

## **Classic Compatibilism, Romantic Compatibilism, and the Claims of Commonsense**

**Abstract:** There are different worries about free will, but the worry concerning determinism has to do with a thesis about the laws, or at least the fundamental laws; roughly, that these laws state sufficient conditions for the occurrence of events. More precisely, determinism is the thesis that the state of the entire universe at any time, together with the laws, is logically sufficient for the state of the universe at any other time. Determinism poses a prima-facie threat to our commonsense belief that we have free will by having the apparent entailment that we never have a choice about anything we do because we are *never* able to do *anything* other than what we actually do. This is the correct way to understand the free will/determinism problem, and the Classical compatibilist is any compatibilist who understands the problem in this way. If we understand the problem this way, it is soluble. Solving the free will/determinism problem is *not* the same as solving the problem of moral responsibility. But it is a big step in that direction, because understanding the compatibilist solution to the free will/determinism problem provides us with the *necessary first steps* for providing a clear-headed account of the bundle of abilities in virtue of which we have, not only the free will that we share with children and other pre-moral creatures, but also the "moral freedom" required for moral responsibility.

"Newspeak... differed from almost all other languages in that its vocabulary grew smaller instead of larger every year. Each reduction was a gain, since the smaller the area of choice, the smaller the temptation to take thought." (George Orwell, *Principles of Newspeak, Appendix to 1984*)<sup>1</sup>

### **1. What is the free will/determinism problem about?**

Let me clarify. I'm talking about the free will/determinism problem, and not about various other problems that arise with respect to free will. And when I say 'free will', I mean free will, not moral responsibility. And I mean determinism, not a vague thesis about causation (eg. that every event has a cause, eg. that every instance of causation is governed by laws), nor any thesis about explanation (including claims that some philosophers make in the name of 'naturalism' or 'mechanism'.)

What do I mean by 'determinism'? I mean what philosophers who are reasonably careful with their terms now standardly mean: Determinism is the thesis that the state of the entire universe at any time, together with the laws, is logically sufficient for the state of the universe at any later time.

Although science no longer tells us that determinism is true, for a long time -- and for most of the history of the free will/determinism problem -- it was assumed that science does tell us that determinism is true. And even now it looks like the odds are pretty good that determinism is close enough to being true, so far as our behavior is concerned. At the very least, we are not *entitled* to assume that it is false, in the ways that might make the

February 2010

kind of difference that the libertarian thinks is required for free will. These facts are significant because if determinism robs us of free will, then we are in the following epistemic situation: While it might be true that we have free will we are not entitled to claim that we *know* that we have free will. But not only do we believe we have free will; we also think that we *know* that we have free will.

What do I mean by 'free will'? Nothing fancy, and therefore not what most free will philosophers mean. I don't regard 'free will' as a term of art, or as a technical term, or as a placeholder for 'whatever sort of freedom is sufficient for moral responsibility'. I mean what ordinary people mean, what our students mean, when they insist that they have free will, that everyone -- well, nearly everyone -- has free will, and has it all the time, or at least most of the time, or at least by the time they are 18. After a bit of questioning, it becomes clear that the one thing that everyone agrees on is that to have free will is to have the ability to *make choices* (by thinking for yourself, on the basis of your own beliefs and desires, by evaluating the consequences of contemplated actions, and so on) and to at least sometimes (and perhaps almost all the time) *really have a choice* about what to do.

Our commonsense belief is that we have free will. We believe that we have free will, and we also believe that we often (though not always) *use* or *exercise* our free will -- we take a moment to think before we act; we choose, and *then* we act. And while we are sometimes in a position where we *make a choice* while being mistaken about what our options are, this is not always, or even usually, the case. Ordinarily, when we make a choice we really do have the choice we think we have. You can leave this room right now, or you can stay. You really *can* do either of these things. It's up to you. You have a choice.

Commonsense either doesn't know, or doesn't take seriously, the thought of determinism. But as soon as a philosopher explains - in sufficient detail so it sinks in - what the thesis of determinism says, commonsense sees the problem: the truth of determinism means the absence of free will.

The reasoning from the thesis of determinism to the conclusion of no free will is usually not very sophisticated and often makes the fatalist's mistake of reasoning from truth to necessity.

"Determinism professes that those parts of the universe already laid down absolutely appoint and decree what the other parts will be. The future has no ambiguous possibilities hidden in the womb. The part that we call the present is compatible with only one totality... necessity on the one hand and impossibility on the other are the sole categories of the real. Possibilities that fail to get realized, are, for determinism, pure illusions; they never were possibilities at all."<sup>2</sup> (William James)

And, moving forward to the 21st century:

February 2010

"The essence of determinism is that everything that happens is the only thing that could possibly happen (given the past) under those circumstances. The category of the possible and the category of the actual are exactly the same. To a determinist, the universe is just grinding along, as a giant machine, with no uncertainty whatsoever. The future and the past are both set in stone, so to speak...Determinism directly contradicts the everyday experience of making choices and having multiple options... To a determinist, choice (in this sense) is an illusion, because only one outcome is possible all along."<sup>3</sup> (Roy Baumeister)

James and Baumeister are apparently reasoning from 'determinism entails that the future is determinate' to 'determinism entails that the future is necessary'. But they are also wrong in a more fundamental way. The free will/determinism problem isn't about whether it's *possible that* our actions turn out differently, or whether it's *possible that* we make different choices; it's about whether *we* are free to choose and act in ways other than the ways we actually choose and act. And questions of our freedom are questions about what *we* are *able* to do, or what is *in our power* to do.<sup>4</sup> If determinism robs us of free will it must be because determinism has the consequence that it is *never*, at any point in our lives, *in our power* either to make different choices or to perform different overt actions.

