An object’s disposition to $A$ in circumstances $C$ is *masked* if circumstances $C$ obtain without the object $A$ing. This paper explores an analogous sense in which *abilities* can be masked, and it uses the results of this exploration to motivate an analysis of agents’ abilities in terms of dispositions. This analysis is then shown to provide the resources to defend a version of the Principle of Alternate Possibilities against Frankfurt-style counterexamples. Although this principle is often taken to be congenial to incompatibilism concerning free action and determinism, the paper concludes by using the dispositional analysis of abilities to argue for compatibilism, and to show why the ‘master argument’ for incompatibilism is unsound.

1. Introduction

Ordinary objects are endowed with countless dispositions to behave in various ways in various circumstances. Plates have dispositions to crack in hot ovens, and to support food when it is placed on them; small rubber balls are disposed to bounce when dropped onto hard surfaces; deciduous trees have the disposition to shed their leaves in winter. But an object’s having a disposition to behave in a certain way in certain circumstances in no way guarantees that it will behave in that way in those circumstances. Even deciduous trees, for example, can be forced to retain their leaves year-round if they are given the right light, the right heat, and the right drugs. Dispositions of ordinary objects can be, and often are, masked by conditions which interfere with the manifestation of those dispositions.\(^1\)

It is not only inanimate objects and plants that have dispositions, of course. Other living organisms, including human agents, have many dispositions as well. I am disposed to wash my hands when I am about to eat; I have the disposition to apply my car’s brakes if I want to come to a stop; and I am disposed to answer the phone when it rings. These dispositions, too, may on occasion be masked. My disposition to wash my hands may not become manifest, even if I am about to eat, if the water

\(^1\)The term ‘mask’ is due to Mark Johnston (1992). A. D. Smith (1977) was the first to raise a case of masking in print. The fact that dispositions can be masked constitutes a significant obstacle to their analysis; see, for example, Johnston 1992, Bird 1998, Mumford 1998, and Fara 2005.
has been shut off. And my disposition to answer the phone when it rings
would be masked if someone had glued the phone’s receiver to its base.

Can something similar be said about an agent’s abilities? That is, can
agents’ abilities be masked in the way that their dispositions can be
masked? It is by no means obvious that they can. One salient difference
between abilities and dispositions is that the attribution of an ability to an
agent, unlike the attribution of a disposition, in no way suggests a ten-
dency or propensity on the part of the agent to perform the action in
question. Inveterate thieves differ from the rest of us in part by being dis-
posed, not just able, to steal when the opportunity presents itself. While it
makes sense to talk of a thief’s disposition to steal being masked by the
sudden arrival of police, it would be odd—to say the least—to speak of
the ability to steal that all of us have being masked in the same way.
Speaking like that would seem to suggest, wrongly, that it is only the pres-
ence of police that stops each of us from exercising our ability to steal.

None the less, I think it is both important and useful to see that there
is a sense in which an agent’s abilities can be masked, in much the
way—although not in the same way—that her dispositions can be
masked. In the next section of this paper I will explain what this sense
is. The remainder of the paper will be devoted to defending the claim
that the notion of a masked ability is an important and useful one. I will
begin by using this notion to motivate an analysis of abilities, one which
endorses the spirit of traditional conditional analyses while avoiding the
well-known objections to them. I will then argue that, with this analysis
in hand, we can defend a plausible version of the Principle of Alternate
Possibilities against Frankfurt’s famous attacks; yet I will also show that
accepting that principle does not lead to incompatibilism about respon-
sibility and determinism. Instead, I will argue, the notion of a masked
ability can be used to develop a compatibilist account of free action in
terms of an agent’s ability to do otherwise.

2. Masked abilities

Whenever an agent is disposed to act in a certain way, she has the ability
to act in that way; and whenever her disposition is masked, so too is
her ability. There is something wrong with saying that I am disposed to

---

1 I am not the first to suggest this. See Smith 2003 and Vihvelin 2004.

2 This strategy is similar to one pursued by Kadri Vihvelin (2004), although the details of my
view are markedly different from hers.

3 Here, and throughout, ‘act’ is to be interpreted strongly, to mean ‘perform an intentional ac-
tion’.

drink coffee when I am tired but that I do not have the ability to drink coffee. But this is compatible with there being occasions on which, for one reason or another, I am tired but do not drink coffee; I might, for example, discover that I have no coffee on hand. These are occasions on which my disposition, and my ability, is masked. Sometimes one finds oneself in a position in which one does not succeed in doing what one is able to do, even if one tries.

These remarks may be taken as partially stipulative of the sense of ‘ability’ with which I am concerned. There is no doubt one sense of ‘ability’ on which it is trivial that one lacks the ability to do whatever one is not in a position to do, and so that one lacks the ability to do what one is disposed to do if that disposition is masked. It is natural to say, for example, that if the door is locked then I am unable to open it; or that if I missed the putt then I lacked the ability, on that occasion, to sink it. But it is also natural to say that I am able to open the door, I just need to find the key first; and that I did have the ability, on that occasion, to sink the putt (I am a reasonable golfer), I just failed, for whatever reason, to exercise it. It is this second sense of ‘ability’ that I wish to focus on.\footnote{Compare the discussion of different senses of ‘ability’ in Mele 2003.}

This second sense of ‘ability’ is a modal one. If, after a long tennis match, a commentator says, ‘She could not defeat her opponent’, the commentator seems to mean only that she did not defeat her opponent (and, perhaps, that she tried hard). The commentator is reporting, that is to say, the non-modal fact that the tennis-player’s trying was insufficient for her winning. If, on the other hand, the commentator says, ‘She could not have defeated her opponent’, then the commentator is doing more than reporting the non-modal facts; he seems to be saying that the tennis-player’s trying would have been insufficient for her winning, across a wide range of counterfactual scenarios. Either of these claims, the non-modal and the modal, might have been expressed by saying ‘she was not able to win’. So it is plausible to suppose that there are, correspondingly, two different senses of ‘ability’ to be distinguished.

