon will occur can be known, with certainty, before that phenomenon occurs.

It is normally assumed that determinism entails predictability and *vice versa*. However, the links between these two notions are quite loose. From the fact that an event is necessitated by its causes it does not follow that we can predict it. We need, as an extra premise, that we have adequate knowledge of those causes. But even then, if the knowledge is enough to facilitate the prediction, it does not have to be true that the event is necessitated. It is the knowing that counts. Conversely, from the fact that an event occurs freely, or indeterministically, it does not follow that we cannot predict it probabilistically or statistically. An indeterministic event might not be wholly random. We might become acquainted with patterns in which freedom is exercised and predict those events which continue, or complete, such a pattern.

de Broglie says quantum mechanics reveals both indeterminism and the failure of exact predictability, at the elementary level. So far as we can tell, using the physics available to us, the universe is ultimately indeterministic. But caution is needed on three points.

Determinism may be retained as a theory, and exact predictions made, on a Newtonian and, by and large, Einsteinian level. In so far as the universe is neither too small nor too large to observe directly, and allows a seemingly clean separation of observer and observed, determinism and exact predictability are safe. Kant was right to say, in the *Critique of Pure Reason*, that the universe which appears to us empirically is deterministic, the hidden universe subject to

freedom, even though quantum phenomena are not strictly 'things-in-themselves' in the Kantian sense, because we can know quite a lot about them. (Kant, 19 .)

Secondly, although we have absolutely no evidence of this at present, it is theoretically possible that the random behaviour of the sub-atomic constituents de Broglie is calling 'elementary' is in fact determined at a still more fundamental, non-physical, level. Their randomness would then be only apparent. Should this turn out to be the case, we would of course have to give up calling de Broglie's phenomena 'elementary'. At the moment, however, it looks as though determinism, and exact predictability, are properties only of *medium sized* physical objects and events.

Thirdly, a huge gap in our understanding is created by the discovery of quantum indeterminism. If the behaviour of the constituents of the atom is random, and only statistically predictable, how can a deterministic and exactly predictable macroscopic universe result? As things stand, despite theories of 'de-coherence', physics has no answer to this question. We do not have a single theory which unifies the Einsteinian theories of relativity with quantum mechanics. Kant would have argued that *the subject*, or the observer, has been left out of the macroscopic picture. If we understood the contribution of consciousness to the appearance of the physical universe we could understand it as relatively enduring. But even a Kantian subject is insufficient to account for a universe unperceived by us. If the universe as a whole presupposes consciousness, it presupposes global consciousness.

de Broglie says

'It seems […] that determinism can no longer be proved by human science, and if we absolutely persist in maintaining its principle, this can only be done as a metaphysical postulate.' (de Broglie 1955: 243)

In fact, determinism was never proven by science. Determinism, like materialism, was always a *philosophical* theory and never a scientific or empirical theory. The precise predictions made within a Newtonian system do not entail that the system is deterministic. Determinism is the thesis that causes *necessitate* their effects: If the cause happens, the effect cannot not happen. As Hume saw, this necessity cannot be detected empirically. Precise prediction is consistent with mere *constant correlation* of event types and so consistent with the falsity of determinism. de Broglie's 'no longer' is therefore inappropriate. Determinism was always a metaphysical postulate.

de Broglie speaks of 'the necessity of replacing the rigorous determinism of the old mechanical or physical theories by a "probability"'. (de Broglie 1955:

240) Prediction is still possible in quantum physics but probabilistic or statistical prediction. For example, in the two-slit experiment, we can still predict the statistical distribution of the protons striking the plate. But as ‘it is impossible to attribute a rigorous determinism to the succession of phenomena, at least on the corpuscular scale.’ (de Broglie 1955: 241) so ‘The apparent determinism in the macroscopic scales must, in small scales, give place to a probability’(de Broglie 1955: 241) In quantum physics we ‘calculate possible eventualities and their respective probabilities.’ (de Broglie 1955: 241) Probabilistic prediction is based on indeterminism.

Although there is no determinism in quantum physics, it would be a mistake to conclude that there is no causality. From ‘causes do not necessitate their effects’ it does not follow that ‘there is no causation’. As de Broglie sees, there are causal relations between quantum phenomena which fall short of deterministic:

‘there certainly exists in quantal physics a “weak” causality in the sense that every effect always has a cause and that the suppression of the cause always involves the disappearance of the effect; but no longer do we succeed in finding there the “strong” causality where the effect necessarily results from the cause and is connected with it by a rigorous determinism.’ (de Broglie 1955: 241-2)

To understand this, we need to draw a distinction between strong and weak causality. In strong causality, if the cause happens then the effect cannot not happen. In weak causality, if the cause happens, the probability of the effect happening is raised (higher than it would be if the cause had not happened) but to some degree short of certainty. de Broglie says:

[1] ‘Weak causality allows us to suppose that the same cause can produce one or other of several possible effects’(de Broglie 1955: 242)

[2] There is ‘only a certain probability that such an effect with be produced and not such another.’ (de Broglie 1955: 242)

One and the same cause is consistent with different possible effects. de Broglie says

‘Physicists no longer succeed in finding anything except this weak causality on the very small scale; it is for the philosophers to examine whether they can be satisfied with it.’ (de Broglie 1955: 242)

Here de Broglie is saying it is a philosophical question whether there really can be objective indeterminism. On the one hand: 'it [quantum physics] does not permit them to be explained as resulting from a causal evolution which works at a still deeper level in the physical world.' Louis de Broglie 'Forward' in David Bohm *Causality and Chance in Modern Physics* (Routledge, London, 1984) p. xiii On the other hand:

'It is possible that looking into the future to a deeper level of physical reality we will be able to interpret the laws of probability and quantum physics as being the statistical results of the development of completely determined values of variables which are at present hidden from us.' (de Broglie, 1984: xiv)

'physical world' or 'physical reality' is the difficulty here. If there is a hidden, deterministic, ground of quantum phenomena it cannot be physical. The quantum world makes the physical world possible. Physical objects are space-time continuants. Space time continuants are, and are constituted by, events. But events that have a beginning, end, and duration, are made of events with no duration. At this level of description we have to give up using the word 'physical'. Quanta are event-like 'packets' of energy, which come in and out of existence. If there is some deterministic cause of this generation and destruction, on pain of circularity, it is not any part of the physical world. Whatever it is has to *explain* the physical world.

de Broglie says 'it [quantum physics] provides no more than laws of probability.' (de Broglie, 1984: xiii) and 'it considers these laws of probability as having a primary character and constituting the ultimate knowable reality.' (de Broglie, 1984: xiii) When de Broglie says the laws of probability have a 'primary' character he means they do no seem to be derived from anything else. They are not 'secondary' (or 'tertiary', and so on) in that they do not depend on a determining cause (which might depend on a further determining cause, and so on). We have to draw a sharp distinction between

(1) ultimate reality

and

(2) ultimate knowable reality.

The Vienna Circle thought the first of these always collapses into the second, because there is no conceptual space in which what is the case could meaningfully exceed what we could, in principle, find out to be the case.

However, meaning is not closely tied to the possibility of verification, not all concepts have to be rooted in empirical concepts, and it would be an astonishing coincidence if reality only contained what human beings could find out about it. 'Verifiable' invites the question 'verifiable for whom?', verifiable for us? (Who are we?), verifiable for some computer? verifiable by consciousness? verifiable by anyone?

If we carry on asking questions where Hume, Kant and the Vienna Circle would have us give up, we can ask: What is ultimate reality? What causes the 'spontaneous' quantum events? Why is there a quantum world? It is by pushing on in this way that human knowledge increases, not by deciding in advance that the questions are unanswerable.

de Broglie says

'Those who have studied the development of modern physics know that the progress of our knowledge of microphysical phenomena has led them to adopt in their theoretical interpretation of these phenomena an entirely different attitude to that of classical physics.' (de Broglie, 1984: xiii)

de Broglie emphasizes that Newtonian and Relativistic laws do not apply at the quantum level. He says

'Whereas with [classical physics] it was possible to describe the course of natural events as evolving in accordance with causality in the framework of space and time (or relativistic space-time), and thus to present clear and precise models to the physicist's imagination, quantum physics at present prevents any representations of this type and makes them quite impossible.'
(de Broglie, 1984: xiii)

de Broglie goes so far as to say

'it [quantum physics] allows no more than theories based on purely abstract formulae, discrediting the idea of a causal evolution of atomic and corpuscular phenomena' (de Broglie, 1984: xiii)

Quanta are not 'causal' if 'causal' means 'subject to determinants' or 'obtaining because of causally sufficient conditions'. Quanta are not 'atomic' if 'atomic' means 'not subject to division' or 'simple'. Quanta are not 'corpuscular' if this means 'physical', or 'possessing shape, size, definite location and speed, and mass'. The mathematics works but does not depict anything that falls under these descriptions. On the other hand, even if there are no truly atomic

phenomena, de Broglie's statement does not rule out deterministic relations between physical objects on the macroscopic scale. We are left with the difficult question of why, if the world is indeterministic and non-physical, it should appear to us determined and physical, on a day to day basis. This is hard because, for example, the functioning human nervous system is part of the macroscopic world which needs to be explained, and so cannot be invoked as its explanation, without circularity.

Interestingly, de Broglie sees the possibility of extending the models of quantum mechanics to biology and philosophy. He says

'the impossibility of establishing on this [atomic] scale a rigorous determinism will […] appear as capable of playing an essential role in the evolution of our ideas on all that relates to life, and will perhaps even bring new elements to many of the traditional problems of philosophy.' (de Broglie 1955: 242)

de Broglie thinks

'it would […] be as impossible to describe life solely by the physico-chemical processes as it is impossible in atomic physics to describe the corpuscle soley by the wave or *vice versa*.' (de Broglie 1955: 241)

Life is not only atoms in motion. Life is not only organic molecules in their mutual relations. If this were the case, we could dispense with the science called 'biology' altogether and proceed with physics and chemistry. In fact, we need not only bio-chemistry but a biology which brings with it the concepts of nutrition, growth, reproduction, adaption, mutation, information, and goal directed behaviour. Descriptions only at the level of the bonding of molecules, or the motion of atoms, leaves out distinctively biological explanation. The essence of biology is not its physical ground.

'The point is that the photon is in a superposition of states going one way or another, the detector is in a superposition of states of being off and on, and the cat is in a superposition of states of being alive and dead.'
Roger Penrose *et. al. The Large, the Small and the Human Mind* (Oxford University Press, Oxford, 1997) p. 72

‘It is quite extraordinary that quantum mechanics enables you to test whether something might have happened but didn’t happen. It tests what philosophers call counterfactuals. It is remarkable that quantum mechanics allows real effects to result from counterfactuals!’
Roger Penrose *et. al. The Large, the Small and the Human Mind* (Oxford University Press, Oxford, 1997) p. 67

Hoyle on Free Will

Hoyle says ‘I do not think we should banish free will in what is really a facile way’ (Hoyle, 1983: 205) There is no good argument against the existence of free will, yet we are confronted with the reality of choice in our everyday lives. Determinism, which is inconsistent with free will, has been refuted by quantum physics. As Hoyle mentions ‘[…] results of a predictable or deterministic kind in which one large scale event was said to be the cause of another.’ (Hoyle, 1983: 198) but then says ‘On an atomic scale things were different, however, because the usual concept of cause and effect dissolved into indeterminacy.’ (Hoyle, 1983: 198) In quantum physics, determinist models have to be replaced by indeterminist models. However, indeterminism, although necessary for free will, is not sufficient. Controversially and radically, Hoyle gives reason to suppose that quantum physics not only shows there is indeterminism but that there is free will. We apply quantum models to the brain. Hoyle says

‘Imagine the quantum event in queston being repeated many times under identical conditions […] It is evident that such a sequence could carry a message, it could carry information. Suppose our brains contain a quantum “experiment”, an experiment repeated many times under similar initial conditions, each with the equivalent of a dot or a dash as its result. The outcome could be a potential message available for permanent storage in the memory, ready to be acted upon, an injection of information that could form the basis of the behaviour that we call free will.’ (Hoyle, 1983: 205)

In the case of quantum experiments, *the same initial conditions can give rise to different results*. That the same conditions produce the same results was thought, for centuries, to be the hallmark of science. We now know it is only a hallmark of Newtonian science, science that does not hold at the quantum level. In the human case also, the same initial conditions can give rise to different results. Human beings are quantum beings.

Hoyle points to a profound puzzle

‘if quantum sequences are ordered in this way, a profound difficulty, with which it is otherwise hard to cope can be resolved. What is it that distinguishes our animate selves from inanimate objects? Certainly not the individual atoms of which we are built. There is no difference between the carbon atoms in a limestone cliff and the carbon atoms in our bodies.’ (Hoyle, 1983: 205-7)

This is exactly the right question to ask. A rock or a pebble has no free will but a human being does. However, rocks, pebbles and human bodies all have a quantum constitution. *What is it about the human case which makes possible free will?* Similarly, we could ask *What is it that distinguishes conscious beings from non-conscious beings?* Human beings and non-conscious physical objects both have a quantum constitution. *What is it about the human case which makes possible consciousness?*

Hoyle sees an intimate connection between free will and consciousness. He says ‘it is the process of recognition of such sequences that constitutes the phenomenon of consciousness.’ (Hoyle, 1983: 207) Hoyle’s ‘recognition of’ is doing a lot of work here. Nevertheless, Hoyle might well be right. Consciousness of what are, in fact, quantum processes might help constitute the self-consciousness necessary for free will; the ‘space of choices’.

III

Free Will and Empiricism

Hobbes draws a distinction between two kinds of motion: voluntary and involuntary. The vital processes of an organism, such as the circulation of the blood, respiration and digestion, count as involuntary because ‘there needs no help of Imagination’ (L 118). An action is voluntary if, and only if, it is caused by the making of a choice which, for Hobbes, is imaginative deliberation. Voluntary motion includes speaking and the movement of limbs. Hobbes might have pointed out (but does not) that breathing may be either voluntary or involuntary and so cuts across his taxonomy. I mean, I may breathe if I choose, or refrain from breathing (briefly); but if I do not choose to breathe, breathing continues in any case through causes other than my own choice. Voluntary actions are caused by the operations of the imagination: ‘the Imagination is the first internall beginning of all Voluntary Motion’ (L 118). If someone is walking, it makes sense to ask: Where? If they are speaking, it makes sense to ask: What? If they move their limbs, it makes sense to ask: Which way? Indeed, the thought of such options precedes and causes the voluntary action, and these thoughts are themselves

motions. It is because thoughts are motions that they may cause actions. Actions are motions and only motions may cause motions. Also, if thoughts cause actions and actions are motions and if only motions may cause motions, then it follows that thoughts are motions. Hobbes's theory of the will is thus logically consistent with his materialism. We must not be misled into thinking that thoughts are not motions because they are invisible. There may be motion that is invisible because it is too minute to be perceptible.34

The thought that causes an action is a trying, an attempt:

'These small beginnings of Motion, within the body of Man, before they appear in walking, speaking, striking, and other visible actions, are commonly called Endeavour' (L 119)

Endeavour may be described in two ways, depending on its relation to the object which causes it. Appetite is the endeavour to obtain something, the desire for it. Aversion is the endeavour to avoid something, the aversion to it. Hobbes assumes the object of an endeavour is always and everywhere identical with the cause of that endeavour and therefore he does not distinguish qualitatively or numerically between the intentional object and the cause of trying.

So, are we free? Hobbes has a particular account of what the will consists in which makes use of the appetite/aversion distinction. The oscillation of appetites and aversions and hopes and fears is called by Hobbes 'Deliberation' (L 127): 'it is called Deliberation; because it is a putting an end to the Liberty we had of doing, or omitting, according to our own Appetite, or Aversion' (L 127). So the liberty to do something partly consists in not having done it and partly in there existing the possibility of doing it. De-liberation is the process that ends this freedom by causing the action or preventing it. Appetites cause actions. Aversions cause omissions. Alternatively, deliberation may end in the thought that the action is impossible (and presumably if this thought is true there was no freedom to perform that action anyway).

So as long as there exists

'this alternate Succession of Appetites, Aversions, Hopes and Fears [… wee retain the liberty of doing, or omitting' (L 127)

By 'omitting', Hobbes means deliberately refraining from action. Arguably there is only freedom to act if there exists the possibility of refraining from acting and there is only freedom to refrain from acting if there is freedom to

act. If there was not freedom not to act then there would be compulsion to act, and that is incompatible with freedom to act. If there was not freedom to act then there would be compulsion not to act, and that is incompatible with freedom not to act.

The will is simply the final thought in the succession of appetites or aversions before the action or its omission: 'in Deliberation, the last Appetite, or Aversion, immediately adhering to the action, or to the omission thereof, is that wee call the Will' (L 127). Hobbes means an act of will, not the faculty of will, or the capacity to will. Our freedom thus consists in our being the causes of our own actions and omissions, or at least our being amongst those causes. A necessary condition of this freedom is the existence of the real possibility of having acted when one omitted to act, or having omitted to act when one acted.

Hobbes thinks the voluntary–involuntary distinction also applies to (non-human) animals. He says, 'Beasts also Deliberate' and 'Beasts that have Deliberation, must necessarily also have Will' (L 127). It follows that the will is not necessarily rational, because animals are not rational yet have a will.

Hobbes's account of freedom contrasts sharply with that of the Scholastics and in one respect with that of Descartes. The quasi-Aristotelian definition of the will offered by the Scholastics was 'Rationall Appetite' (L 127). Hobbes rejects this definition because it makes the irrational exercise of freedom a logical impossibility: 'then could there be no Voluntary Act against Reason' (L 127). Hobbes maintains that people do act freely but irrationally, so the Scholastic account is false and says instead, 'will … is the last Appetite in Deliberating' (L 128).

Hobbes disagrees with Descartes over the status of animal action. Descartes and Hobbes agree that non-human animals are merely highly complex physical objects with no immaterial souls.35 Hobbes, however, maintains, while Descartes denies, that animals have a will; in other words, animals' actions fall on the voluntary side of the voluntary–involuntary distinction. A problem for Hobbes is that in ascribing to animals a will he implicitly credits them with the capacity to make environmental discriminations and even exercise preferences, yet this requires a minimal rationality Hobbes wishes to deny them. It may be that reason and the will are not so readily separable as Hobbes supposes. After all, only a rational creature may act irrationally.

Hume

Hume tries to reconcile freedom and determinism; to show they are not incompatible after all.15 Hume argues that the issue is 'merely verbal' (E I VIII.73), and in common with certain twentieth-century linguistic philosophers argues that the problem will disappear once we have clarified 'liberty' and 'necessity'.

Liberty, or freedom, has to be distinguished from chance, and from something happening without any cause whatsoever. Actions are caused. They are caused by the 'motives, inclinations, and circumstances' of the agent (E I VIII.73). They do not occur by chance because they are performed deliberately and follow from their causes with a certain uniformity. Indeed, there is no chance in nature, according to Hume. If we say something happens by chance we are merely reporting our ignorance of some of its causes. So freedom is not chance for Hume, but is the ability to make choices and act upon them. One's freedom consists in being, or being amongst, the causes of one's own actions: 'By liberty, then, we can only mean a power of acting or not acting, according to the determinations of the will' (E I VIII.73). It follows that a person acts freely if, and only if, their action is at least partly caused by their choice to act, and they could have chosen not to act and not acted. So, 'a person did something freely' means they chose to do what they did and did it, but they could have chosen not to do what they did, and then not done it. In Hume's example: 'If we choose to remain at rest, we may; if we choose to move we also may' (E I VIII.73). According to Hume, liberty is properly contrasted with constraint rather than with determinism. To say that someone acted freely is to say that person was not forced to act in that way and could have either refrained from acting, or acted otherwise. As Hume puts it, 'This hypothetical liberty is universally allowed to belong to everyone who is not a prisoner and in chains' (E I VIII.73). He considers this account of freedom both internally consistent and consistent with our experience as agents.

Is it consistent with determinism? To decide this we must distinguish strong from weak determinism. Weak determinism is the theory that every event has a cause. Strong determinism includes that theory, but adds that causes necessitate their events such that every caused event (i.e. every effect) is inevitable. Hume's theory of action is compatible with weak determinism but incompatible with strong determinism. He accepts that human actions are caused, that they are caused by choices, so he does not thereby provide an exception to the thesis that every event has a cause. (Both kinds of determinist are going to want to know whether our choices have causes.) Hume must reject strong determinism, though, because, as we have seen (pp. 147–8 above), he rejects the theory that causes necessitate their effects. In

strong determinism, if A is the cause of B and A happens, then B not only happens but happens by necessity. If something happens by necessity, it not only happens but could not fail to happen. Hume rejects this whole picture of causation because, for him, causal connections are contingent, not necessary. There is no contradiction involved in the notions that an event could have had a cause other than its actual cause, or that a cause should have an effect other than its actual effect, or that event should be neither a cause nor an effect. Putative causal necessities are psychological expectations derived from the observation of regularities: 'Had not objects a regular conjunction with each other, we should never have entertained any notion of cause and effect' (EI VIII.74). Objective, perceiver-independent causality is only constantly conjoined events.

If nothing happens by necessity then it logically follows that no human actions happen by necessity, and it follows from that that they are not strongly determined. Thus strong determinism is false. Not only does Hume's epistemology of causation imply the falsity of strong determinism, but his theory of freedom is inconsistent with that same doctrine. Hume thinks we may act freely. When we do, we could have done other than what we did. But this can make no sense within the framework of strong determinism. If A causes B then A necessitates B, so B could not not happen given A. It makes no sense for there to be a kind of freedom which allows someone to possibly not do that which they do, nor, for that matter, to possibly do that which they do not do.

Although we perform our actions freely when not constrained, that does not imply that human actions are unpredictable:

A man who at noon leaves his purse full of gold on the pavement at Charing-Cross, may as well expect that it will fly away like a feather, as that he will find it untouched an hour after. (EI VIII.70)

We predict human actions, as we would predict any other kind of event, by observing regularities. We observe consistencies in the exercise of choices and the behaviour they cause. We are not acquainted with any causal conditions which necessitate human actions, so there is no basis for prediction there.

Hume fails to draw one important implication of his arguments. The fact that we are able to predict human actions does not show that we are not free. Causal connections are contingent, so a person may always have chosen and acted otherwise. Contrary to a widespread philosophical assumption, predictability does not imply determinism. We predict a person's behaviour by coming to know the ways in which he exercises his freedom.

Mill

Mill distinguishes libertarianism from determinism as follows:

'Liberty is the thesis that the will is not determined like other phenomena, by antecedents, but determines itself'

and determinism is

'necessity, as asserting human volitions and actions to be necessary and inevitable'.

Libertarianism essentially includes the doctrine that:

'our volitions are not, properly speaking, the effects of causes, or at least have no causes which they uniformly and implicitly obey'. (SL I. 413)

Determinism essentially includes the doctrine that:

'the law of causality applies in the strict sense to human actions as to other phenomena'. (SL I. 413)

Determinism, here, is therefore the twin theses that every event has a cause and causes necessitate their effects, including human actions. If a cause necessitates its effect, then if the cause happens the effect could not fail to happen. Libertarianism is, here, the view that the will is the cause of its own actions, and what is willed could have been otherwise. The will has no cause other than itself.

Clearly, as defined, libertarianism and determinism are radically incompatible. Determinism is the view that every event has a cause, but if this means 'cause other than itself' and if the operations of the will are un-caused (save by itself), both theories cannot be true. Similarly, if causes necessitate their effects, those effects could not fail to happen, given their causes. But the actions of the will must be able not to have happened even if they do happen, otherwise the will is not free. Part of freedom consists in the ability to do otherwise, and this implies the contingency denied in determinism.

Mill is a determinist. However, what he means by 'determinism' turns out to be something rather different from the definition of 'necessity' given

above. His determinism consists in the claim that if we knew enough about a person we could predict their actions with as much certainty as we could predict natural events:

The doctrine called Philosophical Necessity is simply this: that given the motives which are present to an individual mind, and given likewise the character and disposition of the individual, the manner in which he will act may be unerringly inferred; that if we knew the person thoroughly, and knew all the inducements which are acting upon him, we could foretell his conduct with as much certainty as we can predict any physical event. (SL I.414)

The problem with this as a kind of determinism is that from the fact that we are able to predict a person's behaviour it does not logically follow that that person is not free. From a very great knowledge of a person's motives and character we may predict their actions accurately, but this may only show that we are intimately acquainted with the ways in which they exercise their freedom. Notice that the claims of the above passage are logically independent of Mill's twofold definition of 'necessity'. The fact that we are able to predict human actions does not entail, and is not entailed by, either or both the theses that every event has a cause, and causes necessitate their effects.

Mill thinks his determinism is consistent with the experience but not the reality of freedom. I think Mill's predictivism is consistent with the experience and the reality of freedom.

III

Probabilistic Causation

6 CAUSATION

Hume's problem of causation is how an empirical account can be given of the origin of the idea of causation when that idea does not seem to be straightforwardly empirical. To see this, contrast the concept of causation with the concept of red. The concept of red is straightforwardly empirical. An impression of red causes an idea of red, and our idea of red is a faded copy of the impression that caused it. The idea of red represents the

impression of red. In the case of causation, however, this genetic account is less feasible:

When we look about us towards external objects, and consider the operation of causes, we are never able in a single instance, to discover any power or necessary connection, any quality, which binds the effect to the causes, and renders the one an infallible consequence of the other. (E I VII.50)

Our idea of A causing B is that A necessitates B, that A is related to B such that if A happens, B cannot fail to happen. Hume points out that the idea of this putative relation is not empirical. If A causes B, and if A and B are both perceived, then there is no additional perception to be had of the causal relation which obtains between A and B. No such relation is observed over and above the occurrence of A and the occurrence of B. It follows that there exists no impression of the causal relation. Causal relations are not perceived, so they are not empirical. Crucially, this is a putative case of an idea without an impression. There is only an impression of the string of events which we hold to be causally related, not of the causal relation itself. Still less is there any impression of a connection between cause and effect which would make the second the inevitable and infallible consequence of the first: 'We only find, that the one does, in fact, follow the other' (E I VII.50).

Therefore the idea of causation seems, prima facie, to violate the central principle of Hume's empiricism; there are no ideas without impressions. We certainly have an idea of causation, but if Hume is right we have no impression of causation, only of constantly conjoined events. We observe that events of one sort are regularly followed by events of another sort. But that is all. We do not perceive causal relations between them. Hume's task is to show that, nevertheless, the idea of causation has an empirical origin.

That there is no single impression of causation has several important consequences: 'From the first appearance of an object, we never can conjecture what effect will result from it' (E I VII.50). Hume means that no empirical scrutiny of an object will reveal to us what sorts of effects it will have, indeed we may not know what these effects are until we have witnessed the events which follow it chronologically. Knowledge of which causes bring about which effects is completely a posteriori for Hume, and not a priori: 'Were the power or energy of any cause discoverable by the mind, we could foresee the effect, even without experience' (E I VII.50). We need observation to discover causes and effects, so they are not knowable a priori.

Hume also thinks causal relations are contingent. Even if it is true that A causes B, it is logically possible that this should have been false. A could

have had a different effect. B might have had a different cause, B might not have been an effect and A might not have been a cause. Contingent facts are facts that could have been otherwise. If p is contingently true, then p is true, but it is logically possible that p should have been false. 'Contingent' contrasts semantically with 'necessary'. If p is necessarily true, then p is true, but it is not logically possible that p should have been false. Hume is saying that where p reports some causal relation, p is contingent, not necessary. All causal relations might not have been, and given that they are, they might not have been as they are.

So, from the fact that A exists it does not logically follow that A is a cause of B, and if B exists it does not necessarily follow that B is an effect of A. There is nothing intrinsic to an event or an object which could make it either an effect or a cause. The qualities of a physical object such as 'solidity, extension, motion … are all complete in themselves, and never point out any other event which may result from them' (E I VII.50).

All that is directly given through impressions of the outward senses is one event following another in time; 'we only learn by experience the frequent conjunction of objects, without being ever able to comprehend anything like a connection between them' (E I VII.54). It would seem to follow, if the meaning of a word is always derived from some impression, that the expression 'necessary connection' is meaningless (E I VII.58).

Hume provides a psychological account of our idea of causation, to explain its empirical origin. We perceive that many events of type A are regularly followed in time by many events of type B, and the repetitive nature of this process leads us to suppose that a type A event will be followed by a type B. This happens so often that it is psychologically almost impossible for us not to expect a B after an A. This expectation is conditioned into us by our experience of regularities: 'After a repetition of similar instances the mind is carried by habit, upon the appearance of one event, to expect its usual attendant, and to believe that it will exist' (E I VII.59). The idea of necessary connection, or causal efficacy, or causal inevitability, is only the impression of expectation that the second event will follow the first. We mistake the subjective expectation for an objective necessity, but really there is no impression of necessary connection, only of our own expectation. For Hume, the universe is thoroughly contingent, not necessary. That there is a universe at all is a contingent fact, and it is a contingent fact about it that what happens in it does. Although it is ingrained in our psychological habits to foist necessities, forces and inevitabilities on the universe, in empirical reality.

the scenes of the universe are continually shifting, and one object follows another in an uninterrupted succession; but the power or force, which actuates the whole universe, is entirely concealed from us, and never discovers itself in any of the sensible qualities of body. (E I VII.50)

It is possible to analyse the concept of causation more thoroughly than Hume has done. For example, suppose A is the cause of B, then this could mean that A is a necessary condition for B ('that without which not'), so that unless A happens, B does not happen. Or, again, suppose A is the cause of B, then this could mean that if A happens then B happens, perhaps not with logical necessity but, nevertheless, always and everywhere. Perhaps the set of necessary conditions for B, which will be very large, are jointly sufficient for B. It is possible to develop sophisticated accounts of causation on these lines, and, further, to ascribe real causal powers to mind-independent physical objects. It could be, conversely, that Hume's scepticism about the causal relation is an important first step to the abolition of that concept. Hume does not say this, but perhaps there is no causation.10

KANT'S THEORY OF FREE WILL

A QUANTUM DEFENCE

Stephen Priest

Stephen Priest is Senior Research Fellow in Philosophy at the University of Oxford.

Stephen Priest
Faculty of Philosophy
University of Oxford
Woodstock Road
OXFORD OX
ENGLAND UK
01865 426625
stephenpriest99@hotmail.com

Also by Stephen Priest:

The British Empiricists
Theories of the Mind
Merleau-Ponty
The Subject in Question
The God of the Physicists

Hegel's Critique of Kant (ed.)
Jean-Paul Sartre: Basic Writings (ed.)
A Dictionary of Philosophy (ed. with Antony Flew)

Preface

In order to solve the problem of free will and determinism we have to pay far more attention to physics than previously. Quantum physics presupposes indeterminism and, I shall argue in this book, free will.

Stephen Priest Oxford 2017

Acknowledgements

I am grateful to the following for conversations I have had over the years: Sir A. J. Ayer, Renford Bambrough, Hazel Barnes, Graham Bird, Daniel V. Came, Edward Craig, Antony Flew, Lazar Fotev, Benedikt Göcke, P. M. S. Hacker, R.M. Hare, Michael Inwood, Sir Anthony Kenny, Fergus Kerr, Saul Kripke, John Leslie, David Lewis, Grahame Lock, Bryan Magee, A. W. Moore, Thomas Nagel, Alexander Norman, G. E. L. Owen, Robert

Nozick, Terence Penelhum, Sir Karl Popper, Howard Robinson, Roger Scruton, Shogo Shimizu, Kerry Smallman, Richard Swinburne, W. H. Walsh, Bernard Williams, and Timothy Williamson. Some of these people disagree with me profoundly. The world is much poorer for the fact that some of them are dead.

Daniel Came organised a one day conference on an early draft of this book in the Faculty of Philosophy of the University of Oxford. I thank him, Benedikt Göcke, Michael Inwood and Alexander Norman for reading papers about my work that day, and Stephen Mulhall for his comments. Michio Andho organised a one day conference on the manuscript at Keio University, Japan. I thank Kenji Murota and Michio Andoh for their papers and the audience for their comments. I thank Masataka Murama of Hokkaido University and Hitoshi Nagai of Nihon University, Tokyo, and Grahame Lock and the University of Nijmegen for organising seminars on my work. I thank Lazar Fotev and the Faculty of Philosophy at Skopje University, Republic of Macedonia, for their hospitality and for useful discussions. I thank Klaus Müller and Benedikt Göcke for a one day meeting on the ideas in this book at the Wilhelms Universität, Münster, Rheinland-Westfalen.

I am grateful to the Master of Blackfriars Hall, Oxford, and the Fellows and Dominican bretheren, the Principal and Fellows of Wolfson College, Oxford, and the members of the Faculty of Philosophy of the University of Oxford for environments conducive to thought.

Stephen Priest

Contents

I The Possibility of Free Will

The Third Antinomy……..….
Thesis ……..…
Antithesis ……....
An Ambiguity ……....