Still, even if we set aside the fatalist's collapse of truth and necessity, and distinguish possibility from ability, this is not enough to lay to rest the worry about determinism and choice. If determinism is true, it *seems* that *every single case of choice* is an illusion. Not because we never undergo a process that we might call 'deliberating' 'deciding', or 'making a choice', but because when we do so we believe that we have *more than one choice* to make -- that we have options, alternative courses of action, more than one thing we are able to do. And if determinism is true, it is difficult to understand how this can *ever* be true. Since we are human beings and not gods, our abilities are constrained by the circumstances in which we find ourselves as well as the laws. And determinism says that our circumstances together with the laws fix a unique future. How, then, can it ever be true that we are able to do otherwise?

There is more to our view of ourselves as free agents than the belief that we have genuine alternatives and real choices, but this belief is at the core of our commonsense view of ourselves as persons with free will. If we were somehow convinced that we never *really* have a choice about what we do, we would conclude that our belief that we have free will is either outright false, or, at best, that we are radically mistaken about the nature of the free will that we have.

We can agree on this - that is, agree on the centrality of choice and alternatives to free will -- while leaving other questions up for grabs. Do we have a choice only on those occasions when we deliberate, weigh our reasons, make a decision, and *then* act, or do we also have a choice on those occasions where we act spontaneously, or on the basis of a hunch, or what feels right, or seems obvious? Do we have a choice in Buridan's Ass cases? Do we have a choice in cases of coercion, and other cases where we say we have 'no real choice' or 'no reasonable choice'? What about cases where making a particular

February 2010

choice would go so deeply against our character or values that it is 'unthinkable' for us to take it seriously, as one of our options? And so on.

We can also leave open questions about the exact relationship between choice and moral responsibility. We can all agree that moral responsibility doesn't require having a choice at the time of action -- we blame a drunk driver for the accident he causes, even if he wasn't able, at the time of the accident, to do otherwise -- but does it require having a choice between sets of foreseeable (or reasonably foreseeable) outcomes of one's action? Or is it enough, so far as moral responsibility is concerned, that a person had a choice at some time in the past, between 'self-forming actions' which led, eventually, to the formation of her present character? <sup>5</sup>

The answers to these questions don't matter, so far as the free will/determinism is concerned, because the threat posed by determinism is an *indiscriminating* one, which levels all these distinctions.

If determinism is true, it seems that what we actually do (whether we correctly describe this as a decision, a choice, an action, an intentional action, a voluntary action, a free action, an action caused by a character we helped to cause, and so on) is, it appears, *always*, the one and only thing we are able to do.

And since having a choice about something requires being *able* to do otherwise, it *seems* that determinism has the upshot that we never have a choice about anything.

We can sum up this line of thought with the following simple argument.

### **The No Choice Argument for Incompatibilism:**

1. I have free will only if I sometimes have a choice about what I do.
2. I have a choice about what I do only if I am able to do otherwise.
3. If determinism is true, I am *never* able to do otherwise.
4. Therefore, if determinism is true, I never have a choice about what I do.
5. Therefore, if determinism is true, I have no free will.

I will be arguing that the incompatibilist owes us an *argument* in defense of Premise 3.

## **2. Classic Compatibilism**

I define a Classic compatibilist as any compatibilist who understands the free will/determinism problem in the way I have just described. More specifically, the Classic compatibilist is any compatibilist who rejects the third premise of the No Choice argument -- the No Ability to do Otherwise premise.

**Hobart** (Classic compatibilist *par excellence*): "The thesis of this article is that there has never been any ground for the controversy between the doctrine of free will and

February 2010

determinism...that they have been opposed only because of our natural want of the analytic imagination. In so saying, I do not tamper with the meaning of either phrase. That would be unpardonable. I mean free will in the natural and usual sense, in the fullest, the most absolute sense...I mean it as implying responsibility, merit and demerit, guilt and desert. I mean it as implying, after an act has been performed, that one 'could have done otherwise' than one did."

"...that we are free in willing is, broadly speaking, a fact of experience. That broad fact is more assured than any philosophical analysis. It is therefore surer than the deterministic analysis of it, entirely adequate as that in the end appears to be..."<sup>6</sup>

In the old days, when the fundamental laws were thought to be deterministic, every compatibilist was a soft determinist, and thus asserted not only the compatibility of free will and determinism, but the *actual co-existence* of free will and determinism. Indeed, one standard way of arguing for the compatibility of free will and determinism was by arguing that we have good reason to believe that both of these apparently incompatible claims are *in fact* true, and therefore the two are, despite initial appearances, compatible. Since we can no longer assume that determinism is true, we can no longer defend compatibilism this way. (And we would be unwise to defend compatibilism by arguing, as Hobart did, that free will "strictly implies" determinism.) But the Classic compatibilist, as I understand her, is committed to an *existential claim* as well as a compatibility claim. She says that we *in fact* have the free will that commonsense says we have *and* that we would have the free will that commonsense says we have *even if* determinism turned out to be true. (The addition of the belief 'determinism is true' to our set of beliefs about the world should not cause us to revise our belief that we have free will. We in fact have free will *and* there are nearby possible worlds where determinism is true and we have free will.)

I define a Classic incompatibilist as any incompatibilist who sees the problem in these terms. The Classic incompatibilist says that *if* determinism turned out to be true, we would *not* have the free will we think we have; in particular, the beliefs that we have in situations in which we make a choice would *always* be mistaken. And, more generally, the Classic incompatibilist says that determinism strictly implies the absence of free will. (No deterministic world is a world where anyone has free will.) I understand Peter van Inwagen as a Classic incompatibilist.

### 3. Romantic Compatibilism

As I see it, this is how the free will/determinism problem was understood until at least the first half of the twentieth century. Call this the traditional, or classic, understanding of the problem.<sup>7</sup>

But things are different now. There are almost no compatibilists left anymore -- at least, not if we understand a compatibilist as someone who rejects the No Ability to do Otherwise premise of the No Choice argument and asserts that the making and having of

February 2010

choices is compatible with determinism. Almost everyone currently working on free will either accepts the No Ability to do Otherwise premise or at least *doesn't bother* to contest it. But lots of philosophers still call themselves 'compatibilists' (or 'semi-compatibilists' or 'revisionary' compatibilists). What happened?

