

TRADITIONAL COMPATIBILISM REFORMULATED AND DEFENDED

Markus E. Schlosser, University College Dublin, markus.schlosser@ucd.ie

Forthcoming in *Journal of Philosophical Research*

This is the author's copy (which may differ from the final print version in minor details)

ABSTRACT: Traditional compatibilism about free will is widely considered to be untenable. In particular, the conditional analysis of the ability to do otherwise appears to be subject to clear counterexamples. I will propose a new version of traditional compatibilism that provides a conditional account of both the ability to do otherwise and the ability to choose to do otherwise, and I will argue that this view withstands the standard objections to traditional compatibilism. For this, I will assume with incompatibilists that the mere possession of a general ability to do otherwise is not sufficient for having the ability that is required for free will. This concession distinguishes the view from the traditional conditional analysis and from recent dispositional accounts of the ability to do otherwise, and we will see that this concession enables a straightforward response to the counterexamples. This, in turn, will play a crucial role in my response to the strongest version of the consequence argument for incompatibilism.

I. INTRODUCTION

According to what we may call *classical* compatibilism, free will is the freedom to do what one wants to do, which requires nothing but the ability to pursue one's desires and the absence from various kinds of constraint and coercion (Hobbes 1958/1651). These two conditions on free will are widely accepted, but they are also widely thought to be insufficient. Even many compatibilists have acknowledged that free will requires, in addition, the ability to do otherwise. We may call, then, *traditional* compatibilism the view that free will requires also the ability to do otherwise, and that having this ability is compatible with the thesis of determinism (Locke 1975/1689; Hume 1955/1748; Moore 1912; Ayer 1954).

Traditional compatibilism is widely considered to be a dead end, as it appears to be subject to devastating objections. In particular, the traditional conditional analysis of the ability to do otherwise appears to be subject to clear counterexamples (Chisholm 1964; Lehrer 1968; van Inwagen 1983: 114–119). Recently, some philosophers have tried to revive the view by way of replacing the conditional analysis with a dispositional account of the ability to do otherwise (Vihvelin 2004, 2013; Fara 2008). However, opponents of the view have pointed out that

dispositional accounts merely assert what traditional incompatibilists have always denied—namely, that having the ability to do otherwise can be reduced to the possession of a mere disposition, capacity, or general ability (Clarke 2009). Traditional compatibilists might reply, here, that this begs the question. But that, I think, would only result in a dialectical deadlock. In what follows, I will develop a new account of traditional compatibilism that proposes to overcome this impasse by agreeing with the opposition. I will agree, that is, that the mere possession of a capacity or general ability to do otherwise is not sufficient for having the ability that is required for free will. We will see that this concession helps to enable a straightforward response to the most serious objection to traditional compatibilism—the mentioned counterexamples—and that this, in turn, facilitates a response to the strongest version of the consequence argument for incompatibilism (Huemer 2000).

Before we proceed, let me briefly address the motivation for defending traditional compatibilism. There are, by now, many alternative non-traditional versions of compatibilism on offer. So, why defend traditional compatibilism? First, most alternative versions of compatibilism are based on the assumption that Frankfurt-style examples (Frankfurt 1969) show that the ability to do otherwise is not required for moral responsibility. This, however, leaves us either with semi-compatibilism about moral responsibility and not compatibilism about free will (Fischer and Ravizza 1998), or it results in revisionist accounts which simply deny that free will requires the ability to do otherwise (Frankfurt 1971; Dennett 1984; Vargas 2007). More importantly, the argument from Frankfurt-style examples has been rather controversial (Ekstrom 2002; Widerker 2003; Goetz 2005; Vihvelin 2013: Ch. 4). Given this, traditional compatibilism remains an attractive position insofar as it promises to provide a basis for the ascription of moral responsibility that does not depend on the success of the argument from Frankfurt-style examples. Second, should it be true that moral responsibility does not require the ability to do otherwise, it would certainly not follow that this ability is insignificant or irrelevant to free will. To the contrary, it is plausible to hold that the ability to do otherwise is required for free will independently of whether or not it is required for moral responsibility, and it is, in my opinion, clearly of interest in its own right to see whether or not having this ability is compatible with determinism. Third, traditional compatibilism is attractive insofar as it meets the main opponents of compatibilism head on. The main opponents are traditional incompatibilists who hold that free will requires the ability to do otherwise. Traditional compatibilism accepts this conception of free

will. That is to say that, unlike non-traditional varieties of compatibilism, it does not attempt to defend compatibilism by way of proposing an alternative or, as opponents would say, compromised and second-best conception of free will.