Perhaps the existence of these two senses of the expressions ‘ability’ and ‘is able to’ shows that these expressions are context-sensitive, expressing different properties or relations on different occasions of use.\footnote{Compare Lewis (1976) on the context-sensitivity of ‘can’.

\footnote{Compare Lewis (1976) on the context-sensitivity of ‘can’.

\footnote{Compare the discussion of different senses of ‘ability’ in Mele 2003.}}

Or perhaps it shows that these expressions are ambiguous, having different meanings on the different occasions. Or perhaps it shows that these expressions univocally express the same property or relation on
every occasion, but that this property or relation has highly circumstance-dependent conditions of instantiation. I take no stand on these issues here. It is enough that there is a sense of ‘ability’, a modal one, on which one may have the ability to do something even when one has tried and failed to do it.

There are different ways of failing to exercise an ability that one has. I have the ability to smash all the windows in my house, but I routinely fail to exercise that ability. I fail, on these occasions, because I do not even try. When an agent fails to exercise an ability that she has because she does not try to exercise it, I will call the failure a straightforward failure to exercise the ability; otherwise I will call it an ability-frustrating failure.

All of us are constantly exhibiting straightforward failures to exercise abilities that we possess. Just now, for example, I straightforwardly failed to exercise my abilities to recite the alphabet, to eat cold mashed potatoes, and to run through the streets naked. The more interesting cases of ability-frustrating failures arise in the context of masked abilities. Whenever an ability is masked, its possessor exhibits an ability-frustrating failure to exercise it. But the converse does not hold: masking cases are not the only cases involving an ability-frustrating failure to exercise an ability. I have the ability to swim in icy waters; but if I tried to exercise this ability now, I would be bound to fail, since all the water nearby is warm. Or, to use an example of Richard Taylor’s (1963), I have the ability to lift a heavy weight, but if I tried to lift a heavy weight now I would be bound to fail: there is no heavy weight around. These are cases where I straightforwardly fail to exercise the ability because I realise there is no point in trying; but if I were to try, I would fail in the ability-frustrating sense, not because the ability was masked but because the opportunity for exercising the ability would not even arise.

As I will be using the word ‘opportunity’, to say that an agent has the opportunity to exercise her ability to do thus-and-so in certain circumstances is just to say that those circumstances obtain.7 Whenever an agent’s ability is masked, the agent exhibits an ability-frustrating failure to exercise the ability, even though she has the opportunity to exercise it.

But still the converse does not hold. I have the ability to count to ten out loud if I am asked to. Suppose I also have the opportunity: I am asked to count to ten out loud. Suppose, moreover, that I try to count

7 My stipulated use of ‘opportunity’ is thus different from that of J. L. Austin (1956), and will in some instances run counter to ordinary uses. For example, I have the ability to hear sounds when my eyes are closed, yet (in my stipulated use) when my eyes are open I lack the opportunity to exercise that ability.
to ten out loud. Still I might fail to exercise my ability because my ability might be removed before I finish counting. This is exactly what happens when one is given general anaesthesia before surgery. In this case, my ability to count to ten is fragile, with a life so short that it could never be exercised. It is essential to an ability’s being masked that its possession survive an attempt to exercise it.

Cases of masked abilities, then, are cases in which an agent fails to exercise an ability that she has and continues to have, despite both having the opportunity to exercise it and trying to exercise it. Consider, for example, my ability to open my office door when I am standing next to it, and suppose the door is locked. We can distinguish at least three relevant cases with respect to this ability. It might be, first, that I have no key for the door. Perhaps, for example, the lock has been changed, and so the key that is in my pocket is no longer the right key for the door. This would be a case in which I simply lack the ability to open the door. Alternatively, it might be that the key in my pocket is the right one, and that if I were to use the key then the door would open. In this case, provided I know that the key is in my pocket, my trying to open the door would have to include reaching into my pocket to retrieve the key. If I failed to open the door then the failure would not be due to a lack of ability; it would be a straightforward failure, a failure to exercise the ability because I did not try to exercise it. Finally, the case might be one in which although the key in my pocket is the right one, still if I were to use the key right now then the lock would jam, since an insect happens to be crawling through the lock mechanism. This is a case in which I have the ability to open the door, and I have the opportunity to exercise that ability, yet if I were to try to exercise it then I would fail because the ability would be masked.

Cases of masked abilities, like cases of masked dispositions, are relatively commonplace. I have the ability to stack ten pennies one on top of the other in a neat pile. Suppose I also have the opportunity—the pennies are in front of me—and that I try to exercise this ability. Still I might fail, since the ability might be masked by a loud crash outside my window, causing me to flinch and spill the coins. Or consider Austin’s seasoned golfer. When he tries but fails to sink the easy putt, his ability has been masked, perhaps by a sudden gust of wind or a momentary distraction. More fancifully, we can imagine that a certain sorcerer has the ability to frighten his enemies when they are near just by raising his eyebrows. But that ability can be masked by a literal mask: if the enemies do not see the eyebrows then the sorcerer will fail to exercise his ability even if he tries.
We can gather what has been said so far into an explicit definition of what it is for an ability to be masked.

(Masking) An agent’s ability to \( A \) in circumstances \( C \) is masked iff

(i) The agent tries to \( A \);
(ii) circumstances \( C \) obtain;
(iii) the agent retains the ability while trying to \( A \); yet
(iv) the agent does not succeed in \( A \).

I have done little more here than to shuffle around some terminology. I began by insisting that we focus on a sense of ‘ability’ according to which an ability can be masked (in much the way that a disposition can be masked), and I introduced some distinctions in the service of explaining what it is for an ability to be masked. For all I have said so far, it may be that the phenomenon of maskable abilities, in the sense I have elucidated here, has little or no philosophical interest. The remainder of this paper is an attempt to show that this is not so.