Peter van Inwagen's Consequence Argument<sup>8</sup> and Frankfurt's argument against the Principle of Alternate Possibilities<sup>9</sup> is what happened. The Consequence Argument had no immediate effect on compatibilists, but it made incompatibilism respectable again. Frankfurt's argument had no immediate effect on incompatibilists, but it encouraged compatibilists to stop worrying about the traditional free will/determinism problem and to devote all their attention to the problem of moral responsibility.<sup>10</sup> Later, some compatibilists came to be persuaded by the Consequence argument (while others continued to ignore it). And some incompatibilists came to be persuaded by Frankfurt's argument, but offered arguments for a new kind of incompatibilism ("Source incompatibilism"). And the peculiar result is that the debate between compatibilists and incompatibilists about the traditional free will/determinism problem has virtually ceased. The newfangled compatibilists have taken over, and there are almost no Classic compatibilists left.

My main concern in this paper is not with this new breed of compatibilists, but I need a name for them. Since I have already been assigned the label 'Classic compatibilist', I hereby name all the others 'Romantic compatibilists'. A Romantic compatibilist, I hereby stipulate, is any philosopher who calls herself a compatibilist but does not understand the free will/determinism problem in the traditional way, as I have described it. Some Romantic compatibilists do not even talk about determinism, as opposed to some less clearly defined thesis about causation or explanation or "what science tells us about our place in the natural order". Many Romantic Compatibilists are what we might call 'moral compatibilists'; they understand the freewill/determinism problem either *explicitly* in terms of moral responsibility or as the problem of showing that the kind of freedom that "justifies" or "grounds" moral responsibility is compatible with determinism. But whether she explicitly concerns herself with moral responsibility or not, the Romantic Compatibilist rejects the idea that the free will/determinism problem is a problem about choice and the ability to do otherwise. She either accepts, *or at least does not reject*, the No Ability to do otherwise premise of the No Choice argument.

Romantic compatibilists are either explicitly or implicitly committed to a denial of the commonsense belief that we have free will. And because of this, they owe us an *analysis*, or an *account*, or at least a *sufficient condition* of freedom, or moral freedom, or whatever it is that they claim is compatible with determinism (or causation, or explanation, or naturalism, or mechanism, "what science tells us", and so on).

The Classic compatibilist, who affirms the claims of commonsense, is under no similar obligation.

I will explain.

#### 4. Free Will, Determinism, and the Claims of Commonsense

Freedom of will is distinct from freedom of (overt) action; a person doesn't lose her freedom of will by being gagged and shackled.<sup>11</sup> She retains a highly valuable ability, the ability to use her mind in ways that allow her to transcend the constraints imposed on her body.<sup>12</sup> However, our commonsense belief that we have free will is not limited to the belief that we have the kind of freedom that the gagged and shackled prisoner has. We believe that we typically enjoy freedom of action *in addition to* freedom of will. More specifically, we believe that we are often in a position in which we not only make a choice, but also have a choice about what we do, where 'do' means "perform an overt action".

We can decide without acting, and we can act without deciding, but there are also lots of situations in which we do both. So let's set aside the more controversial cases and consider only those cases in which almost everyone would agree that we have a choice about what we do: cases in which we *first* decide, and *then* act on the basis of our decision.

You offer me cheesecake or chocolate mousse; I hesitate for just a moment - if only I could have both - then pick the mousse. I'm at a fork in the road and must decide between the winding scenic coastal route, and the speedy but boring freeway; I take the freeway. The speeches are finally over, and the vote has been called: all in favor, raise their hands; I raise my hand. We are in situations like these daily, situations which have in common the following features:

1. We believe that we have options (alternative courses of action): that is, we believe that there are *at least two different things we are able to do*. (Eat one dessert or another, take one road or the other, raise our hand or keep it lowered). Call these options 'A' and 'B'.
2. We deliberate (or ponder, ruminate, consult our feelings, etc.) between A and B.
3. We decide to do A. (We make up our minds, make a choice, form an intention.)
4. We do A.

We are, of course, sometimes mistaken about what our options are. We ponder the offerings on the menu, trying to decide between the mousse and the cheesecake, not realizing that the mousse is not available today. We contemplate going out for a drive, having forgotten that the car is at the garage.

This shows that we are *not entitled to infer*, from the fact that we deliberate between A and B, that A and B are both options for us. We may be mistaken about what we are able to do. But it seems incredible to suppose that we are *always* mistaken. That is, it seems incredible to suppose that we are *never* in a situation in which, in addition to 1-4, it is also true that:

5. The belief that we had before we made up our mind was a *true belief*. While we were deliberating, we really *were able to do A* and we really *were able to do B*.

Call the belief that we are often, or at least *sometimes*, in situations where our beliefs about our options are correct and 1-5 are all true: **Choice**.<sup>13</sup>

**Choice** is an uncontroversial part of our commonsense view of ourselves as agents with free will. It seems as obviously and undeniably true as other beliefs that only a philosopher would question: the belief that there are tables and chairs, the belief that things continue to exist through time, the belief that I have a pair of hands. **Choice** is a commonsense platitude -- a commonsense belief that is so obvious that we typically don't bother saying it. Like our belief that there are tables and chairs and things that continue to exist through time, the belief that we are often in situations in which we not only decide and act but are *able to do more than one thing* is more obvious and self-evident than any analysis that we can give of it. We do not need to give arguments or provide an analysis in order to defend our belief in a commonsense platitude.

In saying this, I do not mean to suggest or imply any of the following: That we should never be convinced, by argument, to reject a commonsense platitude,<sup>14</sup> that we have a priori knowledge of commonsense platitudes; that Moore's response to the skeptic is correct,<sup>15</sup> that commonsense platitudes are philosophically unproblematic. In particular, I don't mean to suggest that if something is a commonsense platitude, there is no further work left for the philosopher to do. (Compare the debate between perdurantists and endurantists about the two different ways in which it might be true that things continue to exist through time.) But I do claim that we need to be convinced, by a *very good argument*, before we should be persuaded that a commonsense platitude is false.

