

## Free Will Demystified: A Dispositional Account

We believe that free will is or includes the ability to make choices on the basis of reasons, an ability that can be exercised in more than one way. You chose ice cream, but you could have had an apple instead. You chose to speak, but you could have remained silent. You chose selfishly but you could have chosen to do the right thing, for the right reason.

From this commonsense belief about free will it is often only a short step to the incompatibilist view that free will requires the falsity of determinism. The short step consists in what looks like a natural way of spelling out what we mean, when we say things like: “You didn’t have to eat the ice cream; you could have chosen the apple.” We mean “You could have chosen the apple, *given the way things were right up to the point at which you chose the ice cream.*”

This short step is so natural that it’s often thought to be part of commonsense. But here is another short step (and this one will take us all the way): Part of what we mean, when we say “You could have chosen the apple” is: “Your choosing the apple is possible, given the laws of nature.” If we add this assumption, then what we mean, when we say: “You could have chosen the apple” is: “You could have chosen the apple, *given the laws and the past until just before you chose the ice cream.*” But if we grant this assumption, we have granted the key premise in the following argument for incompatibilism:

1. We have free will only if we can choose otherwise.
2. We can choose otherwise only if we can choose otherwise, *given the laws and the past until just before our choice.*
3. If determinism is true, we can *never* choose otherwise, given the laws and the past until just before our choice.
4. Therefore, if determinism is true, we can never choose otherwise.
5. Therefore, if determinism is true, we have no free will.

Since the key premise (2) is based on a claim about the meaning of ‘can’, I will call this the “Meaning Argument” for incompatibilism. So far as I know, no incompatibilist has *officially and explicitly* argued for incompatibilism in this way. However, one often finds something like this argument lurking just below the surface of a philosopher’s explicit argument for incompatibilism. I suspect that something like this argument is the real reason many people, including many philosophers, are incompatibilists.

Since the argument is deductively valid and premise 3 is entailed by the definition of determinism, there are only two ways to resist the conclusion of the argument – by rejecting premise 1 or by rejecting premise 2. I accept premise 1; that is, I agree that free will entails the ability to choose, and I agree that someone with the ability to choose must be able to choose otherwise. But I reject premise 2 and therefore reject the conclusion.

But how can I reject premise 2? Isn’t it just *obvious* that what we mean, when we say “You could have chosen the apple instead” is, “You could have chosen it given the laws and given all the facts about the past until just before your choice”?

I agree that it is easy to talk ourselves into accepting premise 2.<sup>1</sup> But if we accept premise 2, then we must accept incompatibilism. And incompatibilism is not so easy to accept. So we even if we are tempted to accept premise 2, we should not be too quick to accept it.

‘Can’, like ‘must’, can be used in many different ways. Sometimes when we say ‘I can do X’, we mean that our doing X is morally permissible, sometimes we mean that it’s technologically possible, sometimes we mean that it’s physically possible. And so on. But that doesn’t mean that ‘can’ is multiply ambiguous or that it changes meaning from context to context. As David Lewis<sup>2</sup> has observed, ‘can’ has always and everywhere one meaning. When we say “I can do X”, we are saying that our doing X is compossible with certain other facts F. What varies – depending on the intentions of the speaker, the context of utterance, or some combination of the two - is what facts F are relevant. Thus, when we say that doing X is morally permissible, we mean that doing X is compossible

with facts about what's morally permitted; when we say that doing X is physically possible we mean that doing X is compossible with facts about the laws, and so on. The 'can' relevant to free will is the 'can' that we have in mind in contexts in which we raise questions about moral responsibility, and, in particular, contexts in which we raise questions about the justification of choices and the evaluation of agents on the basis of their choices. We hold someone morally responsible for what she does only if we believe that she has free will, and we believe that someone has free will only if we believe that she has the ability to make choices on the basis of reasons. This suggests the following proposal: In these contexts, the relevant facts F are facts about the person's ability to choose for reasons. That is, when we say that someone could have chosen otherwise in contexts where questions of moral responsibility are being discussed, we mean that she could have chosen otherwise given the facts about her ability to choose on the basis of reasons. She in fact chose X, but she had "what it takes" to choose Y instead.<sup>3</sup>

Note that this proposal is neutral with respect to the debate between compatibilists and incompatibilists. The incompatibilist can argue that if determinism is true, then no person ever has the *genuine ability* (as opposed to some inferior, merely apparent "ability") to make choices on the basis of reasons. After all, the ability to make choices for reasons is an ability that can be exercised in more than one way, the incompatibilist might argue, and if determinism is true, then what actually happens is the only thing that can happen, thus no one can ever make any "choice" other than the "choice" she actually makes. So the proposal that the relevant facts F are facts about the ability to choose on the basis of reasons does not beg the question in favor of compatibilism. Nor do I offer it as a defense of compatibilism.

However, I do propose the following way of defending compatibilism. We have the ability to choose on the basis of reasons by having a bundle of capacities which differ in complexity but not in kind from the capacities of things like thermostats, cars, and computers. These capacities are either dispositions or bundles of dispositions, differing in complexity but not in kind from dispositions like fragility and solubility. So my view is

that to have free will is to have a bundle of dispositions. Note that I do not say that to have free will is to “just have a disposition”. To say this is to oversimplify free will in a crucial way.

Let’s give a name to this proposal and define it as follows:

**FWBD:** To have free will is to have the ability to make choices on the basis of reasons and to have this ability is to have a bundle of dispositions.

If **FWBD** is true, then compatibilism is also true. For no one denies that dispositions are compatible with determinism. Indeed, it is precisely because dispositions are uncontroversially compatible with determinism that it has seemed obvious to some people that **FWBD** *cannot* be true. Despite this, I believe that **FWBD** is true. And I believe that compatibilists should accept **FWBD**.

**FWBD** consists of two claims – the claim that free will is the ability to make choices on the basis of reasons and the claim that we have this ability by having a bundle of dispositions. I take the first claim to be uncontroversial; insofar as there is any agreement about what free will is, it is agreement about this. It is the second claim that is controversial. Part of the controversy stems from controversy about the nature of reasons in general and practical reasoning in particular. In what follows, I hope to sidestep this controversy as much as possible. For disagreements about the nature of reasons and reasoning cut across the compatibilist/incompatibilist divide and there are objections to **FWBD** that have nothing to do with worries specific to reasons and reasoning.

Before we consider the objections, let’s begin by noting some similarities between dispositions (of objects and persons) and the abilities of persons.

### **Dispositions and Abilities**

We believe that objects have dispositions (tendencies, causal powers, capacities). A lump of sugar is soluble. Coffee has the power to dissolve sugar. A rubber band is elastic. A thermostat has the capacity to regulate heat. We have empirical knowledge of the dispositions of objects, and we rely on this knowledge in our dealings with things. Avoid drinking arsenic; pack crystal carefully. In relying on this knowledge, we assume that dispositions are relatively stable characteristics of things that typically continue to exist even when they are not manifested. We also believe that something with the disposition to X *can* X even during times that it is not Xing and even if it never X's. A fragile glass is a glass that is liable to break; that is, it is a glass that *can* break, even if it never does.