II. THE TRADITIONAL CONDITIONAL ANALYSIS

According to traditional compatibilism, the ability to do otherwise is compatible with determinism, because having this ability can be analyzed in terms of the truth of certain counterfactual conditionals, and because the truth of counterfactuals is obviously compatible with determinism. This traditional conditional analysis can be stated by way of the following conditional analysis of an agent's ability to act (more precisely, to perform some act-type, *A*):

(TCA) *S* is able to *A* if and only if *S* would *A*, if *S* were to choose (want, intend, or try) to *A*.¹

It has been argued that nothing of substance hangs on whether the antecedent of the counterfactual refers to choosing, wanting, intending, or trying (Berofsky 2012: 84–85, for instance). We will see, however, that formulations in terms of mental actions (such as choosing or trying) and formulations in terms of mental states (such as desire or intention) have both distinct advantages and disadvantages. This is one reason why I will not endorse any one of the particular versions of TCA. I will, rather, propose to split the account into two components: a conditional account of the ability to do otherwise and a conditional account of the ability to choose to do otherwise. First, however, I will turn to some of the standard objections to traditional compatibilism. This will enable me to set the stage, as it were, as my account will partly be based on the lessons that can be gained from a discussion of those objections. I will offer replies to some of those objections in the following section, and I will turn (and return) to the remaining standard objections below (section V).

III. SOME STANDARD OBJECTIONS

A. Freedom of the Will

One traditional objection is that TCA gives at best an account of freedom of action, but not of freedom of the will. In the past, compatibilists often dismissed any talk of 'the will' and they argued that the problem of free will is nothing but the problem of free action, construed as the

freedom of overt behavior. However, most parties in the contemporary free will debate agree that the problem is, at least in part, a problem about the freedom of the will.² This is now usually construed as the freedom of choice, which does not require the postulation of a faculty of ‘the will’. I agree with this construal of the problem, and I acknowledge that a defense of traditional compatibilism must, therefore, include an account of the ability to *choose* to do otherwise.

B. Regress

Some versions of TCA give rise to a regress problem. If the antecedent of the counterfactual refers to a mental action, such as choosing or trying, we face the obvious question of whether or not the agent is able to choose or try to do otherwise. And an application of TCA to choosing or trying would generate an obvious regress (and the apparently absurd notion of choosing to choose or trying to try). One possible solution is to refer, instead, to the possession or acquisition of mental states, such as desires and beliefs, in the antecedent of the conditional (Davidson 1973). In my account, I will not eliminate the reference to mental acts. But I will incorporate this point and avoid the regress by proposing an account of the ability to choose to do otherwise in terms of the agent’s *reasons* (more on this below, in section IV).

C. Austin’s Putt

Consider a skilled golfer who misses an easy putt. Given that she has the required skills, and given that she makes most easy putts, we say that she could have made the putt, even though it is false, evidently, that she would have made the putt, had she chosen (wanted, intended, or tried) to make the putt. This is a counterexample to the claim that the truth of the counterfactual in TCA is necessary for having the ability (Austin 1956).

My response is that we may drop the necessity claim anyway. What is the aim of the conditional analysis? Traditionally, it has been taken to provide an analysis of what it means to be able to do something by proposing a condition that is both necessary and sufficient. There are, however, good reasons to abandon this rather ambitious aim. It is now generally thought that an analysis of virtually any interesting concept that provides necessary and sufficient conditions is bound to fail, as only one fanciful counterexample suffices to defeat it. But no one thinks that we should therefore stop to propose accounts or characterizations of interesting notions, and no one thinks that we should therefore stop to propose necessary *or* sufficient conditions, or even necessary *and* sufficient conditions that are hedged (with *ceteris paribus* clauses, for instance).

Such accounts and conditions can be illuminating and they can make substantial claims that are the target of genuine debate and disagreement.