If we understand the free will/determinism problem as a problem posed by determinism to the claims of commonsense, we are in a better position to understand the response of the historic Classic compatibilists to the problem.

The historic Classic compatibilists viewed the free will/determinism problem in something like the way we think of the problem of skepticism about the external world. We worry about the problem of skepticism not because we are tempted to accept the skeptic's conclusion, but because we worry about whether we have found the right diagnosis of *what is wrong with the skeptic's argument*. The Classic compatibilists (who, unlike contemporary compatibilists, were all soft determinists) thought about the free will/determinism problem the same way. They were not, for a moment, tempted to accept the hard determinist's claim that science tells us that we have no free will; the problem, as they saw it, is say *what is wrong with any argument that takes the truth of determinism as a premise and has the denial of our commonsense belief that we have free will as its conclusion*.

They offered a variety of diagnoses of why we make a mistake if we reason from the truth of determinism to the conclusion that we lack free will: the confusion of causation

and compulsion, of descriptive and prescriptive laws, of 'categorical' and 'hypothetical' or 'conditional' liberty. Sometimes they were too quick with their criticisms, and with their readiness to see the problem as a "pseudo-problem", a "merely verbal" problem, or a "failure of the analytical imagination". But they were right, I think, to insist that the problem is *about* something *non-mysterious and commonplace*, and right also to insist that the hard determinist (or incompatibilist) owes us an argument in defense of the claim that science in fact shows (or might show, consistent with our present evidence) that this commonplace thing does not exist.

## 5. Should We be Classic Compatibilists?

Yes, because Classic compatibilism asserts the claims of commonsense, and in the absence of a good argument, we should assume that commonsense is correct.

Romantic compatibilists say we shouldn't be Classic compatibilists, but their reasons are invariably bad ones, and come down to one of three claims:

### i) Failure of Simple Conditional Analysis of "could have done otherwise"

Classic compatibilism is often associated with a particular claim that was once made about the *meaning* of sentences like 'S could have done otherwise'. The claim was that the "can" that we use when we attribute abilities to agents is always "iffy"; the relevant 'can' sentences are disguised conditionals. More specifically, the correct analysis of "S could have done otherwise" is something along the lines of "if S had chosen (or decided, tried, intended, or wanted) to do otherwise, S would have done otherwise". There was a vigorous debate, ironically conducted just before the advent of possible worlds semantics for counterfactuals, and the debate was decisively lost by the defenders of the Conditional Analysis.

But we shouldn't confuse this kind of *simple* Conditional Analysis with Classic compatibilism. The Simple Conditional Analysis was just one compatibilist solution -- oversimple, as it turned out - to the free will/determinism problem. We can agree that it fails, without agreeing that Classic compatibilism fails. If the abilities that we have in choice situations are dispositions to act,<sup>16</sup> and if David Lewis is right about what dispositions are<sup>17</sup>, then the counterexamples that defeated the Simple Conditional analysis no more count against the claim that our abilities to decide and act are dispositions than counterexamples to a Simple Conditional analysis of disposition count against the claim that fragility is a disposition.<sup>18</sup>

It has been claimed that there are decisive counterexamples to Lewis's Revised Conditional Analysis of dispositions. This may be so. (I don't agree, but this is not the place to defend that claim.) But all that this would show is that Lewis's analysis of dispositions (and thus the abilities that constitute free will, if these are dispositions) is wrong. It doesn't show that Classic compatibilism is wrong. For another analysis might be correct. And even if no other analysis is correct, this does not show that Classic

February 2010

compatibilism is false. For the Classic compatibilist defends the claims of commonsense and this does not require an analysis.

## ii) Frankfurt's argument against the Principle of Alternate Possibilities

Free will was traditionally thought to matter at least partly because free will is a necessary (though not sufficient) condition of moral responsibility. If there is no point, during our *entire lives*, that we have any kind of choice about anything we decide or do, it seems that we can never be morally responsible *for anything*. But Frankfurt notoriously argued that this commonsense belief is false. More specifically, he argued, by way of an ingenious thought experiment, that a person may be morally responsible for his overt action even though he lacks freedom of will as well as freedom of overt action -- even though he can neither decide nor act otherwise. So the debate between the Classic compatibilist and the Classic incompatibilist is *irrelevant*, so far as moral responsibility is concerned.

I have argued, in several different places, that Frankfurt's argument fails. It is based on an underdescribed thought experiment that conflates two kinds of background interveners -- Bodyguards, who intervene *after* the agent has decided or begun to act contrary to the intervener's plan, and Pre-Emptors, who intervene *before* the agent even begins to decide contrary to the plan. A Bodyguard robs the agent of all freedom to successfully act otherwise, but leaves freedom of will (including the freedom to deliberate for the purpose of making a decision and the freedom to decide otherwise, or at least begin to decide otherwise) intact. A Pre-Emptor has the *power* to rob the agent of freedom of will, but since he doesn't exercise his power, he leaves the agent's freedom of will intact. The agent is at *risk* of loss, but he hasn't *actually* lost. If we think otherwise, it's because we aren't paying careful attention to these distinctions or because we are making a fatalist or other modal mistake.<sup>19</sup>

But *even if* Frankfurt's argument succeeds in showing that we may be morally responsible despite never having *any* choice about *anything*, this is beside the point, since the question of whether we would sometimes have choices even if determinism turned out to be true is of interest in its own right.

## iii) Merely verbal dispute

The last objection, by the Romantics against the Classics, is that if we understand the free will/determinism problem as a problem about whether we can ever do otherwise (or, as current jargon puts it, whether we ever have "alternate possibilities"), there is no genuine disagreement between compatibilists and incompatibilists. The compatibilist and incompatibilist are talking about different *senses* of 'can'; the incompatibilist says that determinism entails that no one can ever do otherwise, given the past and the laws; the compatibilist says that determinism is compatible with people doing otherwise, given some more restricted set of facts. But, says the Romantic compatibilist, there is a genuine

February 2010

disagreement between compatibilists and incompatibilists; therefore, we must understand the problem in some other way.