We also believe that persons have dispositions (tendencies, causal powers, capacities). Some people are color-blind; others have normal color-vision; a lucky few can discriminate shades and hues the rest of us cannot see. Some people can understand French; others can't. Some people are easy-going, others are more easily provoked, and some are hot-tempered. And so on. We have empirical knowledge of the dispositions of persons, and we rely on this knowledge in our dealings with persons. Don't provoke Joe, especially when he's been drinking. If you need a cool head in a tough spot, you can count on Mary. In relying on this knowledge, we assume that dispositions are relatively stable features of persons that typically continue to exist even when they are not manifested. We also believe that if a person has the disposition to X then he *can* X even when he is not Xing, and even if he never X's. Suppose that Joe quits drinking and learns to control his temper. He may still be disposed to lose his temper; insofar as we believe that this is true, we will believe that he is someone who *can* lose his temper, even if he never does.

Philosophers don't dispute these commonsense claims. The philosophical questions concern the proper analysis of dispositions; not whether we speak truly when we say that objects and persons have dispositions.

We also believe that persons have abilities (or, as some philosophers prefer to say, "powers" or "agent-causal powers"). Here are some examples: the ability to speak

French, to sing in tune, to do mental arithmetic, to take charge of situations and direct others, to deliberate for the purpose of figuring out what to do.<sup>4</sup>

There are some striking similarities between abilities and dispositions. We have empirical knowledge of both. Abilities, like dispositions, don't typically pop into existence only on the occasion of their exercise or manifestation. Nor do they go out of existence simply because a person is not exercising them. A person may lose her ability to speak French if she doesn't speak French for many years, but she does not lose this ability every time she stops speaking French. Finally, abilities, like dispositions, entail the corresponding 'can' claim. Someone with the ability to play piano is someone who can play piano even when she's not playing it; someone with the ability to speak French can speak French even when she's speaking English instead.

I think these similarities between abilities and dispositions are no co-incidence. I think that these similarities exist because **ABD** is true:

**ABD:** To have an ability is to have a disposition or a bundle of dispositions.

If **ABD** is true, this provides support for **FWBD**, since the ability to choose for reasons is just a special case of an ability. If other abilities, including mental abilities like the ability to add numbers in one's head and the ability to deliberate for the purpose of figuring out what to do, are dispositions or bundles of dispositions, then it's hard to see why we should deny that the ability to choose for reasons is a disposition or bundle of dispositions.

Is **ABD** true? Some say it isn't. Here, for instance, is van Inwagen:

“For a man to have the capacity to understand French is for him to be such that if he were placed in certain circumstances, which wouldn't be very hard to delimit, and if he were to hear French spoken, then, willy-nilly, he would understand what was being said. But if a man can speak French, then it certainly does not follow that there are any circumstances

in which he would, willy-nilly, speak French. The concept of a causal power or capacity would seem to be the concept of an invariable disposition to react to certain determinate changes in the environment in certain determinate ways, whereas the concept of an agent's power to act would seem not to be the concept of a power that is dispositional or reactive, but rather the concept of a power to *originate* changes in the environment.”<sup>5</sup>

Van Inwagen appears to be claiming that the concept of an agent's ability (“power”) is different from the concept of a disposition (“capacity”) in the following way: the concept of a disposition is the concept of something that is compatible with determinism whereas the concept of an agent's ability is the concept of something that is incompatible with determinism – the ability to originate changes in the environment. If van Inwagen is right, and if determinism is true (or close enough to being true), then our commonsense beliefs about the abilities of agents are all mistaken.

Suppose that van Inwagen is wrong and **ABD** and **FWBD** are both true. If so, then we have the makings for a very robust compatibilism. In particular, we can respond to the charge that determinism robs us of free will by robbing us of the ability to choose and/or do otherwise by arguing as follows:

1. Dispositions are compatible with determinism.
2. Abilities are dispositions or bundles of dispositions. (**ABD**)
3. Therefore, the existence of abilities is compatible with determinism.
4. Free will is the ability to choose on the basis of reasons and we have this ability by having a bundle of dispositions. (**FWBD**)
5. Therefore free will (the ability to choose on the basis of reasons) is compatible with determinism.
6. Abilities (like other dispositions) typically continue to exist even when they are not being exercised or manifested.
7. Therefore, determinism is compatible with the existence of unexercised abilities, including the ability to choose on the basis of reasons.

8. Abilities are like dispositions with respect to the entailment from the claim that a person has the ability (disposition) to do X to the claim that the person can do X.
9. Therefore, determinism is compatible with the truth of the claim that persons can choose and do other than what they actually choose and do.

And if **ABD** and **FWBD** are true, then the compatibilist has what she needs to answer the argument for incompatibilism that is the most influential one these days – van Inwagen’s Consequence Argument.<sup>6</sup>

The Consequence Argument goes something like this. Suppose, for the sake of argument that determinism is true. If so, there is some proposition P which is the conjunction of the laws of nature and facts about the remote past such that P entails every true proposition about the future, including propositions about my choices and actions. I cannot do anything that would make it false that P (since I lack causal power over the laws and the past). I cannot do anything that would make it false that P entails that I will choose and do what I will in fact do (since this entailment is a necessary truth and I lack the ability to bring it about that any necessarily true proposition is false). Therefore, I can neither choose nor do *anything* other than what I in fact choose and do.

But if our abilities, including the ability to choose on the basis of reasons, are dispositions, the conclusion of the Consequence Argument is false. Suppose that I am asked to raise my hand if I want to vote ‘yay’. I refrain from raising my hand. Assume normal circumstances -- no broken bones, pathologies or nefarious neurosurgeons manipulating my limbs or brain states. Could I have raised my hand? Yes, because my ability to raise my hand is one of my dispositions and my dispositions don’t cease to exist *merely* because I am not exercising them. Could I have *chosen* to raise my hand? Yes. But note that there are two possible ways in which this might be true. I might have kept my hand down without going through any kind of reasoning process, perhaps without thinking about the matter at all. Even so, my ability to choose for reasons is one of my dispositions, and my dispositions don’t cease to exist *merely* because I am not exercising them. On the other hand, I might have thought it over, and decided, for what seemed to

me good reasons, to refrain from raising my hand. In that case, I exercised my ability to choose for reasons. But just as a malleable object which is bent can be bent in more than one way, the disposition to choose for reasons is a disposition that can be exercised in more than one way. So even though I in fact manifested my disposition to choose for reasons by choosing to keep my hand down, I could have manifested the very same disposition by choosing to raise my hand. The truth of determinism is not relevant to any of this. Therefore, even if the premises of the Consequence Argument are true, the conclusion is false.

So here are the makings for a compatibilist program. Why isn't anyone working on it?

There is a historical explanation.