Proof of the Thesis ……..…
Kant's Observation on the Thesis ……...
Individuating Causal Chains……..
Proof of the Antithesis
The Argument from the Unity of Experience……....
The Empirical Argument……..….
The Rules Argument……..…
Practical Freedom……..…
Transcendental and Practical Freedom……..…
(a)The Argument for Practical Freedom……...
(b) The Argument from Practical to
Transcendental Freedom……...
My Freedom……..…..
Allison, Guyer, Irwin……..…....
Freedom and Self-Knowledge……..…..
Freedom and Pure Apperception……
Bennett on the Third Antinomy…….
Reason and Freedom…….
(a) The Argument for the Freedom of Reason…….….
(b) The Argument that Reason is Not Determined……

II The Agent

Notes

Bibliography

The Third Antinomy

I clarify the role of the concept of an uncaused cause in Kant's putative solution to the problem of free will and determinism in the Antinomies chapter of *The Critique of Pure Reason* and offer some comments.

Kant's most concise yet clearest definition of 'freedom' is the statement that freedom is 'self-activity' (*Selbsttätigkeit*) (A417, B446). I shall understand some action or event, '*e*', to be free on this account if and only if *e* is either the cause of itself, or the cause of some further action or event *e* ' but is not itself the effect of any cause. So, on either construal, *e* has an effect, but on the first its effect is not other than itself, but on the second its effect is other than itself. On the first construal, the event has a cause but that cause is only itself but, on the second construal, the event has no cause. By 'effect' I mean 'caused event' and by 'cause' I mean 'event with an effect'. (I assume that actions are a species of event.) This is a sharpening of Kant's claim (A417, B446-7) that 'die unbedingte Kausalität […] in der Erscheinung 'heißt' die Freiheit', 'unconditioned causality in [the field of] appearance is called freedom'. (Kemp Smith's 'field of' translates nothing in the German here, and the definite article is best omitted in the English.)

Kant says 'Da heißt nun die Bedingung von dem, was geschieht, die Ursache', 'the condition of that which happens is entitled the cause'. Kant does not say what sort of condition he has in mind. He might mean 'necessary condition' such that, if *e* is a 'condition' of *e*:

$-c \rightarrow -e$

or he might mean 'sufficient condition' such that:

$c \rightarrow e$

or he might mean 'both necessary and sufficient condition'. Kant's use of 'the' in 'the condition' does not logically guarantee which condition. He might be using 'the' in a generic sense, roughly equivalent to 'any', so that

$(\forall c, \forall e)\ -c \rightarrow -e$

or

$(\forall c, \forall e)\ c \rightarrow e$

However, I shall read 'condition' here to mean 'both necessary and sufficient condition'. So, some event *e* is a caused event if and only if there exists, or existed, at least one other event *c* such that if *c* failed to occur then *e* could not occur, and if *c* does occur then *e* could not fail to occur:

$(\forall c, \forall e)\ ((\text{-}c \rightarrow \text{-}\ e)\ \&\ (c \rightarrow e))$

Kant's discussions of freedom and determinism presuppose an ontology of events. This is implied, for example, by his use of 'that which happens' (*was geschieht*) (for example at A417, B446). I shall follow Kant in this, but it obviously needs to be born in mind that some necessary or sufficient condition for an event's occurrence might not itself be an event.

Kant thinks 'I am free' is true only if: There is at least one respect in which I am the cause of some event, and at least one respect in which I am not the effect of any event. This provides, so far, a necessary but not a sufficient condition of my freedom because my being the cause of some event might itself be caused, if the respect in which I am uncaused does not preclude this. So, we could try: I am free if and only if there is at least one respect in which I am the cause of some event, and at least one respect in which I am not the effect of any event, and the respect in which I am the cause of an event is just the respect in which I am uncaused. This precludes my being caused to be the cause of at least one of the events of which I am the cause.

On the other hand, I am not free but wholly determined if and only if, even if I am the cause of at least one event: I am caused to be the cause of any event of which I am the cause. (Here 'caused' has to be taken to imply the obtaining of a sufficient condition, or the sentence might not be a definition of 'determinism'.) With these definitions, and criteria for analyzing ‘free’, in mind we may turn to the third antinomy. I state the thesis and antithesis, and then examine the putative proof of each:

Thesis

'Die Kausalität nach Gesetzen der Natur ist nicht die einzige, aus welcher die Erscheinungen der Welt insgesamt abgeleitet werden können. Es ist noch eine Kausalität durch Freiheit zu Erklärung derselben anzunehmen notwendig.' (CPR 409, A445, B473)

'Causality in accordance with laws of nature is not the only causality from which the appearances of the world can one and all be derived. To explain these appearances it is necessary to assume that there is also another causality, that of freedom'

Antithesis

'Es ist keine Freiheit, sondern alles in der Welt geschieht lediglich nach Gesetzen der Natur.' (CPR 409, A445, B473)

'There is no freedom; everything in the world takes place solely in accordance with the laws of nature.'

In understanding both proofs, an important ambiguity needs to be dispelled. Kant is appraising the claim that there exists a kind of freedom from which 'the world of appearances can one and all be derived'. 'die Erscheinungen der Welt insgesamt abgeleitet werden können'. This could mean either an uncaused event which is the initiator of every subsequent causal chain, or more narrowly, an uncaused event which is the initiator of any member of the set of all causal chains. The first sort of freedom putatively pertains to God, or the cause of what exists. The second sort putatively pertains to finite persons, or performers of causally efficacious actions within the world. In what follows, I shall confine the discussion entirely to the question of the freedom of (human) persons and ignore any passages which can only be given the first interpretation. This is because we are only concerned with the freedom of the self as human agent, here, and this is Kant's own expressed concern in writing the Third Antinomy. At CPR 414, A450, B478, he explains:

'Denn wir reden hier nicht vom absoluten ersten Anfange der Zeit nach, sondern der Kausalität nach. Wenn ich jetzt (zum Beispiel) völlig frei, und ohne den notwendig bestimmenden Einfluß der Naturursachen, von meinem Stuhle aufstehe, so fängt in dieser Begebenheit, samt deren natürlichen Folgen ins Unendliche, eine neue Reihe schlechthin an, obgleich der Zeit nach diese Begebenheit nur die Fortsetzung einer vorhergehenden Reihe ist.'

"[...] the absolute first beginning of which we are here speaking is not a beginning in time but in causality. If, for instance, I at this moment arise from this chair, in complete freedom, without being necessarily determined thereto by the influence of natural laws, a new series, with all its natural

consequences, in infinitum, has its absolute beginning in this event, although as regards time this event is only the continuation of a preceding series'

Kant means that if an event were the cause of every subsequent event then that event would necessarily also be the first of those events in time. This is to be contrasted with free actions performed by a finite person, such as Kant himself rising from his chair. Although, if free, such an action is according to Kant's theory a cause of subsequent events but not itself caused, it differs from the cause of all events in (at least) being temporally preceded by earlier events (none of which is *ex hypothesi* the cause of that free action). This is not just incidental to the passage because Kant thinks there is a strong interdependence between being free and being able to use 'I': All and only the self-conscious beings have free will.

Proof of the Thesis

The proof takes the form of *reductio ad absurdum*. Kant assumes the contradictory of what is to be proved and attempts to derive a contradiction. On the assumption that any event is causally necessitated by some temporally precedent event, (including *a fortiori* any event which precedes and causes any event), so there will be 'no completeness of the series' 'keine Vollständigkeit der Reihe' (CPR 410, A446, B474) ‘on the side of the causes that arise one from another' 'auf der Seite der von einander abstammenden Ursachen'. In other words, determinism entails an infinite regress of caused events. This according to Kant is 'self-contradictory'. Kant does not make explicit what the contradiction is but I think the following can be reconstructed: There both was and was not a first event if (*per impossibile*) complete determinism were true. This is because any putative first event would have to turn out to be caused by some temporally prior event. But such a series could never have begun if there had not been just one event of which every subsequent event were a causal consequence, but to claim that conjunction is precisely to require what is logically impossible (or self-contradictory to describe): that there both was, and was not, an uncaused first event.

This argument for the thesis of the First Antinomy leaves room for human freedom but does not entail it. Kant thinks the way out of the regress is to halt it. The way to escape from the contradiction is to deny that there is complete determinism:

'Diesemnach muß eine Kausalität angenommen werden, durch welche etwas geschieht, ohne daß die Ursache davon noch weiter, durch eine andere vorhergehende Ursache, nach notwendigen Gesetzen bestimmt sei, d.i. eine absolute Spontanität.' (A446, B474-5)

'We must, then, assume a causality through which something takes place, the cause of which is not itself determined in accordance with necessary laws, by another cause antecedent to it, that is to say, an absolute spontaneity.'

I shall understand an action or an event's being an 'absolute spontaneity' to mean its being a cause or initiator of a causal chain, but its not itself being an effect of any cause or causal chain. So, *e* is an absolute spontaneity, if and only if there exists at least one event *e′* temporally subsequent to *e* of which the occurrence of *e* is both a necessary and a sufficient condition, but there exists no event *e-1* temporally anterior to *e* such that the occurrence of *e-1* is either a necessary or a sufficient condition (or both) for the occurrence of *e*. (1)

From the fact that there must be postulated an absolute spontaneity to make some sense of the origin of whatever exists, it follows that part of what exists is free action performed by human persons, if human action is itself truly described as an absolute spontaneity. So the argument for the thesis of the third antinomy amounts to a proof of the consistency of 'I am free'.

INSERT PROOF

Kant's Observation on the Thesis

Clearly, from 'p may be imagined' or, even 'p may be coherently be imagined' it does not follow that p. Kant says the thesis falls short of establishing libertarianism as a truth about persons when he says:

'Nun haben wir diese Notwendigkeit eines ersten Anfangs einer Reihe von Erscheinungen aus Freiheit, zwar nur eigentlich in so fern dargetan, als zur Begrifflichkeit eines Ursprungs der Welt erforderlich ist.' (A449, B476)

'The necessity of a first beginning due to freedom, of a series of appearances we have demonstrated only in so far as it is required to make an origin of the world conceivable'

implying that further argument is required to show that 'free' may be truly predicated of human actions. Does Kant have such an argument at (A448, B476-A52, B480)?

Kant analyses the concept of freedom into two sub-concepts; transcendental freedom and empirical freedom, and calls the whole concept thus sub-divided 'psychological freedom'. (A448, B476) He has already explained what transcendental freedom is. (A446, B474) It is a kind of uncaused cause 'without which even in the (ordinary) course of nature, the series of appearances on the side of the causes can never be complete'. This introduces the shift in his concern from an uncaused origin of what exists to the freedom of the human agent. Otherwise he would have omitted 'appearance' and simply written 'series of causes' or 'series of events'.

So, putatively, there is a kind of freedom which is a condition of appearances in particular, not merely of events in general. I take it that any transcendental cause of appearances will not itself be an appearance because, on Kant's view, something's being truly characterised as transcendental strictly precludes it's being truly characterised as empirical, and appearances are empirical. Still, he says the content of the psychological concept of freedom is 'mainly empirical' 'welcher großen Teils empirisch ist.' (A448, B476) I shall interpret this to mean that, more often than not, the predicate 'free' and the noun 'freedom' feature in sentences which may be verified and falsified empirically. Paradigmatically, what we choose to do or omit is empirical. But clearly, on Kant's account that is not the only use of the two terms. The idea of transcendental freedom, itself, does not have any empirical use, yet must be invoked if human action is to be correctly described as free:

'Die transzendentale Idee der Freiheit macht zwar bei weitem nicht den ganzen Inhalt des psychologischen Begriffs dieses Namens aus '…' sondern nur den der absoluten Spontaneität der Handlung, als den eigentlichen Grund der Imputabilität derselben.' (A448, B476)

'The transcendental idea […] stands only for the absolute spontaneity of an action, as the proper ground of its imputability.'

Indeed, 'imputability' here implies that a certain sub-class of events could not be rightly described as actions unless they were transcendentally free. It clearly follows that transcendental freedom is a necessary condition for empirical freedom because, unless actions could be attributed to persons,

there could not be true empirical sentences about those persons and their actions. Agency implies the possibility of free agency.

Although this is the position in the observation on the thesis of the third antinomy, Kant is clearly unhappy about it. He says;

'…ist aber dennoch der eigentliche Stein des Anstoßes für die Philosophie, welche unüberwindliche Schwierigkeiten findet, dergleichen Art von unbedingter Kausalität einzuräumen. (A448, B476)

'This, however, is for philosophy a real stumbling block; for there are insurmountable difficulties in the way of admitting any such unconditioned causality.'

For example, Kant finds it difficult to make sense of the idea of 'a power of spontaneity beginning a series of successive things or states'. 'Kemp Smith's 'spontaneity' translates nothing in the German in this passage. Kant has at B476: 'ein'em' Vermögen […], eine Reihe von sukzessiven Dingen oder Zuständen an 'zunehmen (original; angenommen werden müsse)'', that is, exactly a cause that is not an effect. But, on the argument so far, an uncaused cause is required if the will is free. Kant's solution is to side-step what is in fact a genuine philosophical problem. He simply asserts that the idea of a caused event is just as obscure as that of an uncaused event, yet (because 'cause' is a category) we have '*a priori* knowledge that this latter type of causality must be presupposed' (A448, B476) 'da wir uns eben sowohl bei der Kausalität nach Naturgesetzen damit begnügen müssen, a priori zu erkennen, daß eine solche vorausgesetzt werden müsse'. In other words, we are compelled to think of our experience *as of* causally related events, as one of the conditions of our finding it intelligible. The only genuine question now remaining on Kant's view is whether we may legitimately postulate uncaused causes to similarly make intelligible our empirical belief that we are agents who, therefore, perform actions freely. Thus he concludes:

'Wie ein solches möglich sei, ist nicht eben so notwendig beantworten zu können, da wir uns eben sowohl bei der Kausalität nach Naturgesetzten damit begnügen müssen.' (A448, B476)

'How such a power is possible is not a question which requires to be answered in this case, any more than in regard to causality in accordance with the laws of nature.'

So, 'Am I free?', construed as a transcendental question, is given no answer in the observation on the thesis of the third antinomy. Before turning to Kant's substitute question it should be noted, in Kant's defence, that the concept of an uncaused event is not inherently incoherent. From the fact that some event occurs it does not follow, as a matter of logic, that it was caused to occur. Nothing turns on accepting these as examples but, to illustrate the logic, suppose an event was *the only event that ever happened* or, think of *everything* as an event. It is at least not self-contradictory to assert that an event occurs but is not thus caused to occur.

Kant argues that we may coherently postulate uncaused causes, so that we may think of our wills as free, and his argument takes the form of removing a misunderstanding, and pointing out a parallel with natural determinism. I'll take the parallel with determinism first. The first premise is

(1) 'Ob wir gleich die Möglichkeit, wie durch ein gewisses Dasein das Dasein eines andern gesetzt werde, auf keine Weise begreifen.' (A448, B476)

'We are not in the least able to comprehend how it can be possible that through one existence the existence of another is determined.'

The second premise is

(2) 'Weil aber dadurch doch einmal das Vermögen, eine Reihe in der Zeit ganz von selbst anzufangen, bewiesen (obzwar nicht eingesehen) ist.'

'But the power of spontaneously beginning a series in time is thereby proved (though not understood)'

(Here we assume the reconstructed piece of argument from page 2; otherwise the causal chain could not begin) and the conclusion is:

(3) '…so ist es uns nunmehr auch erlaubt, mitten im Laufe der Welt verschiedene Reihen, der Kausalität nach, von selbst anfangen zu lassen, und den Substanzen derselben ein Vermögen beizulegen, aus Freiheit zu handeln.'

'It is now also permissible for us to admit within the course of the world different series as capable in their causality of beginning of themselves, and so to attribute to their substances a power of acting from freedom'

There are two strands of thinking here. One is: We are compelled to think of the world as containing deterministic causal chains, even though the precise nature of causal connection is obscure to us (for roughly Humean reasons). Therefore, we are entitled, equally, to think the world as containing undetermined events which are initiators, but not effects, of causal chains, even though the notion of 'free', or uncaused, causes might be just as obscure to us as the causal connection. The second strand is that determinism is only thinkable if incomplete. In particular, if there is a deterministic chain then it must have had a beginning. So there was a first event, which initiated all subsequent events, but was not itself initiated by any prior event. But this contains the idea of uncaused cause that is needed for us to think of our own actions as free. So, determinism alone does not preclude freedom of the will, or my rightly answering 'Yes' to 'Am I free?', even when this is a transcendental question. Clearly though, this does not prove 'I am free'.

I turn next to the misunderstanding that Kant warns us against. This is the assimilation of the start of a causal chain to the start of a mere chronological series. If we fail to draw this distinction, we will be led to erroneously suppose an uncaused cause impossible unless it is the first cause in time. Someone who fails to draw the distinction might argue that, because an allegedly free action would not be a first cause but would be only the putative initiator of a causal chain, it would be 'always preceded in the world by some other state of things' (448, B476) 'indem doch immer ein Zustand der Dinge in der Welt vorhergeht' and so would not be itself uncaused. Kant is clearly correct to argue against this way of thinking because, even in the case of a caused event, not every event preceding it in time is a cause of it, even if some are. So, from the fact that some event occurs, and its occurrence is temporally preceded by earlier events, it does not follow that it is caused by any one of those earlier events. This removes one thinkable line of objection to the truth of 'I am free'. (2)

Individuating Causal Chains

Kant draws a distinction between an absolute and a relative beginning (A451, B476) and wishes to show the possibility that the uncaused causes necessary for human freedom might be absolute and not merely relative beginnings. I shall read the distinction this way: An event is an *absolute* beginning if it is an initiator of a causal chain but is not itself an effect. An event is a *relative* beginning if it is the initiator of a causal chain but is temporally preceded by at least one earlier event. Kant wishes to leave room

for an event to be both an absolute and a relative beginning; that is, an event that is a cause but not an effect, and yet which is temporally subsequent to earlier events (none of which are its cause). If human actions are free, then they will turn out to be such events. Kant's example; 'I at this moment arise from this chair' (A451, B479) is designed to be an example of just such an event. Kant claims that if he arises from his chair

'Wenn ich jetzt, völlig frei, und ohne den notwendig bestimmenden Einfluß der Natururs achen, von meinem Stuhle aufstehe, so fängt in dieser Begebenheit, samt deren natürlichen Folgen ins Unendliche, eine neue Reihe schlechthin an.'

'in complete freedom, without being necessarily determined thereto by the influence of natural causes, a new series, with all its natural consequences in infinitum, has its absolute beginning in this event'

There is a specific criterion for the individuation of event series at work here, which is causal or genetic: If two putatively numerically distinct event series *S* and *S′* were in fact causally initiated by one and the same event then, under at least one description, *S* and *S′* are numerically the same event series. If two putatively numerically identical event series *S* and *S′* were in fact causally initiated by numerically distinct events then *S* and *S′* are, in fact, numerically distinct event series. (The second half of the criterion has more plausibility than the first).

As Kant deploys the criterion, in his example, an event series is initiated by an event and that event has no cause, but every event series does have some event as its initiating cause. The 'new series' exists under a different description from the 'natural' event series.

Clearly this is consistent with his view that, if I freely rise from my chair, then that action may be an absolute and not merely a relative beginning, because, presumably, some events preceded my rising from my chair (or any other action I perform), even though possibly my action *qua* that is uncaused.

There then follows a claim that might construed as false:

'Denn diese Entschließung und Tat liegt gar nicht in der Abfolge bloßer Naturwirkungen, und ist nicht eine bloße Fortsetzung derselben, sondern die bestimmenden Natururs achenhören oberhalb derselben, in Ansehung deser Eräugnis, ganz auf.' (CPR 414, A451, B479)

"[...]' this resolution and act of mine do not form part of the natural succession of purely natural effects, and are not a mere continuation of them. In respect of its happening, natural causes exercise over it no determining influence whatsoever.'

From the fact that if I am free (in for example rising from my chair) then I initiate a new causal chain, it may indeed be validly inferred that my decision and action are not causally determined. This follows logically given the sense Kant has allocated to 'free'. But the validity of an inference is no guarantee of the true of its conclusion. For that the truth of the premises is also required, and Kant, it might be objected, has not demonstrated the truth of 'I am free' amongst the premises.

For Kant the postulation of an uncaused cause is a 'requirement of reason'. (A451, B479) We have to make use of the notion to make sense of human decision and action, and of the origin of the universe. But from the fact that we seem psychologically compelled to believe that *p* it does not follow that *p*. The fact that much of my thought and action would be unintelligible to me if I did not at least assume the truth of 'I am free' is itself no guarantee of the truth of that assumption.

Nevertheless, that the assumption is true is a plausible explanation of its being psychologically compelling. At least, if it is held to be false then some account is needed of why it appears so evidently true. (3)

Proof of the Antithesis

If the argument of the thesis proves 'I am free', or at least its consistency, the argument of the antithesis putatively disproves it. It is, in fact, the conjunction of three arguments. If any one of them goes through then 'I am free' is false, in the sense of 'freedom' defined at (A417-B446). I shall call these: The Argument from the Unity of Experience, The Empirical Argument, and The Rules Argument. Each has as its conclusion the impossibility of transcendental freedom.

The Argument From the Unity of Experience

The first premise is

(1) 'Also ist die transzendentale Freiheit '…' eine solche Verbindung '…' nach welcher keinerlei Einheit der Erfahrung möglich ist.' (A447, B475)

'Transcendental freedom "..." renders all unity of experience impossible.'

The second is

(2) 'Transzendentale Freiheit '…' 'reißt' den Leitfaden der Regel 'ab', an welchem allein eine durchgängige zusammenhängende Erfahrung möglich ist' (A447, B475)

‘'Transcendental freedom' abrogates those rules though which alone a completely coherent experience is possible.'

An unstated assumption, implied by (2), is

(3) But there is rule-governed unity of experience.

and the conclusion is

(4) 'Es gibt keine 'transzendentale' Freiheit' (A445, B473)

'There is no 'transcendental' freedom.'

The form of the argument is $(p \rightarrow \text{-}q) \rightarrow (q \rightarrow \text{-}p)$, which is a valid pattern of reasoning. But are the premises true? Let us accept (3) in the sense in which it is accepted by Kant. Then the conjunction of (1) and (2) amounts to the claim that there could be no unity of experience if there were uncaused events. In particular, we are interested in whether if I, and other persons, act freely, experience thereby must be chaotic to us, but two preliminary points need to be dealt with. Arguably, if all, or most, events were uncaused then our experience might be so unpredictable as to be unintelligible and chaotic. But a quasi-Humean reply could be that our experience could exhibit a regularity that was not causal where events of type e^1 are invariably subsequent to events of type *e* but *e*s are not causes of e^1s. Secondly, if there were an uncaused past event, which is the cause of all subsequent events, this need not render our present experience of those events unintelligible. The present epistemological *aporia* about the existence of a first event does not render our present experience unintelligible. But suppose the claim is restricted to the case where, if a person acts freely, their behaviour would be so unpredictable as to be unintelligible to us. This would introduce a strong arbitariness into action which might thwart our attempts to understand others through observation.

There are two replies to this. Firstly, from the fact that I am free (*ie* that my actions and decisions are uncaused causes) it does not follow that other persons cannot predict my behaviour. This is because I might exercise my freedom with a regularity with which other persons might become acquainted, and they might base their predictions on that acquaintance.

Secondly, there is notoriously a difficulty in predicting human behaviour with the certainty that is possible in the study of natural occurrences. This measure of unpredictability is, on the one hand, not so severe as to 'render all unity of experience impossible' and indeed might be due precisely to the exercise of freedom in Kant's sense. So, the Argument From the Unity of Experience, although valid, is not sound because the assumption the second premise rests on is false. Clearly then, it cannot be soundly concluded from this argument that there is no transcendental freedom.

Suppose we read the Argument From the Unity of Experience in a more restricted form, so that it means this:

(1) If all my decisions and actions were free then all my experience would become intelligible to me.

(2) But my experience is not intelligible to me.

(3) Therefore, I am not free.

Kant leaves the proponent of transcendental freedom open to this kind of objection because he does not explain the relationship between decisions and actions. But the account could be supplemented on Davidsonian lines to protect it. The tenet that my decisions' being freely made consists in their being uncaused causes could be retained, and it could be argued in addition that I might find my decisions predictable by being acquainted with which decisions I had made previously. It is true in any case that I do not often know what I will decide before I decide it. Indeed, knowing what I will decide (because of what 'knowledge' means) could be taken to entail deciding anyway. I call this supplement to Kant's theory of freedom 'Davidsonian' because it allows decisions to be described as causes of actions. This strategy requires sacrificing the transcendental freedom of action, because actions are not uncaused causes if they are caused by decisions, but this only amounts to a weakening of the claim 'I am free' because it leaves room for me to be, by my decisions, the uncaused cause of those actions. Complete determinism, which entails the 'there is no freedom'

of the antithesis, requires the falsity of even this weakened claim. So, even when re-written in first person singular form the Argument From the Unity of Experience is not a refutation of libertarianism.

The Empirical Argument

The premises of the Empirical Argument are

(1) 'Also ist die Transzendentale Freiheit '…' 'so dass sie' also auch in keiner Erfahrung angetroffen wird' (A447, B475)

"Transcendental freedom' is not to be met with in any experience;

and

(2) '‘'Transzendentale Freiheit ist' mithin ein leeres Gedankending’'

‘'Transcendental freedom' is therefore an empty thought entity'

and the conclusion, as before, is

(3) 'Es gibt keine 'transzendentale' Freiheit’'

There is no 'transcendental' freedom.

This argument requires as an assumption the version of concept-empiricism which I call the Sense and Sense Argument. According to this doctrine, a concept has meaning if and only if it either admits of empirical (ie ostensive) definition, or may be completely translated into concepts that admit of empirical (ostensive) definition. Without this assumption, there is no reason why, if transcendental freedom is not met with in any experience, it should be 'an empty thought entity'. I understand by 'an empty thought entity' a putatively conceptual item which has no existing referent and, therefore, has no informative use. (A *Gedankending* is just an object of thought, but *prima facie*, according to the German use of the term, it might be informative. It might be part of a *Gedankenspiel*. In this sense, Pegasus might be a *Gedankending*.)

Premise 1 is analytic; It is part of the meaning of 'transcendental' that if something is transcendental it cannot be an object of experience. The inference from (1) to (2) requires the truth of concept empiricism. Then the

conclusion follows directly from (1) and (2), because if 'empty' means 'without any referent', and if 'freedom' is an empty thought entity, then 'freedom' has no referent, which is equivalent to saying there is no freedom.

The doctrine of transcendental freedom is vulnerable to this argument in a way that the doctrine of empirical freedom is not. Although 'free' has empirical uses in sentences whose truth conditions are 'to be met with in experience' when we commonsensically talk of someone being free or not free to do such and such (for example, walk out of jail), 'transcendentally free' has no such uses.

Only two strategies seem plausible in resisting this line of argument. One is to weaken the concept of transcendental freedom, the other is to find reasons for denying concept-empiricism.

Weakening 'transcendental freedom' is difficult just because (1) is analytic. However, it is sometimes argued by libertarians that we are each directly acquainted with our own free will as a kind of experiential datum, that is qualitatively distinct from our experience of event series in the external world. Although clearly this is to collapse transcendental freedom into empirical freedom, and so not to save transcendental freedom from the first premise, if the direct acquaintance is veridical, it entails that there is free will. I am confronted with choice, so there is choice. I experience my initiation of causal chains, so I initiate causal chains, and so on. Of course, the determinist, if they admit the experience, will deny it is veridical.

A more promising version might be this: We are each, at least sometimes, directly acquainted with our own decisions as, or soon after, they come to mind. But, at least sometimes, we are ignorant of any cause which precipitated our deciding one way or the other. We allow also that our decisions are (amongst the) causes of our actions. If so, then there exists an event of which the following is true: It is a cause but is, at least, not known to be an effect. This leaves conceptual room for the possibility that such an event is in fact a cause that is not an effect. This is because there are two reasons why it might appear to be a cause that is not an effect, viz: Either it is false that it is a cause that is not an effect, but we are not acquainted with the cause of it, and so do not know that it is an effect, or else; It is in fact a cause that is not an effect, and that is the reason we are not acquainted with any cause of it. There is nothing to be acquainted with. This argument leaves room for the transcendental freedom of decisions (but not actions). The predicate 'transcendental' shifts from actions to decisions.

The premise that we are introspectively aware of what we are deciding, or have decided, is not a threat to transcendental freedom, because even if what one is thus aware of is, in fact, a transcendentally free event, no-one is

thus aware of it *qua* transcendentally free event, which is impossible in Kantian terms: Nothing transcendental is empirical. Any object of awareness is empirical, so there is no awareness of transcendental freedom.

Yet this argument is still strong enough to pose a threat to The Empirical Argument. In the introspective situation, which way should a strict empiricist choose? When directly acquainted with an event which seems to be a cause but which does not seem to be an effect, should the strict empiricist say it is a cause that is not an effect and that is why there is no experience of the cause of it, or should he say it is really an effect but only appears to be a cause that is not an effect because the cause of it is not revealed to experience? It is clear which way the strict empiricist should choose: He should, provisionally, adopt the view that there is at least one cause that is not an effect but continue to look for causes of that cause by empirical means. Indeed, this is much the relationship between libertarianism in the philosophy of mind and neurology at present. But what the empiricist is thereby forced to concede is very close to transcendental freedom. He is holding as a provisional hypothesis that there is an event that is a cause but which is not itself caused, *even though he would like to falsify that hypothesis empirically*.

It follows clearly that 'transcendental freedom' is not an empty concept in the sense of 'uninformative' or 'without a coherent epistemological function', even on a strictly empirical account of mental efficacy. Indeed, the empiricist needs something very like it to make sense of the situation that might later be amenable to causal explanation.

The criticism of concept empiricism in general would take us too far from the topic of free will, so I do not pursue that here. But even without such criticism, the Empirical Argument leaves room for a weakened version of transcendental freedom and so does not amount to a disproof of 'I am free'.

The Rules Argument

The first premise is

(1) 'Die Freiheit '…' ist zwar eine Befreiung vom Zwange, aber auch vom Leitfaden aller Regeln' (A447, B475)

'Freedom is "..." a liberation from compulsion, but also from the guidance of all rules.'

The second is

(2) 'Wenn diese 'd.i. Freiheit' nach Gesetzen bestimmt wäre, sie nicht Freiheit '…' wäre.' (411, A447, B475)

'If freedom were determined in accordance with laws, it would not be freedom.'

and the conclusion is;

(3) 'Denn man kann nicht sagen, daß, anstatt der Gesetze der Natur, Gesetze der Freiheit in die Kausalität des Weltlaufs eintreten.' (A447, B475)

'It is not permissible to say that the laws of freedom enter into the causality exhibited in the course of nature, and so take the place of natural laws.'

(Kemp Smith's reading of the German is not obviously right here. He might have reversed premise and conclusion. However, I shall go with Kemp Smith.) This argument is designed to rebutt any theory according to which freedom is governed by its own laws, which are putatively in some respects analogous to but in fact incompatible with laws of nature.

If not plausibly synthetic *a priori* truths, the first two premises are either analytic, and so trivially true, or else synthetic and false. Certainly the first conjunct of the first premise is straightforwardly analytic. It is true by definition that freedom involves liberation from compulsion; This idea is essential to Kant's concept of transcendental freedom as uncaused causality. The second conjunct of the first premise is synthetic and false. It is clearly intended as synthetic because 'but also' is a grammatical marker which indicates that new and additional information is to be supplied not already expressed by the first conjunct. It is false because it is open to the same objection as the first two premises of the Argument From the Unity of Experience: From the fact that 'I am (transcendentally) free' is true it does not follow that my decisions and actions will not be rule governed. Of course, these rules will not be the exhibiting of causal connections, nor rules according to which decisions and actions of a certain sort are shown to be consequences of causes of a certain sort. They might, however, be regularities or repetitions in decision or action. Here 'governed' or 'guided' cannot mean 'caused'. A non-Davidsonian libertarian might, in addition, claim that a person's reasons for performing an action were not causes of it; perhaps because they do not necessitate the action. Then it is open to such a

libertarian to claim that one's actions exhibit a rule guided regularity that is not causal: It depends upon the contingent, but uniform, conjunction of roughly similar sorts of reasons being given for roughly similar sorts of actions. If this is right, then 'I am free' need not require 'liberation from the guidance of all rules'. 'All' is an exaggeration.

The second premise is analytic if 'laws' means 'causal laws' because it would be self contradictory to assert that some action is transcendentally free, yet could be subsumed under a causal law. That would entail that one and the same event both did and did not have a cause. But this analytic truth is one that the libertarian would be happy to concede. Construed synthetically however the claim is false, because it is subject to the same sort of objection as the first premise's second conjunct. Freedom has to be distinguished from mere randomness, or arbitrariness, and it is possible that such regularities its exercise exhibits should not turn out to be causal.

The conclusion, without the support of the premises, begs the question against the libertarian. On Kant's criterion for the individuation of event-series, an uncaused cause might 'enter into the causality exhibited in the course of nature' in just the sense of initiating a causal chain in which every subsequent event in the chain was a causal consequence of it, but where it was not itself a causal consequence of any event. This would not amount to the uncaused cause 'taking the place of natural laws'. Rather, the chain thus initiated falls under natural law. On Kant's account, 'uncaused cause' has taken the place of 'caused cause'. If 'I am free' is true, and if this entails that there are uncaused causes, then human decisions are just such initiations of causal chains. (4)

Kant’s Compatibilism

Kant's solution to the Third Antimony is transcendental idealism; the adoption of the distinction between things as they are in themselves, and as they appear to the human subject. Kant will argue that as noumenal selves we are free, but as we appear to ourselves we are determined in accordance with natural law. Here, if anywhere in the *Critique of Pure Reason*, Kant answers 'yes' to the question 'Am I free?'.