More importantly, there is a particular and independent reason to drop the necessity claim. According to TCA, counterfactual differences in choosing (wanting, intending, or trying) should co-vary with counterfactual differences in acting (in the closest possible worlds, assuming the standard possible world semantics). Why should this be? What would guarantee this? If determinism is true, then deterministic causal connections between choosing and acting could guarantee this.³ But there is, first of all, no good reason to think that the causal connections between choosing and acting hold without exception, even if determinism is true. By itself, determinism does not protect against accident, interference, weakness of the will, and so on. Moreover, to assume that the relevant connections are deterministic would beg the question, because incompatibilists would insist that the relevant causal mechanisms must be indeterministic. Traditional compatibilists can and should, I think, concede at this point. They *can* concede, because incompatibilists deny, first and foremost, that the truth of the counterfactual in TCA is *sufficient* for having the ability that is required for free will—the necessity claim is not at the core of the disagreement. They *should* concede this, because the assumption that the relevant causal mechanisms are deterministic begs the question. Further, compatibilism says that we *can* have free will *even if* determinism is true. It is not committed to the truth of determinism, and so it is not troubled by the fact that modern physics gives us reason to think that determinism is false. There is, as far as I know, no clear evidence or reason to think that the causal mechanisms that underlie our abilities, at *macro* levels of description, are genuinely indeterministic (Atmanspacher and Rotter 2011, for instance). But, as this possibility cannot be ruled out, compatibilists should not commit themselves to the assumption that the relevant causal mechanisms are deterministic.

So, there is independent reason to make this concession, and making this concession requires that we drop the necessity claim in TCA. It is important to note that this is not too concessive. At the core of the disagreement is the claim that having the ability to do otherwise is compatible with determinism, and holding this is perfectly compatible with dropping the necessity claim in TCA—the controversy concerns the claim that the truth of the counterfactual is sufficient for having the ability to do otherwise (see also Berofsky 2012: 76). We can conclude, then, that Austin's putt is a counterexample to TCA, but not to traditional compatibilism as such.

The possibility of indeterministic causal mechanisms provides also some reason to weaken the counterfactual itself, such that the agent's choosing (wanting, intending, or trying) to do otherwise should merely render doing otherwise *probable*. This would also not be too concessive, because incompatibilists and libertarians would still deny that the truth of a probabilistic counterfactual is sufficient for having the ability required for free will. Despite this, I will formulate my account in terms of the familiar deterministic conditional. One reason for this is that an account in terms of a probabilistic conditional would raise distracting questions concerning *how* probable an agent's doing otherwise should be, if the antecedent holds. More importantly, compatibilists can acknowledge that the relevant causal mechanisms *may* be probabilistic, but they are still committed to the claim that those mechanisms *may just as well* be deterministic. Indeed, this is at the very core of the disagreement: compatibilists hold that the relevant causal mechanisms may be deterministic and incompatibilists deny this. A formulation of compatibilism in terms of a deterministic conditional is a useful reminder of this important point—and it does not beg the question, given that the truth of the conditional is only a sufficient condition for having the relevant ability.

D. General Abilities Are Not Sufficient

Many incompatibilists have argued that having a mere capacity or general ability to do otherwise is insufficient, because having free will requires also that it is *up to the agent to exercise* that capacity or general ability in one way rather than another (see, for instance, van Inwagen 1983: 10–13; Kane 1996: 58–59; O'Connor 2000: 17). This seems particularly clear for accounts that construe the ability to do otherwise as a disposition (capacity or power).⁴ A sugar cube has the disposition to dissolve when put in water, but it is, of course, not up to the cube to exercise this disposition. So, even if a conditional analysis could capture what it is to possess a disposition, it would still be inadequate as an analysis of the ability that is crucial to free will, because it would capture only what it is to *possess* such a disposition. Similarly, having the ability to do otherwise requires that one possesses the relevant *general* abilities, such as the ability to perform certain movements or to speak a certain language. But the mere possession of such general abilities is insufficient, as it must also be up to the agent to exercise them (see, in particular, Clarke 2009: 336–339, 344).

Traditional compatibilists might respond as follows. According to traditional compatibilism, the kind of ability required for free will can be analyzed in terms of counterfactual conditionals, just like any other capacity or ability. But this does not mean that the view collapses the distinction between the possession and the exercise of an ability. Rather, the exercise of the ability to do otherwise consists in the manifestation of that ability, which is clearly distinct from its mere possession. How does this ability differ from the sugar cube's disposition to dissolve? Well, the sugar cube does not act on the basis of choosing (wanting, intending, or trying), but human agents do. This is, in part, why the sugar cube does not have any abilities to *act* at all, and why human agents do have such abilities. So, traditional compatibilism does provide an account of what it is *for an agent* to *exercise* the relevant abilities. It is only that some may find this difficult to see, because it provides a *reductive* account of agency. It reduces, that is, the agent's exercise of agency to connections between agent-involving states and events—just like the event-causal theory of action reduces the agent's role in the exercise of agency to the causal roles of the agent's mental states and events (Davidson 1963; Enç 2003; Mele 2003).⁵

In the following section, I will propose an account that accommodates both sides. I will agree with incompatibilists that a general ability to do otherwise is insufficient. But I will also provide a decidedly reductive account of why and how it is up to the agent to exercise an ability. As indicated, by 'reductive account' I mean here an account that explains the exercise of agency in terms of connections between agent-involving states and events. I do not mean an account that provides a reductive analysis by providing conditions that are necessary and jointly sufficient, which would be incompatible with the point that we should drop the claim that the truth of the counterfactual is a necessary condition (see section III.C).