### **Brief History:**

At one time, the debate between compatibilists and incompatibilists about free will and determinism was widely believed to turn on the question of whether 'could have done otherwise' is "categorical" or can be given a "hypothetical" or "conditional" analysis. Compatibilists argued that when we say things like "he could have done otherwise" (in contexts where questions of justification of action and moral responsibility are at stake), we mean something like "if he had wanted (or chosen, decided, willed, intended, or tried) to do otherwise, he would have done otherwise". This idea can be traced as far back as Hobbes and Hume, but it was first clearly proposed and defended by G.E. Moore<sup>7</sup>, who remarked that while in one sense of 'can' no one can do otherwise if determinism is true, in another sense – the sense in which a cat can climb a tree but a dog can't —we can often do otherwise even if determinism is true. Moore then proposed that this other sense of 'can' is equivalent in meaning to a conditional; he suggested that what we mean, when we say, "I could have done otherwise" is just "if I had chosen to do otherwise, I would have done otherwise".

What has not been noted is that Moore made two distinct and distinguishable proposals: The first was the proposal that the ‘can’ relevant to questions of free will and moral responsibility – let’s call this “the agent ability ‘can’” -- is a kind of capacity or disposition, to be understood the same kind of way we understand capacities like the cat’s capacity to climb a tree or a ship’s capacity to steam 20 knots. The second was the proposal that the ‘can’ relevant to free will and moral responsibility is equivalent in meaning to a single and ‘simple’ conditional.<sup>8</sup>

These are different proposals, with independent motivations. The first might be true without the second being true, and vice versa. In the debate that followed, a debate which reached its peak in the middle years of the last century, the emphasis was on the second proposal. That is, the debate was generally taken to be over the question of whether the agent ability ‘can’, is equivalent in meaning to a single ‘simple’ conditional. By the mid-seventies everyone agreed that the compatibilists had lost this debate. It wasn’t just that no satisfactory analysis of ‘could have done otherwise’ had been proposed. The consensus in the literature was that the very project of trying to give a Conditional Analysis of ‘could have done otherwise’ is mistaken in principle.<sup>9</sup>

This consensus about the failure of the Conditional Analysis is significant. Without this agreement, van Inwagen’s Consequence Argument would not have enjoyed the success it did.<sup>10</sup> And without this agreement, compatibilists would not have been so ready to turn to other strategies for defending the claim that free will or at least the kind of free will worth wanting or at least moral responsibility can somehow be reconciled with determinism.<sup>11</sup>

I’m not saying that these philosophical investigations are not significant. But I think that the philosophical despair which motivated them was unfounded. The Strawsonian program can be endorsed by a hard determinist. The idea that free will can be understood in terms of higher order desires or in terms of a “real” or “deep” self” can also be endorsed by a hard determinist. In fact, the contemporary free will literature is such that it is very hard to see what the difference is supposed to be between a hard determinist and a compatibilist.<sup>12</sup>

I will argue that the despair was unfounded because Moore's original proposal – that agent abilities are capacities or dispositions – was not properly developed or defended. Given the time period in which these debates took place, this was inevitable. For at that time, everyone assumed that dispositions could be analysed in terms of a single 'simple' conditional about what an object or person would do given the appropriate conditions. We now know that this is not the case. Objects and persons have dispositions and capacities by having intrinsic properties which are the causal basis of the disposition. And some of the arguments that were made against the simple Conditional Analysis of agent ability count equally against a simple Conditional Analysis of dispositions.

### **Dispositions: Why the Simple Conditional Analysis is False**

Consider the following proposal about the correct analysis of dispositions: O has the disposition at time t to X iff if Conditions C obtained at time t, O would X. Thus, O is fragile at time t iff if O were dropped or struck at t, O would break. Call this "the Simple Conditional Analysis" of disposition, for it says that the having of a disposition is equivalent to the truth of a counterfactual conditional.

At one time, it was widely assumed that the Simple Conditional Analysis is the correct analysis of disposition. But the current consensus in the literature is that the Simple Conditional Analysis of disposition has been decisively refuted. According to David Lewis<sup>13</sup>, the refutation has been part of philosophical folklore since 1971 or earlier, but it did not appear in print until 1994.

The refutation is due to C.B. Martin<sup>14</sup> who argued against the Simple Conditional Analysis by arguing that the truth of a counterfactual conditional is neither necessary nor sufficient for the truth of the claim that something has a disposition. His argument was based on two kinds of counterexamples to the Simple Conditional Analysis, counterexamples that have come to be known in the literature as cases of 'finkish

disposition' and 'finkish lack of disposition.' Here are examples of the two kinds of cases:

1. A sorcerer takes a liking to a fragile glass, one that is a perfect intrinsic duplicate of all the other fragile glasses off the same production line. He does nothing at all to change the intrinsic properties of the glass. He only watches and waits, resolved that if ever his glass is dropped or struck, he will quickly cast a spell that changes the glass, renders it no longer fragile, and thereby aborts the process of breaking.<sup>15</sup>

This is a case of a finkish disposition – a disposition which would vanish immediately, on being put to the test.

This case shows that the truth of 'if the glass were dropped or struck at time *t*, it would break' is not *necessary* for the truth of 'the glass is fragile at time *t*'.

2. Another sorcerer says: "I shall make the glass cease to be fragile, but if it is ever struck or dropped, I will make it fragile again." He melts the glass. I throw a stone at it, and, just before the impact, the glass cools and solidifies and the stone breaks the glass. Then the glass melts again. I do the experiment again, and the same thing happens.<sup>16</sup>

This is a case of a finkish lack of disposition – an object which gains a disposition when, and only when, it is put to the test.

This case shows that the truth of 'if the glass were dropped or struck at time *t*, it would break' is not *sufficient* for the truth of 'the glass is fragile at time *t*'.

There is nothing peculiar about fragility that makes it vulnerable to this kind of counterexample. Here is Lewis's argument for the possibility of finkish dispositions:

1. Dispositions are not permanent.
2. Anything can cause anything, so the conditions C in terms of which the disposition is defined might be the very thing that causes the disposition to go away.
3. If the disposition goes away quickly enough, it would not be manifested in Conditions C.
4. If so, then O has the disposition at time t to X but it's false that if C at time t, O would X.
5. Therefore, the truth of 'if C at time t, then O would X' is not *necessary* for the truth of 'O has the disposition at t to X'.<sup>17</sup>

And here is Lewis's argument for the possibility of a finkish lack of disposition:

1. Dispositions are not permanent.
2. Anything can cause anything, so the conditions C in terms of which the disposition is defined might be the very thing that causes the disposition to come into existence.
3. If the disposition comes into existence quickly enough, it would be manifested in conditions C.
4. If so, then it's true that if C at time t, O would X, but it's false that O at time t has the disposition to X.
5. Therefore, the truth of 'if C at time t, O would X' is not *sufficient* for the truth of 'O has the disposition at t to X.'<sup>18</sup>

These kinds of counterexamples refute the Simple Conditional Analysis of disposition. But they don't refute the claim that there are dispositions, nor do they refute a more complex Conditional Analysis of disposition. What these cases seem to show is that objects have dispositions by having intrinsic properties which are the causal basis of the disposition. An object loses or gains a disposition by changing with respect to these intrinsic properties.

Lewis proposes the following Revised Conditional Analysis of dispositions:

O has the disposition at time t to X iff , for some intrinsic property B that O has at t, for some time t' after t, if Condition C were to obtain at time t and O retained property B until t', C and O's having of B would jointly be an O-complete cause of O's X'ing.<sup>19</sup>

For the purposes of this paper, I will assume that Lewis's analysis of dispositions, or something reasonably close to it, is correct.