My reply to this is that the issue isn't what any philosopher -- compatibilist or incompatibilist -- means by 'can', or 'is able to'. The issue is what ordinary people mean.

My Classic incompatibilist opponent, van Inwagen, agrees with me.

"...compatibilists and incompatibilists mean the same thing by 'able'. And what do both compatibilists and incompatibilists mean by 'able'? Just this, what it means in English, what the word means. And, therefore, 'free will', 'incompatibilist free will', 'compatibilist free will' and 'libertarian free will' are four names for one and the same thing. If this thing is a property, they are four names for the property *is on some occasions able to do otherwise*. If this thing is a power or ability, they are four names for the power or ability to do otherwise than one in fact does."<sup>20</sup>

We can grant that ordinary people use 'can' and 'is able to' in different ways, but this is not relevant. According to the Classic compatibilist (and incompatibilist) the free will/determinism problem is a problem about Choice (its truth and compatibility with determinism), and 'can' and 'is able to' mean whatever they mean when we use these words in contexts of choice.

So I agree with van Inwagen's assessment of the situation - he and the Classic compatibilist are talking about the very same thing. He and I have a genuine, and not merely verbal, disagreement. One of us is right and one of us is wrong.

## 6. Classic Compatibilism and the Consequence Argument.

We need reason *not* to be Classic compatibilists. That is, we need *a good reason* to accept the claim that if determinism turned out to be true, **Choice** would be false. Appeal to intuition is not an argument. We already know that taking determinism seriously causes people to doubt their commonsense belief that they often have a choice about what they do. The philosophical problem is to *decide whether this doubt is justified*. We know that it is very easy to reason fallaciously -- from truth to necessity, from causation to compulsion, from descriptive to prescriptive laws -- when thinking about determinism and ability to do otherwise. What we need -- what the incompatibilist needs -- is an argument. A good one.

The only serious argument on the table is the Consequence argument. The job of the Classic compatibilist, at the present time, is to reply to the Consequence argument, and *only* to reply to the Consequence argument. The striking -- indeed shocking - fact about the present literature is that this argument is not seriously discussed by the vast majority of compatibilists.

The version I will discuss is due to David Lewis.<sup>21</sup>

February 2010

He tells us to think of the argument as a reductio. A (Classic) compatibilist is someone who claims that the truth of determinism is compatible with the existence of the kinds of abilities that we assume we have in typical choice situations. Let's call these 'ordinary abilities'. The Consequence argument claims that if we suppose that a deterministic agent has *ordinary abilities*, we are forced to credit her with *incredible abilities* as well.

Van Inwagen doesn't object to Lewis's way of stating his argument. On the contrary, he has said that Lewis's paper is "the finest essay that has ever been written in defense of compatibilism – possibly the finest essay that has ever been written about any aspect of the free will problem".<sup>22</sup>

Van Inwagen now agrees that the Consequence argument fails as a reductio.<sup>23</sup> However, he claims that it has nevertheless succeeded in 'raising the price' of compatibilism.<sup>24</sup>

I disagree. I say that the argument neither succeeds as a reductio nor succeeds in 'raising the price' of the price of compatibilism - that is, the price of commonsense at a deterministic world. What the argument does achieve -- at least on Lewis's articulation of it -- is a clear statement of what the Classic compatibilist position entails. But I think that every clear-headed Classic compatibilist already knew this.

Here, then, is Lewis's statement of the argument.

**The Consequence argument (Incredible Ability version):**

Pretend that determinism is true, and that I did not raise my hand (at that department meeting, to vote on a proposal) but had the ordinary ability to do so. If I had exercised my ordinary ability – if I had raised my hand -- then either the remote past or the laws of physics would have been different (would have to have been different). But if that's so, then I have at least one of two incredible abilities – the ability to change the remote past or the ability to change the laws. But to suppose that I have either of these incredible abilities is absurd. So we must reject the claim that I had the ordinary ability to raise my hand.

## **7. Lewis's Criticism of the Consequence Argument**

Lewis's criticism of the Consequence Argument was published in 1981, two years before the publication of van Inwagen's book, *An Essay on Free Will*. His criticism was impeccable but his timing was bad. Lewis had published *Counterfactuals* (his possible worlds semantics and logic for counterfactuals) only 8 years earlier, in 1973, and counterfactuals were still poorly understood, and apparently not understood at all by some of the critics of Lewis's reply who seemed to think that Lewis had invented "local miracles counterfactuals" for the express purpose of defending a new and bizarre kind of compatibilism - "Local Miracles Compatibilism". There was further confusion due to the fact that Lewis developed his theory of counterfactuals in two stages: the formal logic

February 2010

came first (in 1973); and it was not until 1979 ("Counterfactual Dependence and Time's Arrow") that Lewis proposed a detailed *similarity ranking* for possible worlds, and showed how to apply this similarity ranking in a way that gets the right verdicts for the counterfactuals that we standardly agree are true. It was also well-known, by then, that Lewis hoped to use counterfactuals to provide a counterfactual analysis of causation. All this, of course, was wildly ambitious, and many people were skeptical that Lewis could pull it all off. But -- and this is my main point -- it was natural, and understandable, *back then*, to think that Lewis's theory of counterfactuals is a "package deal" which you can accept only if you accept other parts of Lewisian metaphysics. So many people hesitated, and this may explain why Lewis's critique did not have the effect it should have had.

But now that time has passed and the dust has settled, it's clear that this is not the case. Lewis's theory of counterfactuals is independent of most of his other views. You can accept his theory of counterfactuals (including everything he says in response to the Consequence Argument) without accepting any of the following: Lewis's controversial brand of realism about possible worlds, his counterfactual analysis of causation, his "Best System" version of a Humean account of laws, his thesis of Humean supervenience.