IV. TRADITIONAL COMPATIBILISM REFORMULATED

As mentioned, I propose to split the view into an account of the ability to do otherwise and an account of the ability to choose to do otherwise. For this, I will assume that the ability to choose is the ability to form intentions. The ability to act otherwise will be explained in terms of the counterfactual dependency between actions and intentions, and the ability to choose to do otherwise will be explained in terms of the counterfactual dependency between intentions and the agent's reasons. It should be clear that those dependencies must not be coincidental. Rather, they must be grounded in the fact that the agent would act differently *because* of having a different

intention and in the fact that the agent would intend differently *because* of having different reasons (more on this below). In order to avoid the objection from Austin's putt and in order to accommodate the possibility of indeterministic causal mechanisms, I drop the claim that the truth of the relevant counterfactual is necessary. But I retain the familiar deterministic counterfactuals in order to stress the point that, according to compatibilism, the causal connections between reasons, intentions, and actions may well be deterministic.

I agree with incompatibilists that having a general ability to do otherwise is insufficient. This allows me to assume that we have the relevant general abilities and that having such abilities is compatible with determinism. I can simply assume this, because this is not what incompatibilists deny. Incompatibilists maintain that the possession of general abilities to do otherwise is not sufficient. In making this point, they usually acknowledge that the possession of the relevant general abilities is necessary and that having them is compatible with determinism. They acknowledge this, because they do not mean to dispute that objects can possess unmanifested dispositional properties and that agents can possess general and unexercised abilities in a deterministic world.⁶

Given that both sides agree on this, there is no need to provide an analysis or account which shows that the possession of general abilities is compatible with determinism (see also Vihvelin 2013: Ch. 6). But, some clarifications are in order. I am able to walk, to raise an arm, to speak a language, to lift a bike, and so on. I am not able to run a marathon in two hours, to fly through the sky by flapping with my arms, to speak Russian, to lift a truck, and so on. It is uncontroversial that we possess (or lack) such general abilities. It is uncontroversial, in particular, that we possess (or lack) such general abilities (at certain times), independently of whether or not they are exercised (at those times), and independently of whether or not our actions are causally determined (at those times). I will express this by saying that some actions clearly are in an agent's *repertoire*, while others are not.

The same holds for mental acts. I am able to remember what I did yesterday, I am able to calculate $35+12$ in my head, but I am not able to remember what I did when I was two years old, I am not able to calculate the square root of 3248 in my head, and so on. Some mental actions are in my repertoire, others are not. The acquisition of an intention may be an action, but it need not be. Consider, for instance, the habitual acquisition of the intention to lock the house when one leaves, to eat something at lunchtime, or to pick up the phone when it rings. Arguably, such

intentions are acquired in a ‘nonactional’ manner (see Mele 2003: 198–201). But usually, when the acquisition of an intention is not a mental action, it is nevertheless in the agent’s repertoire to form that intention actively—doing so is, as I shall say, in the agent’s *act* repertoire.

Something is in an agent’s act repertoire as a certain act-type (or under a certain description). For instance, tossing a coin is in my act repertoire, whereas tossing a coin to land ‘heads up’ is not in my act repertoire. A full account of what is in an agent’s act repertoire would have to address a number of further complications. Many things are in my act repertoire, on a given occasion, only if certain means are available, only if I have the right background knowledge, only if I have the right opportunity, or only if the conditions in that situation are conducive. For instance, I can solve certain problems only with the help of a calculator and only if I know how to use the particular calculator that happens to be available. Or, I can express my views publicly on a certain occasion only if that occasion provides the opportunity to do so, and I can improve my personal best over five miles on a given day only if I am in peak condition, only if the weather conditions are conducive, and so on. We could try to formulate conditions that capture all the points just mentioned. But that is not necessary, because none of the mentioned complications has any bearing on the disagreement between compatibilists and incompatibilists. Both sides can agree that all those further conditions are sometimes satisfied, independently of whether or not determinism is true. We can, therefore, agree to restrict our considerations and arguments to cases in which the mentioned further conditions are, as a matter of fact, satisfied.