The distinction between freedom and determinism which Kant draws (at CPR 464, A523, B560 - A533, B561) entails that freedom, like determinism, is a kind of causality:

'Man kann sich nun zweierlei Kausalität in Ansehung dessen, was geschieht, denken, etweder nach der Natur, oder aus Freiheit' (CPR 464, A532, B560)

'When we are dealing with what happens there are two kinds of causality conceivable by us: the causality is either according to nature or arises from freedom'

Freedom is a kind of causality because a free event is a cause that is not an effect. Natural, or determined, causality, in contrast, requires that there be events which are both causes and effects and, arguably, requires every event to be both a cause and an effect. Kant uses this distinction to argue for the sort of freedom he thinks we do in fact possess.

Practical Freedom

His best definition of 'practical freedom' is at CPR 464, A534, B562:

'Die Freiheit im praktischen Verstande ist die Unabhängigkeit der Willkür von der Nötigung durch Antriebe der Sinnlichkeit'

'Freedom in the practical sense is the will's independence of coercion through sensuous impulses'

I shall interpret this to mean: When presented with a choice between two courses of action, X and Y, I may still do X instead of Y, even if the closest possible empirical scrutiny of the causal conditions for my making the choice one way or the other suggested I would do Y. In other words, from no set of empirical sentences, which express a prediction of my performance of an action, does it follows that I will perform that action. Further, it remains logically possible that I should perform an alternative action where the empirical evidence for such a performance ranges from every small to nil. This formulation widens the scope of 'sensuous impulse' to '(empirical) causal conditions'. This includes sensuous impulses in its denotation but, as we shall see, Kant needs, and often subscribes to, the wider formulation.

This reading of 'practical freedom' allows Kant to retain his distinction between the animal will and the human will. Both human and (non human) animal wills are 'sensuous' because they are 'pathologically affected ie by sensuous motive' 'pathologisch affiziert ist (durch Bewegungsursachen der Sinnlichkeit)' (CPR 465, A534, B562). This means that in the case of animals always, but in the case of humans sometimes, a particular action is causally determined by the mental state of the animal, or human. All

sensuous wills are sensuously 'affected' but only a subclass, the animal class, have wills which are pathologically 'necessitated'.

This is an important distinction between people and animals and I take it to be constitutive of people that they are selves, and being a self partly consists in having a will that is not entirely (pathologically) necessitated. The fact that Kant claims that the human will may, nevertheless, be pathologically affected needs to be reconciled with its practical freedom. This can be done in a number of ways.

It could be argued on Kant's behalf that if I am presented with two possible courses of action X and Y then I can always, in principle do X and not Y even if 'affected' to do Y. Even if I feel the 'sensuous impulse' to do Y I may still refrain from doing Y and do X . Even if sensuous impulses raise thr probability of my doing Y, that is consistent with my doing X.

Alternatively, and less satisfactorily as an interpretation, I might sometimes be able to do X but sometimes be affected to do Y rather than do X . This is less satisfactory because it suggests that persons might be pathologically necessitated, not just affected, and that is something Kant denies. Further, it contravenes the principle that I might always do X, even when a study of the empirical determinants of my behaviour suggest I would do Y.

So, I think we have to take Kant literally when he says of the human will 'sensibility does not necessitate its action', that ‘Sinnlichkeit […] Handlungen nicht notwendig macht' (CPR 465, A534, B592). Here 'does not' means 'never does'.

Although this is Kant's theory, it might be objected that such a strong doctrine of practical freedom is intuitively implausible. If the human will is never 'pathologically necessitated' then would it not be the case that I never lose when struggling with weakness of will? Complete freedom from sensuous impulses would seem to make *akrasia* impossible. To meet this objection we have to allow with Kant that the human will is sometimes sensuously 'affected', where this implies that there are empirical causes of my action, and my felt real preference to do X rather than Y is not exercised. I lose the battle of the will, and my behaviour is empirically determined.

Kant can allow this through his distinction between affected in general and necessitated in particular. I take it that if *e* necessitates *e* ' then it follows that *e* affects *e* ', but that if *e* affects *e* ' then it does not follow that *e* necessitates *e* '. This is because if *e* affects *e* ' then *e* (is amongst what) causes *e* ' to happen but it is further true, in the case of necessitation, that if *e* necessitates *e* ' then *e* ' could not not have happened given that *e* happened.

In contrast, if *e* affects *e* ' then it remains logically possible that *e* ' should not have happened even if *e* happened.

The exercise of my freedom, for example, might have prevented the occurrence of *e* ' so long as *e* ' is affected but not necessitated, and that is precisely the way in which Kant has drawn the distinction between human and animal wills. Although my actions might be causes, they are not necessitated but all an animal's actions are both caused and necessitated. It is unique to persons that they may perform some actions freely.

After all, Kant talks of human freedom as a power, and this does not commit him to the view that that power is always and invariably successful when exercised, even though he retains the strong view that it could always, in principle, be so exercised. He says

'Dem Menschen 'wohnt' ein Vermögen 'bei', sich, unabhängig von der Nötigung durch sinnliche Antrieb, von selbst zu bestimmen.' (CPR 465, A534, B562)

'There is in man a power of self-determination, independently of any coercion through sensuous impulses.'

'Self-determination' needs clarifying here. One way of reading it is this: *e* is self determining if *e* is the cause of itself. This is an obscure and barely coherent idea if it implies the logical possibility of something bringing itself into existence. In general, if A causes B it follows that A is not identical with B, if only because causation is an *n*-ary relationship and there cannot be *n*-ary relationships without distinct *relata*. That would make 'A causes itself to exist' incoherent.

But Kant does not need that possibly incoherent notion for his theory of freedom. We can read 'self-determination' in a way that allows persons to be (amongst) the causes of their own actions.

Then 'self' determination refers to the person, so if I am self determining I cause myself to be what I am, or I cause myself to do what I do.

There is a problem about 'I' here, but, in the last resort, the referent of the first person singular pronoun for Kant will turn out to be one's free, noumenal, self.

On yet another reading of 'self-determination', we may invoke the notion of uncaused causality outlined so far. Then if *e* is self-determining then *e* is a cause but not an effect: *e* is determining as a cause, but self-determining in not being caused by anything other than itself, to be thus determining. It is self-determining in the sense of 'autonomous' but not in the sense that it

causes itself to be, or to be what it essentially is. The second two, but not the first, of these readings are close to Kant's use of 'self-determination'. The second in particular is essential to a conceptual link Kant thinks exists between transcendental and practical freedom.

Transcendental and Practical Freedom

Transcendental freedom is an essential part of practical freedom, in the sense that if a person is practically free then that entails that that person is transcendentally free. The concept of transcendental freedom is

'die Idee von einer Spontaneität, die von selbst anheben könne zu handeln, ohne daß eine andere Ursache vorangeschickt werden dürfe, sie wiederum nach dem Gesetze der Kausalverknüpfung zur Handlung zu bestimmen' (CPR 465, A533, B561)

'the idea of a spontaneity which can begin to act of itself, without requiring to be determined to action by an antecedent cause in accordance with the law of causality'

Kant should have added here 'and in no way being determined to action by an antecedent cause' otherwise his formulation leaves it open that an event should be transcendentally free and yet causally determined which, as he admits, is self-contradictory.

This concept of transcendental freedom is deployed in two arguments: one designed to demonstrate the reality of practical freedom, that persons are practically free, and the other to show the impossibility of truly affirming the reality of practical freedom without thereby having truly affirmed the reality of transcendental freedom; that persons are transcendentally free. In Kant's text these two issues are telescoped together, so I shall try to separate them out.

(a) The Argument for Practical Freedom

The premises are:

(1) 'Wenn alle Kausalität in der Sinnlichkeit bloß Natur wäre, so würde jede Begebenheit durch eine andere in der Zeit nach notwendigen Gesetzen bestimmt sein.'

'If all causality in the sensible world were mere nature, every event would be determined by another in time, in accordance with necessary laws'

(2) 'Da die Erscheinungen, so fern sie die Willkür bestimmen, jede Handlung als ihren natürlichen Erfolg notwendig machen müssten'

'Appearances in determining the will would have in the actions of the will their natural effects, and would render the actions necessary'

(3) "Praktische Freiheit' setzt voraus, daß, obgleich etwas nicht geschehen ist, es doch habe geschehen sollen…'

'practical freedom presupposes that although something has not happened it ought to have happened'

and the conclusion is

(4) '…und seine Ursache in der Erscheinung also nicht so bestimmend war, daß nicht in unserer Willkür eine Kausalität liege. (CPR 465, A534, B562)

'its cause (as found) in the (field of) appearance is not therefore so determining that it excludes a causality of our will.'

In (4) 'its' refers back to 'something' in 3. (4), as it appears in Kant's text, is a weaker conclusion than an affirmation of practical freedom. It is the claim that practical freedom is merely not excluded by natural law determinism. But the context of the argument, for example the straight affirmation at CPR 465, A534, B562 that 'there is in man a power of self-determination', makes it clear that Kant would like to draw the stronger conclusion, and perhaps he thinks he has done.

To obtain the stronger conclusion we need to add some assumptions. Suppose we accept the truth of all three premises. Then we need to supplement (3) with this:

(3a) If something ought to happen, it could even if it does not.

The determinist cannot accept (3a) but Kant clearly can. His distinction between affectation and necessitation gives him the distinction between what both happens and had to happen, and what happens but could, in principle,

not have happened (given a particular set of empirical conditions). Then freedom is postulated as what makes this distinction possible.

If this seems unacceptable to the determinist, three lines of objection are open:

(a) Deny the distinction between what happens and what has to happen allowing only what has to happen: what could not not happen.

(b) Try to make sense of 'ought' without implying 'can'.

(c) Postulate something other than the freedom of the will that might operate 'contrary to the force and influence' of 'natural causes'.

But (a) collapses natural law determinism into a kind of metaphysical fatalism. (b) would need much extra argument. (c) would hardly be more acceptable to the determinist than free will itself.

Kant's argument is really quite a strong one. It amounts to: We need 'free' to make sense of 'ought', because 'ought to have' and 'ought not to have' imply 'could' and 'could have done otherwise'.

So, if we ought not to have done X then we could have not done X, or were free not to do X. Similarly, if we ought to have done Y then we could have done Y, or were free to do Y. The determinist either has to give up the language of morals, or find a non-Kantian account of the truth conditions for 'ought' sentences.

(b) The Argument from Practical to Transcendental Freedom

(1) and (2) from the Argument for Practical freedom also feature as the first two premises here. The third premise is the definition of transcendental freedom:

(3) "Transzendentale Freiheit ist' die Idee von einer Spontaneität, die von selbst anheben könne zu handeln, ohne daß eine andere Ursache vorangeschickt werden dürfe, sie wiederum nach dem Gesetze der Kausalverknüpfung zur Handlung zu bestimmen' (CPR 465, A533, B561)

'Transcendental freedom is the idea of a spontaneity which can begin to act of itself, without (requiring) to be determined to action by an antecedent cause in accordance with the law of causality.'

The fourth premise is the definition of practical freedom

(4) 'Freiheit im praktischen Verstande ist die Unabhängigkeit der Willkür von der Nötigung durch Antriebe der Sinnlichkeit' (CPR 465, B534, B562)

'Freedom in the practical sense is the will's independence of coercion through (sensuous) impulses.'

and the conclusion is

(5) 'Die Aufhebung der transzendentalen Freiheit 'würde' zugleich alle praktische Freiheit vertilgen.'

'The denial of transcendental freedom must, therefore, involve the elimination of all practical freedom.'

Is the argument sound? Because soundness requires both validity and the truth of the premises, let us assume the truth of deterministic premises (1) and (2), as before. Let us assume, also, the accuracy of Kant's definitions of 'freedom' in (2) and (4).

Now the only question that remains is that of validity; whether (5) follows from the conjunction of (1)-(4). To decide this, we need to know whether the truth of (1) and (2) precludes the truth of freedom in the sense of (4), and whether the non-existence of freedom in the sense of (4) precludes, logically, the possibility of freedom in the sense of (3). If these two conditions are met then (6) follows, and the argument is valid.

The existence of determinism as defined in (1) and (2) is clearly inconsistent with the existence of freedom as defined in (3) because (3) asserts that freedom is independence of causality, but (2) asserts that if determinism is true then the actions of the will are effects. So, (3)'s freedom is incompatible with (1) and (2)'s determinism. If we give (4) the wider reading indicated above then denying (4) will entail denying (3), at least in this respect.

(3) defines a sort of freedom from causality in general. (4) defines freedom of the will from causality in particular. So, even if determinism is true, there could be some sort of transcendental freedom, but this could not possibly be freedom of the will. (It could, for example, be the sort of transcendental freedom required to give sense to a certain account of the origin of the universe). But, so far as the freedom of the will is concerned,

Kant's conclusion goes through: If I am not transcendentally free, then I am not practically free either.

This accords well with Kant's general thesis that all practical freedom requires transcendental freedom but not all transcendental freedom requires practical freedom. That is because not all freedom is freedom of the will (the putative uncaused first cause for example) but all practical freedom is freedom of the will. That is why 'will' appears where it does in the definition of 'practical freedom' (4) above.

May Kant coherently state that persons are free as they are in themselves but causally determined as they appear to themselves? He intends us to accept this thesis within the framework of transcendental idealism but, before we can, several conceptual obstacles need to be removed. The first is this:

"Praktische Freiheit' setzt voraus […] eine Kausalität '…' 'welche' unabhängig von jenen Naturursachen und selbst wider ihre Gewalt und Einfluß etwas hervorzubringen, was in der Zeitordnung nach empirischen Gesetzen bestimmt ist, mithin eine Reihe von Begebenheiten ganz von selbst anzufangen.' (CPR 465, A534, B562)

'Practical freedom presupposes [...] a causality which, independently of those natural causes, and even contrary to their force and influence, can produce something that is determined in the time order in accordance with empirical laws, and which can therefore begin a series of events entirely of itself.'

The problematic phrase here is 'contrary to their [ie natural causes] force and influence'. If everything phenomenal is causally necessitated, and may be subsumed under natural laws, how can it also be true that the kind of causality present in practical freedom may act 'contrary to' such causal connections and natural laws?

Kant has a solution to this problem but I think its force can only be fully appreciated if we accept the distinction drawn earlier between two sorts of description under which an event may fall, one featuring 'cause', the other featuring 'effect'. First it is useful to distinguish different possible strengths of 'contrary to' as it appears in 'contrary to their force and influence'.

On the strongest interpretation, the exercise of freedom breaks a law of nature. Kant is not interested in such a strong sort of freedom.

A weaker construal is this: If *e* and *e*' are phenomenal events, causally linked such that the occurrence of *e* is both a necessary and sufficient

condition for *e*' (such that unless *e* happened *e*' could not happen, and such that if *e* happened then *e*' could not not happen), then a free event *f* (ie a cause that is not an effect) operates in a way 'contrary to' *e*, and 'contrary to' the connection between *e* and *e*', if and only if the occurrence of *f* would prevent the occurrence of *e*', even if *e* happened.

This is a strong interpretation of 'contrary' because it would seem that if *e* happens then *e*' must happen, but if *f* happens *e*' cannot happen, but it is further asserted that both *e* and *f* happen. What follows from this is logically impossible: that *e*' both does and does not happen. Kant does not argue this through, but it is sufficiently threatening to his compatibilism to warrant a *prima facie* 'yes' in answer to his question:

'ob es ein richtigdisjunktiver Satz sei, daß eine jede Wirkung in der Welt entweder aus Natur, oder aus Freiheit entspringen müsse' (CPR 466, A536, B564)

'Is it a truly disjunctive proposition to say that every effect in the world must either arise from nature or from freedom?'

Here 'truly disjunctive' means 'jointly exhaustive and mutually exclusive', so we should read 'or' as *aut* and not *vel*, such that, if there is freedom, there is not determinism and, if there is determinism, there is not freedom.

Suppose we do not think of the exercise of freedom as preventing (*per impossible*) the occurrence of an event which is causally necessitated, but rather as causing it. Then there is clear overdetermination. *e′* is causally necessitated in accordance with natural law, and e' results from the exercise of noumenal freedom. There are, therefore, two sufficient conditions for the occurrence of *e*'.

Kant's solution is to adopt neither of these courses. Instead he makes use of the idea of practical freedom initiating a causal chain. The three strategies may be represented as follows:

(i) *e* happens and *f* happens so *e*' both does and does not happen

(ii) *e* happens and *f* happens so *e*' does happen

(iii) *f* happens and some event other than *e* and *e*' happens, irrespective of whether *e* and *e*' happen.

Strategy (iii) is expressed in this argument at CPR 467, A537, B565:

(1) 'Obzwar ihre Wirkungen erscheinen, und so durch andere Erscheinungen bestimmt werden können'.

'The effects [...] as intelligible causes appear, and accordingly can be determined through other appearances.'

(2) 'Eine solche intelligible Ursache aber wird in Ansehung ihrer Kausalität nicht durch Erscheinungen bestimmt'

'But its causality is not so determined'

(3) 'Sie ist also samt ihrer Kausalität außer der Reihe, dagegen ihre Wirkungen in der Reihe der empirischen Bedingungen angetroffen werden.'

'While the effects are to be found in the series of empirical conditions the intelligible cause, together with its causality, is outside the series'

(4) 'Thus the effect may be regarded as free in respect of its intelligible cause and at the same time in respect of appearances, as resulting from them according to the necessity of nature'

'Die Wirkung kann also in Ansehung ihrer intelligblen Ursache als frei, und doch zugleich in Ansehung der Erscheinungen als Erfolg aus denselben nach der Notwendigkeit der Natur angesehen werden.'

(4) expresses compatabilism. So, are (1), (2) and (3) true, and do they yield (4)?

I interpret (1) as follows: If *f* is an intelligible cause then *f* is a free event, that is, a cause that it is not an effect. It is also true of *f* that *f* can be known to exist by the exercise of the intellect (for example by following the argument above) but *f* is not a possible object of experience. For this reason, I read 'intelligible' to mean 'only intelligible, and not 'also perceptible'. An effect of an intelligible cause, call it here '*ef*', is a possible object of experience. This gives a restricted reading to 'can be determined through other appearances' and 'appear(s)'.

(2) just confirms the meaning of 'intelligible' as 'merely intelligible, and not perceptible'.

(3) is the claim that a free cause, a cause that is not an effect, is not embedded within an causal chain. Kant's making this claim is rather too

strong for the argument, because what he needs is for *f* to be a cause of subsequent events, which are indeed jointly a series of causally related empirical events, but *f* must not itself be the effect of any chain or event in a chain of antecedent events. *f* is the imperceptible initiator of a chain of perceptible events. The initiator of the chain is not 'outside the series' but its first member.

Should we read 'outside the series' also to mean 'outside the series in time', not only as 'outside the series in respect of being caused'? On the one hand, noumenal reality is not spatio-temporal, so, *a fortiori*, not in time. On the other hand, Kant's, for example, rising from his chair, is chronologically a part of a series of empirical events, even though those events preceding it, including those that seem to be causes of it, are not causes of it, if it is performed freely.

There is no *a priori* objection to the truth of (1), (2) and (3) and Kant has already argued for (1). Also, (2) either qualifies, or is entailed by (1) depending on how we construe 'intelligible'.

(3) is really a conjunction, the first conjunct of which (introduced by 'which') logically entails (1), so long as the series of empirical conditions are appearances, which they are on Kant's account.

But to derive (4) from the conjunction of (1), (2) and (3) it must be read in a special way: It must be possible to seem as though my actions are caused by empirical conditions which are, in fact, events which merely chronologically predate my action and do not, in fact, cause it to happen in any degree whatsoever. For the argument to go through, we have to read 'appearances' as 'non veridical appearances' in 'the effect may be regarded in respect of appearances as resulting from them according to the necessity of nature' otherwise (3) falls into the incoherence of (i) above, or the vacuousness of (ii). Only that reading, whereby it is an illusion that my free actions are empirically determined, allows my actions to be uncaused events: events with empirical effects but without genuine empirical causes. This makes use of Kant's distinction between appearances and things in themselves in a special way. I show now that a different use will not work.

Kant wonders whether

'Es ist also nur die Frage: ob dem ungeachtet in Ansehung eben derselben Wirkung, die nach der Natur bestimmt ist, auch Freiheit stattfinden könne, oder diese durch jene unverletzliche Regel völlig ausgeschlossen sei' (CPR 466, A536, B564)

'An effect, not withstanding its being [...] determined in accordance with nature, may (not) at the same time be grounded in freedom.'

For this to be the case, the dilemma of contradiction or over-determination would have to be eschewed. What is Kant's distinction between 'determined' and 'grounded' here?

If *e* determines *e*' then *e* is sufficient for *e*'. If *f* grounds *e*' then *e* is necessary for *e*'. We may read Kant as saying *e*' is necessitated by earlier causes in the chain in which it is embedded but that chain was initiated by *f*, so e' would not have happened without *f*. Nevertheless, if *e* is necessary for *e*' then *f* cannot be sufficient for *e* 'and if *f* is necessary for *e*' then *e* may be sufficient for *e* '. It would have to be both necessary and sufficient for *e* ' and some free act f would have to be both necessary and sufficient for *e*'. But that seems hopeless. If *e* is sufficient for *e*' then *f* is not necessary for *e* ', and if *f* is sufficient for *e*' then *e* is not necessary for *e*'. (Nevertheless, if *e* is necessary for *e*' then *f* may be sufficient for *e*' and if *f* is necessary for *e* ' then *e* may be sufficient for *e* '.

If there is overdetermination, it may be the case that *e* is both necessary and sufficient for *e* ' and that *f* is both necessary and sufficient for *e* '. It is logically possible that *e* exists if and only if *f* exists.

The solution I have just extracted from (4) obviates this difficulty. It entails that *f* is both necessary and sufficient for *e* but that *e* only appears to be both necessary and sufficient to *e*' but is in fact merely chronologically antecedent to *e*'. That is the most sympathetic construal of his compatibilism: the one that allows him to present a solution to the problem of freedom and determinism. So it does allow as true 'I am free'.

One other objection needs removing: If I am the free initiator of a causal chain, it follows that there takes place a series of empirical events which do not have an empirical cause. Is this intrusion into the phenomenal world logically consistent with the synthetic *a priori* claim that every empirical event has a cause, and does it violate any law of nature?

As far as the first of these is concerned, an exception would have to be made of that empirical event which is the first effect of my free action if the synthetic *a priori* thesis is interpreted to mean every empirical event must have an empirical cause. But there seems no compelling need for Kant to make that stipulation. Even if some empirical events are caused by uncaused free actions, those empirical events are still caused and so do not violate the principle. Similarly with laws of nature, if a free action initiates a causal chain, and that causal chain is empirical, ie, is a possible object of appearance, and is embedded in a nexus of causal relations, no causal law

has thereby been broken. Kant thinks natural laws hold only of phenomena, so, if there is a noumenal and free cause that is not an effect, no natural law is broken by that.

Then there is nothing to preclude the event-components of the resultant causal chain effecting changes in the empirical reality in which it is embedded. This amounts to the fact that it is coherent to suppose that I may intervene in the course of nature, and allows us to give a final and very weak reading to 'contrary' above. This is: Some free action *f* is 'contrary' to some course of events in nature if those events would not have happened if *f* had not happened, or if those events would have happened unless *f* had happened. In this respect (though of course not as an effect) *f* has the same status as any other necessary and sufficient condition for an event or set of events.

My Freedom

I, as I am in myself, am a free agent. I, as I empirically appear to myself, and others, am causally determined. But this raises a problem. If it is true that

'Dieses handelnde Subjekt würde nun, nach seinem intelligibelen Charakter, unter keinen Zeitbedingungen stehen, denn die Zeit ist nur die Bedingung der Erscheinungen, nicht aber der Dinge an sich selbst.'
(CPR 468, A539, B567)

'This acting subject would not in its intelligible character, stand under any conditions of time, time is only a condition of appearances and not of things in themselves.'

then it does indeed follow that

'In ihm würde keine Handlung entstehen, oder vergehen.'

'In this subject no action would begin or cease.'

but then how can such a subject be coherently described as an agent at all? If I am capable of action, then I must be capable of beginning to act and ceasing to act. My actions, in order to count as such, just last a certain time.

One Kantian answer would be that the concept of an agent is an empirical one, and only in so far as I am a possible object of intuition may I be conceived as a (spatio-temporal) agent.

This will hardly do though. The doctrine of transcendental freedom, implied by practical freedom, requires that I be the noumenal cause of effects in the world, as it appears empirically. The conjunction of the conclusion of the Transcendental Aesthetic, that the forms of intuition apply only to phenomena, with the claim that I am a noumenal cause of empirical events commits Kant to the view that there may exist what may be truly called a 'cause' yet to which no spatio-temporal predicates apply. 'Cause' is a category, and the categories only have empirical, that is, spatio-temporal, application. But transcdental freedom requires non-empirical causes. I, as I am in myself, am such a cause. It might turn out that the conjunction of '*x* possesses causal efficacy' and '*x* is not spatio-temporal' is not self-contradictory but first it is worth inspecting the motives which led Kant to this difficulty.

If Kant admits that noumenal causes have spatio-temporal properties, then two gross inconsistencies with the remainder of his philosophy are introduced. Firstly the idea of a noumenon collapses into the idea of an empirical object, because if *per impossibile* noumena were spatio-temporal they would be possible objects of experience. Secondly, if noumenal causes were spatio-temporal they would be embedded in an empirical causal nexus, and freedom in Kant's sense would be impossible.

Kant is fully aware of these two threats. Indeed, the second is his motivation for emphasising the non-temporality of the noumenal at CPR 468-9, A540, B568):

'In ihm würde keine Handlung entstehen, oder vergehen, mithin würde es auch nicht dem Gesetze aller Zeitbestimmung, alles Veränderlichen, unterworfen sein: daß alles, was geschieht, in den Erscheinungen (des vorigen Zustandes) seine Ursache antreffe.'

'In this subject no action would begin or cease, and it would not therefore have to conform to the law of determination of all that is alterable in time, namely that everything which happens must have its cause in the appearances which precede it.'

He could in fact have drawn a stronger conclusion, here, than that implied by 'would not therefore have to', because, if *x* is not temporal, it is not clear how *c* could be empirically determined at all. The stronger conclusion is drawn in this passage:

'So fern es Noumenon ist, nichts 'in ihm' geschieht, keine Veränderung, welche dynamische Zeitbestimmung erheischt, mithin keine Verknüpfung mit Erscheinungen als Ursachen angetroffen wird. ' (B569)

'In as much as it is noumenon nothing happens in it; there can be no change requiring dynamical determination in time, and therefore no causal dependence upon appearances.'

Sensitive to the tension this introduces, Kant adopts two strategies by each of which he hopes there may be noumenal causes with empirical effects.

The first of these relies on the thesis that nothing positive may be known about noumena anyway, and that, if we do try to think them, our thinking is confined to empirical categories:

'Dieser intelligibele Charakter könnte zwar niemals unmittelbar gekannt werden '[…]' aber er würde doch dem empirischen Charakter gemäß gedacht werden müssen, so wie wir überhaupt einen transzendentalen Gegenstand den Erscheinungen in Gedanken zum Grunde legen müssen, ob wir zwar von ihm, was er an sich selbst sei, nichts wissen.' (CPR 468, A540, B568)

'This intelligible character can never, indeed, be immediately known [...] it would have to be thought in accordance with the empirical character just as we are constrained to think a transcendental object as underlying appearances, though we know nothing of what it is in itself.'

Yet, elsewhere, Kant insists that it is known, as a fact about noumenal selves, that they are practically, and hence transcendentally, free. (See for example: CPR 476, A554, B582). Clearly, if I am transcendentally free, and so the uncaused cause of empirical effects, it must be true of me that I am a noumenal cause. This cannot just be a way of thinking of ourselves that we are forced to adopt to make sense of our moral language. If I am free in Kant's sense it must be true that I am a noumenal cause.

Kant realises that the 'no positive knowledge of noumena' strategy fails, so adopts this one;

'Man würde von ihm ganz richtig sagen, daß es seine Wirkungen in der Sinnenwelt von selbst anfange, ohne daß die Handlung in ihm selsbt anfängt.' (CPR 469, A541, B569)

'No action begins in this active being itself; but we may yet correctly say that the active being of itself begins its effects in the sensible world.'

I take it this means that the noumenal self is a cause of a variety of empirical effects, either without itself either beginning, ending, or changing. Is this supposition coherent?

I leave aside the conceptual problem about whether *c* may exist without *x* lasting at all. Let us assume that noumenal selves may exist whether or not they have duration. May a noumenal self, then, initiate spatio-temporal changes without changing itself? This seems to me not logically impossible, that is, its description does not entail a contradiction, but it is rather hard to imagine. We have to imagine that a noumenal self, call it 'x', does not come to exist, or cease to exist, and does not change, yet causes a series of numerically and qualitatively distinct events to occur in the empirical world.

The imaginative difficulties are these. If *x* does not change, there seems to be no possible fact about *x* which would make *x* initiate *e* '… *en*, rather than, say, *e*'...*e'n+1,* or initiate nothing at all.

If *x* is not temporal, may *x* initiate any temporal event series at all? Suppose *x* lasts, then that would allow *x* to last longer than t^1-*t3*, because *x* does not begin or end. If *x* lasts, in other words, *x* is eternal (ie, is, has always been, and will always be). It need not be a difficulty for Kant's theory that there are no facts about *x* that would explain why *x* caused *e*'...' *e'n*, or say, *ee*'...' *en*, rather than *e*'...' *en*. This is not just because if we are ignorant of noumenal facts we will be ignorant *a fortiori* of facts about the causal nature of noumena, but also because the thesis accords well with Kant's notion of free action being a kind of spontaneity.

Suppose 'explains' means 'explains deterministically'. Then, if we could, *per impossibile*, discover noumenal facts which explain why one empirical chain is initiated, or initiated rather than another, or none, then there is a sense in which we could no longer be talking about free actions at all.

Transcendental freedom is a very strong sort of freedom, on which there are no causes of free actions. If there could be a noumenal discovery of why I initiated a certain causal chain, then there would be a severe danger than such an explanation would be a causal one. This need not be a threat to a weaker libertarianism than Kant's, say where my freedom consists in my decisions and motives being (among) the causes of my actions, but Kant could not possible accept this, just because there are no causes of transcendentally free actions. The upshot is that there is no *a priori* objection to my noumenal self being an uncaused cause of a variety of empirical

events, even if such random spontaneity seems too much like chance to accord much with our ordinary intuitions about our freedom of action.

There is a different sort of problem about noumenal causality. If noumenal causes are not possible objects of experience, and if they can be known only intellectually, and if it is further true that the empirical effects of noumenal causes are temporally preceded by earlier empirical events (which are not their causes) what difference could there be between an empirical event which *spontaneously* initiated an empirical causal chain, and an empirical event that was *noumenally caused* to do the same?

The straightforward answer is that, in the first case, there is no noumenal causation and in there second there is. But what are the grounds for drawing such a distinction in our thinking? Kant realises the danger: When we claim the noumenal self 'of itself' initiates a causal chain

'und dieses würde gültig sein, ohne daß die Wirkungen in der Sinnenwelt darum von selbst anfangen dürfen' (CPR 469, 541, B569)

'we should not be asserting that the effects in the sensible world can begin of themselves.'

Clearly, it is a logical impossibility that any kind of effect may begin of itself, but not that any event should begin of itself. On a strongly anti-realist construal there would not appear to be any difference between an unobservable, and merely thinkable, cause of an empirical event, and no cause at all. Instead of facing this difficulty, square on, Kant makes this slide:

'weil sie in derselben Zeit, aber doch nur vermittelst des empirischen Charakters '…' vorher bestimmt, und nur als eine Fortsetzun der Reihe der Natururs achen möglich sind.'

'they ('the effects in the empirical world') are always predetermined through antecedent empirical conditions'

If 'determined' means 'caused' here then that is false, because if an empirical event is the first event in a causal chain, initiated by a noumenal cause, then that event cannot have any empirical cause. But if 'determined' means either of 'appears to be caused' or 'is temporally antecedent to' then the statement accords with the best reading of Kant's theory. Neither account, however, meets the anti-realist objection.

The way to meet it, on Kantian grounds, is to postulate freedom transcendentally, as the conclusion of the 'ought' argument above. Then the difference between 'we are transcendentally compelled to think of ourselves as free', and 'we are free', collapses. In other words, there is no conceptual room to argue for a sceptical difference between how the world is, and how we must conceive it.