If it isn't, some of my arguments will have to be reformulated, but the basic line of argument will be the same. For no one takes the failure of the Simple Conditional Analysis to show that there are no dispositions. Everyone agrees that dispositions are compatible with determinism and everyone agrees that dispositions are real properties of objects – as opposed to being what Ryle used to call 'inference tickets'.<sup>20</sup> And while not everyone agrees that dispositions necessarily have an intrinsic causal basis, everyone agrees that the dispositions that interest us – the dispositions of persons and medium-sized objects -- have an intrinsic causal basis.

### **Abilities: Why the Simple Conditional Analysis is False**

Recall Moore's proposal about the correct analysis of ability: S has the ability at time t to do X just in case S has the capacity (disposition) at t to do X and S has the capacity at t to do X just in case it's true that if S chose at t to do X, S would do X. Call this "the Simple Conditional Analysis" of ability for it says that the having of an ability is equivalent to the truth of a counterfactual conditional.

The Simple Conditional Analysis of ability is prima facie plausible for the same reason that the Simple Conditional Analysis of disposition is prima facie plausible. We typically test disposition and ability claims by observing what the object or person does in the specified conditions C. We test for fragility by dropping or striking the object; we test for water-solubility by immersing the object in water; we test for the capacity to understand a language by speaking to the person in that language. We test for the ability

to do something (eg. to speak a language; eg. to move one's limbs) by seeing if the person succeeds in doing that thing if she decides, chooses, intends, or tries to do it.<sup>21</sup>

But despite the relevance of counterfactual conditionals to the *testing* of ability claims, the Simple Conditional Analysis does not provide the correct *analysis* of ability. Here are two counterexamples:

1. A sorcerer has a peculiar interest in J, who has the ability to speak French. He resolves to make sure that J never succeeds in speaking French. He does nothing at all to change any of J's intrinsic properties. He only watches and waits, resolved that if ever J chooses or tries to speak French he will quickly cast a spell that changes J, removing his ability to speak French before J succeeds in uttering a word of French.

This is a case of a finkish ability – an ability which would vanish immediately, on being put to the test.

This case shows that the truth of 'if S chose (decided, intended, or tried) at time t to do X, S would do X' is not *necessary* for the truth of 'S has the ability at t to do X'.

2. Another sorcerer has a peculiar interest in L. He resolves to make sure that L never moves his left hand absent-mindedly or involuntarily or in any way other than by choosing, deciding, intending, or trying to move it. He says: "I shall remove your ability to move your left hand, but if you ever choose, decide, intend, or try to move it; I will restore your ability." He paralyzes L's left hand. L decides to raise the hand; quick as a flash, the sorcerer removes the paralysis. L puts the hand down; the sorcerer paralyzes it again.

This is a case of a finkish lack of ability – a person who gains an ability when, and only when, it is put to the test.

This case shows that the truth of ‘if S chose (decided, intended, or tried) at time t to do X, S would do X’ is not *sufficient* for the truth of ‘S has the ability at t to do X’.

The lesson to be learned from these counterexamples appears to be the same as the lesson we learned from the parallel counterexamples to the Simple Conditional Analysis of disposition. The Simple Conditional Analysis of ability is false because it does not take into account the fact that persons have abilities *by having intrinsic properties that are the causal basis of the ability*. These intrinsic properties are what we gain when we acquire an ability and what we lose when we lose an ability; they are what persist during the times that the person does not exercise her ability. It is because we believe that J continues to have the intrinsic properties that are the causal basis of his ability to speak French that we believe that he continues to have this ability, even though the sorcerer’s dispositions ensure that he *will not speak French*. And it is because we believe that L lacks some of the intrinsic properties that are the causal basis of his ability to move his left hand that we believe that he lacks the ability to move his left hand, even though the sorcerer’s dispositions guarantee that it is always true that L *will move his left hand, if he chooses*.

Let’s try revising the Simple Conditional Analysis of ability in the way that Lewis suggested we revise the Simple Conditional Analysis of disposition:

**Revised Conditional Analysis of Ability:** S has the ability at time t to do X iff, for some intrinsic property or set of properties B<sup>22</sup> that S has at t, for some time t’ after t, if S chose (decided, intended, or tried) at t to do X, and S were to retain B until t’, S’s choosing (deciding, intending, or trying) to do X and S’s having of B would jointly be an S-complete cause of S’s doing X.

The **Revised Conditional Analysis** gets the right results about the two cases that were counterexamples to the Simple Conditional Analysis of ability. J has the ability to speak French because he has some intrinsic property or set of properties B which is the causal basis of his ability to speak French and because it is true that if he both chose to speak

French *and* retained B for the specified time interval (ie. if the sorcerer does not interfere), then J's choosing to speak French, would, together with B, cause him to speak French and would be a J-complete cause of his speaking French. L lacks the ability to move his left hand because it is false that he has any intrinsic property or set of properties B such that if he chose to raise his left hand *and* retained B during the specified time interval, then B together with his choosing to raise his hand would be an L-complete cause of L's raising his hand.<sup>23</sup>

I have claimed that abilities, including the ability to make choices for reasons, are either dispositions or bundles of dispositions. If I am right about this, and if I am right in claiming that Lewis's Revised Conditional Analysis, or something reasonably close to it, is the correct analysis of dispositions, then the Revised Conditional Analysis of ability, or something reasonably close to it, is the correct analysis of those abilities – let's call them 'basic abilities' - which are dispositions.

Basic abilities are dispositions, but not all abilities are basic abilities. Complex abilities, including the ability to make choices for reasons, are not dispositions; they are *bundles of dispositions*.<sup>24</sup> Which dispositions are included in the bundle? This depends on your views about reasons and reasoning, but at least some elements of this package are relatively uncontroversial. A person has the ability to make choices for reasons only if she has the following dispositions (capacities, causal powers): the disposition to form and revise beliefs in response to evidence and argument; the disposition to form intentions (choose, try to act) in response to her desires (understood broadly as "pro-attitudes") and beliefs about how to achieve those desires; the disposition to engage in practical reasoning in response to her intention to make a rational (defensible, justifiable) decision about what to do and her belief that by engaging in practical reasoning she will succeed in making such a decision.

There is, of course, room for argument about what other items need to be added to the list, as well as argument about whether any bundle of dispositions can add up to the ability to choose for reasons. But much of this argument has nothing to do with the

debate between compatibilists and incompatibilists. Compatibilists and incompatibilists agree that there are possible worlds where creatures like us have free will – and thus the ability to choose for reasons -- in the sense relevant to moral responsibility.<sup>25</sup> Their disagreement concerns whether *determinism* robs us of this ability. That is, incompatibilists agree that there are possible worlds where we, or creatures much like us, have the ability to choose for reasons; their claim is that none of these possible worlds are deterministic worlds. It will be helpful to have a way of characterizing the common ground between compatibilists and their incompatibilist opponents. Earlier I suggested that the common ground lies in the idea that free will is the ability to make choices on the basis of reasons. I now propose that the common ground extends further. Incompatibilists and compatibilists should agree that the ability to choose on the basis of reasons is a complex ability constituted by simpler abilities. To put it more bluntly, the ability to choose for reasons is a bundle of simpler abilities. This gives us the following common ground between compatibilists and incompatibilists:

**FWBA:** Free will is the ability to make choices on the basis of reasons and to have this ability is to have a bundle of simpler abilities.<sup>26</sup>

An incompatibilist who accepts **FWBA**, has, it seems to me, only two ways of resisting my claim that abilities, including the ability to choose for reasons, are dispositions or bundles of dispositions:

She could stamp her feet and insist that *no* abilities are dispositions, that there is a sharp difference between the capacity (disposition) to understand a language, on the one hand, and the ability to speak a language, on the other. The ability to speak a language requires the falsity of determinism; the capacity (disposition) to understand a language does not. This seems implausible.