3. Abilities, dispositions, and conditionals

For an object’s disposition to be masked is for that disposition’s conditions of manifestation to obtain without the manifestation occurring, and without the object’s losing the disposition in question. This is not so for our abilities. As has already been remarked, there are many things that we are able to do that we do not do, not because those abilities are masked but because we are not disposed to do them. This is what helps differentiate us from the inveterate thief in the presence of police. None the less, the above account of masked abilities implies a close relation between an agent’s abilities and her dispositions: necessarily, an agent’s ability to do something (in certain circumstances) is masked if and only if her disposition to do it when she tries to do it (in those circumstances) is in turn masked. This observation motivates an analysis of abilities in terms of dispositions:

The Dispositional Analysis: An agent has the ability to \( A \) in circumstances \( C \) if and only if she has the disposition to \( A \) when, in circumstances \( C \), she tries to \( A \).\(^8\)

---

\(^8\) The ability and disposition ascriptions are to be given their ‘wide scope’ readings in this analysis. It is the ability to \( A \) in circumstances \( C \), and the disposition to \( A \) when, in circumstances \( C \), she tries to \( A \), that are at issue.
So, according to the dispositional analysis, I am able to smash all the windows in my house if and only if I am disposed to succeed in smashing all the windows in my house when I try.

There is something odd about ascribing a disposition to an object when we know, or have good reason to believe, that its conditions of manifestation will never obtain. I doubt that I will ever play the lottery, let alone win it, and so it would be odd to say that I am disposed to celebrate when I win the lottery. Still, odd as it may be to ascribe this disposition to me, I think that it is a disposition I have, and that it is my having of this disposition that, in part, differentiates me from a depressive who is disposed to remain miserable in even the most fortunate of circumstances. Similarly with my disposition to smash the windows in my house when I try. It is odd to ascribe this disposition to me, since we have good reason to believe that its conditions of manifestation will never obtain. It is further odd to ascribe this disposition to me since doing so suggests (but does not imply), falsely, that I have some kind of inclination to smash windows. Yet odd as it is to ascribe it to me, still this is a disposition that I have, and it is my having of this disposition that, in part, differentiates me from a victim of paralysis, or indeed from anyone who lacks the ability to smash windows.

The schematic letter ‘A’ in the dispositional analysis is to be replaced by verb phrases that express actions that are voluntary, in the sense that they are actions that one can in principle try to perform. The analysis is therefore silent about, for example, someone’s ability to become a skilled surgeon (that is not an action) or someone’s ability to try to sing a song (that is not voluntary: one cannot, even in principle, try to try to sing a song). Still, the analysis is not overly restricted. In particular, it is not restricted to only those actions that can be directly performed ‘at will’.

For it is important to recognize that, at least for some actions, there are different ways in which an agent can be said to try to perform an action. Consider, for example, going to sleep. Typically, if one tries to go to sleep by some kind of direct mental attempt, ‘willing oneself to sleep’, one will fail. But there are better ways to try to go to sleep: for me, closing my eyes and doing long division in my head often works. It is this second way of trying—where one tries to perform an action by placing oneself, as best one can, in circumstances appropriate to or conducive to performing the action—that is at work in the dispositional analysis of abilities.

With these points in mind, reflection on individual cases lends considerable plausibility to the dispositional analysis. I am skilled enough

---

9 I thank an anonymous referee for pressing me on this point.
at pool that I am disposed, when I try, to succeed in making bank-shots; unlike beginners, I have the ability to make bank-shots in pool. But I am not sufficiently skilled at chess to beat my computer’s chess program when I try; unlike grandmasters, I am not able to beat my computer.

The dispositional analysis of abilities is also motivated by its similarity to one version of the much-discussed conditional analysis of abilities.

**The Conditional Analysis:** An agent has the ability to $A$ in circumstances $C$ if and only if she would $A$ if she were to try, in circumstances $C$, to $A$.

Like the dispositional analysis, the conditional analysis is supported by reflection on individual cases. My ability to make bank-shots in pool seems to be sustained by the fact that if I were to try to make a bank-shot I would succeed. More generally, it seems, one's abilities reflect counterfactual conditions of success. I can try to do all manner of things, but only those things that I am able to do are such that, if I were to try to do them, I would succeed. The dispositional analysis of abilities respects this intuition by being just the conditional analysis presented in dispositional form.

Notwithstanding its apparent plausibility, however, the conditional analysis of abilities is notorious for being wrong. One reason the analysis is wrong, although this is not a reason that is usually mentioned, is that it rules out the very possibility of masked abilities. If I have the opportunity to exercise an ability that I have, and if I try to exercise it, then the conditional analysis requires that I be successful. But this requirement is too stringent, at least for the kinds of abilities under discussion here. Since one's abilities can be masked, there are occasions on which one tries and fails to exercise an ability one possesses. So the conditional analysis, which requires that whenever one has an ability one would succeed in exercising it if one tried, must be wrong.

This shortcoming for the conditional analysis has no bearing on the dispositional analysis of abilities. Since dispositions can be masked, the dispositional analysis correctly predicts that abilities can be masked as well. Indeed, as I have presented it, the dispositional analysis is motivated by consideration of the fact that abilities can be masked. The failure of the conditional analysis to account for cases of masked abilities is exactly analogous to the failure of conditional analyses of dispositions to account for cases of masked dispositions. If ascriptions of disposi-

---

10 Again, the ability ascription here should be given its 'wide scope' reading.

11 This point is made, with different terminology, by Austin (1956).
tions were understood simply as conditionals, then the dispositional analysis of abilities would be equivalent to the conditional analysis, and so would be incorrect. But ascriptions of dispositions should not be understood simply as conditionals.\textsuperscript{12}

To say that the conditional analysis of abilities fails to account for abilities that can be masked is to point out that there are cases in which the holding of the conditional, ‘if she were to try, she would succeed’, is not necessary for the possession of the ability. The objection to the conditional analysis that is most often discussed, however, is that the holding of this conditional is not \textit{sufficient} for the possession of the ability. (See, for example, Chisholm 1967, Lehrer 1968, and van Inwagen 1983.)