Further, an action is in an agent’s act repertoire only if a relevant *basic* action is in the agent’s act repertoire.⁷ It is, for instance, in my act repertoire to give you a signal only if there is *some* basic action in my act repertoire that constitutes or brings it about that I am giving you a signal. Accordingly, we may speak of an agent’s *basic* act repertoire. A necessary condition for having a certain action in one’s basic act repertoire is to have the relevant intrinsic properties. What are those intrinsic properties? Presumably, they are certain neurophysiological properties that enable us human agents to perform the types of basic actions that we are, as a matter of fact, able to perform. According to empirical theories of motor control, this includes the possession of all the ‘motor schemata’ that make up our ‘motor repertoire’, which we acquire and develop over the years through practice and interaction with the world (see, for instance, Clarke 2010). It is important to note, though, that it is not a philosophical task to specify the intrinsic properties that are necessary for having certain actions in one’s basic act repertoire. Referring to such properties

is analogous to referring, in the abstract, to the intrinsic properties that constitute the ‘causal basis’ of a certain dispositional property (Lewis 1997, for instance). No one thinks that a philosophical theory of dispositions that refers to the notion of a causal basis must specify the relevant intrinsic properties. Moreover, we will see that in the defense of my view I will have to make judgments only about what is clearly *not* in an agent’s act repertoire, and everyone can agree that such judgments can be uncontroversial. If it is clear, for instance, that my physiological properties do not enable me to run at a certain speed, then it is clear that it is not in my act repertoire to run that fast, and if my neurological properties do not enable me to remember a certain number of items, then it is clear that I do not have it in my repertoire to remember that many items. Claims of this kind can be entirely uncontroversial, just like it can be entirely uncontroversial that an object without the relevant physical structure is not fragile or soluble in water, for instance.

Given all this, I propose a conditional account of the ability to do otherwise by way of the following conditional account of an agent’s ability to act:

(CAD) *S* is able to *A* if:

- (1) Doing *A* is in *S*’s act repertoire,
- (2) If *S* intended to *A*, *S* would *A* (because of intending to *A*), and
- (3) *S* is able to intend to *A*.

Further, and to a first approximation, I propose a conditional account of the ability to choose to do otherwise by way of the following account of the ability to intend to act:

(CAC) *S* is able to intend to *A* if:

- (4) Intending to *A* in *S*’s repertoire, and
- (5) If *S* had reason to *A*, *S* would intend to *A* (because of having reason to *A*).

Note that condition 4 does not require that intending to *A* is in the agent’s *act* repertoire. It includes, thereby, habitual and nonactional decisions (such as the decision to lock the house when one leaves, to eat something at lunchtime, and so on). Condition 5 requires that intentions co-vary with the agent’s reasons. This is in need of clarification and amendment. A difference in the agent’s reasons may be construed as a difference in the relevant mental states, such as beliefs, desires, and prior intentions—a difference in what we can call *reason-states* (Williams 1979;

Schroeder 2007). Alternatively, a difference in the agent's reasons may be construed as a difference in facts that provide the agent with reasons for action—a difference in what we can call *reason-giving facts* (Dancy 2000). Obviously, I cannot even begin here to argue for one of the two interpretations. But this does not matter, as nothing of substance hangs on this. In what follows, claims about the agent's reasons may therefore be interpreted in accord with either the indicated internalist or externalist conception. Note, however, that the agent's reason-responsiveness has at least two components according to both conceptions: the responsiveness of actions to the agent's intentions and the responsiveness of intentions to either the agent's reason-states or to reason-giving facts, presumably via the agent's recognition of those facts.

A potential problem with condition 5 is that the closest possible world in which both the antecedent and the consequent are true might be the *only* possible world in which both are true. In such a case, the agent's choice does not seem to be a genuine response to the reason, as the agent's responsiveness seems to lack the required generality. In order to address this, I propose to adopt a component of Fischer and Ravizza's (1998) account of reason-responsiveness. I propose, that is, that there must be some *pattern* of close possible worlds in which *S* intends to *A* because *S* has reason to *A*, and in which the agent's behavior exhibits a minimal degree of rational consistency and coherence concerning the action in question (65–68). To illustrate, if there is a close possible world in which *S* has reason not to vote for candidate *K* because *K* has taken bribes from *three* companies, and in which *S* decides not to vote for *K* because *S* knows that *K* has taken bribes from three companies, then *S* is reason-responsive only if *S* also decides not to vote for *K* in the nearby possible worlds in which *S* has reason not to vote for *K* because *K* has taken bribes from *four* companies, and in which *S* makes this decision because *S* knows that, and so on. I propose, then, to replace the counterfactual in 5 with the following condition:

- (5) There is some pattern, *P*, of nearby possible worlds such that:
 - (a) *S* has reason to *A* (in each world in *P*),
 - (b) *S* intends to *A* (in each world in *P*), because *S* has reason to *A* (in each world in *P*), and
 - (c) In the worlds in *P*, the agent's behavior exhibits at least a minimal degree of rational consistency and coherence with respect to the action in question.