Furthermore, Lewis's criticism of the Consequence Argument doesn't depend on the truth of *his* theory of counterfactuals. (His theory is, I believe, correct, but even if it weren't his criticism would still stand.)

Let me explain. I apologize in advance if this is old news to you. If it were old news to everyone, I wouldn't have to say it and we would all be Classic compatibilists by now. (Or at least I think we should. Randy will explain why we shouldn't.)

The Incredible Ability version of the Consequence argument nicely highlights a point that the better known modal version of the argument glosses over. The argument relies on a claim about *counterfactuals*. The argument says that if determinism is true, then at least one of these counterfactuals is true:

**Different Past:** If I had raised my hand, the remote past would have been different (would have to have been different).

**Different Laws:** If I had raised my hand, the laws would have been different (would have to have been different).

Now I agree that both these counterfactuals strike many people as incredible. But there is a reason for that -- we are not used to thinking in terms of determinism and we are not accustomed to counterfactual speculation about *what would have to have been the case* if anything at a deterministic world had happened in any way other than the way it actually happened.<sup>25</sup>

February 2010

On the other hand, we are good at evaluating counterfactuals, or at least some counterfactuals, and we are especially good at evaluating those counterfactuals that we entertain in contexts of choice, when we ask questions about the causal upshots of our contemplated actions. (What would happen if... I struck this match, put my finger in the fire, threw this rock at that window, raised my hand?) And when we contemplate our options, we *take for granted* the existence of many facts - including facts about the laws and the past.

In other words, when we evaluate counterfactuals in real life, we do so by considering *imaginary situations* which are very like the situation we are actually in, and we do not suppose that there are any *gratuitous* departures from actuality. And to suppose a difference in the past or the laws is a gratuitous difference -- if determinism is false.

So it is no surprise that when our attention is directed to **Different Past** and **Different Laws**, these counterfactuals strike us as incredible, or at least *odd*. But that doesn't mean that they are false, and if determinism is true, then either **Different Past** or **Different Laws** *is* true.<sup>26</sup>

So the first point is that we all need a theory of counterfactuals, and if determinism is true, the true counterfactuals will include either **Different Past** or **Different Laws**. (And if they don't include them, the **Incredible Ability** argument fails at its first step.)

The second point is that the *details* of the correct Classic compatibilist solution to the free will/determinism problem will turn on the details of the correct theory of counterfactuals. The study of counterfactuals is an area in which there has been an enormous gain in understanding in the last 40 years and while so far this understanding has been confined mostly to philosophers who specialize in the logic of counterfactuals and causation, I am hopeful that philosophers in general, and free will philosophers in particular, will eventually be better informed on the subject. And when this happens, I predict that Classic compatibilism will, once again, be the default view among philosophers.

(There will, I predict, be one of those annoying Gestalt shifts from: 'But what you are saying is incredible and therefore *clearly false*' to 'But what you are saying is so obviously true, it's not worth saying; no one ever denied *that*.')

But I digress.

If David Lewis's theory of counterfactuals is correct, or even more or less correct<sup>27</sup>, then the relevant counterfactuals about the past and laws, at a deterministic world, are:

**1. Same Past:** If I had raised my hand, the past would still have been exactly the same until shortly before the time of my decision.

**2. Slightly Different Laws:** If I had raised my hand, the laws would have been ever so slightly different in a way that permitted the occurrence of a lawful divergence from actual history shortly before the time of my decision.

February 2010

On the other hand, if Lewis's theory is wrong, and counterfactuals are always evaluated by holding the laws constant, then the relevant counterfactuals, at a deterministic world, are:

1. **Same Laws:** If I had raised my hand, the laws would still have been exactly the same.
2. **Completely Different Past:** If I had raised my hand, the past would have been different all the way back to the Big Bang.

We've got to choose. We need a theory of counterfactuals that applies at deterministic worlds, and our choice is limited to a theory that accepts **Slightly Different Laws** or **Completely Different Past**. Which theory we choose has nothing to do with the free will/determinism problem and everything with how we evaluate counterfactuals (in standard contexts).

Having sorted this out, I will now explain Lewis's critique of the **Consequence Argument** in a way that doesn't require you to accept the truth of Lewis's theory of counterfactuals:

Lewis's response to the **Consequence Argument** goes as follows: The argument trades on an equivocation between two counterfactuals.

- (C1) If I had raised my hand, the laws (or the past) would have been different.
- (C2) If I had raised my hand, my decision or action would have *caused* the laws (or the past) to be different

There is a corresponding equivocation between two ability claims:

- (A1) I have the ability to do something such that if I did it, the laws (or the past) would have been different.
- (A2) I have the ability to do something such that if I did it, my decision or action would have *caused* the laws (or the past) to be different.

The problem with the argument, says Lewis, is that it **equivocates between these two ability claims**. To count as a reductio against the compatibilist, the argument must establish that the compatibilist is committed to **A2**. But the compatibilist is committed only to **C1** and thus only to **A1**. The compatibilist is committed only to saying that if determinism is true, we have abilities which we would exercise *only if* the past (and/or the laws) had been different in the appropriate ways. And while this may sound odd, it is no more incredible than the claim that the successful exercise of our abilities depends, not only on us, but also on the co-operation of factors outside our control. Since we are neither superheroes nor gods, we are always in this position, regardless of the truth or falsity of determinism.

To sum up: The Consequence Argument was intended as an argument from premises that we must all accept -- premises about our lack of control over the past and the laws -- to the conclusion that if determinism is true, we don't have the free will commonsense says we have. The **Incredible Ability** version of the argument tried to show that if we attribute ordinary abilities to deterministic agents, we are forced to credit them with incredible past or law-changing abilities as well. But no such incredible conclusion follows. All that follows is something that we must accept anyway, as the price of our nongodlike nature: that the exercise of our abilities depends partly on circumstances outside our control.

Philosophical arguments go on forever, of course, but at this stage in the dialectic I think that we are entitled to conclude that the Consequence Argument fails, both as a *reductio* and as a price-raiser of Classic compatibilism. In the absence of any *other* argument, we are entitled to declare victory for Classic compatibilism.