Alternatively, the incompatibilist could agree that *some* abilities are dispositions or bundles of dispositions, but insist that the bundle of abilities which constitute free will

includes at least one ability which is not a disposition.<sup>27</sup> This is perhaps not as implausible but has an ad hoc look to it nevertheless.

### **Arguments against the Conditional Analysis of Ability Reconsidered:**

If what I have said so far is even roughly correct, compatibilists made a big mistake when they gave up on the idea of giving a Conditional Analysis of ability claims. In this section, I will take a look at the objections that were most influential in the literature. I will argue that these objections do not hold against my proposal that compatibilism be defended by **ABD**, **FWBD**, and the Revised Conditional Analysis of disposition and ability. But before I do this, I would like to make a few general observations which help explain why compatibilists gave up such a promising research program.

First, the debate was cast in terms of the question of whether attributions of agent ability are “categorical”, on the one hand, or “hypothetical”, “conditional”, or “constitutionally iffy”, on the other. We are now in a position to see that this was a false dilemma. The possibility of finkish abilities and finkish lack of ability shows us that persons have abilities by having intrinsic properties that are the causal basis of the ability. If the Revised Conditional Analysis is correct, then the correct analysis of ability quantifies over these intrinsic properties *and* uses an ‘if’ clause. So attributions of agent ability are both “categorical” and “conditional”.

Second, the debate was conducted before the development of possible worlds semantics for counterfactual conditionals.<sup>28</sup> Most participants understood that these conditionals are not material conditionals, but the truth-conditions for subjunctive and counterfactual conditionals were not well understood. The prevailing view was that these conditionals are metalinguistic assertions in disguise. When we say something like “if I had sold those shares last week, I would be rich now” or “if I had struck that match, it would have lit” what we really mean must be filled out by something like “...there is a valid argument from the antecedent and some other true statements to the consequent”. But this left all sorts of things unclear, including the difference between the indicative conditional “if

Oswald did not kill Kennedy, then someone else did” and the counterfactual conditional “if Oswald had not killed Kennedy, then someone else would have”. In his seminal “The Problem of Counterfactual Conditionals,”<sup>29</sup> Goodman identified “the problem of cotenability” as an insoluble problem for the metalinguistic account, and this led to skepticism about the prospects of giving an account of objective truth-conditions for *any* counterfactuals.

Finally, defenders of the Conditional Analysis took the relevant analysandum to be “S could have done X” where “X” is an intentional action which S does by *moving her body*, and they assumed that the relevant ability is a single ability, definable in terms of a single conditional. This was due to the behaviorist and positivist climate of the times in which these debates were conducted; the idea was to reduce the (obscure and intractable) problem of free will and determinism to the (less obscure and hopefully tractable) problem of *freedom of action* and determinism. This wasn’t a crazy strategy. After all, the incompatibilist is someone who believes that if determinism is true, we can neither choose nor *do* anything other than what we actually choose or do. If the Conditional Analysis of ‘could have done otherwise’ is correct, then the latter claim is false, and this is a significant victory for the compatibilist. But of course there is a difference between the abilities which constitute free will and the abilities which constitute freedom of action. In ignoring this difference, the defenders of the Conditional Analysis left themselves vulnerable to objections which can easily be avoided if we do not attempt this reductionist strategy.

Let’s turn, finally, to the historical objections.

Some of the objections were curious, yet revealing. For instance, it was objected that the fact that S *does* A implies that S *can* do A, but doesn’t imply any conditional statement and therefore “S can do A” cannot be equivalent in meaning to any conditional statement.<sup>30</sup> This objection is correct, but irrelevant. The truth of any proposition P entails that it is logically and metaphysically possible that P, so if S does A, this entails that it is logically and metaphysically possible that S does A. (That is, S’s doing A is compossible

with the truths of logic and metaphysics.) But this is obviously not the relevant ‘can’.

That S does A also entails that it is physically possible that S does A. But this is also not the ‘can’ relevant to free will and moral responsibility. What’s needed, for this to be an objection to the Conditional Analysis, is that S’s doing A entails that S has the *ability* to do A. But this entailment does not hold. There are many different ways to do something. One way is by having the ability. Another way is by accident or lucky fluke. Yet another way is by having one’s brain and body moved, puppet-like, in the appropriate ways by a sorcerer. Doing something by accident or lucky fluke does not entail having the ability to do it; doing something due to direct manipulation by someone else does not entail having the ability to do it. This shows that the fact that S does A does not entail that S has the ability to do A.<sup>31</sup>

Another objection was that having the ability to do something is consistent with trying (choosing, deciding, etc.) and failing to do that thing, and therefore the truth of the conditional ‘if she had tried (chosen, decided, etc.) to do X, she would have done X’ is not *necessary* for the truth of ‘she has the ability to do X’.<sup>32</sup> The basic point made by this objection is correct. Just as we may get *lucky* and succeed in doing something we *don’t* have the ability to do, we may also get *unlucky* and fail at something we *do* have the ability to do. This objection succeeds against the Simple Conditional Analysis of ability, but has no force against the claim that abilities are dispositions.<sup>33</sup> A fragile glass doesn’t break *every* time it’s dropped, even when it’s dropped in the right kind of way, and even when no finking sorcerer is present. A flammable match may fail to light, even when struck in the presence of oxygen and without a finking sorcerer present. There is room for disagreement about how these sorts of cases should be handled, but for our purposes the point to be noted is just this: No one thinks that cases like this show that there are no dispositions.<sup>34</sup>

Another influential objection was the infinite regress objection. The objection was that while it is true that we often use conditionals like “if S had chosen to do X, S would have done X” as if they are equivalent to “S could have done X”, this is because we are presupposing that S *could have chosen* to do X. But this means that our analysis is

incomplete. To complete our analysis we have to provide a conditional that is equivalent in meaning to ‘S could have chosen to do X’. And any attempt to provide a Conditional Analysis of “S could have chosen to do X” leads to a vicious infinite regress.<sup>35</sup>

The correct answer to this objection is, I think, this: Yes, when we say, of someone, “if she had chosen to do X, she would have done X”, we usually assume that she could have chosen to do X. But that’s not because we think that having the ability to do X *requires* having the ability to choose to do X but, rather, because we think that people typically have the *ability to choose* whether or not to do what they do *in addition to* having the *ability to do* what they do. That is, we assume that persons have the bundle of abilities which constitute the ability to make choices on the basis of reasons as well as abilities to do various things with their bodies (at ‘will’, for no reason other than impulse or ‘mere’ desire). There is no regress because someone (an animal, a young child) may have abilities of the second kind without having any or many abilities of the first kind.