Allison, Guyer, Irwin

In *Kant and the Claims of Knowledge* Paul Guyer levels a series of criticisms at Henry Allison's *Kant's Theory of Freedom* (Allison, 1990). I offer a way of reconciling the positions of Allison and Guyer. Guyer quotes Allison:

'to conceive of oneself (or someone else) as a rational agent is to adopt a model of deliberative rationality in terms of which choice involves both a taking as and a framing or positing. Since these activities, as expressions of spontaneity, are themselves merely intelligible (they can be thought but not experienced) it is necessary to attribute an intelligible character to the acting subject.' (Allison, 1990: 38)

and Guyer reports Allison's position:

'Because of their spontaneity, any acts seen as products of practical freedom (action according to a principle) must also be seen as products of transcendental freedom (action free of determinism)' (Guyer, 1992: 103)

Guyer accepts from Allison that

'Kant may well suggest such a position in the first Critique (see A 546-8 / B 574-6).'

but Guyer criticises Allison for endorsing Kant's argument, for the following reasons:

'1' 'Even so, it is clearly a poor argument. Conceptualisation and judgment may well be 'spontaneous' rather than 'receptive' in that they are no products of sensory stimulation alone but emerge only from internal cognitive processing of sensory stimuli; but that does not imply that there is

anything nonnatural or nondeterministic about such cognitive processing.' (Guyer, 1992: 103)

'2' 'although incorporating a desire into one's maxims is 'active' in a way that the mere occurrence of the desire is not, it hardly follows that this 'activity' itself cannot be ex plained within a naturalistic framework.' (Guyer, 1992: 103)

'3' 'Without a great deal more argument, the contrast between 'spontaneity' and 'passivity' is only an unredeemed metaphor that cannot support an inference to the real existence or even the mere possibility of a breach with determinism. But this contrast is all that Allison's defense of the spontaneity thesis amounts to; he does not attempt to prove that mental 'activity' is really incompatible with determinism.'
(Guyer, 1992: 103)

Guyer attacks Allison for adopting a two-aspects interpretation on Kant on phenomena and things-in-themselves:

'1' 'No disturbing metaphysics but only the comfortable conceptual distinction between two points of view is therefore necessary to justify Kant's assertion of the compatibility of empirical determinism and transcendental freedom.' (Guyer, 1992: 104)

'2' 'Kant doubtless suggests such an interpretation of his transcendental idealism at various points in the first *Critique* (e.g. B 307), but there are also places where his arguments actually imply the stronger view, not merely that our *concepts* of things in themselves exclude space and time but that the things themselves are not spatial and temporal (e.g. A 26 / B42). (Guyer, 1992: 104)

'3' 'It would seem to be the latter rather than the former interpretation that is necessary to make sense of his claims about noumenal freedom.' (Guyer, 1992: 104)

If the phenomenal self is not the noumenal self, and if the phenomenal self is determined but the noumenal self is free, nothing is thereby both determined and free. The inconsistency between freedom and determinism is therefore avoided, or remains only formal. The price of this reconciliation is high: I am two numerically distinct selves. Nevertheless, Guyer might not be right

that endorsement of the two-worlds ontology is necessary for a reconciliation of transcendental freedom with determinism. I agree with Guyer that the situation seems hopeless. However, consider this alternative strategy:

[A] Assume the double-aspect ontology (following Allison, Bird (1962) *et. al.*) and reject the two-worlds ontology.

[B] Assume that if *p* appears to be the case *p* might be true or *p* might be false.

[C] Consider what is *qua* appearance:

If an agent or action is considered:
(1) It appears determined.
(2) It is not determined.
(3) It is really free (even though not *qua* appearance, or *pace* appearances).

If anything that is not an agent or action is considered:
(1) It appears determined.
(2) It is determined.
(3) It is not really free (at any level).

[D] Consider what is as thing-in-itself:

If an action or agent is considered:
(1) It does not appear free but must be thought of as free.
(2) It is free.
(3) It is not determined (on any level).

If anything that is not an agent or action is considered:
(1) It does not appear so, *a fortitori*, does not appear free.
(2) It is not free.
(3) It is not determined.
(4) It is only minimally thinkable (eg. ‘that it is’ is thinkable, 'what it is' is not).

The interpretations amassed under 'C' and 'D' are mutually consistent, and so provide a coherent reading of Kant’s compatibilism which does not ascribe to him the two-worlds ontology. As with most compatibilisms, something

large has to be given up. In this case, it is the view that determinism is any more than apparently true of human agents. The price of this solution is that Kant is not really a determinist about human action and so not really a compatibilist about free will and determinism. On the other hand, his theory comes out as consistent and, to that extent, possibly true. If we take noumenal freedom seriously, the reading makes the theory sound quite Kantian. It is a reading of the *Critique of Pure Reason* which grounds the concept of freedom in the moral philosophy.

In this dispute, Terence Irwin agrees with Guyer:

'Now if an event is determined, it is true of it under all descriptions that it is determined, even though only some true descriptions […] show why it is determined.' (Irwin, 1984 quoted in Allison, 1990: 38, quoted in Guyer, 1992: 105)

and Guyer defends Irwin against Allison's reply. Here I adjudicate between, on the one hand, Guyer and Irwin and, on the other hand, Allison.

[1] Allison is wrong to argue that Irwin's attack

'rests on the assumption that being causally determined is a necessary or essential property of any occurrence of which it is predicated' (Allison, 1990: 44)

and to argue that

'the concept of causality is merely an epistemic condition, more specifically, a condition of the representation of temporal change.' (Allison, 1990: 44)

[2] Guyer is right to argue that:

'Irwin's point is simply that, if determinism is in fact true at all of humans as they actually are, then it remains true of them no matter how they are described, for leaving determinism out of some description does not make it false.' (Guyer, 1992: 105)

Irwin has pointed to a problem which appears fatal to the two-aspect solution to the problem of freedom. However, this problem can be circumvented. It requires giving up the assumption that agents, or actions, are determined on the Kantian view. Of course this invites the objection:

They are determined. They are *phenomenally* determined, therefore they are determined. But what does 'phenomenally determined' mean here? It means 'appear to be determined' and that is consistent with 'not determined'. This solution involves giving up the assumption that the inference from:

(1) *a* is phenomenally *F*

to

(2) *a* is *F*

is valid. As a matter of Kantian exegesis, this inference should be given up. One point of Kant's reiteration of the claim that there are appearances not just things in themselves is that appearances might not be veridical. We should not, conversely, assume they are always and everywhere non-veridical. The inference from

(1) *a* is phenomenally *F*

to

(2) *a* is not *F*

is also invalid.

Guyer says:

'Allison is forced to maintain that determinism must be true of our representations of ourselves but *false* or at least possible false of ourselves as we really are.' (Guyer, 1992: 105)

But we can give up the assumption that determinism is *true* of us at any level. Then the two-aspect solution is viable. Guyer also says:

'So Allison cannot maintain his interpretation of transcendental idealism yet answer Irwin's objection' (Guyer, 1992: 105)

The solution is to give up an assumption that allows the Irwin objection to be formulated. Give up: 'We are determined'. (5)

Freedom and Self-Knowledge

The distinction between phenomena and noumena is relevant to a view Kant holds about the epistemology of freedom; about how it is possible for one to know of oneself and others that we are free. Kant does not take the implausible view that people base their belief philosophical arguments for freedom, yet, rightly notes that most people are intuitive libertarians. Kant has two accounts of how it is possible for us to know we are free. One to do with pure apperception, the other to do with rationality.

(1) Freedom and Pure Apperception

There is a sense in which Kant thinks first person singular knowledge of freedom is epistemologically prior to third person knowledge of freedom. This is because he thinks there is an asymmetry between one way of obtaining knowledge that a person may possess with regard to himself and others, and another sort of access to knowledge that a person may possess only with regard to himself. This is not, at all, to say that there are two sorts of freedom. 'Freedom' means the same in 'I am free' as in 'He, she is free'. It is just to say that on one of Kant's accounts of knowledge, knowledge of sentences of the second form is only possible on condition there is knowledge of sentences of the the first form. I thus can only know ‘He is free’ is true if I know myself to be free. (This is part of what has come to be called Kant's 'Cartesian starting point'). Here is one means of acquiring knowledge about oneself and others. We could call it the observational route, and freedom cannot be known to exist only by its exercise:

'Wenn wir alle Erscheinungen 'des Menschens' Willkür bis auf den Grund erforschen könnten, so würde es keine einzige menschliche Handlung geben, die wir nicht mit Gewissheit vorhersagen und aus ihren vorhergehenden Bedingungen als notwendig erkennen könnten. In Ansehung dieses empirischen Charakters gibt es also keine Freiheit.' (CPR 474, A550, B578)

'If we could exhaustively investigate all the appearances of men’s wills there would not be found a single human action which we could not predict with certainty, and recognise as proceeding necessarily from its antecedent conditions. So far, then as regards this empirical character there is no freedom.'

Here is the first person singular privileged access to the fact of one's own freedom through pure apperception;

'Allein der Mensch, der die ganze Natur sonst lediglich nur durch die Sinne kennt, erkennt sich selbst auch durch Apperzeption, und zwar in Handlungen und inneren Bestimmungen, die er gar nicht zum Endrucke der Sinne zählen kann, und ist sich selbst freilich eines Teils Phänomen, anderen Teils aber, nämlich in Ansehung gewisser Vermögen, ein bloß intelligibeler Gegenstand.' (CPR 472, A546, B574)

'Man, however, who knows all the rest of nature solely through the senses, knows himself also through pure apperception; and this indeed, in acts and inner determinations which he cannot regard as impressions of the senses. He is thus to himself on the one hand phenomenon, and on the other hand, in respect of certain faculties the action of which cannot be ascribed to the receptivity of sensibility, a purely intelligible object.'

Despite his ‘empirical realism’, it is therefore not at all Kant's view that the observational route can yield everything that can possibly be known about a person, yet the various scientific methods of studying human beings are all constrained by it:

'Nach diesem können wir doch allein den Menschen betrachten, wenn wir lediglich beobachten, und, wie es in der Anthropologie geschieht, von seinen Handlungen die bewegenden Ursachen physiologisch erforschen wollen.' (CPR 474, A550, B578)

'It is only in the light of this (empirical) character that man can be studied - if that is to say we simply observing, and in the manner of anthropology seeking to institute a physiological investigation into the motive causes of his actions.'

First I make some remarks about Kant's statement of the observational route. It takes the form of an argument:

(1) If we could, in principle, obtain complete empirical knowledge of a person then we could predict that person's actions with certainty.

(2) We could, in principle, obtain complete empirical knowledge of a person.

(3) Therefore, that person is not free.

The argument, although frequently deployed by determinists, is invalid. The faulty inference is from (1) and (2) to (3). Let us accept the truth of (1) and (2) and the validity of the inference from (1) to (2) to (3).

The reason why it does not follow that a person is not free from the fact that that a person's actions may be predicted, even in detail, is that the basis of the prediction might, in fact, be a regularity with which that person exercises their freedom of choice. It then begs the question if the determinist insists that he has discovered all the empirical determinations of that persons behaviour. 'Appearances of mens wills' is vague in Kant's formulation but there is nothing about it which precludes those appearances being the regular effects of free actions. Indeed, this is really the conclusion Kant wants because he thinks there is both empirical knowledge of the appearance of determinism, and knowledge of freedom through pure apperception.

If I know I am free, then, this cannot be on the basis of my observing myself or others. In this sense 'freedom' is not an empirical concept (even though Kant allows it has uses in empirical sentences). Rather, through pure apperception, which is the capacity I have for rational, non-empirical, thought about myself, I can know that there exist 'inner determinations' which I cannot regard as sense impressions, and certain 'faculties of action' which do not belong to sensibility.

Kant is claiming that I have direct intellectual knowledge of the operations of my will. This is completely distinct from the findings of inner sense, of introspection. Pure apperception is rational or intellectual and contains no experiential dimension. Is there any evidence that such a capacity for knowing freedom exists?

In Kant's text there is no basis for a distinction between knowledge of the reality of freedom through the exercise of the faculty of reason in general and pure apperception in particular. I do think of myself as free but this is because of the conclusion of an argument about moral language, and the possibility of imperatives that extends to all persons, not just myself. I conclude that I in particular am free as an inference from the generalisation that persons are free. I do not acquire this knowledge on the basis of a private intellectual faculty that reveals to me only my own freedom.

There is then no evidence provided by Kant for the claim that I do have or exercise any such faculty. Pure apperception has not been given any convincing role in acquiring knowledge of freedom independently of reason in general.

Bennett on the Third Antinomy

I offer a reading of Kant on freedom which differs in several important respects from that offered by Jonathan Bennett Chapter 10 of his *Kant's Dialectic*. I point out some contrasts, and present grounds for preferring my reading.

Bennett thinks 'both sides of the third antinomy assume that whatever happens is caused to happen' (KD 184) but this leaves no room for Kant's concept of an ‘absolute spontaneity’, the idea that there logically might be an event which is a cause but not an effect, and, without crediting Kant with this concept in the thesis and proof of the third antinomy, little sense can be made of the doctrine of transcendental freedom. Indeed, it is because Bennett does not detect this notion in Kant's text that he thinks the thesis argument 'offers no positive doctrine of freedom' (KD 184) and ends on page 187: 'I confess defeat; I cannot discover how the thesis argument of the third antimony is supposed to work'.

What has gone wrong? Bennett notes correctly that Kant is discussing two putatively distinct forms of causality, one that operates in accordance with natural law, and one that allegedly operates freely, and is thus not constrained by natural law. This is what is right about Bennett’s reading: that freedom is a kind of causality. But it cannot be inferred from that fact, even when conjoined with the fact that there is causality in accordance with natural law, that there is no event that was not caused to happen. The slide is made from 'there are different sorts of causality, among them freedom' to 'there are no uncaused events', yet, clearly, complete determinism cannot be logically derived from an enumeration of the sorts of causality that might exist.

There are two passages where I think Bennett comes close to extracting a Kantian doctrine of freedom as 'spontaneity' from the thesis. He seizes on Kant's 'cause (which is) sufficiently determined *a priori*' (KD 185) but then confesses 'I do not know what that means, but it cannot mean the same a sufficient cause' (KD 185) and the reason he gives is 'for the latter points rather to the idea of a effect which is sufficiently determined'. What Bennett needs, in making sense of Kant's concept of freedom, is a distinction between two semantically distinct descriptions under which one and the same event might fall. Then we can say an event is both cause and effect: the cause of certain events which are chronologically subsequent to it, and the effect of certain events which are chronologically antecedent to it. Then it will still follow that 'caused cause' and 'caused effect' will not mean the same, but if we accept a distinction between sense and reference this need not matter much. There is, at least, then no incoherence introduced into the

idea of one and the same event being a necessary and sufficient condition for the occurrence of a second event, and some earlier event's occurrence being a necessary and sufficient condition for the occurrence of that first event.

This formulation (which I think is close to Kant's) leaves open two logical possibilities, one of which is essential to Kant's formulation of the thesis argument, but which is almost wholly neglected by Bennett, the other of which is not considered by either Kant or Bennett. The first possibility is:

(1) There might be causes which are not effects.

and the second is:

(2) There might be effects which are not causes.

We may leave aside the second as of little consequence for the free-will issue (even though it might be a useful concept for making sense of the end of the universe). The first, though, must be true if 'I am free' is true, on the Kantian interpretation of 'freedom'. Bennett perhaps allocates too much scope to 'nothing' in Kant's '[...] nothing takes place without a cause sufficiently determined *a priori*'. (KD 185, CPR 410, A446, B474) 'Nothing' here covers 'no event which is in principle a possible object of experience' and not 'nothing whatsoever'. This interpretation is born out by Kant's immediately prefacing it with; 'the law of nature is just this'. We know from the Transcendental Deduction, and the Principles chapter, that laws of nature hold for the phenomenal world and not for things in themselves, or things as they are in themselves. This leaves room for events considered at the level of things in themselves not to be effected, even if it is an all pervasive feature of the phenomenal world, of our experience, that causal relations invariably obtain between the events occurring in it. It also leaves open the logical possibility which Kant theory of freedom requires: that a noumenal cause may have a phenomenal effect, but I might not thereby myself be the effect of any cause. If we read Kant this way, we need not find his use of *a priori* puzzling here. Bennett says:

'The phrase 'a priori' seems not to carry the Kantian technical sense of 'independently of all experience, but rather the pre-Kantian sense of 'in advance' or 'independently.' (KD 185)

If we take seriously the claim that 'cause' is a category, then the sentence 'every event has a cause' is *a priori* in just the sense of *a priori* that Bennett

precludes here but which Kant intends. It is an *a priori* fact about appearances: about what our experience must be like, given the categorical framework we do posses. If *p* is *a priori*, in this sense, *p*'s truth is necessary for experience, so nothing in experience can refute *p*. But clearly this is in no way inconsistent with Kant's theory of transcendental freedom because it is an important part of 'transcendental' to mean 'not a possible object of experience' or 'not an item amongst appearances'.

So, if we accept a distinction between phenomena and things in themselves, and between the sense of 'cause' and the sense of 'effect', but assume a common referent for the two concepts, then Kant's doctrine of transcendental freedom may be extracted from the third antinomy in a way that is not open on Bennett's reading.

The other passage where Bennett comes close to such a reconstruction is on page 184 (KD) when he says the thesis contains mention of 'an alleged causality in which the cause does not antedate the effect'. This can be read in conjunction with Bennett's claim a little later that 'there is another kind of causality - that of freedom - in which an event's cause does not precede the event itself'. (KD 187) The first formulation leaves it open, as a logical possibility, that there could be an effect which predates its cause. Whether this is a possibility need not concern us here, as it does not concern Kant. It can also be read to mean much the same as the second claim however. Then it means: There could be a cause that was simultaneous with its effect. The trouble with reading this notion into the third antinomy is that it forces Kant to make an assumption which is quite absent from the text. Kant is not at all dealing with the notion of two numerically distinct but simultaneous events, one of which is the cause of the other, yet Bennett makes this constitutive of that kind of causality that is freedom. Now, it is difficult to see why anyone should make this part of the concept of freedom, and in fact Kant does not. From the fact that there exist two numerically distinct events *e* and *e′*, where *e* is the cause of *e′,* and yet where both *e* and *e′* endure only from some earlier time *t* to some later time *t′*, it does not follow that something constitutive of freedom in Kant's (or anyone's) sense has thus been described. On the contrary, it might be consistent with the strongest sort of determinism that some cause and some effect may be simultaneous. Why then does this in any way come close to Kant's intention?

Suppose an event is both a cause and an effect, in the admissible and Kantian sense advocated so far. Suppose, further, that two putatively numerically distinct events, a cause *c* and an effect *ef*, are, in fact, one and the same event '*e*', then it follows that *c* and *ef* last just as long as each other. Indeed, if it is a necessary condition for two putatively numerically distinct

events really being one and the same event that they exist over all and only the same time, then it will be necessarily true that *c* and *ef* exist simultaneously if *e*and *ef* are numerically identical. Bennett's account captures the possible simultaneity of cause and effect in Kant's theory, but not their possible identity. Still, I should not wish to make this identification of cause and effect constitutive of Kant's concept of freedom. On the contrary, that concept requires there to be some causes that are not identical with any effects. Noumenally, I am one of those.

A further strong contrast between Bennett's and the present interpretation is this. Bennett thinks the discussion of the possibility of a first event, in the thesis, is entirely irrelevant to the question of human freedom:

'If the concept of freedom solves problems about humans, then why invoke cosmology in its defence? If it cannot satisfactorily solve such problems, how could cosmology rescue it?' (KD 189)

There are two connections that Kant would wish to insist upon that Bennett neglects. The first is, by demonstrating the possibility of a first event that is a cause, Kant shows the logical possibility of an uncaused cause. This gives him, as coherent, just the concept he needs to demonstrate the possibility of human freedom. Admittedly, Kant need not have pursued this strategy. It does not follow that if there was an uncaused cause, which was the first event, then there are subsequent uncaused events which are the free acts of humans, initiating causal chains. But Kant does not claim the obtaining of that strong condition anyway. Kant's procedure in the third antinomy is rather the one outlined by Strawson:

'Once causality through freedom is admitted, it is permissible to entertain the thought that other causal series, occurring within the course of the world's history may be similarly initiated through freedom' (BS 208)

Or, as I should say, once the logical possibility of one uncaused cause has been demonstrated, then the possibility of uncaused causes or the coherence of 'uncaused cause' has been demonstrated. If we wish, we may take the discussion of the origin of the universe as a thought-experiment to make more palatable to us the logic of Kant's discussion of the freedom of the self.

The second connection relies on accepting transcendental idealism. Kant's phrase 'condition of appearances' is ambiguous between an empirical realist and a transcendental idealist reading. Empirically, a first cause is a condition of appearances, if those appearances could not exist unless

ultimately caused to exist by a first cause. But, transcendentally speaking, series of appearances could not exist (at least not *qua* appearances) unless they were presented to a perceiving subject, and for Kant this means a noumenally free, rational, subject.

Bennett has some interesting remarks on the solution to the Third Antinomy. Bennett thinks Kant's compatibilism reconciles freedom and determinism at the cost of making freedom vacuous. This is because if concepts

'are tools for the orderly management of intuitions, then the theory of noumenal freedom does not even make sense' (KD 194)

and

'That is the reconciling endeavour in a nutshell: the two sorts of causality cannot conflict because they cannot meet. One is confined to the empirical realm, whereas the other is banned from it' (KD 194)

Admittedly, if 'cause' is a category, and categories have only an empirical use, then it makes no sense to speak of noumenal, or purely rational, causality. But Kant is not so much of an empiricist as that. It is a pervasive feature of the *Critique of Pure Reason* that much of the language which ostensibly only has the legitimate role of making our experience intelligible to us is, in fact, given the extra role of exhibiting the transcendental conditions for that experience. Kant thinks freedom is amongst those conditions: In particular we need it to make sense of our moral experience. But it is the second quotation which points up the difference between Bennett's interpretation and mine. On my reading it is false that the two sorts of causality cannot meet. This is because Kant thinks there are noumenal causes with empirical effects: chains of empirical events which would not have happened without being caused by noumenal, free, rational, acts of will. Whether or not this theory is true, or in the last resort fully intelligible, is a separate matter from whether Kant subscribed to it, and the textual evidence is that he did. Bennett says about (A 572/B 580):

‘Kant implies that intelligible causality produces or mediates natural causality; but this – of which I can make nothing – seems not to be his considered view.’ (Bennett, 1974: 193 fn 15)

Intelligible causality is noumenal causality. Natural causality is determinism in the empirical world. I suggest that Kant means: The obtaining of intelligible causality is a necessary condition for the obtaining of natural causality. In particular, the way anything is, in itself, is necessary for its empirical appearance. I am not suggesting that Kant's intermittent endorsements of noumenal causality are consistent with his inclusion of 'causation' in the list of categories, any of which has only an empirical use. Therefore this perhaps *should* not be 'his considered view'. Nevertheless, if Kant allows a glimpse of the noumenal anywhere, it is in his theory of freedom: Noumenal ontology makes the phenomenal world possible. The *epistemological* order of priorities is the reverse: The intelligibility of noumenal freedom depends upon the applicability of the categories to phenomena.

Bennett calls up in support of his interpretation Kant's statement that

'The inevitable consequence of obstinately insisting upon the reality of the empirical world is to destroy all freedom' (CPR 565)

Again, Kant is not as empiricist as Bennett would like. We may, if we wish, read this with Bennett as asserting the existence of a neo-rationalist world of noumena, which are more real than empirical appearances. If we adopt a more restricted reading, however, as above, then Kant may be read as warning us not to assume too dogmatically that human actions are the inevitable outcomes of deterministic chains of empirical events which predate them. It might be that they just postdate them. (6)

3. Kant on the Freedom of Reason

Kant thinks that

'Daß diese Vernunft nun Kausalität haben, wenigstens wir un seine dergleichen an ihr vorstellen, ist […] klar' (CP.R 472, A547, B575)

'our reason has causality, or '[...'] we at least represent it to ourselves as having causality'

He has an argument for this, but before examining it we should decide what it means to say reason has causal efficacy.

Kant is talking about practical reason, not theoretical reason, as this statement a few pages later makes clear:

'Wenn wir aber eben dieselben Handlungen in Beziehung auf die Vernunft erwägen, und zwar nicht die speculative, um jenen Ursprunge zu erklären, sondern ganz allein, so fern Vernunft die Ursache ist, sie selbst zu erzeugen.' (CPR 474, A550, B578)

‘When we consider [...] actions in relation to reason - I do not mean speculative reason, by which we endeavour to explain their coming into being, but reason in so far as it is itself the cause in producing them.'

We need to locate this notion of reason within the theory of practical freedom, as interpreted so far. On that theory, if I am practically free then I am transcendentally free, and if I am transcendentally free then I am a cause but not thereby an effect of empirical events. If we now ask; 'what is that cause?' we can answer 'reason in its practical sense'. It would be tempting to straightforwardly substitute 'reasons' here and develop further the quasi-Davidsonian identification of reasons with causes, but that would be tangential to Kant's present concern. Rather, I freely cause empirical events to happen because I believe they ought to happen, and it is the existence of that imperative which is identical with the cause which is my free action. It is in that sense that reason itself has causal efficacy. If on Davidsonian lines we wished to say that is my reason for initiating a certain empirical chain, then there is nothing on Kant's theory to preclude our doing so.

Kant's argument for the causal efficacy of reason is this:

(1) 'Das Sollen drückt eine Art von Notwendigkeit und Verknüpfung mit Gründen aus, die in der ganzen NAtur sonst nicht vorkommt'
(CPR 472-3, A547, B575)

'‘Ought’' expresses a kind of necessity and of connection with ground which is found nowhere else in the whole of nature.'

(2) 'Diese Naturbestimmungen betreffen nicht die Bestimmung der Willkür selbst' (CPR 473, A548, B576)

'Natural conditions [...] do not play any part in determining the will itself.'

(3) 'Dieses Sollen nun drückt eine mögliche Handlung aus, davon der Grund nichts anders, als ein bloßer Begriff ist' (CPR 473, A547, B575)

'This ‘ought’ expresses a possible action the ground of which cannot be anything but a mere concept.'

(4) 'und nach denen sie so gar Handlungen für notwendig erklärt, die doch nicht geschehen sind und vielleicht nict geschehen werden.' (CPR 473, A548, B576)

'actions [might be] necessary even though they have never taken place, and perhaps never will take place.'

(5) 'Daß die Vernunft in Beziehung auf sie Kausalität haben könne'

'It (is) at least possible for reasons to have causality with respect to appearances.'

(6) 'Daß diese Vernunft nun Kausalität habe, wenigstens wir uns dergleichen an ihr vorstellen, ist aus den Imperativen klar, welche wir in allem Praktischen den ausübenden Kräften als Regeln aufgeben' (CPR 472, A547, B575)

'That our reason has causality, or that we at least represent it to ourselves as having causality is evident form the imperative which in all matters of conduct we impose as rules upon our active powers.'

Premise (1) is a version of Hume's claim that from no set of non-evaluative empirical sentences may a value judgement be logically derived or, as it is often put, that it is not possible to derive an 'ought' from an 'is'. The kind of necessity Kant claims in (1) is neither logical nor natural. This is made clear by (4), because if *p* is logically necessary then *p* could not fail to be true, but (4) suggests there might be necessities in some weaker, non-logical, sense which do not in fact obtain. Similarly, if the necessity were natural, then the 'grounds' of this necessity would be found in nature, and Kant explicitly denies this. The necessity is, in fact, imperative, the kind of necessity that may be logically derived from sentences containing 'ought'. Kant captures a feature of ordinary language use when he says 'ought' expresses this necessity, because we often say 'it is imperative you do X ', 'it is essential you do X '. (1) is a claim that the reasons why moral imperatives are true cannot be empirical. This clears the way for them to be purely rational and *a priori*.

(2) must be read to mean that, when a person wills, what they will has no empirical cause. This is part of the doctrine of practical freedom, and need not be taken to preclude the possibility of *akrasia*.

(3) divides into two. When Kant says ‘‘ought’ expresses a possible action’ he means there are embedded in moral imperatives descriptions of actions that may be performed, so that in sentences of the form 'A ought to θ', 'θ' designates some performance that A is able to execute. The second part of (3) is that the 'ground' of this possible action is a 'mere concept'. There are two ways of taking 'ground' here, both of which Kant intends. 'Ground' can mean 'non-empirical cause', and 'ground' can mean 'reason for'. The inference he wants, from (1), is that the non-empirical cause, and the reason for an action, may be identical with a 'mere concept'. If (1) and (2) are true then it is true that the 'ground' of a free action is not empirical. But does it follow that it is merely conceptual? This depends on whether 'empirical' and 'merely conceptual' designate jointly exhaustive categories, and, in particular, whether if *x* is not empirical then *x* is merely conceptual. It, at least, follows that if *x* is not empirical then it is possible that *x* is merely conceptual and, in fact, this all Kant needs for his purpose in (3) if we assume that if *x* is a reason (as well as a cause) then *x* must either be, or contain, some conceptual item. If 'ground' just meant 'cause' then there is nothing to force the conclusion that the ground must be a concept, but as 'ground' also includes the sense of 'reason', it is difficult to deny that if *x* is a ground in that sense then *x* must make sense, and if x makes sense, *x* contains concepts. Then 'mere' goes though easily as 'non-empirical'. The concept Kant primarily has in mind is 'ought', but if the ground of a free action is a reason, as well as a cause, then it would be safer to make that reason a sentence: either an imperative or a moral indicative, so that meta-sentences about a persons' free actions could be constructed with the form: '*a θ*'s because *a* ought to *θ*'. Here '*a* ought to *θ*' is a moral imperative, but it needs to be embedded in an indicative sentence, as a clause introduced by 'because', or some close synonym, in order for it to count as a sentence about the reasons for a free action. If someone wants 'the ground' in the sense of 'the reason' for a free action, then any such inquiry will take the logical form of 'Why did *a θ*?' and any coherent reply to that question could not be a mere concept, nor a straightforward moral imperative, but it would have to include a certain concept: 'ought' (or a close synonym), embedded in a moral imperative, embedded, in turn, in an indicative sentence as the 'because' clause. We can allow Kant this necessary background to 'mere concept' because there could not be concepts without a semantic role, and this

particular role is suggested by the logic of 'ground' as 'reason', and by ordinary language uses of 'ought'.

(4) is a truth about the imperative variety of necessity. It expresses the claim that, from the fact that 'X ought to happen', it does not follow that either 'X has happened' or 'X will happen'. Kant might well have added that it does not follow that 'X is happening' either, because that is also true. The omission is, no doubt, a slip. (4) is just a report about the logic of 'ought'.

Kant derives three conclusions from these premises, two weak and one strong. The weak ones are (5), and, 'we represent (reason) to ourselves as having causality', from (6). The strong one is 'our reasons has causality', also from (6). The first two go though without much difficulty. If there are non-empirical causes for actions, and these are accurately reported by indicative sentences in which imperatives are embedded, and if practical reasons exhibit the logic of 'ought', then it is true that 'it is possible for reason to have causality', and we think of our reason as having causality. These conclusions follow from our ordinary use of moral language; the first from the coherence of that language, the second from its role in how we think of ourselves as moral agents. The difficulty lies with the strong conclusion. From the fact that the members of some set of sentences, *S*'...' *Sn*, are each internally coherent, and also mutually consistent, it does not follow that any of them is true. Kant has yet to show that reports of moral imperatives are truth valued at all, and, still further, that a class of them take the value 'true'. Secondly, it has not been shown, either, what justifies our conception of ourselves as moral agents. Kant claims there are no empirical features descriptions of which might feature as premises of such a justification. But a great deal of further argument is required to show that the language or morals has a genuine application. To say that we need it to make sense of our moral behaviour is to be beg the question. There might be some alternative account of that behaviour under which it did not turn out to be moral, in Kant's sense, or even moral at all.

Reason and Freedom

There are two further important facts about ourselves as moral agents with which Kant wishes to acquaint us. These are: Reason is not deterministic, and reason is not determined. He has an argument for each of these. This is the one designed to convince us that we still perform our actions freely, even though they are caused by reason:

(a) The Argument for the Freedom of Reason

(1) '…da sie 'die Vernunft' im empirschen Charakter derselben (der Sinnesart) ganz genau bestimmt und notwendig ist. Dieser ist wiederrum im intelligibelen Charakter bestimmt'. (CPR 475, A551, B579)

'(the) empirical character (of reason) is itself determined in the intelligible character.'

(2) 'Die Handlung nun, in so fern sie der Denkungsart, als ihrer Ursache, beizumessen ist, erfolgt dennoch daraus nicht nach empirischen Gesetzen.'

'The action, in so far as it can be ascribed to a mode of thought as its cause , does not *follow* therefrom in accordance with empirical laws.'

(3) 'D.i. 'nicht' so, daß die Bedingungen der reinen Vernunft, sondern nur so, daß deren Wirkungen in der Erscheinung des inneren Sinnes vorhergehen.'

'It is not *preceded* by the conditions of pure reason, but only by their effects in the [field of] appearance of inner sense.'

(4) 'Die Kausalität der Vernunft im intelligibelen Charakter entsteht nicht, oder hebt etwa zu einer gewissen Zeit an, um eine Wirkung hervorzubringen.' (CPR 475, A551, B579-80)

'The causality of reason in its intelligible character does not, in producing an effect, *arise* or begin to be at a certain time. For in that case '[...]' its causality would be nature not freedom.'

(5) 'Wenn Vernunft Kausalität in Ansehung der Erscheinungen haben kann; so ist sie ein Vermögen, durch welches die sinnliche Bedingung eienr empirischen Reieh von Wirkungen zuerst anfängt' (CPR 475, A552, B580)

'If reason can have causality in respect of appearances, it is a faculty *through* which the sensible conditions of an empirical series of effects first begin.'