## 8. The Rest of the Story

It doesn't follow, of course, that this is the end of the story, or that it wouldn't be nice -- really nice -- to have an account of the abilities that are the truth-makers of the facts in virtue of which we have the freedom of action as well as the freedom of will that we take for granted in contexts of choice.

I have sketched the beginnings -- but only the beginnings -- of an answer to this question in my "Free Will Demystified: A Dispositional Account"<sup>28</sup>. What I said in that paper was intended to address the question of the abilities in virtue of which we have moral freedom, and, in particular, those abilities to deliberate and decide that the person in a Frankfurt story clearly and *uncontroversially* continues to have. But it may be adapted to accommodate freedom of action as well as freedom of will.

The basic idea is this: the most fundamental facts about our abilities -- to perform overt as well as mental actions -- are either constituted by, or at least structurally similar to, intrinsically based dispositions like fragility. Since 'intrinsically-based disposition' is clunky, I now prefer the term 'narrow ability' -- the idea being that narrow abilities are the abilities that are shared by duplicates governed by the same laws. My claim, then, is that the most basic facts about our abilities are the facts about our narrow abilities. To have a narrow ability to do something is to have the skills required to do that thing plus 'what it takes' to use these skills at a particular time. When you fall asleep, you retain your skills but you are temporarily 'disabled' from using most of them, so you temporarily lack most of the narrow abilities that you have when awake. Having the narrow ability to do something is necessary but not sufficient for being able to do that thing in the sense relevant to choice. A skilled and conscious pianist with unbroken limbs, etc. has the narrow ability to play piano, but if she's locked in a piano-less room, she lacks what I call the 'wide ability' to play. Think of wide ability as narrow ability plus favorable surroundings, where favorable surroundings are the kind of thing that we have in mind

February 2010

when we speak about having the means and opportunity to do something, of their being no obstacles to our doing it, and so on.

The narrow/wide ability distinction is a technical distinction, but I think it maps closely enough to ordinary ways of talking (eg. the ability vs. ability plus opportunity distinction) and to familiar philosophical distinctions (eg. 'general' ability vs. 'particular' or 'all-in' ability) to be a useful starting point for formulating and criticising accounts of freedom of action or freedom of will. And, since the distinction is neutral with respect to the free will/determinism problem, we can use it to formulate that problem in a more precise way. Determinism *appears* to rob us of the free will we would otherwise have by having the upshot that we never have a choice about anything we do. But in real life there is more than one way in which we may fail to have a choice -- the fault might be in us (we lack one of the relevant narrow abilities to decide or to act) or the fault might be with our surroundings (in which case we've got the narrow ability, but lack the wide ability, to do otherwise). It follows, then, that if determinism robs us of free will, this might be because determinism has one of three different consequences:

1. We always lack the **narrow ability to decide or to choose** whether to do something.

Determinism is compatible with the agency we exercise when we perform overt actions but not with the agency we exercise when we decide or choose.

Determinism is like always being asleep or hypnotized?

2. We always lack the **narrow ability to act (decide, choose, etc.) otherwise**.

Determinism is compatible with deciding/choosing as well as acting, but not with deciding/choosing or acting *otherwise*.

Determinism is like always suffering from a series of different and constantly changing compulsive desires or pathological fears that render you temporarily unable to decide/choose/act in any way other than the way you actually act?

3. We are always in **surroundings that are unfavorable** and so we always lack the **wide ability to do otherwise**.

(Determinism is like always being surrounded by invisible force fields that would prevent the success of any attempt to do other than what we actually do?)

It's not *conceptually impossible* that we should be in one of these situations -- perhaps very powerful alien controllers could put us in one of these situations. And perhaps future discoveries about the brain will give us reason to believe that we are in one of these situations, though I doubt it. But it seems to me *highly implausible* that the mere fact of determinism would have the upshot that any of these scenarios obtain. Any argument that determinism would have this consequence has to pass a *very high* bar in order to succeed. So I think that we've got good reason to believe, not only that Classic

compatibilism is currently undefeated by argument, but that it will continue to remain undefeated.

---

<sup>1</sup> George Orwell, 1984, Harcourt Brace and Company, 1949.

<sup>2</sup> William James, "The Dilemma of Determinism", *An Address to the Harvard Divinity Students*, 1884, reprinted in *The Will to Believe and other Essays*, New York: Dover Publications, 1956.

<sup>3</sup> Roy Baumeister, "Determinism is not Just Causality", *Psychology Today*, June 23, 2009.

<sup>4</sup> Incompatibilists have, perhaps, not always been as clear as they should have been on this point. However, according to Peter van Inwagen, this is how the free will/determinism problem should be understood, and I agree. Peter van Inwagen, "How to Think about the Problem of Free Will", *Journal of Ethics*, 2008 12, 337-341.

<sup>5</sup> Robert Kane, *The Significance of Free Will*, OUP, 1998.

<sup>6</sup> R.E. Hobart, "Free Will as Involving Determination and Inconceivable Without It", *Mind* 63 (1934), 1-27.

<sup>7</sup> The philosophers who are usually identified as 'classic compatibilists' - roughly, most compatibilists of the pre-Frankfurt era - defended a variety of different accounts of free will or free action, accounts which are not extensionally equivalent: the claim that free will consists in the absence of constraint or impediments to our desires (Hobbes); the claim that we act freely insofar as we act without being compelled or coerced (Ayer); the claim that free will implies a person's possession of a law-governed power and means the absence of any interference with his exercise of that power (Hobart); the claim that to have free will is to have an ability to act that can be analysed in terms of counterfactual conditionals along the lines of 'if he had chosen to do X, he would have done X' (Moore). I think that what these different accounts have in common is that they are attempts to show that the truth of our commonsense beliefs about choice are compatible with determinism. But if I'm wrong about this, my main points in this paper remain unaffected. My concern is not primarily with how compatibilists have traditionally understood the free will/determinism problem; my concern is with the problem itself and how I think it should be understood.