Another category of objections consisted in counterexamples to the claim that a particular Conditional Analysis succeeds in providing a *sufficient* condition for the truth of “S could have done X”. There are many different sorts of cases, and not all cases work against every (Simple) Conditional Analysis, but the basic structure of these cases goes like this:

A person suffers from some pathological condition which either impairs her capacity to make a rational decision concerning some particular kind of action (eg. claustrophobia, eg. a pianist overcome by stage-fright ) or impairs her capacity to make *any* rational decision about what to do But if she (somehow, miraculously) decided (or chose, formed the intention, etc.) to perform the relevant action, she would succeed. Susan Wolf offers the following example of the second sort of case: A person attacked on a dark street is too too paralysed by fear to consider, much less choose, whether to scream. But if she had chosen or tried to scream, she would have screamed.<sup>36</sup>

The intuition we are supposed to have about these cases is that, due to the psychological incapacity, the person *cannot* perform some bodily action X (eg. go into a small enclosed

space, play the piano, scream). According to the Conditional Analysis, however, the person can perform action X. So the Conditional Analysis is false.

I agree that there is a temptation to respond to these cases by agreeing that the person cannot do X. But if we succumb to this temptation, we will be vulnerable to the vicious infinite regress objection. So I think we had better avoid this temptation. I also think that there is a better answer. I think the better answer, to the question, “Can the person do X?” is “Yes and no.” Yes, she has the ability to go into the tiny enclosed space, play the piano, scream. (She knows how to walk; her legs aren’t paralysed. She knows how to play piano; her fingers aren’t broken. She knows how to scream; she doesn’t have laryngitis.) No, she lacks the ability to bring herself to do those things; she cannot use her reasoning ability to bring it about that she intentionally does these things. Due to her phobia (stage fright, panic, etc.), she is unable to choose or try to act according to her own conception of what counts as a good reason for acting.

These kinds of cases are significant, and they do show us something. They show that it was a mistake, on the part of the defenders of the Simple Conditional Analysis, to try to reduce or replace the question of free will with the question of freedom of action. They show that it’s possible for someone to be unfree due to her inability to choose on the basis of reasons. But they don’t show what the objector claims they show – that the ability to do X entails the ability to choose to do X (which in turn entails the ability to choose to choose to do X, and so on to infinity).

It may be instructive to compare these pathological cases with two different techniques doctors can use to ensure that the patient remains immobile after a delicate surgery. The first technique is by using drugs that work in the way that general anesthesia works – by depriving the person of the capacity for conscious thought. The second technique is by using drugs of the first kind together with drugs (“paralytics”) that work by depriving the person of the ability to move her body. Regardless of which technique is used, we can be sure (so long as the drugs work) that the person *will not* raise her hand (or move any other part of her body). But in the first case, if the person became momentarily conscious, and

chose or tried to raise her hand, she would. In the second case, even if the person became momentarily conscious and chose or tried to raise his hand, she would not. In the second case, there is an *additional* safeguard against movement.

Can the person in either of these cases raise her hand? I think that in the first kind of case, we should answer ‘yes’ and ‘no’. Yes, because she’s not paralysed; she still has the *ability to raise her hand*. No, because she *lacks the ability to choose*, on the basis of reasons, to raise her hand; the general anesthesia is the functional equivalent of mental paralysis. (That is, the person retains reasoning and choosing skills, but these skills are temporarily impaired due to the general anesthesia.)

In the second case, however, I think we want to answer a flat “No”. The unconscious paralysed person cannot raise her hand for the same reason a conscious paralysed person cannot. If she chose or tried to raise her hand, she would fail, no matter how hard she tried, no matter how many times she tried. And her failure would be due to a fault in her. Due to her paralysis, she no longer has some of the intrinsic properties which are the causal basis of the ability.

This brings us, finally, to the objection that in the end proved decisive against the Simple Conditional Analysis -- Keith Lehrer’s argument<sup>37</sup> that “S can do X” cannot be equivalent to *any* conditional of the form ‘if Condition C, S will do X’.

His argument went like this: It is logically possible that the following are all true:

- i) if Condition C, S will do x;
- ii) not-Condition C;
- iii) if not-Condition C, S cannot do X.

ii) and iii) entail, by Modus Ponens, that S cannot do X. Therefore, Lehrer concluded, this logically possible case is a counterexample to any Conditional Analysis whatsoever.

Lehrer backed up his “consistent triad” argument with the following example:

“It is logically possible that as a result of my not willing, not choosing, or not undertaking some action, I might lose any of my powers. Suppose that, unknown to myself, a small object has been implanted in my brain, and that when the button is pushed by a demonic being who implanted this object, I become temporarily paralysed and unable to act. My not choosing to perform an act might cause the button to be pushed and thereby render me unable to act.”

If this story sounds familiar, it should. It is a story about a finkish lack of ability; it is, in all basic respects, the story I told earlier about L and the sorcerer.

In claiming that i) to iii) are a consistent set of propositions, Lehrer was, in effect (though he did not realize it) pointing out that a person may suffer from a “finkish lack of ability”. The story he told is the story of someone who currently lacks the ability to move his arm (because he is paralysed) but of whom it’s true that if Condition C obtained (eg. if he chose to move his arm) he would very quickly acquire the ability to move his arm and immediately exercise this ability.

Lehrer’s case refutes the Simple Conditional Analysis of ability, but, as we have seen, it does not refute the claim that abilities are dispositions. Nor does it refute the Revised Conditional Analysis of ability.

### **Frankfurt’s Argument and the Dispositional Analysis of Free Will:**

The proposal that free will is the ability to make choices for reasons and that to have this ability is to have a bundle of dispositions (**FWBD**) can be used to shed light on a famous argument made by Harry Frankfurt in 1969<sup>38</sup>, during the last years of the debate about the Simple Conditional Analysis of ‘could have done otherwise’. Frankfurt argued that this debate was *irrelevant* to the question of whether moral responsibility is compatible with

determinism because it rested on a widely shared but false assumption. He called this assumption ‘the Principle of Alternate Possibilities’ and characterized it as follows:

**PAP:** A person is morally responsible for what he has done only if he could have done otherwise.

Frankfurt’s argument for the falsity of **PAP** was based on a simple thought experiment. You are invited to tell a story about a person, Jones, who chooses to do, and succeeds in doing, some action X. Tell the story in a way that convinces *you* that Jones is morally responsible for doing X. If you are an incompatibilist, you may specify that Jones is an indeterministic agent who can choose and do otherwise, given the actual past and the laws. If you are a compatibilist, you may specify that Jones has all the abilities you think are required for moral responsibility. Now add to your story the following facts: Standing in the wings is another person, Black, with mysterious powers. Black is interested in what Jones does. In particular, Black wants Jones to choose and to do X and Black has it in his power to prevent Jones from choosing or doing anything other than X. But due to a happy co-incidence, Jones chooses and does exactly what Black wants him to do, and Black never intervenes.