(6) 'Gesetzt nun, man könnte sagen: die Vernunft habe Kausalität in Ansehung der Erscheinung; könnte da wohl die Handlung derselben freu heißen.' (CPR 475, A551, B579)

'Granted, then, that reason may be asserted to have causality in respect of appearance, its action can still be said to be free.'

Premise (1) is intended to stall the objection to reason's freedom that it may be empirically determined, and hence itself be determining of its effects. Far from this being the case, Kant wants to claim, pure (non-empirical) reason itself determines its own 'empirical character'. There is, *prima facie*, an inconsistency in Kant's claiming that pure practical reason has any empirical dimension at all, but this inconsistency is an illusion. The claim is merely that there exists empirical evidence for the operations of pure practical reason, and this is not to claim that what is utterly non-empirical in itself is also empirical:

'So hat denn jeder Mensch einen empirischen Charakter seiner Willkür, welcher nichts anders ist, als eine gewisse Kausalität seiner Vernunft, so fern diese an ihren Wirkungen in der Erscheinung eine Regel zeigt, darnach man die Vernunftgründe und die Handlungen derselben nach ihrer Art und ihren Graden abnehmen, und die subjektiven Prinzipien seiner Willkür beurteilen kann.' (CPR 474, A449, B577)

"[...] the will of every man has an empirical character, which is nothing but a certain causality of his reason, so far as that causality exhibits in its effects in the (field of) appearance, a rule form which we may gather what in their kind and degrees are the actions of reason, and the grounds therefore.'

So, the only causal connection thus asserted is between the operations of pure practical reason and its effects.
(2) entails that this causal relation is not empirical. This accords with the doctrine of transcendental freedom, because, if it is true that a non-empirical cause has an empirical effect it is not clear that we should call the resulting causal connection itself 'empirical'. This is Kant's view, on the ground that the causal link, itself, cannot be subsumed under any empirical generalisation, for example, it cannot be captured by natural law. That is the sense in which reason's effects do not follow from it. There remain two residual senses in which the effect does clearly follow though. Firstly, the empirical event is, in some sense, still caused by reason. Secondly, although we cannot say that reason's causal efficacy is followed by an empirical effect, we may say that the agent's decisions, motives, and reasons, are formulated antecedently (in time) to the occurrence of what is, in fact, an empirical effect of pure-practical reason. This is the claim spelled out in (3).

Here 'their effects in the [field of] appearance of inner sense' I interpret as precisely the introspectively available motives, intentions, and reasons of the agent in acting. These are, themselves, further effects of pure practical reason and not the real causes of the agent's behaviour if he acts freely. So, although empirical motives, and decisions, for an action precede that action in time, pure practical reason does not precede it.

This is not duplicated by but is confirmed by (4) which is the claim that pure practical reason, despite its causal efficacy, is not itself in time. This is the force of its not arising, or beginning to arise, at a certain time. We may leave aside the question of whether it makes sense, in Kantian terms, to speak of pure practical reason lasting (eternally). Here he wishes just to establish that pure practical reason is not an event, or a happening, that may be available to, for example, inner sense, as a decision might be held to be. Kant rightly notes that if reason did possess causal efficacy, at some specifiable time 't', then it would have to be located within nature. This follows taking one conclusion of the Transcendental Aesthetic as an assumption*:* that there is only one time, and putatively distinct times are in fact parts of one and the same time.

(5) is an attempt to explain the sort of causal efficacy reason has over empirical events. 'Through' is crucially ambivalent but this much can be said about it: Reason may initiate empirical causal chains; it is 'through' it that an 'empirical series of events first begins'. Kant needs to give 'through' a content, and a theory of causal efficacy is just missing here, but the outline of one may be reconstructed. Two clues to this theory may be gleaned from the function it must perform in the theory of freedom in general. Firstly, the causation must not be empirical, and it must, therefore, not be the sort of causation that obtains in the empirical world. Secondly, the theory must leave room for *akrasia*, that is for a person's actions to be empirically determined against their will, even if in principle they could have done other than what they were empirically caused to do.

It seems to me that the only realistic option open to Kant here is a probabilistic one; that the exercise of my freedom makes more likely the occurrence of a particular empirical event than its non-exercise otherwise would have done. If we try to analyse the efficacy of the rational in terms of necessary conditions, then that precludes any sort of action performed freely being empirically determined. If we interpret this causality in terms of sufficient conditions its efficacy becomes deterministic, and the exercise of the will could never fail to be successful, no matter what the empirical pressure on the agent to act to the contrary. Kant does not anywhere, so far as I can see, overtly subscribe to a probabilistic account of the causal

efficacy of the rational will, but it is the only one that convincingly fills the lacuna left by premise (5).

Assuming (1)-(5) does (6) follow? In (6), I interpret 'it's action' as 'action caused by reason' and not as 'operation of reason itself'. If 'free' means 'not empirically determined' then that actions caused by pure practical reason are free follows from the premises. If 'free' means 'uncaused' then the conclusion does not go through because, *a fortiori*, if my actions are caused by pure-practical reason they are caused. If we ask 'Am I free?', and if the answer depends upon whether my actions are free, then 'I am free' is true on the first interpretation' but false on the second: unless that is, as Kant will wish to argue, I am free if I act from pure practical reason. Clearly also 'its action' could be read to mean 'reason's own operations', in which case a shift has been made to the question of whether reason is itself determined, and to that we should now turn.

(b) The Argument that Reason is Not Determined

(1) 'Vernunft 'ist' selbst keine Erscheinung und gar keinen Bedingungen der Sinnlichkeit unterworfen […]' (CPR 476, A553, B581)

'Reason is not itself an appearance, and it is not subject to any conditions of sensibility.'

(2) 'Auf 'd.i. die Vernunft' kann also das dynamische Gesetz der Natur, was die Zeitfolge nach Regeln bestimmt, nicht angewandt werden.'

'The dynamical law of nature, which determines succession in time in accordance with rules is not applicable to it 'reason'.

(3) 'So findet in ihr 'd.i die Vernunft' selbst in Betreff ihrer Kausalität, keine Zeitfolge statt.'

'Even as regards its causality there is no[t] time sequence in it.'

(4) 'Aber von der Vernunft kann man nicht sagen, daß vor demjenigen Zustande, darin sie die Willkür bestimmt, ein anderer vorhergehe, darin dieser Zustand selbst bestimmt wird.'

'Of pure reason we cannot say that the state wherein the will is determined is preceded and itself determined by some other state.'

(5) 'D[ie] reine […] Vernunft, welche mithin frei handelt, ohne in der Kette der Naturursachen, durch äußere oder innere, aber der Zeit nach vorhergehende Gründe, dynmamisch bestimmt zu sein.'

'Reason therefore acts freely; it is not dynamically determined in the chain of natural causes through either outer or inner grounds antecedent in time.'

The first conjunct of (1) is analytic, but the second is question-begging unless we may assume that that which is itself non-empirical is not subject to empirical causes. The trouble with accepting this is that Kant importantly allows that empirical events may have non-empirical causes. If there may be empirical effects of noumenal and rational causes, as the theory of freedom requires, a special argument is needed to show that the converse is not equally permissible; that there may be noumenal and purely rational effects of empirical causes. This is a conclusion that would have appalled Kant, but one that he fails to pre-empt.

(2), unpacked, is the claim that reason is non-temporal, natural laws are applicable only to what is temporal, so natural laws are not applicable to reason. That is unproblematic.

(3) follows from (2), if reason's causality is a part of what reason is, and there is no difficulty by 'some other state'. That would seem not to follow if, in principle, there could exist non-empirical, and non-temporal determinants of reason: acts of God, for example. From the fact that reason is not empirically determined it does not follow that reason is not non-empirically determined and we know, *ex hypothesi*, that there exist non-empirical determinants. If there are no non-empirical determinants of reason then (5) goes through: If reason is not empirically determined then it follows that reason is free. This is because of the meaning Kant gives 'free' in (5): 'not dynamically determined in the chain of natural causes'. Here natural causes are empirical causes.

The point about inner and outer in (5) is meant to preclude a person, if they act freely, being empirically determined to do so, either by their own inclinations, motives etc (which are 'inner'), or by some stimulus which is external to them (and so 'outer'). We may accept the inner-outer distinction here as corresponding to 'what may be detected by inner and outer sense' respectively. (7)

Notes

(1) Bennett comments on (B 474) ‘'…' the phrase is not ‘sufficient cause’ as both Schopenhauer and Kemp Smith imply, but rather ‘cause 'which is' sufficiently determined '*a priori*'’. I do not know what that means.’ (Bennett, 1974: 185) Both Bennett and Kemp Smith allude to Schopenhauer *The World as Will and Representation* trans. E. F. J. Payne (New York, 1958) pp. 497-8 (Kemp Smith, 1979: 493) Kant means *a cause which could in principle be discerned as the type of cause it is*. It is true *a priori* of such a cause that it is in principle discoverable because empirical. (It is legitimate to translate *bestimmt* as ‘determined’ but ‘determined’ (like *bestimmt*) is ambiguous between, on the one hand, ‘causally determined’ (‘causally necessitated’) and, roughly, ‘discriminated’ or ‘made out (as the thing or sort of thing it is). We need ‘determined’ in the second sense to make sense of ‘cause 'which is' sufficiently determined '*a priori*'’. Bennett is right that Schopenhauer and Kemp Smith have not read Kant accurately at this point but does understand what they have omitted.

By ‘cause’ Kant still means ‘sufficient condition’ here (even though that is not the meaning of the controversial phrase). A cause is contrasted with a state of the world which ‘in nowise follows from’ an earlier state in the Antithesis-argument. If we read Schopenhauer and Kemp Smith as right and Bennett as wrong then Thesis and Antithesis do come out as clearly mutually inconsistent, as Kant intends them. We do not have to go with their misreading because in the Antinomies Kant thinks of causes as sufficient conditions *anyway*. By going so far as to reject ‘cause’ as ‘is a sufficient condition for’ Bennett is forced to conclude: ‘I confess defeat: I cannot discover how the Thesis-argument of the third antinomy is supposed to work.’ (Bennett, 1974: 187)

(2) Bennett says:

‘According to 'Lewis-White' Beck, the Thesis-argument ‘is largely a repetition of the Aristotelian-Thomistic proof of the impossibility of an infinite series of causes and hence of the necessity of a first cause.’ (Bennett, 1974: 187)

There is a mistake here. Both Aristotle and Aquinas think there is no sound refutation of the claim that the regress of past events might be infinite. By ‘prime’ in ‘prime mover’ Aristotle means ‘necessary condition that has no necessary condition except itself’ and Aquinas thinks that an infinite God is needed to everlastingly sustain a universe that cannot be proven to have begun. That there was a first event (as implied by Genesis) is a matter of

faith, not proof, according to Aquinas. Indeed, Aquinas thinks he has a proof that there cannot be such a proof.

(3) Bennett rightly reports Kant's view that 'there is another kind of causality 'than natural causality' -that of freedom- in which an event's cause does not precede the event itself' (Bennett, 1974: 187) but does not report Kant's point that it is not simultaneous with it either. The causality of freedom is noumenal causality and any putatively noumenal causality is atemporal causality and so not simultaneous with anything. Of course Kant, in his better moments, thinks this barely coherent given that causation is a category and the categories only have empirical uses. (Bennett's footnote ascribes Kant's solution to the Third Antinomy to 'A' 560-86 but it is at A 460-86. This is no doubt a typographical slip. (Bennett, 1974: 187 fn 10))

(4) The answer to Bennett's question 'If the concept of freedom solves problems about humans why invoke cosmology in its defence?' (Bennett, 1974: 189) is that one putative cosmology entails that there was an uncaused cause (because a first cause that was not cause of itself). Of course, Kant thinks the problem of the origin of the universe insoluble (because it presupposes the erroneous doctrine of 'transcendental realism') but 'uncaused cause' might be exhibited as free from contradiction even if devoid of cosmological reference. It might then be given at least *prima facie* legitimate application in analyzing 'free'.

Bennett is too hard on Kant when he writes ''…' the alignment of Thesis/Antithesis with practical/theoretical is totally baseless.' (Bennett, 1974: 189) The freedom entailed by the conclusion of the Thesis is necessary for morality, and so practical reasoning. By 'theoretical' Kant means reason in its applications to the empirical world, which, according to the conclusion of the Antithesis, is deterministic.

(5) Bennett says:

'If Kant is right that thinking involves concepts and that these are tools for the orderly management of intuitions, then the theory of noumenal freedom does not even make sense.' (Bennett, 1974: 194) Bennett needs 'essentially' after 'thinking' and 'only' before 'tools', then his claim goes through. (I assume *p* 'does not make sense' here if *p* cannot be true or false.)

(6) Kemp Smith is right to argue that 'What '…' is to be said regarding the truth of the antithesis? If the principle of natural causality be formulated as

asserting that every event has an *antecedent* cause determining it to exist, then certainly, free, spontaneous and self-originating causality is impossible' (Kemp Smith, 1933: 494) if we read 'antecedent' to mean 'chronologically antecedent'. His claim becomes doubtful if we read 'antecedent' to mean 'prior' in the sense of 'necessary'. That some cause is a necessary condition for some effect does not *per se* entail that it is chronologically antecedent, even if we supply the assumption that all causes are events. Kemp Smith is right to say 'this antinomy is chiefly concerned with the problem of freedom, *i.e.* of spontaneous origination *within* the world.' (Kemp Smith, 1933: 495) Kant is trying to explain how our actions can possibly make a difference within the empirical causal order.

(7) The account of Kant on freedom I have offered provides a way of overcoming the dispute between Allison and Guyer on incorporation. (Allison, 1990: 50 ff., Guyer, 1992: 102 ff.) Allison thinks practical freedom presupposes transcendental freedom. There is some intuitive plausibility in Allison's claim that:

''incentives' do not motivate by themselves causing action but rather by being taken as reasons and incorporated into maxims.' (Allison, 1990: 51)

partly because that *a* has some incentive to θ does not look like a sufficient condition for a's θ-ing. We sometimes fail to act despite the existence of incentives to act and, sometimes, despite our knowledge of those incentives to act. What does Allison mean by 'incorporation'? He provides us with a necessary but not a sufficient condition for some act's being an act of incorporation when he says:

''An' act of incorporation as an expression of spontaneity, cannot be regarded as a causal consequence of antecedent conditions'. (Allison, 1990: 52)

Guyer says 'But this is clearly false' (Guyer, 1992: 105) Guyer's reasons are:

'1' 'the fact that a particular agent holds a certain principle and *has justifed* or *excluded* action on a given incentive on the basis of holding this principle is just as temporal a fact as the fact that a certain desire occurs in the agent. Holding a principle may be a disposition rather than an event, but it is still temporal.' (Guyer, 1992: 105)

'2' ‘*applying* or *deciding* to apply a certain principle to a certain desire may be as much of a temporal event as the occurrence of the desire itself.’
(Guyer, 1992: 105)

'3' ‘the incorporation of '…' incentives into our maxims or principles hardly removes our rational agency from the sphere of temporal discourse by itself.’
(Guyer, 1992: 105)

I fully agree with Guyer that all that is temporal. However Guyer thinks the choice we are left with is exhaustive:

‘We must either accept this result, whatever its implications for freedom, or defend the unpleasant doctrine of timeless events of noumenal choice.’
(Guyer, 1992: 105)

(I take it this is ‘unpleasant’ because ‘timeless event’ looks contradictory.)

I suggest there is a third possibility. Kant, like any coherent libertarian, is committed to the thesis that if I act freely I might have not acted. The possibility of not acting is not something temporal. I suggest it is this feature that Kant thinks is presupposed by actions (rather than mere events). It is a timeless modal property, not any kind of phenomenal or empirical property. With this assumption spelled out, Allison’s view goes through, immune to Guyer’s objections. Practical freedom presupposes transcendental freedom.

That there is free will is part of Rational Psychology (endorsed by Plato, Descartes, Christian philosophy, a priori metaphysics) but <u>not</u> repudiated by Kant (and possibly not recognized by him, as such, to be part of Rational Psychology). Is Kant’s compatibilism logically defensible? Freedom is *spared* in Kant’s attack on Rational Psychology. Is he entitled to spare freedom? His arguments for freedom should be the model for arguing about consciousness.

Existentialist Freedom

Sartre is usually misunderstood as having an exaggerated view of human freedom, no doubt because of the claims in the Existentialism and Humanism lecture that; there is no determinism, we are free, we are

freedom, we are condemned to be free. The only sense in which we are not free, it seems, is that we are not free not to be free. After the war Sartre caused outrage by saying that the French people had never been so free as during the Nazi occupation. In his play Men Without Shadows (Morts sans sepulture, 1946) French resistance fighters confront their own freedom in being tortured by Nazi collaborators. How can this be?

In Being and Nothingness Sartre draws a crucial distinction between freedom and power. Although my freedom is absolute my power may be severely constrained. There is no situation in which I do not have a choice, no matter how unpleasant. Indeed in Sartre's examples, the reality of choice is frequently agonising; a resistance fighter under torture may choose to betray comrades or remain silent for a moment longer. Freedom, for Sartre, is not comfortable it is a capacity to choose that never leaves us so long as we exist. Scientific determinism is a theoretical abstraction by the side of the lived reality of human dilemmas. Even if scientific determinism were true, it would be of no practical help to us in making our commitments.

Sartrean freedom cannot be understood without understanding the situation. (Sartre calls his volumes of literary, political and philosophical essays that appeared from 1947 Situations.) A human being is not separable from the human condition. A person divorced from the totality of their situations is an intellectual abstraction that can only be partly achieved. I am what I am only in relation to my situations. The totality of situations is the world and the kind of being that I have is being-in-the-world. What I make myself, is inseparably bound up with my projects, with my surroundings as I take them to be. Situations obtain in hierarchies: Sartre's being about to smoke depends upon the existence of smoking as a practice in mid-twentieth-century France. Keeping an appointment depends upon friendships or meetings. These in turn depend upon the existence of human beings, their projects and situations. All of these depend fundamentally upon being-in-the-world; the situation of all situations.

Sartre's concept of a situation is anti-Cartesian. Descartes thinks a person could in principle exist in abstraction from their physical and social environment and it makes sense to specify someone's mental states without reference to the ways in which those states are embedded in the world, without reference to what they are typically or paradigmatically about. Sartre's use of 'situation' and 'being-in-the-world' is sharply opposed to this picture. As a mental and physical agent what I do only makes sense if I am existentially related to an external and public world populated by other people who are similar agents.

In our unreflective taken for granted living we do not think of the situation as constituted by our freedom. It is my acquiescence in authority, rather than any objective constraint, that determines my behaviour. Once I recognise my freedom to disobey, to rebel, I am deconditioned. The fixed cognitive contribution of my acquiescence is stripped the world and the possibility of my changing it is opened up.

In Sartre's existentialism, human being and human situation form a mutually dependent totality. The relations between a human being and his or her situation are dialectical or reciprocal. The situation presents the agent with a range of possibilities. The agent acts to realise some of these possibilities and this action alters the situation and thereby presents a new range of possibilities. Agency constitutes both the agent and the situation. The situation only exists as a situation for some agent. The agent only exists as an agent in some situation so to be in a situation is to choose oneself in a situation. It follows that the relation between agent and situation is very close. The reciprocal relation is not only causal. It is not even only constitutive. Agent and situation may only be adequately understood as two aspects of one reality. Sartre does not put it this way, but it is as though the agent is the inside of the situation and the situation is the outside of the agent.

In order to reconcile this dialectical relation between agent and environment with Sartre's absolute libertarianism we need to invoke his distinction between freedom and power. Although our freedom is absolute, our power is limited. Although there is no situation in which we do not have a choice, there is no situation which does not limit our power. Sartre spells this out clearly in the 1947 essay Cartesian Freedom (La Liberte Cartesienne in Situations I) when he insists that the situation of a person and their powers can neither increase or limit their freedom. Although what I can do is limited by where as well as when I am, that I can do something rather than nothing is in no way affected. I retain the dispositional property of being a choosing agent even though which choices I may exercise varies from situation to situation. Clearly some choices may be unpleasant to me but, logically, an unpleasant choice is nevertheless a choice. The expression 'I had no choice' is misleading.

The theme that freedom is unimpaired by constraints on power pervades Sartre's literature. Sometimes his characters are horribly constrained: the tortured resistance fighters in Men Without Shadows, Mathieu and his comrades trapped in the clock tower in the 1949 volume of The Roads to Freedom; Iron in the Soul. As their power is reduced their awareness of freedom increases.

In Sartre's existentialism, the recognition of freedom is a lonely first person singular phenomenon for which recourse to others provides no respite. For example, also in Iron in the Soul, Sartre has Odette shift swiftly from the first person plural thought 'What ought we to want?' to the first person singular thought 'What ought I to want ?' (185) against the background 'situation' of the May 1940 invasion of France. Odette is expressing the ethical tenet of Being and Nothingness that 'It is I who sustain values in being'. Sartre did not write 'It is we who sustain values in being'. For all his repudiation of Descartes in 'Cartesian Freedom' the primacy and inescapability of the first person singular exercise of, and confrontation with, freedom remains thoroughly Cartesian.

Sometimes, the existence of freedom is depicted as dependent upon its acknowledgement or recognition by the agent. For example, in The Flies Sartre has Zeus say of Orestes 'Orestes knows that he is free' and Aegistheus replies
'He knows he is free? Then to say hands on him, to put him in irons, is not enough' (1). Although, as we shall see, Sartre thinks there is a pervasive human tendency to deny one's own freedom, it is the fact of a person's freedom not their knowledge of it that makes freedom unconstrained. Freedom is
entailed by knowledge of freedom but not vice versa. An agent aware of their freedom can act authentically.

Sartre endorses Heidegger's view that we are 'thrown' into the world. We are but we did not choose to be. Seemingly inconconsistently with this, he says in Being and Nothingness that in a sense I choose to be born. Clearly, any kind of Platonic pre-existence is out of the question here. Sartre thinks it is false that we predate (and post-date) our empirical existence. Drawing a distinction between existence and essence, Sartre means that what my birth is, or is to me, largely depends on how I freely think of it. Its significance is the significance I bestow upon it. Freedom does not predate existence. Freedom is existence, and in it existence precedes the essence we freely choose.

In Being and Nothingness a person is their freedom. Sartre identifies the upsurge of freedom, choice, and the person himself, as one and the same being. One existent is subsumed under three descriptions. I do not have my freedom. I am it. The will has no role is the exercise of Sartrean freedom. The moment the will operates, the decision is already taken. Sartre's libertarianism entails that human actions are unpredictable. The only respect in which I am not free is that I am not free not to be free. I am not able not to choose.

We could refrain from action, or omit to act. Would this not be a way of escaping one's own freedom? Sartre's position is that refraining from action presupposes the choice not to act. This is what refraining is. There exists an infinity of actions I am not performing. I am only refraining from doing some of them. In Iron in the Soul Sartre has Ivich and Boris agree about the French soldiers caught up the May 1940 invasion of France 'they chose to have this war' (p69). They did nothing to prevent it.

Sartre's believes those who live in the developed countries are causally responsible for the death, by starvation and malnutrition, of those who live in the third world. To fail to save life is as causally efficacious and as morally culpable, as to actively take life. This kind of reasoning leads Sartre to justify political violence by, or on behalf of, oppressed groups, for example in the Preface he wrote for Frantz Fanon's The Wretched of the Earth (Les Damnes de la Terre 1961) and to support the Baader Meinhof gang in the early 1970's. Sartre denies that the distinction between our acts and our omissions marks a distinction between what we are and are not responsible for.

How is Sartre's libertarianism to be reconciled with his post-war Marxism ? A human individual retains the capacity to choose whatever their situation, whatever the constraints on their power. Our power is constrained because we are alienated. He endorses the view of the early Marx that members of capitalist society are psychologically estranged from their work, the products of their work, nature, and each other. This alienation is an obstacle to the construction of what Sartre would call a free society: a society we would freely choose rather than historically inherit. Our dispositional capacity to choose continues to ontologically differentiate us from naturally occurring objects and artefacts even though we are denied the power to create a free society in Sartre's sense.

Freedom is exercised in history. It is not an option for me to freely act in the situation of a later nineteenth century German coal miner if I am a mid-twentieth century French intellectual. My historical location opens for me a range of actions I may perform but there is an infinity of actions which are closed. There is a dialectical dependency between freedom and truth. There is no truth without freedom and no freedom without truth. A human being is free but at the same time in bondage; a chooser whose power is politically and historically constrained. It is only at the moment of death that a human being is complete. Before death a brave person could become a coward or a coward could become brave. Only death brings an end to freedom.

Jean-Paul Sartre Altona and Other Plays: Altona, Men Without Shadows, The Flies (Penguin, in association with Hamish Hamilton, Harmondsworth, 1962) p. 292.

RESPONSIBILITY

Sartre maintains that ethical values are invented, not discovered. He thinks there is no consciousness so no divine authority on the distinction between right and wrong, and it is an act of bad faith to endorse a pre-established value system such as Christianity, humanism, or Communism. Rather, each person is radically free to create their own values through action. Ethics is something that exists only within the world of things human. Indeed, in the Existentialism and Humanism Lecture (Chapter I above) he says there is no universe except the human universe and we cannot escape human subjectivity. We cannot look outside our lives to answer the question of how to live. We can only do that by freely choosing how to live.

Superficially, Sartre might appear to be a naive relativist about morality. Relativism in morality is the thesis that it makes no sense to speak of some actions as right and some wrong, only of some individual or some society holding them to be right or wrong. Relativism embodies a mistake. From the obvious and uncontroversial historical truth that value systems vary from person to person and from society to society it is invalidly concluded that these systems cannot themselves be right or wrong. It is important to refute relativism because, although it is sometimes misidentified as a liberal and tolerant doctrine, it in fact precludes our condemning individuals or regimes that practice genocide, torture, arbitrary imprisonment and other atrocities. On the relativist view these practices are, so to speak 'right for them but wrong for us'; a putative claim that makes no sense.

Sartre's moral philosophy opens a conceptual space between absolute consciousness-given morality on the one hand and naive relativism on the other. He insists that values belong only to the human world, and that we are uncomfortably free to invent them, yet he provides us with strict criteria for deciding between right and wrong.

The essential concept in the establishment of this middle path is responsibility. To say that someone is responsible for what they do is to say that they do it, they could have refrained from doing it, and they are answerable to others for doing it. (This last component of 'responsibility' is apparent in the word's etymology. It means 'answerability'.) It is a consequence of Sartre's theses that existence precedes essence in the case of

humanity, and people have an ineliminable freedom, that we are responsible for what we are. We are nothing else but what we make of ourselves. It follows that everyone is wholly and solely responsible for everything they do.

Responsibility for Sartre includes another, crucial, dimension. In choosing for myself I am implicitly choosing for others. By joining a trade union, by joining the communist party, by getting married, by becoming a Christian, by fighting in the French resistance, by anything I do, I am implicitly prescribing the same course of action to the rest of humanity. To put it another way, all my actions are recommendations. By acting I set an example for all similarly placed others to follow. I am obliged at every instant to perform actions which are examples.

This implicit recommendation to others is called in moral philosophy 'universalisability', and finds its most sophisticated expression in Kant's ethical works; the Groundwork of the Metaphysics of Morals (Grundlegung zur Metaphysik der Sitten, 1785) and the Critique of Practical Reason (Kritik der Praktischen Vernunft, 1788). Kant, like Sartre, tries to found an objective morality that does not rely on theological premises. In Sartre's texts universalisability admits of two interpretations; one causal, the other logical.

On the causal interpretation we take literally Sartre's notion of setting an example. By joining a trade union I may cause others to join a trade union, and so my responsibility is in a direct sense a responsibility for what I make others do, not just for what I do myself.

On the logical interpretation, in order to be consistent we have to accept that persons similarly placed to ourselves should do as we do. A person is only one person amongst others and it would be inconsistent to maintain that one person but not others should follow a course of action where all those people are similarly placed. There would be something incoherent about someone who freely chose to join a trade union, or who became of convert to Christianity, but disapproved of people making just those choices. Of course Sartre accepts that that may happen. One form of religious or political commitment might be suitable for one person but not another but, prima facie, if it is right for one person to do something then it is right for any similar and similarly placed person to do the same. Being just one person rather than another cannot make a moral difference.

The causal and the logical interpretations are mutually consistent. For any action, say joining the French resistance, it may both be causally efficacious in encouraging others to join, and exhibit the rule that if I join then I do not

judge similarly placed others to be under no obligation to join, on pain of inconsistency.

Consistency is a condition for ethics according to Sartre. Acting immorally, that is, in a way that cannot be universalised, results in incoherence. Following Kant, Sartre says the act of lying implies the universal value which it denies. Not only is there no lying without truth telling but lying cannot be universalised. The implicit recommendation to everyone to lie could never be adopted. If there was no truth telling the distinction between lies and truth would break down and there could be no lying either. Because consistency is a constraint on morality what cannot be universalised is immoral. In fashioning myself I fashion humanity as a whole.

Universalisability provides us with a test to distinguish between the rightness and the wrongness of our actions. If an action cannot be consistently universalised then it is immoral. If the action can be consistently universalised it is not immoral. In trying to resolve a moral dilemma, we have to ask what the consequences would be of everyone adopting our action as a rule.

Realising the full burden of our responsibility to humanity provokes in us the deepest sense of dread and anxiety. This discomfort why we plunge ourselves into bad faith. Facing our freedom requires facing our responsibility. We can hardly bear to face our responsibility so we deny our freedom.

We are free and responsible despite our refusal to accept these objective facts about us. They endure through our pretence so we are in anguish.

In this way Sartre emerges as a moral objectivist despite his rejection of theological premises for ethics. His moral philosophy is in many ways a humanistic transformation of Christian ethics. To take one conspicuous example, instead of being responsible before consciousness a person is responsible before humanity. Instead of consciousness watching our every action everything happens to each person as though the whole human race was watching what they are doing. Sartre's humanity, like Christian humanity, is a fallen humanity but Sartre's fall is a secular fall. We are not fallen from any perfect natural state; we fall short of our own possibilities of acting freely and responsibly. To admit this freedom is to become committed (engage).

BAD FAITH

The reality of our freedom is so unbearable that we refuse to face it. Instead of realising our identities as free conscious subjects we pretend to ourselves that we are mechanistic, determined objects. Refusing to freely make ourselves what we are, we masquerade as fixed essences by the adoption of hypocritical social roles and inert value systems. This denial of freedom is called by Sartre 'bad faith' (mauvaise foi). Almost a secularisation of the Christian Fall, bad faith is pervasive.

It is depicted in merciless detail in Sartre's fiction and in the chapter on bad faith from Being and Nothingness, reprinted below, which contains the locus classicus: a cafe waiter whose exaggerated movements and affected manner make it clear that he is playing at being a cafe waiter. Several kinds of bad faith are displayed by the waiter. He behaves mechanically as though he were a thing rather than a person. He is acting a role, playing a part. His relationship to himself is as false as that of an actor to his part in a play. His behaviour is a display before others, a set of routines which make him comfortable in his own eyes and in the eyes of others.

In another of Sartre's examples, the soldier at attention is in bad faith when he turns himself into a mechanical soldier-thing with a fixed unseeing gaze. A woman on a date with a certain man for the first time is presented with the moment of choice. The man takes her hand. For her to leave her hand in his is to choose a sexual direction for the relationship. To withdraw it is to reject this possibility. Instead of choosing, instead of exercising her real choice, she refuses to face it, leaving her hand rests neither accepting nor rejecting: a thing.

Sartre's philosophical literature is strewn with characters in bad faith: Goetz in The Devil and the Good Lord, Hugo in Dirty Hands, the bourgeoisie of Bouville in Nausea and Kean in the play of that name. In the biographies Baudelaire and Saint Genet: Comedian and Martyr the self-justifying bourgeois hypocrisy of the nineteeth century poet is contrasted with the recognition of freedom by the thieving and streetfighting proletarian homosexual playwright from Brest. Genet's disreputable and criminal behaviour is eulogised by Sartre as a model of good faith; the real exercise of freedom. Baudelaire's law abiding conformity is condemned as the denial of freedom; bad faith. Sartre's distinction between the moral and the immoral cuts across socially acceptable legal and ethical mores. The moral is the free and authentic; the immoral the conformist, the obsequious, the inauthentic.

Sartre says bad faith is a lie to oneself. This raises the philosophical paradox of self-deception because I know I am free but I hide my freedom

from myself . In some sense, I both know and do not know I am free. How is this possible?

Sartre rejects one solution straight away: the psychoanalytical idea that there exists both a conscious and an unconscious mind. The Freudian allows that we know something unconsciously but remain ignorant of it and deny it consciously, and so dissolves the paradox of self-deception. Sartre cannot possibly follow this route because it is a central tenet of his theory of consciousness that no unconscious exists. Sartre's phenomenology implies that every mental state is necessarily a conscious state.

Sartre's solution is as follows. The respect in which I know I am free is different from the respect in which I do not know I am free. I know that I am free in that I have the capacity to make choices. However I mask this capacity from myself by the adoption of everyday roles, by conforming to the fixed image others have of me, by pretending to be a mechanism or a thing. I am fully possessed of the propositional knowledge of my own capacity to act freely but behave rigidly to prevent the realisation of that capacity. I pretend I am not free.

In bad faith I am in relation to myself as the actor is Hamlet. We are all actors. An actor knows he is an actor but in so far as he performs he is not his real self. In bad faith I know I am free but adopt a role which masks my freedom. Bad faith is a representation for others and for myself. Paradoxically, human reality is what it is not (its authentic self-defining project) and is not what it is (its hypocritical social role).