One might worry that the similarities between the conditional analysis and the dispositional analysis make the latter vulnerable to a similar problem. For example, suppose Alice has a neurosis which makes her terrified of spiders. Her neurosis is so strong that there is no way she could stand knowingly being in the same room as a spider, let alone touching one. So Alice could not try to lift a spider. But Alice is as strong and dextrous as the rest of us, so that if, \textit{per impossibile}, she were to try to lift a spider, she would succeed. The holding of this conditional, with its impossible antecedent, is hardly grounds for thinking that Alice has the ability to lift a spider. Yet the conditional analysis requires that Alice does have this ability, and so it is a mistaken analysis. Alternatively, suppose Betty is in a coma and is therefore unable to eat her breakfast. Since the only possible circumstances in which Betty tries to eat her breakfast are ones in which she is not in a coma, if Betty tried to eat her breakfast she would succeed (van Inwagen 1983, p. 119). The conditional is true while the corresponding ability ascription is false.

Is the dispositional analysis equally threatened by cases like these? Is Alice, with her neurosis, \textit{disposed} to lift a spider when she tries? Or is Betty, in a coma, disposed to eat her breakfast when she tries? Surely not. We have already remarked that there is something odd about ascribing a disposition to something when we know, or reasonably believe, that its conditions of manifestation will never obtain. But in the cases of Alice and Betty, it is not just that we reasonably believe that the conditions of manifestation will never obtain, but in addition we are told that they \textit{could} never obtain, in some non-trivial sense of ‘could’. It is not merely odd to ascribe to an object a disposition with impossible manifestation conditions: it is wrong to do so. An object’s dispositions are a matter of

\textsuperscript{12}The literature on dispositions has reached consensus on this point. See, for example, Johnston 1992, Martin 1994, Bird 1998, Mumford 1998, and Fara 2005; although see Gundersen 2003 and Choi 2006 for dissenting opinions.
what it is prone or inclined to do in various actual and counterfactual situations. But objects are not prone or inclined to do anything in situations that could never obtain. Am I disposed to get light-headed when I hold my breath for more than thirty minutes? No, no more than I am disposed to recite poetry in such impossible circumstances.

Note that the kind of impossibility at issue here is not metaphysical impossibility, or even nomological impossibility. It is nomologically possible for Alice to try to lift a spider, since it is nomologically possible for her to lose her neurosis; similarly, it is nomologically possible for Betty to come out of her coma. It is notoriously difficult to pin down the sense of ‘cannot’ in which neurotics cannot try to confront their fears, or in which comatose people cannot try to do anything, or more generally in which people who are in circumstances that prevent them from doing such-and-such cannot do such-and-such, even if those circumstances are only temporarily in place. What matters for present purposes is just that in whatever sense of ‘cannot’ it is correct to say that Alice cannot try to lift a spider, or that Betty cannot try to eat her breakfast, if an agent cannot be in circumstances \( C \) then that agent lacks the disposition to \( A \) in \( C \), for any type of action \( A \). And this claim seems to hold not only for agents but for dispositions of objects quite generally. If a rubber ball is nailed to the wall, and so cannot (in the relevant sense of ‘cannot’) be dropped onto the floor, it is no more disposed to bounce when it is dropped than it is disposed to melt when it is dropped; it simply lacks any dispositions to behave one way or the other when it is placed in conditions that it cannot be placed in.

One could perhaps insist, as an alternative, that dispositions whose conditions of manifestation cannot obtain behave like counterfactual conditionals with impossible antecedents, at least on the standard semantics of counterfactuals: such dispositions are a trivial limiting case, possessed vacuously by every object. But to insist this would be to ignore the fact that the sharing of dispositional properties makes for genuine resemblance among objects. The claim that we are all alike in being disposed to laugh in the presence of unicorns, and in being disposed to cry in the presence of unicorns, defies common sense. Worse, what is possible for one agent may not be possible for another. Jack the Ripper was disposed to murder East London prostitutes in 1888, Queen Victoria was not. Are all of us like Jack the Ripper in this regard, simply because it is impossible for us to do anything in 1888? Surely not.

Ascriptions of dispositions with impossible antecedents are not true, and so compulsives like Alice and comatose people like Betty are not disposed to behave in any way in situations that, because of the circum-
stances they find themselves in, cannot obtain. The traditional problem for conditional analyses of abilities is not faced by the dispositional analysis. My defense of that analysis has not, of course, been exhaustive. But I hope enough has been said to make it seem plausible. It is time now to put the analysis to work.

4. The ability to act otherwise

Incompatibilists about free action and determinism typically hold that if an agent acted freely then she had the ability to act otherwise. According to the dispositional analysis of abilities, this is to say that if an agent acted freely then she was disposed to act otherwise if she tried to act otherwise. Although, for reasons that will emerge, I am a compatibilist, I think that this view about the relation between free action and the ability to act otherwise is correct.

The view has been challenged by those who think that the ‘Principle of Alternate Possibilities’ is false. The Principle of Alternate Possibilities is the claim that an agent is morally responsible for her action only if she could have acted otherwise. If that claim is false, then there are cases in which an agent was morally responsible for her action even though she could not have acted otherwise. On the assumption that one is only morally responsible for one’s free actions, and on the assumption that ‘could have acted otherwise’ ascribes an ability to act otherwise, if the Principle of Alternate Possibilities is false then there are cases of agents who act freely without having the ability to act otherwise.