<sup>8</sup> Peter van Inwagen, *An Essay on Free Will*, OUP, 1983.

<sup>9</sup> Harry Frankfurt, "Alternate Possibilities and Moral Responsibility", *Journal of Philosophy* 66 (1969).

<sup>10</sup> cf. "How, if at all, is responsibility possible?...What kinds of beings must we be if we are ever to be responsible for the results of our wills?" (Susan Wolf, *Freedom within Reason*, OUP, 1990, p. 4.)

---

<sup>11</sup> cf. Winston's heroic though ultimately doomed attempts to engage in philosophical argument with his torturer O'Brien in Orwell's *1984*, *ibid.* Granted, this is possible only because Winston remains able to talk. And, granted, other examples of persons who have lost almost all powers of control over their bodies -- Stephen Hawking, for instance, or the historian Tony Judt -- are also not examples of persons who have lost *all* freedom of (overt) action. But it is easy to imagine the next step in the sequence -- someone whose mind remains intact but who no longer has *any* power to control the movements of his body. And in such a case it seems clear that the person retains freedom of will while lacking all freedom of (overt) action.

<sup>12</sup> For an extraordinary account of the experience of a freedom so circumscribed, see Tony Judt, "Night", *The New York Review of Books*, Vol.57, #1, Jan. 14, 2010.

<sup>13</sup> Peter van Inwagen, my Classic incompatibilist opponent, agrees with me about this way or understanding the problem: "Consider some important choice that confronts you...Consider the two courses of action that confront you. I'll call them simply A and B. Do you really not believe that you are able to do A and able to do B?...It seems clear to me that when I am trying to decide which of two things to do, I commit myself, by the very act of attempting to decide between the two of them, to the thesis that I am able to do each of them." "The Mystery of Metaphysical Freedom", in *Metaphysics: The Big Questions*, ed. by van Inwagen and D. Zimmerman, Oxford: Blackwell Publishing, 1998.

<sup>14</sup> Here I disagree with some of the stronger claims about the evidential status of commonsense made by recent defenders of commonsense against the skeptic. For good discussion, see Thomas Kelly, "Moorean Facts and Belief Revision, or Can the Skeptic Win?", *Philosophical Perspectives* 19, *Epistemology*, ed. by John Hawthorne, forthcoming and William Lycan, "Moore against the New Skeptics", *Philosophical Studies* 103 (1), 2001, 35-53.

<sup>15</sup> Nor am I making *any* claim about knowledge. The free will/determinism problem isn't a problem about whether we *know* that we have free will or whether we *know* that free will is compatible with determinism; it's a problem about whether we *have* free will and whether our having free will is compatible with determinism. (It's a problem within metaphysics, not epistemology.) If the epistemic skeptic is right, then we don't know that we have hands, let alone that we have free will. My intent in calling **Choice** a commonsense platitude is only to point out that **Choice** is as firmly embedded within commonsense as our belief that we have hands, that the world didn't begin to exist five minutes ago, that there are tables and chairs and other minds. It has whatever epistemic status these other beliefs have.

<sup>16</sup> As G.E. Moore suggested in his *Ethics*, OUP, 1911, Chap. 6.

<sup>17</sup> David Lewis, "Finkish Dispositions", *The Philosophical Quarterly* 47 (1997), 143-158.

<sup>18</sup> As my remarks in the final section of the paper make clear, I think it's an oversimplification to say that the abilities that we have in choice situations are dispositions. Insofar as we have the *wide ability* to do otherwise by having the *narrow*

*ability* to do so, our abilities might be constituted or realized by our dispositions. But having a (narrow) disposition, or even a bundle of dispositions, is at best necessary but not sufficient for having a choice about what overt action to perform.

<sup>19</sup> See my "Freedom, Foreknowledge, and the Principle of Alternate Possibilities", *Canadian Journal of Philosophy* 30 (2000), 1-24, and "Foreknowledge, Frankfurt, and Ability to Do Otherwise: A Reply to Fischer", *Canadian Journal of Philosophy* 38 (2008), 343-372. Bodyguards are the interveners I called 'conditional interveners'; Pre-Emptors are the interveners I called 'counterfactual interveners'.

<sup>20</sup> "How to Think about the Problem of Free Will", *ibid.*

<sup>21</sup> "Are We Free to Break the Laws?", *Theoria* 47 (1981), 113-121

<sup>22</sup> "How to Think about the Problem of Free Will", *ibid.*

<sup>23</sup> In "Free Will Remains a Mystery" (*Action and Freedom*, ed. James Tomberlin, *Philosophical Perspectives* 14, 2000, 1-19.), van Inwagen concedes that the rule *Beta*, which he used in the modal version of the Consequence Argument, is invalid.

<sup>24</sup> "Freedom to Break the Laws", *Midwest Studies in Philosophy*, 28 (2004), Blackwell, 334-350.

<sup>25</sup> Unless we are familiar with the logic of counterfactuals. If we aren't, we may be seriously confused. For instance: "To a determinist, there are no counterfactuals. Nothing that didn't happen could possibly have happened. Everything that did happen was the only possible thing that could have happened at that point in time and space, given the causes." Ray Baumeister, "Just What Exactly is Determinism?", *Psychology Today*, Feb. 15, 2009.

<sup>26</sup> In saying this, I assume that counterfactuals have truth-conditions (as opposed to assertibility conditions and that determinism doesn't have the consequence that there are no true counterfactuals. While there are philosophers who reject both assumptions, both assumptions are also widely accepted. For a useful guide to the now extensive literature on counterfactuals, see Jonathan Bennett, *A Philosophical Guide to Conditionals*, OUP, 2003.

<sup>27</sup> I count the theory of counterfactuals endorsed by Bennett in his book, *ibid*, as a Lewis-like theory.

<sup>28</sup> "Free Will Demystified: A Dispositional Account" *Philosophical Topics* 32 (2004) , *Agency*, John Fischer, ed., 427-450.