According to Frankfurt, this story shows that someone may be morally responsible for what he has done even though he could not have done otherwise. The addition of Black to the story makes it true that Jones could not have done, or even chosen to do, anything other than X. But, Frankfurt argued, the addition of Black does not affect Jones’ responsibility for his action. After all, though Black could have intervened, he didn’t. Jones does X for his own reasons, exactly as he would have done had Black never existed. So this story shows that **PAP** is false.

If **PAP** is false, then, even if determinism robs us of the ability to do (and to choose to do) otherwise, it does not necessarily rob us of moral responsibility. So a great deal is at stake here. Many philosophers have been convinced by Frankfurt; others have not. A huge literature continues to discuss Frankfurt’s thought experiment and argument; the

debate about whether Frankfurt is right about the irrelevance of “alternate possibilities” to moral responsibility has displaced the older literature about the Conditional Analysis of ‘could have done otherwise’.

Frankfurt’s argument is supposed to show that someone may be morally responsible for what he did even if he could not have done (or chosen to do) otherwise. And it is supposed to show this no matter how we understand ‘could have done otherwise’. Let’s test his argument with **FWBD**.

Let’s begin by spelling out Frankfurt’s argument for the conclusion that Jones is morally responsible *even though he could not have done or chosen to do otherwise*.

1. In the first stage of the thought experiment, Jones chooses and does X and has whatever you think it takes to be morally responsible for doing X, including the ability to choose and/or do otherwise.
2. In the second stage of the thought experiment, Jones can neither choose nor do otherwise (due to Black).
3. Stage 1 Jones and Stage 2 Jones are alike with respect to all their intrinsic properties and they perform the same action for the same reasons (since Black never actually intervenes).
4. Two persons who are alike with respect to all their intrinsic properties and who perform the same action for the same reasons are alike with respect to their moral responsibility for their action.

---

Therefore Stage 2 Jones is morally responsible for doing X.

The first and third premises are entailed by the thought experiment. The fourth premise says that moral responsibility for action supervenes on the intrinsic properties of a person, her reasons, and her action; this is intuitively plausible.

The controversial premise is the second one. Can Black really make it true that Jones can neither choose nor do otherwise without ever laying a finger on him? In the literature it

is usually assumed that Black can do this if determinism is true; the controversy lies in whether he can do it if determinism is false.

Since the thought experiment is supposed to work for all of us, we can stipulate that Stage 1 Jones has the ability to choose on the basis of reasons whether or not to do X, and that he has this ability by having a bundle of dispositions (**FWBD**). We learned, from our discussion of finkish cases, that objects and persons have dispositions by having intrinsic properties which are the causal basis of the disposition, and that dispositions are altered or removed by altering or removing the intrinsic properties that are the causal basis of the disposition. Since Black never actually intervenes, he does not tamper with any of Jones' intrinsic properties. It follows that Black does not tamper with any of Jones' dispositions, and thus also follows that Black does not tamper with or remove any of Jones' abilities. Since Stage 1 Jones has the ability to choose, on the basis of reasons, whether or not to do X, Stage 2 Jones has the same ability. And since Stage 1 Jones could have chosen otherwise, this is also true of Stage 2 Jones.

So if **FWBD** is true, Frankfurt's argument against the Principle of Alternate Possibilities fails. Indeed, the dispositional analysis of free will allows us to explain why Frankfurt's argument fails. Frankfurt's argument fails because Black is a fink - a superfink.<sup>39</sup> Black's presence makes it the case that *all* of Jones' abilities, including the abilities which constitute free will, are finkish. Black leaves all of Jones' abilities intact, but Black's power and intentions ensure that if Jones ever begins or tries to exercise any of his abilities in any way contrary to Black's intentions, he will immediately lose that ability. This is not a good thing. We would not want to live at a world where our abilities are vulnerable in this kind of way. But, it turns out, this way of being vulnerable does not rob us of our moral responsibility. This is a significant and interesting discovery; it shows that we are more willing to live with moral luck than some philosophers have thought. But to be *at risk* of losing our abilities is not the same thing as *actually* losing our abilities. And because of this difference, Frankfurt's argument fails.<sup>40</sup>

---

<sup>1</sup> See, for instance, pp. 87-89 of John Searle, *Minds, Brains, and Science* (Harvard University Press, 1984) and pp. 47-51 of Thomas Nagel, *What Does it All Mean? A Very Short Introduction to Philosophy* (Oxford University Press, 1987).

<sup>2</sup> David Lewis, “Scorekeeping in a Language Game”, *Philosophical Papers* Volume I (Oxford University Press, 1983) and “The Paradoxes of Time Travel”, *Philosophical Papers* Volume II (Oxford University Press, 1986).

<sup>3</sup> In ordinary language, “ability” is used in two different ways. Does someone with a broken leg have the ability to ride a bike? That depends. Speaking one way, we might agree that she has the ability. (She took lessons, she knows how; she has the necessary skills and competence.) Speaking another way, we might deny that she has the ability. (Her leg is broken; her bike-riding skills are temporarily *impaired*.) In what follows, I will use ‘ability’ in the second, stronger way and I will sometimes characterize it as “she has what it takes”. (Where the intended contrast is with ‘opportunity’, as in ‘she has the ability to ride a bike, but lacks the opportunity; her *bike* is broken.)

<sup>4</sup> Some of these examples of abilities, together with the parallel examples of capacities, are taken from Peter van Inwagen, *An Essay on Free Will*, pp. 8-11 (Oxford: Clarendon Press, 1983).

<sup>5</sup> *An Essay on Free Will*, p. 11.

<sup>6</sup> *An Essay on Free Will*, Chapter III.

<sup>7</sup> G.E. Moore, *Ethics*, Chapter 6 (Oxford University Press, 1911).

<sup>8</sup> By a ‘single’ conditional, I mean something like “if he had chosen to do X, he would have done X” (in contrast to something like: ‘if he had chosen to do X, he would have done X and if he had had different beliefs or desires, he would have chosen to do X’) as an analysis of ‘he could have done X’. For the

---

contrast between ‘simple’ conditionals and more complex conditionals, see the discussion, below, of finks and the contrast between the Simple and Revised Conditional Analysis of disposition and ability.

<sup>9</sup> Keith Lehrer was, I believe, the first compatibilist to argue this in “Cans Without Ifs” in *Freedom and Determinism* (New York: Random House, 1968), reprinted in Gary Watson, ed., *Free Will* (Oxford University Press, 1982). See also Susan Wolf, *Freedom within Reason* (Oxford University Press, 1990); Bernard Berofsky, “Ifs, Cans, and Free Will: The Issues” in Robert Kane, ed, *The Oxford Handbook of Free Will* (Oxford University Press, 2002); and Michael McKenna, “Compatibilism”, *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu>, among many others.