Those who think that the Principle of Alternate Possibilities is false typically think this because of so-called ‘Frankfurt-style’ cases, cases of a kind first described by Harry Frankfurt (1969). Jones has decided to set fire to a building. He calmly gathers the necessary equipment and, without hesitation, he proceeds to set fire to the building. At least with regard to the non-modal facts, there is nothing in the history of Jones’s action to suggest that this action was not free, and that it was not one for which Jones was morally responsible. Yet it appears that Jones could not have acted otherwise, since his mental activity was being closely watched by an evil scientist, Black. If, contrary to actual fact, Black had detected any wavering on Jones’s part, any tendency to veer from his earlier decision to set fire to the house, Black would have intervened, and would have caused Jones to remain firm in his decision. Jones could not have acted otherwise, it appears, because if he had been about to decide to act otherwise, Black would have ensured that he none the less failed to decide to act otherwise.
The crucial feature of this Frankfurt-style case (and of others like it), with regard to Jones’s alleged inability to act otherwise, is that if Jones were to try to act otherwise he would fail.\textsuperscript{13} All by itself, however, this conditional fact does not establish that Jones lacks the ability to act otherwise. We have already seen reason to reject the conditional analysis of abilities, so it is not in general required for the possession of an ability that one succeed in exercising it if one were to try.\textsuperscript{14} Why think it is required in the Frankfurt-style case? Why think, that is, that Black’s counterfactual intervention does anything more than to \textit{mask} Jones’s ability to act otherwise?

I have the ability to open my office door. Both my hands and the door are in good working order; my cognitive capacities, like my capacity to remember that turning and pushing a handle is the way to open a door, are functioning just fine; my sensory-motor capacities are normal. If anyone ever has the ability to open a door, I have the ability to open my office door. But if I were to try to open the door I would fail: I have forgotten that it is locked. The fact that the door is locked does not \textit{remove} my ability to open the door, but it does \textit{mask} that ability. Or, rather, it \textit{would} mask that ability if I were to try to open the door. If there are ever cases of maskable abilities—and I have insisted that we focus on a sense of ‘ability’ according to which there are—then this is such a case.

My predicament with respect to my ability to open my office door, however, is \textit{exactly the same}, in all relevant respects, as Jones’s predicament with respect to his ability to act otherwise than he in fact acts. As a matter of fact, Jones sets fire to the house; as a matter of fact, I do not open my office door. If Jones were to try to act otherwise he would fail, thanks to the interference of the evil scientist; if I were to try to open my door I would fail, thanks to the interference of the lock. Just as the presence of the lock would mask my ability if I were to try to exercise it, so, surely, the presence of the scientist would mask Jones’s ability if he were to try to exercise it.\textsuperscript{15}

\textsuperscript{13} In some versions of Frankfurt-style cases Black would intervene before Jones was in a position to try to act otherwise, and so in these cases it is not true that if Jones were to try to act otherwise he would fail. I ignore this complication in what follows, what I say below applies equally to these variant cases.

\textsuperscript{14} This point is made, in the context of Frankfurt-style cases, by Michael McKenna (1997).

\textsuperscript{15} I am imagining that in the counterfactual scenario Black’s intervention only causes Jones to remain firm in his decision, leaving everything else as it in fact is, as far as that is possible. In an extreme case Black might intervene by rendering Jones unable to do anything except set fire to the house. In such a case Black would not \textit{mask} Jones’s ability, he would \textit{remove} it, effectively making that ability ‘finkish’ (cf. Martin’s (1994) discussion of ‘finkish’ dispositions). These differences in the manner of Black’s counterfactual intervention do not significantly affect the arguments being given. I thank an anonymous referee for bringing this point to my attention.
I have been suggesting that the evil scientist in our Frankfurt-style case plays the role of a potential masker of Jones’s ability to act otherwise, not a remover of that ability. This suggestion is bolstered by the dispositional analysis of abilities. To say that Jones has the ability to act otherwise, according to that analysis, is to say that he is disposed to act otherwise if he tries. Jones’s possession of that disposition is perfectly compatible with his finding himself in a situation which prohibits the manifestation of that disposition. Compare the case of a fragile glass, disposed to break when struck, which is carefully protected by packing material. The packing material does not remove the glass’s disposition, any more than closing your eyes removes your disposition to see what is in front of you. Instead, the packing material ensures that if the glass happened to be struck, it would not break despite its being disposed to break when struck. Similarly, the evil scientist does not remove Jones’s disposition to act otherwise when he tries. Instead, the scientist ensures that if Jones tried to act otherwise he would fail despite his disposition to succeed. So Jones has the ability to act otherwise, despite being in a situation which prohibits him from exercising that ability. There is, at least in this case, no counterexample to the Principle of Alternate Possibilities.

The attitude towards Frankfurt-style cases that I have been advocating, the attitude of taking them to be cases in which a responsible agent’s ability to do otherwise would be masked if she were to try to do otherwise, is similar in many respects to the attitude recommended by Michael Smith (2003). According to Smith, the lesson to be drawn from the fact that dispositions can be masked is that in assessing certain claims about what an agent could or could not do, in the sense of ‘could’ relevant to considerations of moral responsibility, we are entitled to ignore, or to abstract away from, possibilities involving maskers. Moreover, according to Smith, the role of the evil scientist in Frankfurt-style cases is that of a masker, a masker of Jones’s ‘rational capacities’, capacities to form and maintain particular kinds of beliefs and desires.

---

16 The example is from Johnston 1992. The packing material lines the insides of the glass, providing sufficient support to prevent its breaking if struck.

17 Compare Vihvelin (2004), who construes abilities as ‘bundles of dispositions’ and argues that Frankfurt-style cases show that the bundle of dispositions required for having free will (construed as the ability to make choices on the basis of reasons) can be ‘finkish dispositions’.

18 I do not take myself to have argued that the Principle of Alternate Possibilities, construed as the thesis that an agent is morally responsible for her action only if she had the ability to act otherwise, is unrestrictedly true. I only take myself to have shown that this principle is not refuted by the kinds of Frankfurt-style cases I have discussed.
the very kinds that, on Smith’s view, are relevant to Jones’s responsibility for his action.\textsuperscript{19}

Despite the similarities, there are important points of difference between Smith’s account and my account of Frankfurt-style cases. On Smith’s account, Jones is morally responsible for his action because it is the product of Jones’s rational capacities, capacities to acquire and maintain the right beliefs and desires. And, according to Smith, Jones’s possession of these capacities, and the fact that his action is a product of them, is not impugned by the presence of the evil scientist. The evil scientist serves only to mask these capacities. On my view, however, the evil scientist does not (merely) mask Jones’s rational capacities, his capacities governing believing and desiring. Instead (or in addition) the scientist masks Jones’s ability to act in a way other than the way he in fact acts.\textsuperscript{20} Thus where Smith says, ‘abstracting away from [the evil scientist’s] presence, [Jones] could have done otherwise’ (p. 103)—in the sense that Jones had the rational capacities to form beliefs and desires of a kind that, abstracting away from the evil scientist’s presence, would have led him to do otherwise—I claim simply that Jones could have done otherwise, in the sense that he had the ability to do otherwise.