Sartre distinguishes between two kinds of people in bad faith. One kind he calls ‘cowards’ ('les laches'). They hide from their freedom, in a facade of solemnity or with deterministic excuses. Those who deny not only their own freedom but that of others Sartre calls ‘swine’ (' les salauds').' In Nausea, for example Roquentin concludes his tour of the portraits of the bourgeois officials in the city museum with the comment 'you bastards' (‘salauds’). They felt they had the natural or consciousness-given right to exist, to occupy their social location of wealth and privilege and suppress the freedom of others. The denial of freedom is immoral because it is inauthentic and hypocritical. Freedom brings with it a heavy and terrible responsibility described in the last chapter (IX). Bad faith is also therefore an evasion of responsibility.

Reprinted here are the chapter on Bad Faith from Being and Nothingness and an extract from Anti-Semite and Jew. Sartre regarded anti-semitism as a paradigm case of bad faith.

Merleau-Ponty

1 Freedom and Necessity

Merleau-Ponty says 'It is inconceivable that I should be free in certain of my actions and determined in others' (PPT 434). (1) Now, this is not obviously right. Whether true or not it seems at least coherent to suppose that some of the actions are performed freely while others are necessitated by causes. This does not seem to be a self-contradictory supposition.

However, there is at least one definition of each of 'freedom' and 'determinism' on which it comes out as contradictory to hold that one and the same action is both free and determined. If an action is performed freely then it was possible not to perform it, and if that action was determined it was not possible not to perform it. It would be incoherent to maintain that one and the same action both could have not been performed and could not have not been performed. On the plausible supposition that logically impossible states of affairs are inconceivable, such an action is inconceivable.

There seems no incoherence and so no inconceivability in the idea that some token action could have not been performed but some, numerically distinct, token action could not not have been performed by the same agent.

Merleau-Ponty does have the makings of an argument to show that it is at least false that some of a subjects actions are free and some determined.

Merleau-Ponty maintains that freedom is a presupposition of action. He talks for example about 'freedom being anterior to all actions' (PPT 437). He does not make explicit the reasons for this presupposition but an argument may be brought in its favour.

There exists a distinction between two kinds of event; events which are happenings and events which are the doings of human subjects. One, and perhaps the only, way of marking the distinction between on the one hand what only happens and on the other hand what both happens and is done is to say that all and only what is done is performed freely.

Irrespective of the force of this distinction Merleau-Ponty is right to maintain that it is inconceivable that actions should be determined if freedom is a pressuposition of action. This inconceivabily rests on the inconceivabily of the compatibility of freedom and determinism when those two theories are defined as mutually inconsistent.

The view that it is false that I am free in some of my actions, but not all, because I am determined in others relies on establishing that I am free in all my actions.
I am free in all my actions, freedom and determinism are mutually inconsistent theories, therefore, it is not the case that I am determined in any of my actions. As it stands the argument is just valid.

It is sound so long as freedom and determinism are incompatible and the first premise can be proven.
However, Merleau-Ponty's formulation is less clear-cut. He Says;

1. 'Once I am free, I am not to be counted among things' (PPT 435).

2. 'I must then be uninterruptedly free' (PPT 435)

.

therefore

3. 'One cannot be to some extent free' (PPT 435)

.(3)

I read this this way:

1. I am free and if I am free I am not a thing (and only things are determined).

2. Therefore any action of mine is free (and there is not a time when I am not free).

therefore,

3. I am not free in some of my actions but not others.

Construed this way, with surpressed premises added, Merleau-Ponty has a valid argument. Again, its soundness is only as strong as its premises and the first premise in particular needs argument.

Merleau-Ponty sometimes talks as though the subject is free because the subject is not a thing ('une chose') but he sometimes talks as though the subject is not a thing because the subject is free. For example he says

'In order to be determined ... by an external factor, it is necessary that I should be a thing' (PPT 434). (4)

The implication is that I am not a thing and so not determined. However, 1. above was 'Once I am free, I am not to be counted among things' (PPT 435) the implication this time being that I am free and not determined and so not a thing.

There is no incoherence in this, just circularity. It could be that if x is free then x is not a thing and if x is not a thing then x is free. However, in the argumentative order of priorities, Merleau-Ponty either needs to show that the subject is not a thing and derive the freedom of the subject from that, or, he needs to show that the subject is free and derive the conclusion that the subject is not a thing from that.

It is the first course that Merleau-Ponty adopts.

He takes it that he has shown by his phenomenology of the body (see chapters 2 and 3 of the present work) that a person is a concscious bodily subject, not a mechanical physical object. In Merleau-Ponty's view it is not so much being physical which in Merleau-Ponty's view is incompatible with being causally determined as being a subject and being conscious. Merleau-

Ponty regards consciousness as being essentially, or necessarily, free. He says:

'Nothing determines me from outside, not because nothing acts on me, but, on the contrary, because I am from the start outside myself and open to the world' (PPT 456).
(5)

Here Merleau-Ponty draws an important distinction between causation and determinism. Something 'acts on me' (PPT 456) ('me sollicite' PP 520) if and only if something other than me causes some action of mine with a probability more than zero but less than one, but something 'determines me' (PPT 456) ('me determine' PP 520) if and only if something other than me causes some action of mine but with a probability of one, that is, with necessity.

Merleau-Ponty rightly implies that causation is not logically sufficient for determinism, only logically necessary.

Merleau-Ponty characterises our existential situation in two ways. He says 'I am from the start outside myself' and he says he is 'open to the world' (PP 520). We know that these two characterisations carry the implication that I may be caused but not determined to act.

Merleau-Ponty has no argument for the view that the causes I may be subject to cannot be determining causes. A priori if A causes B then there seems no incoherence in the supposition that if A happens then the probability of B is one. Determinism may be false, but Merleau-Ponty has not proved that the subject is of such a nature that it can be caused but not determined.

When he says he is 'open to the world' he means he is a consciousness that exhibits intentionality. When he says he is 'outside' himself he means he is an agent whose behaviour is teleological, or goal directed.

Some considerations may be adduced for the view that a teleological consciousness may not be wholly determined, even if caused. For example a teleological consciousness alters its own behaviour in the light of perceived and changing goals and so to that degree is self-determining. If A is self-determining then that is potentially logically inconsistent with the complete determination of A by causes outside A. If every action of A is an

event necessitated by some cause extraneous to A then that leaves no logical room for A to be, or to be even among, the causes of A's behaviour.

The concept of overdetermination is not incoherent. Suppose E has two causes, C1 and C2, each of which is sufficent for E and each of which occurs. If C1 occurs then E occurs and if C2 occurs then E1 occurs but if either C1 but not C2 occurs or C2 but not C1 occurs then E nontheless occurs.

A potential incoherence in the present case arises from the special circumstance of A's being (or being among) the causes of A's action.

Suppose C is some external cause of A's action. But suppose further that A is the cause of A'sactions. It follows that A may be the cause of some action of A's only if A is the cause of exactly what C causes to happen. It is incoherent to suppose that C and A are both the cause of a single event and thereby cause something qualitively distinct to happen.

Where overdetermination is true then Merleau-Ponty may coherently maintain that there is both external causation and self determination of an action. The inconsistency arises only where the subject determines his or her actions in directions other than effected by external causes. Merleau-Ponty could avoid this by pointing out that the external causes of self-determining actions are not determining.

2 Self-Determination

We should examine more closely Merleau-Ponty's notion of freedom as self-determination. He says: 'we have indeed always the power to interrupt' (PPT 438).(6) We have the ability to intervene in the course of our own actions.

Suppose I perform some set of (mental or physical) acts then, on this view, it is within my power to halt these at any point and to intitate (or resume) some other set of actions. It is just this possibility Merleau-Ponty has in mind when he says this power to interrupt 'implies in any case a power to begin ' (PPT 438).(7)

That we have the power to intervene in the course of our own actions. It must at least be an empirical and not on a priori fact because it is not

something I can know about myself independently of my experience of myself. If all knowledge is either a priori or empirical, but not both, and if we know we intervene in our own actions, then we must know this empirically.

Merleau-Ponty attempts to refute a number of objections to the possibility of human freedom. They are the ideas that my actions determine me, my temprament determines me, the fact that I am subject to weakness of will, and finally, that there are physical constraints on my action.
On the first point, Merleau-Ponty says:

'If, as if often said, motives incline me in a certain direction, one of two things happens: either they are strong enough to force me to act, in which case there is no freedom, or else they are not strong enough, and then freedom is complete and as great in the worst torments as in the peace of one's home' (PPT 435).

(8)

It is wrong that one's motives are, or are amongst the causes of one's actions should be considered a constraint on one's freedom, If I act from my own notives, rather than, for example, those of another or from some external cause then that is arguably part of what my being free consists in. There can be little doubt, however, that when Merleau-Ponty says 'Motives incline me in a sense' he means the agent's own motives.

Merleau-Ponty next presents two possibilities which he takes to be mutally exclusive and collectively exhaustive. The first possibility is motives 'have the force to make me act' and in that case 'There is no freedom' and the second possibility is that 'they do not have it' and in that case 'It [freedom] is complete'.

Suppose some motive is a cause of some action of mine, then Merleau-Ponty is saying if the motive occurs then either it necessitates or does not necessitate the action. If it necessitates the action then I do not have the option not to perform that action. If the motive does not necessitate the

action then I have the option either to perform it or not to perform it. Construed that way Merleau-Ponty has said something true.

To see that this is right, consider the case where if the motive obtains then it necessitates the action. Then the action cannot not occur given the motive and it would be contradictory to maintain that the action does not occur, on the grounds that the agent refrains from performing it. However, if the motive does not necessitate the action then there is no contradiction in the supposition that the agent performs the action or refrains from performing the action.

These arguments are sufficient to show that Merleau-Ponty has presented two possibilities that are mutually exclusive. Logically, he is right that if my motives are determining then I am not free, but if my motives are not determining then I may be free. Merleau-Ponty jumps to a conclusion too hastily when he says freedom is complete. Clearly, my motives being non-determining is logically consistent with my actions being determined by causes other than my motives.

We should examine now the question of whether the two possibilities are collectively exhaustive. They are if and only if, if I am not determined then I am free and if I am not free then I am determined (or both, but that has been ruled out). (Two possibilities P, Q, are collectively exhaustive if and only if if not P then Q and if not Q then P (or both P and Q)).

Suppose some motive is a cause that determines some action of mine (it has 'la force de me faire agir' PP 497) and so if that motive obtains then the probability of my acting (by that action) is one. Suppose, however, conversely, no motive of mine determines any action of mine, and that my freedom is complete ('elle est entiere' Phen 497). It is logically right that if my action occurs with the probability of one then that action does not occur with any probability less than one. However, if I am free then that action occurs with some probability less than one. This is because if an action is performed freely then it could have not been performed and that is only possible if the probability of its occurence is less than one. Clearly, in any case, anything that happens either happens with a probability of one or a probability of less than one. This is sufficient to show that Merleau-Ponty is right to maintain that the two possibilities are collectively exhaustive.

When Merleau-Ponty says that if his freedom is complete it is 'as great in the worst torments as in the peace of one's home' he is suggesting that one's freedom may admit of degrees.

The form of the argument here is this: If I am free then my freedom does not admit of degrees. My freedom does admit of degrees, therefore, I am not completely free.

Merleau-Ponty takes it as intuitive that there exist degrees of freedom; that a person under torture is not as free as a person in the peace of their home. That is intuitive. The question is whether the idea of degrees of freedom is coherent.

Suppose the probability of A's doing x is less than one but greater than zero, then it follows that A's doing x is not determined and that (on the present analysis) is a necessary condition for A's doing x freely. It is logically consistent with the probability of A's doing x being less than one that A not do x, and it follows from that that A is free. To model the idea of degrees of freedom we need the idea of an action of A's being partly caused by A and partly caused by causes external to A. This is an idea Merleau-Ponty accepts. Then we have two causal chains issuing in x: one chain external to A and one chain internal to A.
If the probability of x's occurence is greater given both chains than just the chain external to A then A's freedom to do x is greater than if the probability of x given the external chain were not raised by the internal chain.

Merleau-Ponty next considers the possibility that my actions are determined by my temprament and so I am not free. Prima facie it is a mistake to construe this as an objection to freedom. If my temprament is part of what I am, and if my being free partly consists in my being (or being among) the causes of my own actions, and if my temprament causes (some of) my actions, then to that extent I am free, not determined.

Merleau-Ponty's reply to the objection is:

'My temprament exists only for the second order knowledge that I gain about myself when I see myself as others see me. In so far as I recognise it, confer value upon it,

and in that sense, choose it' (PPT 435).
(9)

There are several claims here. One is the quasi-Hegelian thesis that part of what I am is a product of a self-reflection that employs concepts gleaned from a third person or 'objective' perspective on the person. One idea is that one's self-conception partly consists in thinking of oneself as others perceive and think of one. Another is: this is all that my temprament consists in. Both claims need argument.

It seems empirically right that the self-conception each of us has is partly modelled on the conception we each have of others, partly on the conception others have of us and partly on the conception we have of the conception others have of us. These thoughts have been examined at great length by Hegel and thoughts rather like them have been expressed by Nagel.
(10)

Merleau-Ponty's other claim that my temprament 'exists only ' for this knowledge is harder to prove because my having third person or 'objective' epistemological access to my own temprament is logically consistent with two further facts or putative facts: I may have first person singular access to my own temprament (or at least part of it) and, my temprament may cause whatever kind of epistemological access I have to it. The existence of temprament is logically independent of knowledge of it. Merleau-Ponty wishes to rule out these latter two possibilities but he has no argument which does so.

Nor does Merleau-Ponty have an argument for the neo-Sartrean view that I 'choose' my temprament.

'I recognise it, confer value upon it, and in that sense choose it' (PPT 435)
(11)

From the fact that I recognise and value my temprament it does not logically follow that I choose it. That is consistent with, and arguably entails its prior existence.

If I choose my temprament then Merleau-Ponty is right to maintain that my temprament may still be among the causes of my actions, but he misses the point that they may be free. If A chooses A's temprament freely then A might not have had that temprament (because A may have chosen otherwise). Given that, A's action may not have had A's actual temprament as its cause. If A freely chooses the cause of A's own actions then that rules out their being externally determined.
Then, if A perfoms them, they are performed freely.

Merleau-Ponty next considers weakness of will as a putative objection to freedom. Normally the philosophical problem of weakness of will is thought to be the difficulty of making sense of the fact that we sometimes fail to do what we want to do when we are able to do it and even when we think it is right to do it. Merleau-Ponty has rather a different problem in mind when he uses the expression 'weakness of will' (PP 436) ('l'impuissance de la volonte' PP 498):

'And indeed, although I can will myself to adopt a course of conduct and act the part of a warrior or a seducer, it is not within my power to be a warrior or seducer with ease and in a way that 'comes naturally'; really to be one, that is' (PPT 436).
(12)

He is considering the case where someone tries to be a particular sort of person by sheer act of will. There is an important insight here. There would seem to be very little that I can do just by an effort of will. It is a quasi-Wittgenstein thought that I cannot move my arm off the table just by an act of will: I move my arm by moving my arm.(13) Merleau-Ponty makes an analogous point here: by willing myself to be a person of a certain kind I do not thereby become that kind of person. As Merleau-Ponty puts it:

'Neither should we ssek freedom in the act of the will, which is, in its very meaning, something short of an act' (PPT 436) (14)

Merleau-Ponty is right not to identify freedom with a kind of inner trying. Freedom is doing with the possibility of not doing and that is consistent with his conception of the body-subject as being-in-the-world.

Merleau-Ponty finally considers the putative physical obstacles to freedom. He maintains that the existence of freedom is a necessary condition for the existence of constraints on freedom. For example he says:

'Even what are called obstacles to freedom are in reality deployed by it. An unclimable rock face, a large or small, vertical or slanting rock, are things which have no meaning for anyone who is not intending to summount them, for a subject whose projects do not carve out such determinate form from the uniform mass of the in itself and cause an orientated world to arise - a significance in things' (PP 436).
(15)

Although these are Sartrean thoughts that Merleau-Ponty presents, Merleau-Ponty does not avail himself of a very important Sartrean distinction which is of use to him here. Sartre distinguishes my freedom from my power.

For Sartre, my freedom is my capacity to make choices and according to Sartre everybody is always free in that sense. However, Sartre thinks there are many things, includings physical things, that are not within my power. (Many commentators on Sartre miss this distinction and ascribe to him an exaggerated view of freedom).
(16)

Merleau-Ponty agrees with Sartre that the physical world is perceived by the subject as a set of means and obstacles to the fulfilment of that subject's projects. The world is perceived, in the Heideggerian sense, as 'ready to hand'. (17)

In what sense do constraints on freedom presuppose freedom? When he says the constraints are 'deployed' by freedom he means the subject constitutes the objects as constraints and does so freely, that is, with the possibility of not doing so. Merleau-Ponty recognises that from the fact that a physical constraint is constituted by consciousness it does not follow that

it is not a physical constraint. These constraints on freedom, the one's that freedom deploys are, however, the only constraints on freedom. He says:

'There is, then, ultimately nothing that can set limits to freedom, except those limits that freedom itself has set in the form of its various initiatives' (PPT 436).(18)

It follows that the subject is partially self-constraining. The exercise of freedom ironically generates its own constraints. This is true of every action of the subject, not just physical action but also perceptual synthesis. This is why Merleau-Ponty says 'the subject has simply the external world that he gives himself' (PPT 436).(19)

If, as Merleau-Ponty maintains, the external world is constituted by the projects of the subject then a kind of Kantian constitutive idealism is true, albeit of a highly pragmatic kind. This raises the question of what the world is really like, independently of its constitution through the projects of human subject. Merleau-Ponty's answer is this:

'The world in itself ... without freedom, would be merely an amorphous and unnameable mass' (PPT 439).
(20)

Freedom is a condition for synthesis and synthesis is a condition for the world being as it appears to a subject so freedom is a condition for the empirical world.

Interestingly, although Merleau-Ponty is a phenomenologist he is committed here, more or less overtly, to the existence of a Kantian thing-in-itself. Husserl and Sartre, following Hegel, claim that phenomenology is possible without the thing-in-itself.(21) This view is certainly mistaken and Merleau-Ponty is more consistent than the other phenomenologists when he sees the need to postulate some reality that is conceptually constituted. If we ask 'what reality?' then that question is as metaphysically awkward for the phenomenologists as for Kant. Of Merleau-Ponty's answer one may ask 'How could this be known ?'

3 Compatibilism

Merleau-Ponty's considered view on freedom and determinism is a version of compatibilism. He holds that we are caused, but free, and not determined but constrained:

'What then is freedom? To be born is both to be born of the world and to be born into the world. The world is already constituted, but also never completely constituted; in the first case we are acted upon, in the second we are open to an infinite number of possibilities. But this analysis is still abstract, for we exist in both ways at once . There is, therefore, never, determinism and never absolute choice' (PPT 453).
(22)

In the subject's being-in-the-world there is a reciprocal relationship between subject and world. The world constitutes the subject but the subject constitutes the world. The subject makes the world what it is and the world makes the subject what it is.
This is consistent with human freedom because in constituting the world it is open for the subject to act otherwise. Although the world constitutes the subject the world does not necessitate the actions of the subject. Although there is causation, there is no determinism.

Merleau-Ponty says his account is 'still abstract' (PPT 453) 'encore abstraite'(PP 517). Two things are abstract if they may be separated in thought but not in reality. Although we may conceive of the world as acting on the subject without thereby conceiving of the subject as acting on the world (and vice versa) we are not thereby conceiving a real possibility. The reciprocal dependence between subject and world is existential . It is instantiated in the existence in the world of a human subject: in being-in-the-world.

It is the inextrability of the subject's being from being-in-the-world, and the inseparability in reality of the constitution and the constituting of the subject that has this result:

'It is impossible to determine precisely the 'share contributed by the situation' and the 'share contributed by freedom' (PPT 453).(23)

Ee should not confidently stipulate this a priori . If we can devise sufficiently precise criteria for the individuation of causes and effects the contribution of subject and world may be empirically discerned as separate.

Merleau-Ponty rightly thinks his compatibilism is consistent with his thesis that a person is a bodily subject. He says 'I am never a thing and never a bare consciousness' (PPT 452).(24)

However, caution is needed here. Consciousness with freedom, the physical world with determinism should not be too closely assimilated. The existence of consciousness (or 'pure consciousness' ('conscience nue') is logically consistent with determinism, and the existence of the physical world (or even only a physical world) is logically consistent with the existence of freedom.
(25)
er and above my thoughts, but I am certain that I am thinking at least so long as I

ble theses about the relation between the cogito and time. Each includes a differe

4 The Refutation of Determinism

Merleau-Ponty's arguments against determinism in La Structure du Comportement are deeply embedded in his attempted refutations of behaviourism and materialism in psychology and philosophy . Perhaps because of this, he tends to assimilate these three views when they need to be carefully distinguished.

Determinism is the conjunction of the theses that every event has a cause and caused events are made inevitable by their causes.

Behaviourism in psychology is the thesis that the appropriatly scientific subject matter of psychology is behaviour. Behaviour is scientifically

explicable because it is publically observable and environmentally determined.

Behaviourism in philosophy ('analytical behaviourism' or 'logical behaviourism') is the thesis that mentality may be reduced to behaviour: either overt and occurrent bodily and verbal behaviour or dispositions to behave. Here 'reduced' means: a's are reducible to b's if and only if any sentence or set of sentences about a's may be translated, without loss of meaning, into a sentence of set of sentences about b's.

Materialism is the philosophical thesis that everything that exists is physical.(26)

Merleau-Ponty thinks, rightly, that Behaviourism in psychology logically entails determinism and he thinks, wrongly, that materialism in philosophy logically entails determinism. He also tends to assume that behaviourism in psychology is a kind of materialism which, in a very loose sense, it is. It only deals with physical subject matter. It would be wrong, however, to hold that it entails materialism.

To understand the kind of determinism that Merleau-Ponty seeks to refute we need to understand what he means by 'cause'. He says

'The cause is the necessary and sufficient condition of the effect considered in its existence and in its nature' (SCT 161).(27)

This means that C is the cause of an effect E if and only if if not C then not E and if C then E. Merleau-Ponty misses that if C is necessary for E then E is sufficient for C and if C is sufficient for E then E is necessary for C and that if C is necessary and sufficient for E then E is necessary and sufficient for C. This is damaging to his definition of 'cause' because it obliterates the distinction between cause and effect. On some ontology it might be right to obliterate this distinction but on his ontology there is a distinction between cause and effect.

There are several ways of retrieving Merleau-Ponty's position. A distinction could be drawn between on the one hand the conditions necessary for the occurrence of E considered singularly, call them the members of the set 'C': {c1...cn}, and, on the other hand, the conditions necessary for the occurrence of E considered jointly, call them just the set 'C'. Then we may advance the thesis that each member of {c1...cn} is singularly necessary for E but {c1...cn) are jointly sufficient for E. Then we read ' cause' as 'necessary and sufficient condition' as ambiguous between 'each of the conditions considered singularly' and 'all of the conditions considered

jointly'. Then 'necessary condition' is ascribes a property to each member of {c1...cn} and 'sufficient' ascribes a property to all of {c1...cn}.

Alternatively, Merleau-Ponty could simply drop the claim that the 'cause is a necessary condition' but retain the claim that it is 'sufficient' 'suffisante'. This would have two advantages. It is the putative fact that causes are sufficient conditions for their effects that is necessary for the determinist thesis that caused events are inevitable. So, by adopting this strategy, Merleau-Ponty rules out an essential component of determinism. If he drops the thesis that causes are necessary conditions for their effects then he logically drops everything that is entailed by this, including the consequence that effects are sufficient conditions for their causes. This drops too much because Merleau-Ponty thinks there are causes, but not determining causes.

Merleau-Ponty says that a cause causes both the existence and the nature of an effect. If C is the cause of E then C is both necessary and sufficient for E being and being what E is.

Merleau-Ponty's objections to determinism are of three main kinds: He argues that the essentially situated being of human beings, their being-in-the-world, is inconsistent with determinism. He argues that there are no on-one correlations between single human actions as effects, and occurrences in the external world as causes and, finally, he argues that to regard human actions as the results of determining external causes is to wholly disregard the role of what is internal to the organism in making its actions what they are. These three lines of argument are not always clearly distinguished from one another in Merleau-Ponty's text.

The relations between a body-subject and the world in being-in-the-world which are prima facie causal are not to be understood as mechanical 'mechanique(s)' (SC 174) but as dialectical 'dialectique(s)' (SC 174)

'In describing the physical or organic individual and its milieu, we have been led to accept the fact that their relations were not mechanical but dialectical' (SCT 160).(28)

We need to know now exactly what the mechanical/dialectical distinction consists in.

Some cause, C, and some effect, E, are a mechanical cause and a mechanical effect if and only if both, C determines E but E does not determine C, and every component of C that is causally efficacious on E stands in a one-one mapping with some resulting component of E. Merleau-Ponty means that if C determines E then E does not determine C in the case

of mechanical causation when he says that 'the dependence is uni-directional' (SCT 160-1) 'la dependence est a sens unique' (SC 174) and he asserts the one-one mapping between causal parts and resultant parts when he says

'The cause and the effect are decomposable into real elements which have a one-to-one correspondence'
(SCT 160).(29)

Some elements, A and B, are dialectically related if and only if both A determines B and B determines A and if neither A nor B is analysable into discrete causally efficacious components which stand in a mutual one-one correlation. If it makes sense to speak of dialectical relations as causal this is because they are holistic not atomistic, and reciprocal not one-way.

It is also part of the concept of a 'dialectical relation' that dialectical relations are dynamic not static. This means that if A determines B and B determines A such that A and B make each other what they are, then A is not only altered by B but A is altered by B qua determinant of B and B is not only altered by A but B is altered by A qua determinant of A. It follows that relata that are dialectically related perpetually alter one another qua determinants of one another, so long as the dialectical relation obtains. It further follows that it makes sense, derivatively, to speak of A making A what A is via B or B making B what B is via A because if A determines B qua determinant of A then A determines A by determining B qua determinant of A and if B determines A qua determinant of B then B determines B by determining A qua determinant of B.

It follows from these definitions that no relation is both mechanical and dialectical because it would be contradictory to affirm either that some relation was both one-way and reciprocal or both constituted by relata that admitted of analysis into causally efficacious components and did not, or both. It follows that if a relation is mechanical it is not dialectical and if a relation is dialectical it is not mechanical. It has been shown, then, that the two kinds of relation are mutually exclusive (but not that they are collectively exhaustive of relation types or even causal relation types).

5. The Dialectic of Freedom

We need to examine next Merleau-Ponty's grounds for repudiating mechanistic explanations of human behaviour and endorsing a kind of dialectic.

In his view there is a sense in which it is false that identifiable parts of the human being respond to external stimuli even though it makes sense to speak of human actions as responses to stimuli. Merleau-Ponty is aware that a person's eyes may be made to dilate by the shining of a light into them, or a person's blood-pressure may be raised by diet or stress but he does not count responses of this type as actions. Action, or behaviour, is something performed by the whole person and not by parts of the person. This is what Merleau-Ponty means when he says 'physical stimuli act upon the organism only by eliciting a global response' (SCT 161).(30)

Although Merleau-Ponty fully allows that actions performed by the whole person may have external causes and even allows that 'a global response will vary qualitatively when the stimuli vary quanitatively' (SCT 161) (31), nevertheless, such external causes are not deterministic causes. They are what he calls 'occasions' (SCT 161) 'occasions' (SC 174).

Merleau-Ponty does not define 'occasion' when he says of such putative external causes 'they play the role of occasions rather than that of cause' (SCT 161)(32) but a definition may be suggested. As we have seen, a deterministic cause is a sufficient condition for an effect. Whether a given effect happens depends upon whether a deterministic cause of it happens. There is nothing incoherent in the idea that any given effect should have more that one deterministic cause because the concept of overdetermination is free from internal contradiction. Deterministic causes determine whether their effects happen because if a deterministic cause of a event happens that event happens but if no deterministic cause of an event happens then that event does not happen. If a deterministic cause happens then something sufficient for an effect has happened. If no deterministic cause has happened then nothing sufficient for an effect has happened.

I suggest that if some event is correctly called the 'occasion' of some effect then it determines not whether that effect will happen but when it will happen. Occasions are triggers or preciptants of events which are the effects of causal chains that lie elsewhere. They just determine the timing of events and it does not much matter whether we decide to apply or withold the name 'causes' from occasions so conceived. (Merleau-Ponty contrasts occasions with causes).

According to Merleau-Ponty it is the human agent who is the cause of his or her own actions and external causes at most determine the timing of those actions. External causes do not determine whether or how an agent will act

and in so far as they are causes of actions at all, this depends upon how they are 'taken' by the agent: the meaning the agent attaches to them. For example, Merleau-Ponty says

'The reaction depends on their vital significance rather than on the material properties of the stimuli. Hence, between the variables upon which the conduct actually depends and this conduct itself there appears a relation of meaning, an intrinsic relation' (SCT 161).(33)

It follows that for Merleau-Ponty no token physiological input to the organism is causally sufficient for the occurrence of any token action of that organism qua agent. Why should we believe this ?

Merleau-Ponty regards deterministic psychology as a scientific failure. It has proved impossible to predict human actions on the basis of a knowledge of external stimuli. Merleau-Ponty presumably thinks that if there were one-one correlations between stimuli types and behaviour types then it would have been possible to predict human behaviour given knowledge of its causes. The fact that such predictions have not been forthcoming suggests that the determinisitic picture is a failure.

Given the failure of deterministic psychology, Merleau-Ponty may be read as providing an explanation of that failure. If something internal to the organism mediates between physiological input and behavioural output then we would not expect to be able to predict human actions given knowledge only of stimuli external to the organism. What is internal to the organism makes a difference to physiological input, behavioural output, and the relations between them.

The way the agent reacts to an external stimulus depends on the 'significance' 'signification' or 'meaning' 'sens' (SCT 161, SC 174) the agent bestows on it. What is that ?

A stimulus is never presented to an agent qua physiological stimulus, but always qua part of the human world, or what Merleau-Ponty calls 'L'ordre humain' (SC 175) 'the human order' (SC 162). Following Husserl, Merleau-Ponty says the agent perceives the environment as made up of 'use-objects' (SCT 162) '<<objets d'usage>> (Gebrauchobjekte)' (SC 175). The world is a human constituted world, not a purely natural world. It follows that human beings are not made what they are by purely natural causes. Self and world are mutually constituting.

V

Free Will and Neuroscience

How is Mental Causation Possible?

The problem of Overdetermination:

A neurological event is sufficient for a physical event.

A mental event is sufficient for that physical event.

This gives overdetermination because:

If P then P1

If M then P1

so if not M, then if P then P1 and if not P then if M then P1 which is overdetermination.

The solution is:

A mental event can override a physical cause of a physical event.

If P then P1 unless M, and if M then P2.

ie M cancels the effect of P, even though P happens. How is this done?

An analogy: A ship is steered by a helmsman, using wheel, rudder. The helmsman can also control the speed of the vessel. Suppose the helmsman were to do nothing. The ship would move anyway, propelled by the current. Mental causes = the helmsman. Neurological causes = the current.

Breathing goes on anyway. One can overide this and decide to breath faster, slower, more deeply etc.

Thinking goes on any way. I can choose to think this or that. I can decide to do this or that. Moving my arm. I can deliberately move it. My heartbeat: I cannot control this (without special training in yoga).

The neurological is (empirically) necessary for the mental.

The neurological is sometimes sufficient for the physical.

The neurological is not always sufficient for the physical. Sometimes it is overridden by the mental.

M -> P

-P -> -M

We know the physical is necessary for the mental.
It follows that the mental is sufficient for the physical.

I can mentally intervene in the course of my actions.

Suppose there is mental causation, then how is this possible? This is the Q. here. (We are not doubting that there is mental causation here.)

Do we know beyond a shadow of a doubt that the brain is a deterministic system? Suppose some probabilisitic account of its causal connections is right. Then we have:

Neurological event P1 causes neurological event P2 with some probability greater than zero but less than one. Suppes in Probabilistic Metaphysics argues that causation is probabilistic. Apply this to neurological connections. Popper in A World of Propensities and other places argues that probabilities are objective (not just degrees of belief and ignorance or Baysian). Also, the attempts to model the brain in terms of quantum theory might well yield a probabilisic or partly indeterministic account of neurological causal connections.

Suppose the neurology of the brain is indeterministic (but obviously not utterly random). Then we could have a chain of causally related neurological events:

P1 makes P2 likely, P2 makes P3 likely etc.

P1 is not associated with any mental event. P2 is necessary for mental event M2, so M2 is sufficient for P2. M2 clinches it that P2 happens – mental causation.

M2
suff
I
v
P1 prob->P2 prob-> P3

We know that the neurological is necessary for the mental. It follows as a matter of logic that the mental is sufficient for the neurological. If the neurological is not determined by the neurological then there is not overdetermination if the mental determines the neurological. The mental has a causal role. Intuitively: we can intervene in our own actions.

What makes M2 happen? We do not know that there are sufficient conditions for mental events. Intuitively (or scientifically) we assume that nothing happens without sufficient conditions for its occurrence.