The condition of reason-responsiveness, as expressed by CAC, is not unusual or extravagant. In the philosophy of action, it is common to hold that there is a necessary connection between acting *intentionally* and acting *for reasons* (Anscombe 1957; Davidson 1963). Note, further, that CAC does not require that an agent actually responds to reasons (in order to have the ability to intend to *A*). All it says is that an agent is able to intend to *A* *if* it is the case that the agent responds to reasons in some pattern of nearby possible worlds. This means that the conditions in CAC may be satisfied in cases where the agent chooses knowingly against his or her reasons, and even in cases where the agent chooses knowingly to do something that is not supported by reasons at all (under both interpretations of having a reason).

The conjunction of CAC and CAD provides us with an account of the ability to choose and do otherwise—CACD, for short. On this account, the ability to choose to do otherwise is accounted for in terms of the agent's having reason to do otherwise. This provides a straightforward solution to the regress problem, and it accommodates the point that free will is not just freedom of action, but also freedom of choice. CACD is a version of traditional compatibilism, because it provides an account of the ability to do otherwise that is fully compatible with the thesis of determinism. It is not a conditional *analysis*, but it can nevertheless be called a conditional account, as counterfactual conditionals play a central role.⁸ It assumes, with incompatibilism, that having the general (or repertoire) ability to do otherwise is not sufficient for being able to do otherwise in the sense required for free will, which is why condition 1 is by itself not sufficient. The agent must also be able to exercise his or her repertoire abilities, on a certain occasion, in more than one way. This is captured by the conditional in 2, which accounts for the ability to exercise one's repertoire abilities in more than one way. On this account, an agent possesses this ability if 2 holds, in conjunction with 1 and 3, and *exercising* this ability would consist in nothing over and above the following: *actually doing* otherwise because of having the intention to do otherwise (and, perhaps, because of having reason to do otherwise).

Is it also *up to the agent* to exercise this ability in one way rather than another? CACD provides sufficient conditions for having the ability to choose and do otherwise. As pointed out, it provides a reductive account in the sense that it reduces an agent's exercise of agency to connections between agent-involving states and events (in particular, the agent's intentions, reason-states, or recognitions of reason-giving facts). This *is* an account of why and how it is up

to the agent to choose and act otherwise, which *is* an account of why and how it is up to the agent to exercise the relevant abilities in one way rather than another.

Still, one might object that something is missing here, as the agent's choosing and acting is dependent on antecedent states and events and as the agent has, apparently, no power over those antecedents. This is another standard objection to traditional compatibilism, which I will address in the following section, where we will turn (and return) to the remaining objections.

V. TRADITIONAL COMPATIBILISM DEFENDED

A. Power over the Antecedent (and Regress)

According to traditional compatibilism, the agent's doing otherwise is dependent, counterfactually, on the agent's choosing (wanting, intending, or trying) to do otherwise. This raises the question of whether the agent is able to choose (want, intend, or try) to do otherwise. Moreover, it may seem that the agent is not able to do otherwise, because it seems that the agent has no power over the antecedent (or, more precisely, no power over whatever the antecedent of the conditional refers to).

Let us ask, first, what it means to have power over the antecedent. What is at stake here, it seems, can only be a power that we can have by virtue of being able to *choose* and *act*. What else is the power of an *agent* supposed to be? Once we take this into account, we can see immediately that this raises two rather different questions for CACD, because there is a crucial difference between the antecedent of the conditional in CAD and the one in CAC: the former refers to the agent's *intention* and the latter refers to the agent's *reasons*. It makes sense to ask whether or not the agent has power over the antecedent in CAD, because an agent can have voluntary control over his or her intentions (insofar as the formation of an intention can be a mental action). But it makes no sense to ask whether or not the agent has power over the antecedent in CAC, because we are not supposed to have any voluntary control over our reasons. We should, rather, be *responsive to reasons*. So, the agent has power over the antecedent in CAD, given that the agent is able to intend to act (as required by condition 3). But the agent has no power over the antecedent in CAC insofar as the agent has no voluntary control over his or her reasons. And this is just how it should be. The agent should be responsive to reasons and it would make no sense to require that the agent should have power or voluntary control over what those reasons are.