Despite this important difference between the two accounts, they are similar enough that a recent criticism of Smith’s account, given by Cohen and Handfield (2007), would, if it were cogent, apply equally to my own. Cohen and Handfield concede that Smith’s view provides a reasonable account of Frankfurt-style cases. They would concede, I take it, that in my example Jones could have done otherwise, in the sense of having the ability to do otherwise, and that the evil scientist serves only to mask, not to remove, that ability. But they argue that another case of Frankfurt’s, the case of the willing addict (Frankfurt 1971), is not amenable to Smith’s treatment, and so shows that the Principle of Alternate Possibilities is false.

Frankfurt’s willing addict is a person who, although addicted to a narcotic substance, is perfectly happy with their addiction. Unlike many addicts, who struggle daily against the force of their addiction, the willing addict has no interest in struggling. She wants to take the drug, and she regularly acts on that desire; the fact that in addition to her straightforward desire for the drug she is also compelled by her addiction to take it is of no concern to the willing addict. Cohen and

\textsuperscript{19}I take it that ‘capacity’, as Smith uses the term, is just another word for what I am calling ‘ability’.

\textsuperscript{20}I take no stand on the questions whether Jones has the rational capacities Smith thinks he has and whether the evil scientist, in the Frankfurt-style cases, masks those capacities.
Handfield claim that the willing addict, in virtue of being willing, is responsible for taking the drug when she takes it, even though, because of her addiction, she could not do other than to take the drug, and they therefore conclude that the case constitutes a counterexample to the Principle of Alternate Possibilities.

Cohen and Handfield further argue that the Principle of Alternate Possibilities cannot be salvaged by the claim that the willing addict does have the ability to avoid taking the drug, an ability which is masked by her addiction. For the willing addict’s addiction is a matter of her internal psychology, and so is an intrinsic property of the addict. Following Sungho Choi (2005), Cohen and Handfield claim that intrinsic masks—properties whose possession masks an object’s disposition (or, in this case, ability) that are intrinsic properties of the object—are impossible. So, Cohen and Handfield conclude, the only way to rescue the Principle of Alternate Possibilities in light of this example would be to insist that the willing addict is not responsible for taking the drug. According to Cohen and Handfield, this is implausible.

Cohen and Handfield, and with them Choi, are surely wrong to think that intrinsic masks are impossible. I have the disposition to get a stomach ache when I drink lemon juice; but this disposition can be masked by my taking a certain kind of antacid, a drug that so changes my stomach’s constitution that, for a short while, ingesting lemon juice would not cause me pain. Taking the drug causes me to acquire an intrinsic property—having a certain kind of stomach—the possession of which masks the disposition in question. Taking caffeine functions similarly with respect to my ability to go to sleep when I am tired. Too much of it causes a change in electrochemical activity in my brain: it causes me to acquire intrinsic properties which prevent me from falling asleep, even when I am tired.

These examples of intrinsic masks, however, are all of them examples of very temporary properties. It does seem wrong to think that an object could possess a disposition or ability that would be masked by the simultaneous possession of a long-standing, permanent, intrinsic property of that object. It would be absurd, for example, to say that a brick is disposed to roll downhill, but that this disposition is permanently masked by the brick’s rectangular shape. And it is not just permanent intrinsic masks which seem impossible: any kind of property of an object, intrinsic or otherwise, is better called a remover than a

21 Note that it is not (only) the extrinsic process of taking the drug which masks the disposition. If my stomach were to change intrinsically in the relevant way all on its own, still the disposition would be masked.
masker of a disposition if that property is permanently instantiated.\textsuperscript{22} Even a perfect sphere would not be disposed to roll downhill if it were permanently fastened to the floor by iron rivets.

The addiction of Frankfurt’s willing addict is a long-term, relatively permanent psychological property: one does not become an addict overnight, and addiction is notoriously a condition which, once acquired, is not easily lost. For \textit{that} reason, not because it is an intrinsic property (if indeed it is an intrinsic property), Cohen and Handfield are right to say that the addiction cannot plausibly be regarded as masking the addict’s ability to resist taking the drug. Because she is an addict, she lacks this ability; and so if she is responsible for taking the drug then her situation constitutes a counterexample to the Principle of Alternate Possibilities, even when that principle is interpreted in light of the dispositional analysis of abilities.

Yet why think that the willing addict \textit{is} responsible for taking the drug when she takes it? As far as assignment of moral responsibility is concerned, her situation seems just like that of a culpably lazy bystander in a variation of a case described by Peter van Inwagen (1983). Suppose I look out of my window and see a robbery taking place. I consider phoning the police, but I decide against it because I think it would be too much bother. Unbeknownst to me, I could not have phoned the police even if I had wanted to—my phone was disconnected years ago when (out of laziness) I stopped paying the phone bill.\textsuperscript{23} Van Inwagen argues, correctly in my view, that in this case I am not responsible for failing to phone the police, since phoning the police was not something that I was able to do. I am, of course, responsible for many other things: for deciding not to phone the police, for preferring my own small comforts to the potential relief of someone else’s significant hardship, perhaps for being the kind of person who does not try to help in such circumstances. But for failing to phone the police I am excused, since I was, in fact, powerless to do otherwise.