Suppose the reverse picture applies. Suppose the occurrence of some prior mental event M1 makes the occurrence of M2 likely but there is not determination of the mental by the mental. Suppose then there is some physiological input that causes in turn some neurological event that clinches it that M2 occurs. Then we have:

M1 prob-> M2 prob –>M3

P1 prob-> P2 prob -> P3

but, say, P1 determines M2.

But then, if P1 determines M2, and M2 determines P2 then P1 determines P2 (via M2). But then the brain is deterministic after all in a sense. Maybe the brain needs the mental to complete it as a deterministic system.

It is better to think of these psycho-physical connections as webs rather than chains. Look at neural nets.

The Indeterminate Mind: How is Mental Causation Possible?

(Ted Honderich has refuted the Davidsonian program on mental causation: The pears make the scales go down because they weigh so many pounds, not because they are green.)

How is the Problem of Other Minds Possible?

One's own mind is unusual.
You have a mind.

Mind and Matter: A Difference of Degree?

Thoughts are mental.
They have: intentionality, phenomenology, syntax, semantics
They are events or sets of events.

Brain processes are physical.
They have spatio temporal locaction.
They are events or sets of events.

Experiences are both mental and physical.

Eg a pain has both phenomenological content and spatial location.
No semantics.
No syntax.

Awareness
Semantics
Intentionality
Phenomenology (content) Syntax (form) [equiprimordial]
Neurology
Cellular biology

Matter shades into mind.
Matter blurrs into mind.
Nothing mental is physical and nothing physical is mental. (Even though people are both mental and physical.)
Analogy: nothing black is white. Nothing white is black. There is grey in the middle.

Thought: mental
Experience: both mental and physical – not as mental as thought and not as physical as matter.
Matter: physical.

Mind and matter are on the same continuum. (Of course, leaves many questions unanswered.)

It is because the difference between mind and matter is one of degree that there is no interface problem.

PSYCHO-PHYSICAL CAUSATION

Psycho-physical causation takes place if a mental event causes a physical event or a physical event causes a mental event or both. Although not beyond metaphysical doubt, it is empirically plausible and essential to common sense that psycho-physical causation takes place. I shall assume that it does take place and offer a solution to the problem of how it can take place.

Three obstacles to psycho-physical causation seem intractable:

1) If mental events are not physical events it is hard to see how causal relations could obtain between events of such radically different types.

2) If mental events are physical events and only cause physical events because they are physical events then they do not cause mental events because they are mental events.

3) If mental events cause physical events because they are mental events, whether they are physical events or not, then the physical universe is not a closed system and the third law of thermo-dynamics is false.

We do not know that mental events are physical events but we do know that physical events are necessary for mental events. That mental events are physical events although in principle falsifiable is not verified and possibly not verifiable so should not be assumed.

If mental events are causes only because they are physical events then they are not causes because they are mental events. I wish to explain how a mental event can be a cause of a physical event because it is a mental event irrespective of whether it is a physical event.

That physical events are necessary for mental events does not entail that mental events are physical events but that mental events are physical events does entail that physical events are necessary for mental events so that physical events are necessary for mental events is consistent with mental events being physical events. However, it might turn out that the truth about psycho-physical causation is inconsistent with mental events being physical events.

We know that physical events are necessary for mental events because it is empirically confirmed that thought and experience depend on neural prerequisites. Understanding psycho-physical causation depends on understanding what follows from this empirical dependency: neural dependence entails mental causation.

The Neural Dependency Hypothesis

(NEURODEP) Mental events depend on neural events.

the hypothesis of neural dependency, has the following logical and epistemological features:

It has the logical form -p -> -q or, quantifying over events: -ExMEx -> -ExNEx

It is an empirical claim: it may be confirmed or refuted by observation.
It is true because it is confirmed empirically.
It is a contingent truth not a necessary truth. It is logically contingent not logically necessary because there is no contradiction in the supposition that mental events have no neurological prerequisites (Indeed, this might be true

in cases of artificial intelligence).<- NB except in the people who made the machine. It is metaphysically contingent not metaphysically necessary because it is false that mental events have neorological prerequisites in all possible worlds. Crucially, the hypothesis of neural dependency logically entails mental causation

Mental Causation. Mental events are sufficient for neural events.

The theorem of mental causation has the following logical and epistemological features:

It has the logical form p->q or, quantifying over events: Ex MEx -> Ex NEx

It is a metaphysical claim. It cannot be confirmed or refuted by observation.

It is true because it is entailed by a truth.

It is a contingent not a necessary truth. It is logically contingent not logically necessary because there is no contradiction in the supposition that mental events are not sufficient for neurological events. It is metaphysically contingent not metaphysically necessary because it is false that mental events are sufficient for neural events in all possible worlds. (Indeed, this might be true in cases of artifical intelligence) <- NB suff for neurology of people who made the machine

The hypothesis of neural dependency entails mental causation because if neural events are necessary for mental events then mental events are sufficient for neural events. The logical form of the hypothesis and its entailment is:

-Q -> -P -> P -> Q (axiom)
-Q -> -P (confirmed empirical hypothesis)

P-> Q

or, quantifying over events:

-Ex NEx -> -Ex MEx -> ExMEx -> ExNEx (axiom)
-Ex NEx -> -Ex MEx -> ExMEx (confirmed empirically)

ExMEx -> ExNEx

(problem of scope: mental events have many prerequisites: mental event is sufficient for all of them but does not cause all of them: only those simultaneous with it)

Overdetermination ? (because physical event has physical suff causes ?)

Problem of violating 3rd law of thermodynamics. Distinguish between:
1. Amount of energy in the physical universe remains constant.
2. Physical universe is a closed causal system.

These two views are not the same.

If there is mental causation then 1. is not threatened but 2. has to be discarded.

1. is not threatened because a physical cause may usually be sufficient for a physical effect beut in some cases - eg neural cases, the physical cause may only privide a 50 % chance of one event happening and a 50% chyance of a different event happening. Eg: NE -> either NE1 or NE2. Whether NE1 happens or NE2 happens the amount of energy in the universe is the same. However, some mental event is sufficient to tip the probability in say NE1's favour. The mental is sufficeint for the physical in this case (which is out theory). Clearly then the physical system is not in that case a closed causal system. But science does not require that.

Suppose some kinds of mental event are sufficient for physical events.
eg. Suppose a thought, desire, belief etc is sufficient for a neurological event, then the neurology is necessary for the mental state (propositional attitude)

Suppose some kinds of mental event are necessary for some kinds of physical event.
eg. Suppose a neurological event (eg c-fibre firing) is sufficient for a sensation, a pain, then the pain is necessary for the c-fibre firing. (Could there empirically be c-firbre firing without pain. I doubt it. Logically, yes.)

Then:

If M is a propositional attitude then M is sufficient for P and P is necessary for M.

and

If M is a sensation then M is necessary for P and P is sufficient for M.

Could mental causation be the solution to the mind-body problem?

(1) Hard
(2) Logic throughout esp. modal logic
(3) Quotations from physicists (cf ms. A Theory of Free Will)
(4) Philosophers mainly to FN. refs in text to (x, 1979)
(5) c. 120 pp

'The point is that the photon is in a superposition of states going one way or another, the detector is in a superposition of states of being off and on, and the cat is in a superposition of states of being alive and dead.'
Roger Penrose *et. al. The Large, the Small and the Human Mind* (Oxford University Press, Oxford, 1997) p. 72

'It is quite extraordinary that quantum mechanics enables you to test whether something might have happened but didn't happen. It tests what philosophers call counterfactuals. It is remarkable that quantum mechanics allows real effects to result from counterfactuals!'
Roger Penrose *et. al. The Large, the Small and the Human Mind* (Oxford University Press, Oxford, 1997) p. 67

Planck

Will

Accepting that:

'All changes in nature are in reality simply the transformations of one form of energy into another' (Planck, 1950: 175)

Planck says

‘[…] the most adequate formulation of this law creates the impression in every unbiased mind that nature is ruled by a rational, purposive, will.’ (Planck, 1950: 177)

Planck thinks the existence of change is grounded in *will*. This matters for the existence of God, because if God makes the natural world possible, then nature is indeed grounded in will as Planck says, because God is or has a will; an infinite will. What is this rational, purposive, will?

Anything has a *will* if its actions are caused by its own desires or intentions: if it is capable of doing what it wishes. Anything has *free will* if it has a capacity to choose: if, in some situation, it could either act or refrain from acting. It is controversial whether the will in this sense is free, perhaps essentially free, or whether free will is illusory. If the will is itself the capacity to choose, then ‘free will’ looks pleonastic.

Some will is *rational* if an agent has a reason for performing the actions willed.[5] Some will is *purposive* if an agent has a goal, aim, or intention in acting. It might make no sense to speak of action that does not presuppose free will, and at least some minimum reason or purpose in acting, for example, if the possibility of refraining distinguishes actions within the set of happenings.

Without mentioning Schopenhauer, Planck is logically committed to an essentially Schopenhauerian picture. The German philosopher, Arthur Schopenhauer (1788-1860), argues that without the existence of a cosmic will, there would be nothing. Schopenhauer thinks the German philosopher Immanuel Kant (1724-1804) is right that there is a way the world really is, in itself, independently of any experience of it, but Kant is wrong in assuming a plurality of *things-in-themselves*, and wrong in restricting will to human will. Schopenhauer rethinks the Kantian *thing-in-itself* (*Ding-an-sich*) monolithically as Will (*Wille*).[6] Accepting that the world might not fundamentally be as it appears, as *presentation*, (*Vorstellung*), Schopenhauer argues that the world as

---

[5] 'Rational', 'rationality' etc. are overused words in philosophy and social science and, in the absence of definition, it is not not clear that they mean much. As a predicate of belief (in the sense of *what* is believed):

p is rational iff (*def.*) there exist reasons for *p*
p is rational iff (*def.*) *p* is free from contradiction
p is rational iff (*def.*) if *p* is more likely true than false

As a predicate of some agent:
x is rational iff (*def.*) *x* acts in *x*'s interests
x is rational iff (*def.*) *x* acts in *x*'s perceived interests

[6] The usual translation of Kant’s *Ding-an-sich* as ‘thing-in-itself’ is misleading. It has generated a long debate about whether Kant is committed (as a matter of logic or biography) to some set of objects extra to those we perceive. But ‘an sich’ means ‘as such’. Kant means that there might be a way in which the objects we perceive are, in abstraction from what we may know about them.

it is in itself is *one*, and is *Will*.[7] Schopenhauer also maintains that, at a fundamental level, the world as Will is identical with the will each of us finds inside ourselves.[8]

Anticipating Planck, Schopenhauer draws a distinction between force (*Kraft*) and will (*Wille*) where 'the concept of force […] is abstracted from the realm governed by cause and effect' (Schopenhauer, 1844, 2010: 136). Following Kant, Schopenhauer argues, that the concepts of force and causation are restricted to the world of appearances; to the world as it is presented to us. Schopenhauer says:

'Until now people have subsumed the concept of will under the concept of force. I will do precisely the opposite, and let every force in nature be known as will.' (Schopenhauer, 1844, 2010: 136)

If Kant and Schopenhauer are right, it is a category mistake to abstract the concepts of force, or causation, from appearance and apply it to that which grounds appearance.[9] This is the error of determinism. The freedom of the will at the basis of the world entails a profound contingency: Whatever happens might not have happened. Nevertheless, the Will makes everything that happens happen.

Something like the Planck-Kant-Schopenhauer thesis is entailed if God creates the world. If God creates, God creates freely. But why suppose nature is ruled by a will?

One answer is this: Science explains particular events, often with predictive certainty, but not *why anything happens at all,* not why there are *any* happenings. Irrespective of whether the whole event series began, or is infinite in past time, *why there are any such things as events* stands in need of explanation.

Any will that is *rational* and *purposive*, as Planck claims, presupposes freedom. But does anything else have to be true of this Will if it makes anything

---

[7] (Schopenhauer, 1844, 2010) Schopenhauer's term *Vorstellung* is frequently translated 'representation'. But Schopenhauer is at pains to repudiate any representational or quasi-Lockean theory of perception. Translating *Vorstellung* as 'presentation' captures Schopenhauer's meaning: The world as it fundamentally is, is not how it is presented through sense experience.

[8] Schopenhauer, wittingly or unwittingly, is thereby logically committed to a version of the Hindu identity of *Atman* with *Brahman*, of the soul with the world-soul.

[9] The term 'category mistake' was coined by Gilbert Ryle. 'It represents the facts [...] as if they belonged to one logical type or category (or range of types or categories) when they actually belong to another.' (Ryle, 1949: 16) For example: 'The march-past was not a parade of battalions, batteries, squadrons *and* a division; it was a parade of the battalions, batteries and squadrons *of* a division'. (16-17) Whether sentences expressing category mistakes are meaningless, ungrammatical, or false (and so either true or false) is examined by Olfra Magidor (Magidor, 2013).

that happens happen? What would *the cause of everything that happens* have to be like?

VII

Moral Cosmology

What is Ethics?

Planck says 'all the problems of ethics are outside the field of natural science' (Planck, 1950: 182) Is Planck right?

There is a distinction between:

(1) *what is*

and

(2) *what ought to be*

between all and only *what there is*, *what exists*, and *what there morally should be*, all and only *that which is ethically right*.

Suppose, as the early Wittgenstein thinks, 'The world is all that is the case' (*Die Welt ist alles, was der Fall ist*). (Wittgenstein, 1922:1) There is then only *what is the case* whether the totality of facts (*Tatsachen*) as Wittgenstein himself thinks, or things, or events, or anything else. But then there seems no ontological room for what *ought to be* or *ought not to be* the case. What there should, or should not be, in some way *ontologically exceeds* what there is. What ought or ought not to be, is mystical, mysterious, and naturally inexplicable. The early Wittgenstein says 'Ethics is transcendental'.[10] (6.421) For this reason, science could not just *tidy up* the problems of ethics, given time. Even if science

[10]'Die Ethik ist transzendental' (6.421), where: *x* is transcendental *iff x* exists, but *x* is not part of 'the world' (the totality of facts).

(by some criterion) were complete, ethics would not have been touched. Moral problems are of a wholly different order from scientific problems. No amount of answering *Why does this happen?* answers *Ought this to happen? This will happen* logically, semantically, and ontologically, always and everywhere falls short of *This ought to happen*.

But if no amount of science answers moral questions, what does? Planck says:

'The law of causation is the guiding rule of science, but the Categorical Imperative – that is to say, the dictate of duty – is the guiding rule of life.' (Planck, 1932, QQ: 151)

The Categorical Imperative is the central tenet of Kant's moral philosophy: the purely formal rule governing what *ought* to be done as opposed to what *is* done, in one formulation:

'I ought never to act except in such a way *that I can also will that my maxim should become a universal law*' (Kant, 1785: 70)

Kant's *Wille* is usually translated 'will'. However, Kant's moral philosophy makes more sense if *Wille* is translated 'intention'. In deciding the right thing to do, ask, roughly: *What if everyone did that?* Kant thinks that if the proposed course of action is immoral it will be self-undermining when universalised in this way, if moral, coherently universalisable. In his examples: Implementing a universal recommendation to *make promises with the intention of breaking them* would destroy the convention of promising breaking. Acting on a universal recommendation to *commit suicide* would eliminate all moral choosers. Ethics is therefore grounded in logic. Kant insists that the consequences of actions have no bearing on their morality or immorality. In particular, considerations of pain or happiness, essential to utilitarianism, are irrelevant.

Kant's ethics may be criticised. For example, bare logical consistency does not look sufficient for morality, or immorality, even if logical consistency is necessary condition for the plausibility of any claim whatsoever.[11] But, even if utilitarianism is flawed, moral concepts might not be coherently wholly freed from their ground in experience: A world without pain is a world without ethics. Nevertheless, influenced by Kant, Sartre suggests at least an oddness, if not logical inconsistency, in joining a Trade Union, the Communist Party, or

[11] Intuitively, any putative claim entailing p & $-p$, to that extent, expresses nothing whatsoever. (Aristotle, 1971: Book Gamma Ch. 4)

marrying, while maintaining that people should not do that, or converting to a religion but recommending that others do not, or even not recommending that others do. Our actions are at least implicit recommendations to similarly placed others. (Priest, 2001: 191-203)

About any guide to how to live life, Planck says:

'Here intelligence has to give way to character, and scientific knowledge to religious belief. And when I say religious belief here I mean the word in its fundamental sense.' (Planck, 1932, QQ: 151)

Religious belief 'in its fundamental sense' is not the *opinion that* God exists, but *reliance on* God as that ultimate spiritual reality which is already known to exist. Belief is then the belief entailed by knowing and faith is 'belief in'; confidence, or trust. Lack of faith or loss of faith is lack of trust, or loss of trust. Planck suggests human beings cannot live by science, only by 'religious belief' in these senses.

Planck might be right about that, but he endorses Kant's categorical imperative which, it might be objected, grounds morality not in God, but in logic. If Kant's deontology provides satisfactory logical criteria for distinguishing right from wrong, theological criteria would seem unnecessary. As a matter of biography, this is not Kant's view. Kant thinks at least the *postulation* of the existence of God necessary to guarantee the categorical imperative. Planck and Kant are therefore closer than this line of criticism might suggest.[12]

Why does Planck think that, in ethics, scientific knowledge should 'give way' to religious belief? Like the early Wittgenstein, Planck thinks that because there cannot in principle be a scientific explanation of ethics, ethics are not ultimately grounded in the universe but outside it. Planck says:

'And if we ascribe to God, in addition to His omnipotence and omniscience, also the attributes of goodness and love, recourse to Him produces an increased feeling of safety and happiness in the human being thirsting for solace. Against this conception not even the slightest objection can be raised from the point of view of natural science, for […] questions of ethics are entirely outside its realm.' (Planck, 1950: 185)

[12]Kant says the existence of God is a postulate of Pure Practical Reason (*Reine Praktische Vernunft*): Human beings have to assume the existence of God, as a necessary condition for making sense of themselves as ethical beings.

Value is irreducible to fact, so, as Wittgenstein says, *transcendent*. Once all the scientific facts are in, understanding value has not begun.

If Planck is right, the explanation of ethics requires a personal God after all, or *malgré lui*, one with 'the attributes of goodness and love' who we have 'recourse' to. As in Kant's philosophy, there cannot be any scientific proof or disproof of the existence of God (so it would be a category-mistake to think the existence of God a 'scientific hypothesis'). As in Kant's philosophy, Planck's proof of the existence of God is *ethical*.

The early Wittgenstein says 'Ethics and aesthetics are one' (*Ethik und Ästhetik sind Eins*) (Wittgenstein, 1922: 6.421) because equally transcendent. Planck thinks both ethics and aesthetics are grounded in spirituality. Planck is perhaps thinking of the cathedrals of the Western middle ages, and the painting of the Renaissance, when he writes: 'the mightiest benefits ever enjoyed by art were the result of its becoming a servant to religion.' (Planck, 1950: 161) If he is right, this explains why the greatest art-works are produced during the historical periods of greatest spiritual awareness: Divinity *shines through* a genuine work of art. Planck would have perhaps endorsed Theodore Adorno's critique of much that passed for art during the twentieth century: 'The darkening of the world makes the irrationality of art rational: radically darkened art.'[13] Appropriation by atheistic materialism, whether global-capitalist or state-socialist, reduced art to anti-art, the real aesthetic to arbitrary banality or desecration, to the grey and soulless repetitions of the machine age.

Planck, if not Adorno, thinks that hope of human improvement depends on preserving religious awareness. Planck perhaps has in mind the massive destruction wrought by the atheistic ideologies, Nazism and Stalinism, when he writes:

'the victory of atheism would not only destroy the most valuable treasure of our civilization, but – what is even worse – would annihilate the very hope for a better future.' (Planck, 1950: 156)

It is frequently assumed that religion causes war but Planck hints that the opposite might be the case. Meditation, calm, inwardness, and contemplation, are revelatory of inner peace, and conducive to outer peace. Yet it is hard to verify empirically the claim that religion has prevented wars. Events that *did not happen* do not show up in the investigation.

Ethics and Meaning

[13] (Adorno, 2013: 27) Theodor W. Adorno (2013). Aesthetic Theory p.27, A&C Black

Like Max Planck, Einstein distinguishes *what is* from *what ought to be*; fact from ethical value. Although excellent in providing facts,

'[…] the scientific method can teach us nothing else beyond how facts are related to, and conditioned by, each other. The aspiration toward such objective knowledge belongs to the highest of which man is capable, and you will certainly not suspect me of wishing to belittle the achievements and the heroic efforts of man in this sphere. [...] Yet it is equally clear that knowledge of what *is* does not open the door directly to what *should be*.' (Einstein, QQ: 105-6)

Even if science were complete, even if every fact were known, there would still be complete ignorance of value:

'One can have the clearest and complete knowledge of what is, and yet not be able to deduct [ie deduce] from that what should be the goal of our human aspirations.' (Einstein, QQ: 106)

Einstein uses 'deduct' (meaning 'deduce') because values do not *follow logically* from facts. Purely factual premises do not yield evaluative conclusions.[14] Where 'p' is purely factual, or expresses no value judgement, any putative logical inference from:

(1) It is the case that *p*

to

(2) It ought to be the case that *p*

is fallacious. Einstein says:

'Objective knowledge provides us with powerful instruments for the achievement of certain ends, but the ultimate goal itself and the longing to reach it must come from another source.' (Einstein, QQ: 106)

The point of doing anything, cannot be explained factually, so the point of doing science cannot be explained scientifically. The *means to* an end is not the same as the *justification of* that end. Science cannot answer: *What does it all mean? Why have any goal? What is the point of it all?* In the history of

[14] (Searle)

thought, two kinds of answer have been salient in answering the question of the meaning of existence.

On one view, self-conscious human beings *invent* meaning. What human beings *decide* is worth living for, or *choose* to live for, is worth living for. After all, if 'worth living for' means 'worth human beings living for', who else could decide this but human beings? This sort of view is articulated by the atheistic French existentialists; Simone de Beauvoir, Maurice Merleau-Ponty and, with most originality, Sartre. Human beings are defined as the self-defining beings. We, as opposed to stones or motor cars, are the beings who, individually or collectively, *make ourselves what we are*. Meaning is *invented* not discovered.[15]

On the other view, human beings, being self-conscious beings, *discover* meaning. Meaning is not invented. Meaning generates a regress: If the point of *this* is *that* then what is the point of *that*? and so on. To halt the regress there must be an ultimate meaning, or source of all meaning. Meaning is a value, so is not to be found in the world of facts, so meaning exceeds anything in space-time. The source of value therefore lies outside the universe. These conditions may only be met by some ultimate source of meaning which is not spatio-temporal. But this is what religious people call 'God'. The goal of existence, including human existence, is then union with God; acquaintance with ultimate meaning. This kind of view is articulated by St. Augustine, and St. Thomas Aquinas.[16]

Einstein is aware of both kinds of view. He says:

'And it is hardly necessary to argue for the view that our existence and our activity acquire meaning only by the setting up of such a goal and of corresponding values. The knowledge of truth as such is wonderful, but it is so little capable of acting as a guide that it cannot prove even the justification and the value of the aspiration toward that very knowledge of truth. Here we face, therefore, the limits of the purely rational conception of our existence.' (Einstein, QQ: 106)

Then Einstein says religion is essential to values:

'[…] mere thinking cannot give us a sense of the ultimate and fundamental ends. To make clear these fundamental ends and valuations, and to set them fast

[15] (de Beauvoir, Merleau-Ponty, Sartre, Mackie)

[16] The two kinds of view are far from jointly exhaustive. In Utilitarianism, right and wrong are grounded in happiness and pain. Renford Bambrough argues for a moral objectivity which does not depend on theological premises. On Blackburn's 'quasi-realism' moral properties might have a degree of objectivity comparable to Lockean secondary qualities.

in the emotional life of the individual, seems to me precisely the most important function which religion has to perform in the social life of man.'
Albert Einstein (QQ: 106)

Although '[...] invents *x*' and '[...] discovers *x*' are mutually exclusive predicates, the two models of ethics and meaning might not be as incompatible as they appear. The values that human beings create might be more or less adequate. People *seek* significance in work, creativity, friendship, giving to others, in sex, alcohol, tobacco, in other drugs. However worthy or ignoble, in these pursuits human beings are clutching at the *transient*. True or lasting significance may only be found in the *permanent*. Moral knowledge then depends on spiritual knowledge; discerning *the permanent in the transient*. Human beings then *invent* their own values and, perhaps, *discover* God's values.

The concept of free will, and the doctrine of Deism, both offer strategies for reconciling the two moral pictures.

Deism is the view that God created, or creates, a self-determining world; a world that freely makes itself what it is. God creates the world's *existence*. The world creates its own *essence*. To state something's essence is to answer the question:

What is it?

To state something's existence is to answer the question:

Is it?

in the affirmative. Something's *essence* is *what it is*, its nature: those characteristics singularly necessary and jointly sufficient for its being of the type it is, or for its being the very thing it is. Something's *existence* is its *being rather than not being*. It is philosophically controversial whether anything could *be anything* without *being*, or whether there could be *being* without *being anything*. On Deism, what the world makes itself be, subsequently might or might not meet with divine approval.[17]

On influential Judaeo-Christian views, God has given humanity freedom of choice.[18](WHICH VIEWS? FN LUTHER. FREE WILL BEFORE THE FALL.) There is *a spontaneity of actions in the now* rather than chains of cause and effect stretching into the past or future. The causal chains, 'behind' and 'in front of' oneself, are merely an abstract picture. They do not exist outside the

[17]Flood in Genesis.
[18]

human mind. The constantly changing present is the lived, existential, reality. KANT FIGURATIVE USE OF IMAGINATION. If human beings have freedom of choice, they have freedom to choose their own values. Again, what human beings choose, then might or might not meet with divine approval.[19]

Despite his determinism, Einstein insists that human beings are choosers.[20] Without freedom of choice there is no responsibility. If someone is responsible for, say, stealing something, this implies they could have refrained from that stealing. Freedom entails the possibility of refraining. Like Sartre, Einstein is interested in the responsibility freedom allows; how choice should be exercised. If Sartre and Einstein are right, despite real social pressures or lack of economic opportunity, human beings are responsible for everything they do. We remain choosers. How should anyone cope with the burden of such freedom? Einstein thinks the highest role is *service*. He says '[…] the high destiny of the individual is to serve rather than rule' (Einstein, QQ: 107) Instead of thinking, for example, *How could I be superior?* one should think *How can I help?* But, in that case, who or what should a human being serve? Einstein's sustained answer is: their fellow creatures.

Einstein thinks the Judaeo-Christian tradition exhibits the highest values:

'The highest principles for our aspirations and judgments are given to us in the Jewish-Christian religious tradition. It is a very high goal which, with our weak powers, we can reach only very inadequately but which gives a sure foundation to our aspirations and valuations.' (Einstein, QQ: 106)

Einstein considers religion the reverse of ethically damaging. Einstein is against deification in politics, but even then, on religious grounds:

'There is no room […] for the divinization of a nation, of a class, let alone of an individual. Are we not all children of one father, as it is said in religious language?' (Einstein, QQ: 106)

Einstein asks 'What is the meaning of human life, or, for that matter, of the life of any creature?' (Einstein, 1954: 11) Einstein says 'To know an answer to this question means to be religious.' (Einstein, 1954: 11) But could not my life be meaningful if I am *not* religious? Einstein thinks this is not right. The existence of purpose and meaning exceeds the physical world. The physical world is the set of physical events, or physical objects in motion: billions of atoms in empty

[19]Apple in Garden of Eden in Genesis.

[20]Kant argues that the phenomenal world is deterministic, governed by Newtonian laws, but we have free will as we are in ourselves.

space. Purpose and meaning are not billions of atoms in empty space. Purpose and meaning exceed even the psycho-physical world. Purpose and meaning are transcendental. They exceed, or point 'outside' the world. *Why are we here?* may mean either: *What is the cause of our being here?* or *What is the purpose of our being here?*, *What is the point?* But even a thorough knowledge of the *causes* of our existence does not provide any *reason* or *purpose* for our existence. Reasons are more than causes.

Einstein deals quickly with those, such as the Logical Positivists, who think the question of the meaning of life meaningless: 'You ask: Does it make any sense, then, to pose this question?' (Einstein, 1954: 11) 'I answer: The man who regards his own life and that of his fellow creatures as meaningless is not merely unhappy but hardly fit for life.' (Einstein, 1954: 11)[21]

Einstein's own service is to provide humanity with some of the most creative physics of all time: for better or for worse. Without Einstein the atomic bomb would be historically impossible, but so would the harnessing of nuclear energy for peaceful purposes. Einstein himself was appalled by the destructive power of the atom bomb. It bears out, perhaps, his view that science does not provide humanity with values, but only religion can do that. EINSTEIN ON BOMB QUOTATION.

Ethics

de Broglie thinks

'Scientific discoveries and the applications of which they are capable are in themselves neither good nor bad; all depends on the use which we make of them.' (de Broglie 1955: 217)

de Broglie's 'and the applications' suggests that (for example) a machine gun, flamethrower, or atomic bomb is intrinsically neither good and bad but, in itself, morally neutral. A weapon has ethical significance only in so far as it is used. Is this right?

On one level, 'in itself', in its own right, a weapon is only a physical object: a constellation of atoms in a particular pattern, occupying a portion of space-time. At this level of description, de Broglie's moral neutrality looks plausible. It is only in the hands of a conscious being, with certain intentions, that the constellation of atoms becomes a weapon.

On another level, described functionally, a weapon has the essential purpose of injuring or killing. A rifle or an atomic bomb has this function as a

[21]Logical Positivism. Priest.

disposition, potential, or propensity, even when no one is operating it. The disposition is exercised, realised, when the weapon is used. At this level of description, de Broglie's moral neutrality looks implausible. If the world is inundated with machines the function of which is to kill and maim, the expected result should be widespread killing and maiming.

We could read de Broglie as saying that ethical distinctions only apply within the human world, and the physical world is, in itself, neither moral nor immoral. As Shakespeare puts it:

'Nothing is good or bad, bu thinking makes it so' REF.

This might be right as a matter of ontology but, by definition, no-one inhabits the morally neutral physical world. Although weapons are designed by scientists, and sold and distributed by governments and arms dealers, responsibility begins not with the private individual, the politician or the member of the armed forces but with the inventor, and then the manufacturer. Scientists then have a profound and inescapable moral responsibility for what they do *as scientists*.

Because technology is used to devastating effect by humanity on humanity, de Broglie says

'To be able to survive the appropriate progress of his attainments, mankind of tomorrow will have to find in the development of his spiritual life and in the uplifting of his moral ideal, the wisdom not to abuse his increased forces.' (de Broglie 1955: 217)

de Broglie thinks morality depends upon 'spiritual life'. What *ought* or *ought not* to be the case exceeds, or transcends, *what is* the case. The world of morals is 'beyond' the world of facts, so inexplicable just by citing facts. The physical belongs only to the world of facts, so morality is not physical. Morality pertains to consciousness, or as de Broglie puts it, 'spirit'. de Broglie says:

'This is what Henri Bergson has splendidly expressed in one of his last works by saying; "Our enlarged body clamours for an addition to the spirit."' (de Broglie 1955: 217)

Bergson sees the spread of technology as an extension of the human body. de Broglie sees extreme danger in this, because Human technological development has outstripped human spiritual development. Humanity has developed the technological capacity to destroy itself, but perhaps not the moral capacity to

save itself. Life has, perhaps, produced the means to destroy all life, in a kind of mass biological suicide which would end history. But life has, it seems, not produced the ethics to stop this.

Consciousness, or 'spirit', is not just intelligence but will, so de Broglie says

'In the work of science man has been able to show the force of his intelligence; if he wishes to survive his own successes; it is necessary that he should now show the wisdom of his will.' (de Broglie 1955: 267)

'Will' is ambiguous between 'will power' and 'the capacity to choose' or 'the capacity to exercise preferences'. We may read de Broglie in both senses: To avoid nuclear destruction, annihilation in Quantum War, humanity will need to exert self-restraint, which implies will-power, and the living of more spiritual lives, which implies changing human preferences.

de Broglie asks

'Shall we be able to acquire this addition to the spirit as rapidly as the advances of science will develop? Undoubtedly on that depends the future destiny of humanity.' (de Broglie 1955: 217)

de Broglie endorses the essentials of Henri Bergon's philosophical integration of science and spirituality. de Broglie calls Henri Bergson's *The Two Sources of Morality and Religion* 'his great work' (de Broglie QQ: 122) de Broglie quotes Bergson's remark: '"Now, in this excessively enlarged body, the spirit remains what it was, too small now to fill it, too feeble to direct it."' Henri Bergson quoted by de Broglie (QQ: 122) If de Broglie is right, human survival depends on the recognition of consciousness and will as the basis of quantum physics and ethics. This requires a synthesis of science and spirituality, through Bergson, or otherwise.