But one might object, further, that the agent must nevertheless be able, in some sense, to have different reasons, and that it is therefore not clear that the regress has been stopped. However, all we need to note here is that this sense of ‘being able’ is clearly not the sense in which an agent is said to be able to choose and act (otherwise). It is, rather, the capacity to have or to recognize different reasons. Metaphysically speaking, this is tantamount to having unactualized mental capacities—in particular, either the capacity to have different reason-states or the capacity to recognize different reason-giving facts (see section IV). As pointed out, incompatibilists and libertarians do not want to deny that we can have unactualized capacities if determinism is true—they deny, rather, that we are able to choose and do otherwise if determinism is true. Given this, an agent’s lack of the capacity to have (or recognize) certain reasons would have to be explained in terms of *particular* facts and circumstances, not in terms of the metaphysical thesis of determinism.

Note, further, that it would clearly be a mistake to assume that if an agent’s control in the exercise of agency is explained in terms of antecedent factors, then the agent must have the *same* kind of control over the antecedents. It would even be a mistake to assume that the agent must have *some* kind of control over the antecedents. Once again, the view provides a reductive account of agency. It explains an agent’s control in the exercise of agency in terms of the roles of the agent’s mental states and events. It is part and parcel of this account that the agent need not and does not have control over the relevant mental states and events themselves. This account may be deficient, of course (see note 5). But one cannot show that it is deficient simply by pointing out that the agent does not have control over those factors which are supposed to explain what the agent’s control in the exercise of agency consists in.

B. Counterexamples

According to one of the most influential objections, traditional compatibilism is subject to clear counterexamples (Chisholm 1964; Lehrer 1968; van Inwagen 1983: 114–119). Consider the following two representative cases (adopted from Fara 2008: 851):

(E1) Alan has a neurosis which makes him terrified of spiders. His neurosis is so strong that there is no way he could stand knowingly being in the same room as a spider, let alone touching one. Alan could not even choose (want, intend, or try) to lift a spider. But Alan is

as strong and dexterous as the rest of us, so that if he were to choose (want, intend, or try) to lift a spider, he would do so.

(E2) Betty is in a coma and is therefore unable to eat her breakfast. Since the only possible circumstances in which Betty chooses (wants, intends, or tries) to eat her breakfast are ones in which she is not in a coma, it seems true that Betty would eat her breakfast, if she were to choose (want, intend, or try) to do so.

In both cases, it seems clear that the agent is unable to perform the action. But, it may well be that the agent would perform the action, if he or she were to choose (want, intend or try) to do so. So, the traditional conditional analysis gets it wrong, and clearly so. The truth of the counterfactual conditional is not sufficient for having the ability to perform the action. Does CACD face the same problem?

Let us begin with condition 3. Are the agents in E1 and E2 able to intend to perform the action in question? In order to answer this, we need to consider the conditions in CAC. Concerning condition 4, it seems clear that the agents do not have it in their repertoire to form the relevant intention. In E1 it is clear, by hypothesis, that Alan altogether lacks the ability to form the required intention, at the time in question. In E2 it is clear that Betty does not have it in her repertoire to form any intentions, at the time in question, because she is in a coma. In both cases, condition 4 is not satisfied. It does not follow that 3 is not satisfied. But it means that there is no reason to suppose that 3 is satisfied in either case. Moreover, some reflection on the relationship between actions and intentions will show that condition 1 is not satisfied.

On virtually all theories of action, a certain movement is an action only if the agent performs that movement with the right intention. Given this, it is in an agent's act repertoire to perform a certain action only if it is in the agent's repertoire to form an intention to do so. So, if it is not in an agent's repertoire to intend to do *A*, then it is, *ipso facto*, not in the agent's act repertoire to *A*. Consider, for instance, the performance of a basic action such as raising one's arm. In order to have it in one's repertoire to raise one's arm it is not sufficient that one has the motoric skills and the physical strength to raise one's arm. One must have it in one's repertoire to produce this movement intentionally, and in order to have this ability, one must have it in one's repertoire to intend to raise one's arm.

Given, then, that the agents in E1 and E2 do not have it in their repertoire to form the required intention, it follows that it is not in their repertoire to perform the action in question. Condition 1 is not satisfied in both cases, and so CACD does not yield the result that the agents are able to perform the actions in question. E1 and E2 are not counterexamples to CACD. Note that this response to the counterexamples does not presuppose that doing otherwise is in the agent's repertoire. It merely exploits the fact that the action in question is clearly *not* in the agent's repertoire for reasons that have nothing to do with the thesis of determinism.