<sup>10</sup> Van Inwagen was well aware of this: “If someone produced a simple, easily graspable conditional analysis of ability and if this analysis seemed right and if there were no known counter-examples to it, this would vitiate my argument. If someone produced a plausible argument for the correctness of some conditional analysis – however complex that analysis might be – this would vitiate my argument. But none of these things have ever been produced...” (p. 125, *Essay on Free Will*, *ibid.*)

<sup>11</sup> See, for instance: P.F. Strawson, “Freedom and Resentment”, *Proceedings of the British Academy*, vol. 48 (1962), 1-25; Harry Frankfurt, “Alternate Possibilities and Moral Responsibility”, *Journal of Philosophy* 46 (1969) and “Freedom of the Will and the Concept of a Person”, *Journal of Philosophy* 48 (1971); Daniel Dennett, *Elbow Room: The Varieties of Free Will Worth Wanting* (Bradford Books, 1984) R. Jay Wallace, *Responsibility and the Moral Sentiments* (Harvard University Press, 1994).

<sup>12</sup> John Martin Fischer, who sees this point more clearly than most, has coined the term “semicompatibilist” to describe those who believe that determinism is incompatible with “alternative possibilities” and the “freedom to do otherwise” but is compatible with moral responsibility. (*The Metaphysics of Free Will*, Blackwell, 1994.) I suspect that many of the philosophers who call themselves ‘compatibilists’ are in fact ‘semicompatibilists’.

<sup>13</sup> Lewis, “Finkish Dispositions”, *The Philosophical Quarterly* 47 (1997), 143-158.

<sup>14</sup> Martin, “Dispositions and Conditionals”, *The Philosophical Quarterly* 44 (1994), 1-8.

<sup>15</sup> This example is from Lewis, *ibid*, p. 147.

<sup>16</sup> This example is adapted from Martin, *ibid*, p. 2.

<sup>17</sup> Lewis, *ibid*, p. 144.

<sup>18</sup> Lewis, *ibid*, p. 144.

<sup>19</sup> Lewis, *ibid*, p. 157. I have replaced Lewis’s “were to undergo stimulus s” and “to give response r” with “if Condition C were to obtain” and “to X”.

<sup>20</sup> Gilbert Ryle, *The Concept of Mind* (Penguin, 1949), pp. 112-130. For an excellent collection of some of the older literature on dispositions, see Raimo Tuomela, ed., *Dispositions*, (D. Reidel, 1978).

<sup>21</sup> There is, of course, a difference in the way in which we test ability claims, the difference to which van Inwagen alludes when he says that a person’s ability to speak French does not entail that there are any circumstances in which he would “willy-nilly, speak French”. We cannot test for ability simply by placing the person in some *external circumstance* and then watching to see if she exercises the ability. A person may have reasons for refusing to exercise an ability that she in fact has. However, it does not follow that abilities can’t be defined in terms of what a person would do in response to some *internal* condition; for instance, in response to wanting or choosing or trying to do some action. And this seems right, as a general characterization of what distinguishes abilities from other dispositions.

---

<sup>22</sup> Why “set of properties”? Because to have the ability to do X is to have, not just the skills or competence required to do X, but for these skills to be in good working order, not impaired due to broken limbs, paralysis, laryngitis, and so on. (See note 3.)

<sup>23</sup> Note that property set B includes all the properties which are the intrinsic causal basis of the capacity to understand French. You can’t speak a language unless you can understand it, though you can understand a language without being able to speak it. Given this, we can test the ability to speak a language even in finkish cases, by testing to see if the person can understand the language. If she can understand it, this provides us with some evidence that she can also speak it. I believe that similar points can be made about other abilities and their corresponding capacities.

<sup>24</sup> The bundles is not necessarily limited to basic abilities. Depending on how we draw the line between basic abilities and other dispositions, some elements in the bundle may be ‘mere’ dispositions.

<sup>25</sup> I argue this in “Compatibilism, Incompatibilism, and Impossibilism”, in *Contemporary Debates in Metaphysics*, ed. Ted Sider, John Hawthorne, and Dean Zimmerman (Blackwell, 2008). I argue this by distinguishing between the incompatibilist (someone who thinks that *determinism* robs us of free will) and the impossibilist (someone who thinks that free will is impossible for creatures like us, regardless of whether determinism is true or false). I count Galen Strawson as an impossibilist.

<sup>26</sup> Even the agent-causation theorist can agree with this.

<sup>27</sup> Which ability? That depends on the incompatibilist. Traditional agent-causation incompatibilists will presumably insist on the ability to be a causal originator, where this is understood as the ability to be the irreducible substance-cause of events.

<sup>28</sup> Possible worlds semantics for counterfactuals was developed independently by David Lewis and Robert Stalnaker in the late 60’s and early 70’s. David Lewis, *Counterfactuals* (Harvard University Press, 1973); Robert Stalnaker, “A Theory of Conditionals” in N.Rescher, ed., *Studies in Logical Theory* (Blackwell, 1968); Robert Stalnaker, *Inquiry* (Bradford Books, 1984).

<sup>29</sup> *Journal of Philosophy* 44 (1947), 113-128.

<sup>30</sup> This objection is reported by Bernard Berofsky, “Ifs, Cans, and Free Will: The Issues”, in Robert Kane, ed., *The Oxford Handbook of Free Will* (Oxford University Press, 2002), p. 184. Berofsky cites M.R. Ayers, *The Refutation of Determinism* (London: Methuen, 1968) as the source of the objection.

<sup>31</sup> Compare dispositions. The fact that something broke entails that it is logically, metaphysically, and physically possible that it breaks, but it does not entail that it is fragile.

<sup>32</sup> This point was most famously made by J.L. Austin, “Ifs and Cans” in Austin, *Philosophical Papers* (Oxford University Press, 1970), p. 218. Austin’s example was a golfer who misses a very short putt and then kicks himself because “I could have holed it”.

<sup>33</sup> And whether it has force against the Revised Conditional Analysis of dispositions and abilities depends on how that analysis is spelled out. Lewis seems to think that it’s not a problem; others disagree.

<sup>34</sup> In the literature on dispositions, these kinds of cases are known as cases of ‘masking’. For discussion, see Lewis, *ibid*, and also Mark Johnston, “How to Speak of the Colors”, *Philosophical Studies* 68 (1992), 221-263.

<sup>35</sup> R.M. Chisolm, “J.L. Austin’s Philosophical Papers”, in Bernard Berofsky, ed., *Free Will and Determinism* (New York: Harper & Row, 1966). See Berofsky (2002), *ibid*, pp.184-188, for discussion.

---

<sup>36</sup> *Freedom Within Reason*, *ibid*, p. 99.

<sup>37</sup> "Cans Without Ifs", 1968, *ibid*.

<sup>38</sup> "Alternate Possibilities and Moral Responsibility", *Journal of Philosophy* 66 (1969), 829-839.

<sup>39</sup> There are different ways of telling "Frankfurt-style" stories, as they have come to be known in the literature; I do not say that Black is a fink in all of them. However, if **FWBD** is true, then Black does not and cannot succeed in depriving Jones of the ability to choose otherwise while leaving his moral responsibility intact. For discussion and criticism of other kinds of Frankfurt-style stories and arguments, see my "Freedom, Foreknowledge, and the Principle of Alternate Possibilities", *Canadian Journal of Philosophy* 30 (2000), 1-23.

<sup>40</sup> I am grateful to Janet Levin, Terrance Tomkow, Joe Campbell, the Philosophy Department of North Carolina State University, and the audience at the Bellingham Summer Philosophy Conference for helpful discussion and comments.