\textsuperscript{22}Compare Fara’s (2005) distinction between ‘entrenched’ and ‘transient’ finickishness, and Pereboom’s (2001, p. 28) discussion of a brain tumor patient. ‘Permanent’ is being used here in more than a temporal sense. In a merely temporal sense of ‘permanent’, the gravel covering a driveway is a permanent feature of the driveway provided it has been around for a long time. But in another sense of ‘permanent’, the sense at work in the text, gravel is not a permanent feature of driveways since it is easily removed.

\textsuperscript{23}It is here that the case becomes a \textit{variation} of van Inwagen’s. Van Inwagen’s description of his example suggests that the phone lines are disconnected only temporarily. I have changed the example to make it more analogous to the case of the willing addict, and to preclude the suggestion that, in the example, I do have the ability to phone the police, an ability which is masked by the disconnected phone.
Similarly, Frankfurt’s willing addict is responsible for many things related to her taking the drug. She is responsible for being indifferent to her addiction, and for taking the drug *willingly*. She is responsible for failing to struggle with her addiction. Perhaps (depending on the circumstances) she is responsible for having become addicted to the drug in the first place. But she can hardly be said to be responsible for taking the drug since, because she is an addict, she is powerless to do otherwise.  

Perhaps there is a sense of ‘responsible’ according to which a person is responsible for anything they do that is a foreseeably likely consequence of something else for which they are responsible. If this kind of ‘derivative’ responsibility is to count as responsibility *tout court*, then my claim that the willing addict is not responsible for taking the drug should be replaced by the claim that she is not *directly* responsible for taking it, since her taking the drug now is a foreseeably likely consequence of her having taken it all those years ago before she became addicted. This hardly weakens the case for the Principle of Alternate Possibilities, however, since if derivative responsibility counts as responsibility *tout court* then that Principle must itself be restricted, if it is to have any hope of being true, to the claim that an agent is *directly* responsible for her action only if she had the ability to act otherwise. Mele’s (2006) case of the drunk driver nicely illustrates this: a drunk driver who kills a pedestrian might have been so intoxicated that he was not able to do other than kill the pedestrian; yet he is derivatively responsible for killing the pedestrian, since the killing was a foreseeably likely consequence of his drinking so much before driving, an action for which he was responsible. If derivative responsibility is a kind of responsibility then the Principle of Alternate Possibilities should not be understood as applying to it; and so the fact that the willing addict is derivatively responsible for taking the drug does not threaten that principle, as long as she is not *directly* responsible for taking the drug.

It might be thought, none the less, that the view that the willing addict is *not* directly responsible for taking the drug runs counter to our intuitions about the appropriateness of *blaming* the addict for taking the drug (as opposed to merely blaming her for, say, having become addicted to the drug in the first place). Instead of a willing drug

---

24I therefore disagree, not only with Cohen and Handfield, but also with Frankfurt (1971) about this case.

25See Mele 2006 for an elucidation of the distinction between direct and derivative responsibility. I thank Randolph Clarke, Alfred Mele, and Derk Pereboom for (each independently!) calling this distinction to my attention.

26I thank Gideon Rosen for pressing this point.
addict, consider a willing serial killer, a character addicted to her horrific activities in just the same way, and with just the same force, as the drug addict is addicted to taking the drug. She is driven inexorably towards killing, and, when the opportunity arises, she does not struggle against her urge to kill—instead, she delights in it. It is hard to imagine a more repulsive and morally reprehensible being. Can we really say that she is not (directly) responsible for killing, as the Principle of Alternate Possibilities would have us say? After all, we blame the willing serial killer for her killing, and surely such blame is appropriate.

Disgusting as this character is, however, it seems just wrong to blame her for the killing. She is to be blamed for delighting in her condition; for allowing herself to remain in a position for opportunities to kill to present themselves (as opposed to turning herself in); and, presumably, for earlier killings that preceded the addiction’s stranglehold. But to blame her, and so to hold her responsible, for an action that she was powerless to prevent, is surely inappropriate, and motivated only by an unjustified tendency to assign blame whenever a harm is done. At the very least, much more would need to be said for examples like these to be taken to refute the Principle of Alternate Possibilities.

5. Compatibilism

The Principle of Alternate Possibilities is typically taken to be congenial to incompatibilism about free action and determinism. The reason for this is fairly straightforward. Many compatibilists concede that determinism precludes the ability to act otherwise. If, as determinism would have it, the laws of nature and facts about the past together entail that I did so-and-so, then, these compatibilists will agree, I cannot have been able to do other than so-and-so. But—the compatibilists go on to point out—we should not conclude that determinism is incompatible with acting freely, since we should not suppose that having the ability to act otherwise is required for acting freely. It is here that the Principle of Alternate Possibilities enters the dispute. Given that principle, many incompatibilists insist, acting freely must require the ability to act otherwise. For one acts freely only if one is morally responsible for one’s action (at least if the action is one for which the question of moral responsibility arises), and so it just follows, given the Principle of Alternate Possibilities, that one acts freely only if one is able to act otherwise.27

27This is a gross simplification of a complex and nuanced topic, but I take it that, as simplifications go, this one would not be regarded as especially controversial.
If abilities to act otherwise are understood in terms of the dispositional analysis, however, there is no reason to think that determinism is incompatible with free action, *despite* the truth of the Principle of Alternate Possibilities, and *so despite* the fact that one acts freely only if one is able to act otherwise.

Suppose determinism is true, and suppose that, at some time \( t \), I take a sip of coffee. It follows from determinism and the facts that, at \( t \), it was nomologically impossible for me to do other than take a sip of coffee, and that, at \( t \), it was nomologically impossible for me to *try* to do other than take a sip of coffee.\(^{28}\) Does it follow, given the dispositional analysis of abilities, that, at \( t \), I lacked the ability to do other than take a sip of coffee? Does it follow, that is to say, that, at \( t \), I lacked a certain disposition, the disposition to do other than to take a sip of coffee when I try?