Mystery

de Broglie says

'Mystery surrounds us; as Puvis of Chavannes has symbolically represented it in the vast fresco which adorns the amphitheatre of the Sorbonne, we are placed as at the centre of a small clearing surrounded on all sides by an immense and gloomy, unexplored forest. de Broglie (QQ: 120)

The mural 'The Sacred Grove: Beloved of the Arts and Muses' by the Burgundy painter Pierre Puvis de Chavannes (1824-1898) may be understood on many levels.[22] The clearing is *the known*, or the availability of the known, the gloomy forest the unknown, or the hiddenness of the unknown. The clearing is c*onsciousness*, or the zone opened up by immediate awareness. The forest is then whatever is not consciousness, or not immediately available to consciousness. The clearing is *presence*, the forest absence. The clearing is *one's own soul*, the subjective space, the light, the absolute interiority which is one's own existence, the forest the external world, or everything not one's own soul. The forest is the hidden, surrounding the soul. The forest is *Being*, the clearing Nothingness. The forest is *thing-ness*, the clearing *no-thing-ness*.

If one is *in* the centre of the clearing then it is one's body which is there. In his fundamental ontology, Heidegger describes as the *Lichtung* the clearing in the forest in which Being (*Sein*) is disclosed. (Heidegger, 1927) *Dasein*, the manner of being which is human *being*, or one's own being, is the site of this disclosure. Read this way, one is identical with the clearing, and the forest is everything which is not oneself. Whatever the merits or demerits of these readings, de Broglie thinks science uncovers only part of reality, which, in itself, or as it is, remains essentially hidden. The unknown makes the known possible.

But, what is spirit?

Spirit, in de Broglie's sense, entails the existence of consciousness. The modern preoccupation with physical objects, caused by commodification and industrialisation, has led to a philosophical playing down of the significance of consciousness. Consciousness, where acknowledged, has been misconstrued as only a set of neurological states of the brain, or as a mere epiphenomenon of matter. Yet, de Broglie suggests that only *the thought of* consciousness misconstrues it in these narrow ways. de Broglie quotes Bergson to express this insight:

'"Let us add that this increased body awaits a supplement of the soul and that the mechanism demands a mysticism."' Bergson quoted by de Broglie (QQ: 122)

The *supplement* of the soul is consciousness's recognition of itself as already expanded consciousness. Phenomenologically, consciousness is an infinite

[22] Pierre Puvis de Chavannes

subjective space; the human body of its owner at its centre. de Broglie adopts the expression 'supplement of the soul':

'Confronted by the dangers with which the advances of science can, if employed for evil, face him, man has need of a "supplement of the soul" and he must force himself to acquire it properly before it is too late.'
Louis de Broglie (QQ: 125)

de Broglie is optimistic that nuclear war may be averted by a resurgence of mysticism, renewing spiritual values. SUPPLEMENT IN BERGSON & DERRIDA. ETHICS GROUNDED IN MYSTICISM. He says

'Often the feeling of the immanence of a danger gives birth in the heart of men to sentiments or mysticisms which can serve to avoid it.' de Broglie (QQ: 125)

Nevertheless,

'It is the duty of those who have the mission of being the spiritual or intellectual guides of humanity to labour to awaken in it this supplement of the soul.' de Broglie (QQ: 125)

Priests, theologians, and mystics are the spiritual guides, scientists and philosophers the intellectual guides.

de Broglie sees science and religion as continuous, as, in a sense, sharing faith, and sharing the disclosure of the same hidden order. He refers to

'that bold act of faith which leads us to bear witness to the existence of a certain correlation between the sucession of natural phenomena on the one hand, and the pictures or reasonings which our mind is able to conceive on the other hand.' (de Broglie 1955: 206)

de Broglie thinks scientific conjecture and hypothesis formulation, require faith in the sense of *confidence*. It is correct that scientific hypotheses are confirmed or refuted empirically, and religious claims are metaphysical, but the distinction between physics and meta-physics is not always sharp. Physics becomes metaphysics in postulating unobservable realities, and in asking fundamental questions such as *Was there a first event? What are the ultimate constituents of matter?* Religion admits of an empirical dimension, when theologians advance cosmological evidence for the existence of God, or endorse mysticism; the experience of God.

To the extent that 'scientists have come to feel more and more keenly that there exists in nature an order, a harmony'(de Broglie 1955: 206) knowingly or unknowingly, scientists have a spiritual sense. de Broglie asks the question:

'What, in fact, is the *raison d'être* of this [scientific] effort, what mysterious attraction acting on certain men urges them to dedicate their time and labours to works from which they themselves often hardly profit?' (de Broglie 1955: 207)

He answers:

'Evidently this is one of the aspects of the dual nature of man, so often put into relief by thinkers and philosophers; restrained by our organic constitution and our different emotions in the lower sphere of our daily occupations, we also feel ourselves urged on by the appeal of the ideal, by more or less precise aspiration towards spiritual values.' (de Broglie 1955: 207)

This duality between physical and spiritual is Plato's distinction between, on the one hand, the constantly changing, imperfect, world of *doxa*, belief or opinion, and appearances, and, on the other hand, the timeless, perfect, immaterial world of *episteme* (knowledge), truth, and Forms (*eidos*). Human beings know they perpetually fall short of a perfection of which they are dimly aware. Morally imperfect, we know there is moral perfection. As the physical world is geometrically approximate, there exists a perfect geometrical reality. Rather like Kant, in his moral arguments for the existence of God, de Broglie thinks the existence of the ideal of moral perfection, and its clear absence from the empirical world, shows the reality of spirituality. (Kant)

A quasi-Platonic duality is essential to the Christian doctrine of the Fall. Human beings are hypocritical; conceal their faults, inadequacies, and lies. 'Personality' derives from the Latin *per sona*, 'sound through', which refers to the mask worn by actors in the Greek and Latin theatre. Instead of honestly being what they are, human beings pretend to be what they are not.[23] In moments of insight into the Fall, they are suddenly aware of their tawdry hypocrisy, and ashamed. de Broglie thinks 'from those sentiments even the worst of us do not entirely escape.' (de Broglie 1955: 207) In Neo-Platonic Christian terms: Everyone is fallen, but has glimpses of perfection.

According to de Broglie, the aim of science is to disclose order: 'This goal […] is without doubt to succeed in penetrating further into the knowledge of natural harmonies'(de Broglie 1955: 207) or 'to come to have a glimpse of a reflection of the order which rules the universe' (de Broglie 1955: 207) Science

[23]Sartre's 'bad faith' (*mauvaise foi*) is a secularisation of the Christian doctrine of the Fall.

reveals ‘some portions of the deep and hidden realities which consititue [the universe].’ (de Broglie 1955: 207) de Broglie says

‘Even the scientists or philosophers who, pragmatic in tendency, have reduced to a utilitarian role the value of scientific theories, as for example, the eminent physicist Pierre Duhem, have had to recognise that these theories establish between the phenomena a “natural classification”, allowing us to sense the existence of an “ontological order” which is beyond us.’
(de Broglie 1955: 207-8)

What does de Broglie mean?

A ‘natural classification’ is not of our own making. It is a mathematical order that is real, or ‘out there’, and would exist whether or not conscious beings exist. Mathematical laws are discovered, not invented, yet are not observable in the manner of physical events. Yet mathematical laws make possible everything discovered by science.

Pragmatism is a cluster of views: Truth and falsity are constituted by the possibility of practical tests (not just *detected by* practical tests). The meaning of a claim is its ‘cash value’. The value of anything is its practical value. Pragmatism implies the functionalist thesis that what something *is* is exhausted by what it is *for*. (REF) Pragmatism is a philosophical systematisation of technological and commercial attitudes which have increasingly prevailed in the West over the last half millennium.

‘Utilitarian’, in de Broglie’s sense, means ‘useful’ or ‘of practical use’. In moral philosophy, utilitarianism is the theory that ‘good’ means ‘conducive to the greatest happiness of the greatest number’. de Broglie might be hinting at this here, but utilitarian moral philosophy massively exceeds pragmatism in its ethical claims. (Mill)

‘“an ontological order” which is beyond us’ means a pattern which exists whether detectable or not, an order that exists ‘out there’. The ontological order is, in this sense, objective order, not order imposed by human purposes. It obtains *anyway*. Pragmatists are disinclined to admit such an order, because they think reality does not exceed our practical dealings with it.

de Broglie is saying pragmatists have to recognise an objective order *as a result of scientific practice*, despite their contrary intentions. de Broglie says ‘All those who dedicate their efforts to pure science admit, whether they agree with it or not, the existence of such an order.’ (de Broglie 1955: 208) So even scientists disavowing ‘ontological order’ are logically committed to it by doing science.

de Broglie suggests that philosophy has much to learn from the new physics:

‘Already, contemporary physics, by introducing its new ideas on space and time, on the impossibility of following the determinism of elementary phenomena, on the “complementary” character of certain pictures apparently contradictory (such as that of wave and corpuscle), on the inability to discern elementary particles, already contemporary physics, I say, offers to the meditations of philosophical minds entirely new themes of which, at the present hour, we are far from having perceived all the consequences.’
de Broglie (QQ: 120)

de Broglie does not overtly endorse the existence of the personal God essential to Judaeo-Christianity. On the other hand, he offers refutations of the materialism inconsistent with the existence of God. He argues that the physical world presupposes consciousness, that ethics has a transcendental source, and the reality of free-will is inescapable. In arguing in these ways, de Broglie is in fact arguing for three essential tenets of Judeo-Christianity entailed by the Catholicism he neglected to practice.

Stapp

Ethics

Stapp argues that ethics would be impossible if materialism were true:

‘The proposition, foisted upon us by a materialism based on classical physics - that we human beings are essentially mechanical automata, with every least action and thought fixed from the birth of the universe by microscopic clockwork-like mechanisms - has created enormous difficulties for ethical theory.’ (Stapp, 2007: 117)

If materialism had turned out to be true, if it had not been falsified by twentieth century quantum science, as Stapp suggests, there would have been no ontological room for morality or immorality. This sort of argument would have been sound:

(1) If anything exists then it is physical.
(2) Morality is not physical
(3) Therefore, morality does not exist.

But it is practically impossible to dispense with moral distinctions: Not only could life not be lived, science would be impossible. With no sense of value there could be no 'point' in pursuing this research project rather than that, seeking funding, building weapons, medical facilities, energy installations. Like mental causation, morality is not a postulate but an ineliminable presupposition. Morality exists. Therefore, the following argument is sound:

(1) Morality exists.
(2) Morality is not physical.
(3) Therefore, it is false that if anything exists then it is physical.

'If anything exists then it is physical' is a statement of materialism, so the argument is a refutation of materialism.

Appendix I: The Problem of Freedom and Prediction

Appendix II: Aristotle and the Sea Battle: The Problem of Future Contingents

Is the future inevitable?

sea-battle argument. Argument criticised by *Aristotle in *de Interpretatione* Book IX. Either there will be a sea-battle tomorrow or there will not. These possibilities are mutually exclusive and jointly exhaustive. It seems, therefore, to be a necessary truth that there will be a sea-battle tomorrow or there will be no sea-battle. The future is inevitable now, it is just that we are presently ignorant of which future events will happen. The argument putatively establishes the fatalistic conclusions that everything that happens happens necessarily, and it was always true that the those events that do happen would happen.

In criticism, Aristotle argues that from the necessity of the disjunction (*p* v –*p*) it does not follow that the disjunct *p* is necessary nor that the disjunct –*p* is necessary. Although it is necessary that a sea-fight take place tomorrow or not, it is not necessary that it should take place tomorrow, and is it not necessary that it should not take place tomorrow. Therefore the argument does not validly establish that future events are inevitable. This is consistent with its being a necessary truth that either there will or will not be a sea-battle tomorrow. *future contingents

Premises:

(1) Every proposition is either true or false,

so

'There will be a sea battle tomorrow' is either true or false.

(2) One of a pair of contradictories must be true and the other false (P v -P)

so either

'There will be a sea battle tomorrow'

is true or

'There will be no sea battle tomorrow'

is true.

(3) 'Any given predicate must either belong to the subject or not' (179)
(Fa v -Fa)

so

either there will or will not be a sea battle tomorrow.

(4) 'The predicate cannot both belong and not belong to the subject at one and the same time with regard to the future' (179)

so if there is a sea battle tomorrow the proposition

'There will be a sea battle tomorrow' is true.

If there is no sea battle tomorrow the proposition

'There will be no sea battle tomorrow'

is true.

Conclusions:

(1) 'There are no real alternatives' (180)
(2) 'Everything takes place of necessity and is fixed' (180)
(3) 'If a thing is [...] now, it was true before to say that it would be [...]' (180)
(4) 'If it was always true to say that a thing is or will be, it is not possible that it should not be or not be about to be' (180)
(5) 'All, then, that is about to be must of necessity take place'
(180)
(6) 'A man may predict an event ten thousand years beforehand, and another may predict the reverse; that which was truly predicted at the moment in the past will of necessity take place' (181)
(7) 'Of that which takes place, it was always true to say that it would be'
(181)

These conclusions seem inconsistent with these claims:
(1) The future is 'open'.
(2) Future events are contingent events.
(3) Human beings are free to determine the future by their actions.
(3) Future tense assertions are neither true nor false when asserted.

Aristotle's solution:
'A sea-fight must either take place tomorrow or not, but it is not necessary that it should take place tomorrow, neither is it necessary that it should not take place, yet it is necessary that it either should or should not take place tomorrow' (182)

Aristotle on the Future. Aristotle’s discussion of whether the future is fixed (now) or open (now) can be found in his *De Interpretatione* (which means ‘Concerning Understanding’) Book IX. His argument runs as follows:

Premise 1: This is the claim that any proposition is either true or false. On this view, certain alternatives are ruled out e.g. the possibility that P is neither true nor false, or both true and false, or that there is a difference of degree between truth and falsity. This premise may simply be stipulative, part of his definition of 'proposition'.
Not every sentence expresses a proposition: for example commands and questions are obeyed or disobeyed. Propositions are what is expressed by an indicative sentence. (What meanings of sentences and propositions are is an unsolved philosophical problem.) Aristotle concludes that 'There will be a sea battle tomorrow' is either true or false.

Premise 2: Of a pair of contradictories (i.e. of the form both P and not P), either the proposition or the negation of the proposition must be true. So either 'There will be a sea battle tomorrow' or 'There will be no sea battle tomorrow' is true. In order for this premise to work, there must be a clear distinction between there being a sea battle and there not being a sea battle. Again, we may just stipulate that.

Premise 3: Any given predicate must either belong to the subject or not. A predicate is a word or string of words that is used adjectivally; to ascribe a property to something. The subject is the word or set of words that denotes or picks out what ever the predicate is used to ascribe the property to. A predicate ascribes a property, but it is an unsolved problem in the philosophy of language how much should be built into a predicate. Does something have to be a word, or a sense in order to be a predicate? Also, the term 'subject' is ambiguous: it can denote the word that is the subject of the sentence or the object that word denotes.
Aristotle argues that any given predicate must belong to a subject or not: 'a is F' is either true or false. Again, he assumes that our concepts apply cleanly to the world, and that a distinction can be drawn between something's being F and its not being F.

Premise 4: A predicate cannot both belong and not belong to the subject at one and the same time with regard to the future.

Behind these Aristotelian premises are certain axioms of Aristotelian logic, which can be found in Aristotle's *Metaphysics* (in Book Gamma). There, Aristotle defends certain laws, but it is not clear whether these laws are semantic or metaphysical. These laws include:

A is A. Nothing differs from itself. This principle strikes some as completely obvious, but others see it as extremely controversial and conservative.

Nothing is both A and not A (Aristotle means in exactly the same respect).

Everything is either A or not A.

These principles have become enshrined in modern propositional logic as

$P \rightarrow P$
$- (P \& - P)$
$P \vee - P$

But it not clear that Aristotle is talking about axioms that hold between truth-valued items, or whether he is simply talking about psychology (about limits on our imagination - we cannot imagine A not being A, for example). On the latter view, the laws are contingent facts, and if they are indeed only psychological principles they may be false in reality.
Some people argue that these principles only hold because of language, and that they are necessary preconditions for language. But construed ontologically, they seem to be necessary conditions for existence.

Aristotle's conclusions

Conclusion 1: There are no real alternatives. If we make a prediction ('*x* will happen', for example) then it is either true or false (at the present time). There will only be one future, as there was only one past. We in the present are simply ignorant of the one future.

Conclusion 2: Everything takes place of necessity and is fixed. Predictions are true now. Therefore, what will happen will necessarily happen. It is not an option for the future to happen in any other way than it is true now that it will happen.

Conclusion 3: If a thing is now, it was true before to say that it would be. A prediction may be true at the time, rather than being neither true nor false till the predicted event happens or fails to happen.

Conclusion 4: If it was always true to say that thing is or will be (timelessly true); it is necessarily true. On this view, 'P is necessarily true' means 'P is

true and could not be false' (It is not clear what 'could' or 'can' means here. It is also an unsolved problem what 'necessarily P' says that 'P' does not say already. Wittgenstein argued that 'the world is all that is the case', but if we say that something is not only the case, but is necessarily the case, what extra information has been provided? A similar problem arises for moral philosophy when we say something 'ought' or 'ought not' be the case)

Conclusion 5: All that is about to be must of necessity take place. If 'P' is future tense, and 'P' is true, then it is a necessary truth. It follows from conclusion 5 that all truths are necessary truths

Conclusion 6 & 7: Predictions may be true, whenever they are made. Of that which takes place, it is always true to say that it would be

Each of these Aristotelian conclusions seems inconsistent with the following claims:

[1] The future is now open.

[2] Predictions are neither true nor false when they are made

[3] Future events are contingent events

[4] Human beings are now free to determine the future.

If sound, Aristotle's argument seems to rule out human freedom.

The argument against freedom given here does not simply hinge on a premise about predictability. We need a stronger premise than the bare fact that x's behaviour is predictable to conclude that x is not free. Predictions may be based on tendencies, patterns in which people exercise their freedom. It is fallacious to assume that as we can predict someone's behaviour they are not free (as Ayer does).

Future tense assertions are neither true nor false when asserted

Aristotle's Solution

Aristotle argues if we say 'P v – P', then this disjunction is necessarily true. But what we cannot conclude from this is 'necessarily P' or 'necessarily –

P'. Though the whole disjunction is necessarily true, the disjuncts taken singularly are not necessarily true. The necessary truth of the disjunction is consistent with the future being open. Whether the future is fixed or not seems to be another unsolved problem.

It seems an odd fact that the future exists at all. Why is there a future? Supposing that we follow McTaggart in distinguishing between an A series (past, present, future) and B series (earlier than, simultaneous with, later than), and follow Russell in saying that the A series is introduced to the universe by conscious beings, then it seems that Sartre must be right to say that the future is always *somebody's* future. This raises the question of why events should be fixed before me (the past) or open after me (the future).

References are to Aristotle *De Interpretatione* partly reprinted in Richard Gale (ed) *The Philosophy of Time* (Sussex/New Jersey) p.179-182.

Aristotle's Categories and *De Interpretatione* trans. J. L. Ackrill (Oxford, 1963)

Aristotle's conclusion is that the future is not fixed in advance. In the present, at the moment when anything is happening, there is indeterminancy. Reality is poised, ready to go one way or another. The difference between the present and the future is an openness to possibilities.

5. Aristotle, De Interpretatione. An argument for the conclusion that the future i) exists and ii) cannot be changed. In a sense the future exists already; we simply haven't arrived at it yet. A fatalist argument. Aristotle produces a second argument to show that the future i) does not exist (yet) and ii) we can affect it through our actions and choices.

6. Aristotle advances a number of premises for his argument that the future exists, eg. "propositions must be true or false." The proposition "There will be a sea-battle tomorrow" must be either true or false; it is not both true and false, and it is not neither true nor false. The second premise is that if we take this statement and negate it, then either the statement or the negation of the statement must be true. The third premise is: "Any given predicate must either belong to the subject or not." A predicate is a word or set of words used to ascribe a characteristic or property. The subject is that which is noted by that. An object either possesses or lacks a property.

7. Aristotle argues that it follows that if it is true that there will be a sea-battle tomorrow, then there will necessarily be one. This might be a fallacious inference. The conclusion of Aristotle's argument is: "There are no real alternatives." There are only possibilites tomorrow: Either there will be a sea-battle, or not. We think the future is open because largely we are ignorant about it. But the future is as fixed as the past; it is only going to happen in one way, as the past has only happened one way. This being given, we are simply ignorant of the one way in which the future is going to happen.
8. Aristotle's subsidiary argument. If something is true now, it has always been true. Obviously it is now true that we are in the lecture theatre, but we have not always been here. What has always been true is that at this time we would come into this lecture theatre. "A person may predict an event ten thousand years beforehand..." Taking this view of the present, we should take the same view of the future. A statement made about the future now is either true or false. What happens is inevitable, including everything that happens in the future. What happens cannot fail to happen.
9. Aristotle proceeds to attack this position. He says it leads to an impossible conclusion. Deliberation and action affect the future; in other words, they are causes. If it is true that present actions can affect the future, according to Aristotle, this implies that future events are contingent. A contingent event is one that is not inevitable. By an action is meant "free action," and by that is meant an action which someone can refrain from performing. If we could refrain from our actions, it is a contingent fact that we perform our actions. Therefore future events, being caused by our actions, are contingent. "Events may either take place or not take place."
10. Aristotle proceeds to criticise his original fatalist argument. "It is not of necessity that everything takes place." Aristotle accepts that the statement "There will be a sea-battle tomorrow" is either true or false; he does not accept that it is inevitable that there will be a sea-battle tomorrow, or that it is inevitable that there will not be a sea-battle tomorrow. Consider the prediction and the negation of the prediction; one of them take by itself is not inevitable or necessary. However, says Aristotle, the original argument was correct in that if we take both prediction and negation together, it is inevitable that one will be correct and the other false.

Bibliography

Adams, Robert (1987). ‘Must God Create the Best?,’ in *The Virtue of Faith and Other Essays in Philosophical Theology*. New York: Ox ford University Press, 51-64.

Aquinas, Thomas (1945). *Basic Writings of Saint Thomas Aquinas* (2 vols.). New York: Random House.

----- (1993). *Selected Philosophical Writings*, ed. T. McDermott. Ox ford: Ox ford University Press.

Aristotle (1985). *Nichomachean Ethics*, translated by Terence Irwin. Indianapolis: Hackett Publishing.

Augustine (1993). *On the Free Choice of the Will*, tr. Thomas Williams. Indianapolis: Hackett Publishing.

Ayer, A.J. (1982). ‘Freedom and Necessity,’ in Watson (1982b), ed., 15-23.

Baker, Lynne (2000). *Persons and Bodies: A Constitution View*. Cambridge: Cambridge University Press.

Bourke, Vernon (1964). *Will in Western Thought*. New York: Sheed and Ward.

Campbell, C.A. (1967). *In Defence of Free Will & other essays*. London: Allen & Unwin Ltd.

Chisholm, Roderick (1982). ‘Human Freedom and the Self,’ in Watson (1982b), 24-35.

----- (1976). *Person and Object*. LaSalle: Open Court.

Clarke, Randolph (1993). ‘Toward a Credible Agent-Causal Account of Free Will,’ in O'Connor (1995), ed., 201-15.

----- (1995). ‘Indeterminism and Control,’ *American Philosophical Quarterly* 32, 125-138.

----- (1996). ‘Agent Causation and Event Causation in the Production of Free Action,’ *Philosophical Topics* 24 (Fall), 19-48.

----- (2003). *Libertarian Accounts of Free Will*. Ox ford: Ox ford University Press.

Dennett, Daniel (1984). *Elbow Room: The Varieties of Free Will Worth Having*. Cambridge. MA: MIT Press.

Descartes, René (1984). *Meditations on First Philosophy* [1641] and *Passions of the Soul* [1649], in *The Philosophical Writings of Descartes*, vol. I-III, translated by Cottingham, J., Stoothoff, R., & Murdoch, D.. Cambridge: Cambridge University Press.

Donagan, Alan (1985). *Human Ends and Human Actions: An Ex ploration in St. Thomas's Treatment*. Milwaukee: Marquette University Press.

Dilman, Ilham (1999). *Free Will: An Historical and Philosophical Introduction*. London: Routledge.

Double, Richard (1991). *The Non-Reality of Free Will*. New York: Ox ford University Press.

Edwards, Jonathan (1957) [1754]. *Freedom of Will*, ed. P. Ramsey. New Haven: Yale University Press.

Ekstrom, Laura (2000). *Free Will: A Philosophical Study*. Boulder, CO: Westview Press.

Farrer, Austin (1958). *The Freedom of the Will*. London: Adam & Charles Black.

Fischer, John Martin (1994). *The Metaphysics of Free Will*. Ox ford: Blackwell.

----- (1999). 'Recent Work on Moral Responsibility,' *Ethics* 110, 93-139.

Fischer, John Martin and Ravizza, Mark. (1992). 'When the Will is Free,' in O'Connor (1995), ed., 239-269.

Fischer, John Martin (1998) *Responsibility and Control*. Cambridge: Cambridge University Press.

Frankfurt, Harry (1969). ‘Alternate Possibilities and Moral Responsibility,’ *Journal of Philosophy* 66, 829-39.

----- (1982). ‘Freedom of the Will and the Concept of a Person,’ in Watson (1982), ed., 81-95.

----- (1988). *The Importance of What We Care About*. Cambridge: Cambridge University Press.

----- (1992). ‘The Faintest Passion,’ *Proceedings and Addresses of the American Philosophical Association* 66, 5-16.

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

----- (1997). ‘Freedom, Responsibility, and Agency,’ *The Journal of Ethics* 1, 85-98.

----- (2002) ‘Reasons Ex planations of Action: Causalist versus Noncausalist Accounts,’ in Kane, ed., (2002), 386-405.

Hobbes, Thomas and Bramhall, John (1999) [1655-1658]. *Hobbes and Bramhall on Liberty and Necessity*, ed. V. Chappell. Cambridge: Cambridge University Press.

Honderich, Ted (1988). *A Theory of Determinism*. Ox ford: Ox ford University Press.

Howard-Snyder, Daniel and Frances (1994). ‘How an Unsurpassable Being Can Create a Surpassable World,’ *Faith and Philosophy* 11, 260-8.

Howard-Snyder, Daniel and Moser, Paul, eds. (2002). *Divine Hiddenness: New Essays*. Cambridge: Cambridge University Press.

Hume, David (1977) [1748]. *An Enquiry Concerning Human Understanding*. Indianapolis: Hackett Publishing.

Kane, Robert (1995). ‘Two Kinds of Incompatibilism,’ in O'Connor (1995), ed., 115-150.

----- (1996). *The Significance of Free Will*. New York: Ox ford University Press.

Kane, Robert, ed., (2002). *Ox ford Handbook on Free Will*. New York: Ox ford University Press.

Kant, Immanuel (1993) [1788]. *Critique of Practical Reason*, tr. by Lewis White Beck. Upper Saddle River, NJ: Prentice-Hall Inc.

Kretzmann, Norman (1997). *The Metaphysics of Theism: Aquinas's Natural Theology in* Summa Contra Gentiles I. Ox ford: Clarendon Press.

Leibniz, Gottfried (1985) [1710]. *Theodicy*. LaSalle, IL: Open Court.

Libet, Benjamin (2002). 'Do We Have Free Will?' in Kane, ed., (2002), 551-564.

MacDonald, Scott (1998). 'Aquinas's Libertarian Account of Free Will,' *Revue Internationale de Philosophie*, 2, 309-328.

Magill, Kevin (1997). *Freedom and Experience*. London: MacMillan.

McCann, Hugh (1998). *The Works of Agency: On Human Action, Will, and Freedom*. Ithaca: Cornell University Press.

Mele, Alfred (1995). Autonomous Agents (New York: Ox ford University Press).

Morris, Thomas (1993). 'Perfection and Creation,' in E. Stump. (1993), ed., 234-47

Murray, Michael (1993). 'Coercion and the Hiddenness of God,' *American Philosophical Quarterly* 30, 27-38.

----- (2002). 'Deus Absconditus,' in Howard-Snyder amd Moser (2002), 62-82.

Nozick, Robert (1995). 'Choice and Indeterminism,' in O'Connor (1995), ed., 101-14.

O'Connor, Timothy (1993). 'Indeterminism and Free Agency: Three Recent Views,' *Philosophy and Phenomenological Research*, 53, 499-526.

-----, ed., (1995). *Agents, Causes, and Events: Essays on Indeterminism and Free Will*. New York: Ox ford University Press.

----- (2000). *Persons and Causes: The Metaphysics of Free Will*. New York: Ox ford University Press.

----- (forthcoming). 'Freedom With a Human Face,' *Midwest Studies in Philosophy*, Fall 2005.

Pettit, Philip (2001). *A Theory of Freedom*. Ox ford: Ox ford University Press.

Pereboom, Derk (2001). *Living Without Free Will*. Cambridge: Cambridge University Press.

-----, ed., (1997). *Free Will*. Indianapolis: Hackett Publishing.

Pink, Thomas (2004). *Free Will: A Very Short Introduction*. Ox ford: Ox ford University Press.

Plato (1997). *Complete Works*, ed. J. Cooper. Indianapolis: Hackett Publishing.

Quinn, Phillip (1983). 'Divine Conservation, Continuous Creation, and Human Action,' in A. Freddoso, ed. *The Existence and Nature of God*. Notre Dame: Notre Dame University Press.

Reid, Thomas (1969). *Essays on the Active Powers of the Human Mind*, ed. B. Brody. Cambridge: MIT Press.

Rowe, William (1993). 'The Problem of Divine Perfection and Freedom,' in E. Stump (1993), ed., 223-33.

----- (1995). 'Two Concepts of Freedom,' in O'Connor (1995), ed. 151-71.

----- (2004). *Can God Be Free?*. Ox ford: Ox ford University Press.

Sartre, John Paul (1956). *Being and Nothingness*. New York: Washington Square Press.

Schopenhauer, Arthur (1999) [1839]. *Prize Essay on the Freedom of the Will*, ed. G. Zoller. Cambridge: Cambridge University Press.

Scotus, John Duns (1986). ‘Questions on Aristotle's Metaphysics IX , Q.15’ in *Duns Scotus on the Will and Morality* [selected and translated by Allan B. Wolter, O.F.M.]. Washington: Catholic University of America Press.

----- (1994) [1297-99]. *Contingency and Freedom: Lectura I 39*, tr. Vos Jaczn *et al*. Dordrecht: Kluwer Academic Publishers.

Shatz, David (1986). ‘Free Will and the Structure of Motivation,’ *Midwest Studies in Philosophy* 10, 451-482.

Smilansky, Saul (2000). *Free Will and Illusion*. Ox ford: Ox ford University Press.

Strawson, Galen (1986). *Freedom and Belief*. Ox ford: Clarendon Press.

Strawson, Peter (1982). ‘Freedom and Resentment,’ in Watson (1982), ed., 59-80.

Stump, Eleonore (1996). ‘Persons: Identification and Freedom,’ Philosophical Topics 24, 183-214.

-----(1997). ‘Aquinas's Account of Freedom: Intellect and Will,’ *The Monist* 80, 576-597.

-----, ed., (1993). *Reasoned Faith*. Ithaca: Cornell University Press.

van Inwagen, Peter (1983). *An Essay on Free Will*. Ox ford: Ox ford University Press.

----- (1994). ‘When the Will is Not Free,’ *Philosophical Studies*, 75, 95-113.

----- (1995). ‘When Is the Will Free?’ in O'Connor (1995), ed., 219-238.

Wainwright, William (1996). ‘Jonathan Edwards, William Rowe, and the Necessity of Creation,’ in J. Jordan and D. Howard-Snyder, eds., *Faith Freedom, and Rationality*. Lanham: Rowman and Littlefield, 119-133.

Wallace, R. Jay (2003). ‘Addiction as Defect of the Will: Some Philosophical Reflections,’ in Watson, ed., (2003b), 424-452.

Watson, Gary (1987). ‘Free Action and Free Will,’ *Mind* 96, 145-72.

----- (2003a). ‘Free Agency,’ in Watson, ed., 1982b.

-----, ed., (2003b). *Free Will*. 2nd ed. Oxford: Oxford University Press.

Wegner, Daniel (2002). *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.

Widerker, David and McKenna, Michael, eds., (2002). *Moral Responsibility and Alternative Possibilities*. Aldershot: Ashgate Publishing.

Wolf, Susan (1990). *Freedom Within Reason*. Oxford: Oxford University Press.

$\Diamond(\forall x\ )\ (\forall p)\ \mathrm{K}x\ p \rightarrow \in p$
$(\forall x\ )\ (\forall p)\ \mathrm{K}x\ p(\mathrm{t}^1) \rightarrow p(\mathrm{t}^2)$
$(\exists \mathrm{x}\ )\ (\exists)$ - & v $\in \rightarrow \leftrightarrow$
$\Diamond(\forall x\ )\ (\forall p)\ \mathrm{K}x\ p \rightarrow \in p$
$(\forall x\ )\ (\forall p)\ \mathrm{K}x\ p(\mathrm{t}^1) \rightarrow p(\mathrm{t}^2)$
$(\exists \mathrm{x}\ )\ (\exists)$ - & v $\in \rightarrow \leftrightarrow$

(1) Hard
(2) Logic throughout esp. **modal logic**
(3) Contemporary philosophers mainly to FN. refs in text to (x, 1979)
(4) There is considerable material about free will in *The God of the Physicists* (inc. Hegel on S-C applied to freedom). Rewrite this, here, with rigour. (Use everything you have on free will in this book.)

Eccles, J.C. (1986). 'Do mental events cause neural events analogously to the probability fields of quantum mechanics?' Proc. R. Soc. London. B 227, 411-428.