Given any of the plausible recent accounts of dispositions, this simply does not follow. Consider, first, David Lewis’s influential view of dispositions (Lewis 1997). On Lewis’s view, I have, at \( t \), the disposition to do other than to take a sip of coffee when I so try if, and only if, I have some intrinsic property such that, if I were to try at \( t \) to do other than take a sip of coffee, and if I were to retain that property, then my trying and my having the property would together be a cause of my succeeding, a cause that is complete as far as my properties are concerned. ‘An unlovely mouthful!’ says Lewis—but a true one, in the circumstances. For I do have the requisite intrinsic property, some property of psychomotor control, such that if I were to try, say, to scratch my nose instead of taking a sip of coffee, my trying and my having the property would together cause me to scratch my nose; and, as far as my properties are concerned, this cause is complete: no further participation by me is required for me to to scratch my nose.

So if Lewis’s view of dispositions is correct then the dispositional analysis of abilities implies that determinism is compatible with my having the ability to act otherwise. Consider next my own view of dispositions (Fara 2005). On my view, I am, at \( t \), disposed to do other than take a sip of coffee when I so try if, and only if, I have at \( t \) some intrinsic property in virtue of which I (generally, usually, normally) do other than take a sip of coffee when I try. Again, some intrinsic property of psychomotor control does the job. It is in virtue of this property that (generally, usually, normally) I scratch my nose when I try, I snap my fingers when I try, I hum a tune when I try, and so on. I have the disposition, and so I have the ability to act otherwise—even on the supposition that determinism is true.

\(^{28}\)I am supposing that my sip was voluntary.
Many incompatibilists about free action and determinism are convinced of their view by some version of what has variously been called the ‘master argument’, the ‘consequence argument’, and the ‘transfer argument’. To take just one version of this argument, suppose again that determinism is true and that, at \( t \), I take a sip of coffee. Since determinism is true, the laws of nature and facts about the remote past together entail that I took a sip of coffee at \( t \). Since I am unable to change the laws, and I am unable to change the remote past, it seems that I am unable to do anything but take a sip of coffee at \( t \).

At the heart of this argument is a principle which licenses its last step. One version of this principle — the only version I will consider here — is van Inwagen’s principle

\[
(\beta) \quad Np, N(p \rightarrow q) \vdash q
\]

where \( N \) abbreviates ‘\( \neg \) and no one is, or ever has been, able to make it the case that not-\( \neg \)’,\(^{29}\) Much has been written about this principle, and its status is highly contested. Given the dispositional analysis of abilities, however, principle \((\beta)\) is quite clearly false.

Suppose, for example, that an experienced golfer is presented with a particularly easy putt. Being experienced, the golfer has the ability to sink the putt. But imagine that the ability is, in this instance, masked by a sudden gust of wind. The golfer tries to sink the putt but fails. Replacing \( p \) in principle \((\beta)\) by ‘there is a sudden gust of wind’, and \( q \) by ‘the golfer misses’, we have a counterexample to the principle. No-one is able to make it the case that there was not a sudden gust of wind,\(^{31}\) and no one is able to make it the case that there was a sudden gust of wind but the golfer did not miss — the golfer being experienced, we may suppose, only in wind-free putting. None the less, the golfer is able to make it the case that she does not miss — she has the ability to sink the putt.

Principle \((\beta)\) is in general false because of the possibility of masked abilities. Masking conditions for an ability are conditions which, if they were to obtain, would prevent the ability from being exercised, a fact which typically no one is able to do anything about. In those cases

\[^{29}\text{See, for example, Ginet 1980, van Inwagen 1983, and Fischer 1994.}\]

\[^{30}\text{See van Inwagen 1983. I have reinterpreted the operator \( N \) here, to make it clear how an analysis of abilities might bear on principle \((\beta)\). Van Inwagen's original interpretation of \( 'N' \) reads: \( 'a \) and no one has, or ever had, any choice about whether \( a \).'}\]

\[^{31}\text{People do, of course, have abilities to prevent sudden gusts of wind in general, since they have abilities to erect windproof fences and the like. Such irrelevant complications could be avoided, at the expense of succinctness, by incorporating into the sentence replacing \( p \) in principle \((\beta)\) a reference to the time at which the golfer took her shot and a description of the golf course (there were no other people around, no fast-acting Erect-a-Fence kits, etc.).}\]
where no one has the ability to prevent the masking conditions from obtaining, therefore, principle ($\beta$) will have a false instance. And if principle ($\beta$) is false, then the master argument for incompatibilism is unsound. The dispositional analysis of abilities has rescued compatibilism.

The incompatibilist might wish to resist all that has been argued in this paper by insisting that the sense of ‘ability’ with which I have been concerned is not the sense that is properly at issue when considering questions about freedom and responsibility. So, the incompatibilist might contend, even if determinism is compatible with the ability to act otherwise in my sense of ‘ability’, it is not compatible with the kind of ability to act otherwise that is required for free action.

It is worth being aware, however, of the commitments incurred by such a view. The argument given above shows that principle ($\beta$) and its variants, the key ‘transfer’ principles in arguments for incompatibilism, will have false instances given any sense of ‘ability’ according to which abilities can be masked (by conditions that no one has the ability to prevent from obtaining). If the incompatibilist is to resist this argument, therefore, she must maintain that the sorts of abilities involved in free action are abilities that cannot be masked. On this conception of ability, whenever one tries but fails to do something (as happens to most of us much of the time), it turns out that, on that occasion, one did not have the ability to do the thing in the first place. (So why try?) And if free action requires the ability to act otherwise in this sense, then one can never act freely if, just before the time of action, one tries to do something else instead. This kind of incompatibilism therefore carries the burden of imposing requirements on free action that are implausibly hard to meet.32

---

32 I thank Randolph Clarke, Adam Elga, John Maier, Michael McKenna, Alfred Mele, Derk Pereboom, Gideon Rosen, Michael Smith, and members of the Corridor Reading Group for extremely helpful comments on previous drafts. I also thank Delia Graff Fara for her unending support and advice over the years in which this paper was written.

Department of Philosophy
Princeton University
212 1879 Hall
Princeton, NJ 08544
USA
fara@princeton.edu

Michael Fara
References