C. Lehrer's Objection

A closely related objection is due to Keith Lehrer (1968). Lehrer put forward an argument against any attempt to analyze the ability to do otherwise in terms of conditionals. This argument is based on two examples. The first is structurally identical to E1 and E2. We have just seen that CACD avoids those counterexamples in a straightforward fashion, and so we can put Lehrer's first example aside. The second example goes as follows (1968: 32):

(E3) 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 became temporarily paralyzed 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.

In E1 and E2, the agent is unable to choose (want, intend, or try) to perform the action. In E3, the agent retains this mental capacity. If the agent chooses not to act, then he is temporarily unable to act. But it remains true that he would act, if he decided to act. This defeats TCA. It shows that the truth of the counterfactual in TCA is not sufficient for having the ability.

The problem here is very similar to what Lewis (1997) called a 'finkish lack' of a disposition. Lewis argued that this problem can be avoided if it is required that the object or agent in question possesses the relevant causal basis: a set of intrinsic properties that would cause the disposition's manifestation in response to the occurrence of the relevant stimulus condition. If we assume that an agent has a certain ability only if the agent has the relevant intrinsic properties, we can apply the solution to the present case. Consider an occasion when the agent chooses not to act. By hypothesis, this may render him unable to act due to a temporary paralysis (caused by pushing the button). But to cause temporary paralysis is to remove the causal basis of the ability

to act. Given this, we could avoid Lehrer's objection by adopting the condition that the agent must possess the relevant causal basis (see Vihvelin 2004, 2013: Ch. 6; Berofsky 2012: 82). But that is not necessary, because CACD already requires that the agent possesses the relevant causal basis. As noted (section IV), the performance of an action is in one's act repertoire only if the relevant basic action is in one's act repertoire. A necessary condition for that is that the agent possesses the relevant neurophysiological properties, which is tantamount to the claim that the agent possesses the relevant causal basis. Hence, in E3, the action is not in the agent's act repertoire, and so my account does not entail that the agent is able to act. E3 is not a counterexample to CACD either.

D. The Consequence Argument and the Overall Dialectic

Another important objection consists in the consequence argument for incompatibilism. The conclusion of this argument is that no one is ever able to do otherwise if determinism is true. The most widely discussed formalized version of this argument employs the following rule of inference, known as rule β (van Inwagen 1983):

$$(\beta) \quad Np, N(p \supset q) \vdash Nq.$$

Van Inwagen defined the operator N in terms of 'having no choice' (1983: 93). Others have construed this operator in terms of a lack of power. On this interpretation, β transfers 'powerlessness' (Fischer 1994; Kane 1996). This rule is crucial, as it helps to establish a connection between past states of affairs, the laws of nature, and our actions. Given that we have no choice or power over the past and no choice or power over the laws of nature, it is not difficult to see how β enables an argument to the conclusion that no one ever has the ability to do otherwise if determinism is true.

Compatibilists have been quick to point out that the conditional analysis can be used for the construction of counterexamples to rule β (and to similar rules or principles).⁹ Incompatibilists, in turn, have rejected counterexamples to β that deploy a conditional analysis. This begs the question, they have argued, as it presupposes that the ability to do otherwise is compatible with determinism. Compatibilists have argued that *this* begs the question, because it does not provide any independent reason to reject the conditional analysis. Both sides, it seems, can defend their view, but neither side can claim a real victory. As far as that goes, the debate has settled at a

‘dialectical stalemate’, which consists, at its core, in a direct conflict between the plausibility of β (and similar rules or principles) and the plausibility of the conditional analysis (Fischer 1994: 83–85; Kane 1996: 50–52; Pereboom 2001: 36–37).

More recently, it has been shown that β is subject to counterexamples that do not presuppose compatibilism (McKay and Johnson 1996).¹⁰ This was a clear victory for traditional compatibilism. But, Huemer (2000) proposed an ingenious interpretation of the N operator, which shows that there is a valid interpretation of β after all. This yields the strongest version of the consequence argument that I know of, and it shifts the burden of proof back to traditional compatibilists. Unfortunately, I cannot offer a counterexample to Huemer’s argument that does not entail the truth of compatibilism. But it is not difficult to see that traditional compatibilists can achieve at least a dialectical stalemate against Huemer’s argument as easily as against van Inwagen’s version of the argument. To see this, we need to consider Huemer’s interpretation of the N operator. Np , he suggests, can be interpreted as: p is the case, no matter what anyone does. More precisely (2000: 538):

$Nsp =_D p$ is the case, and for each action, A , that S can perform, if S were to perform A , it would still be the case that p .

Let P be a true proposition describing the state of the universe at some time in the past, and let L be a true proposition describing the laws of nature. For the sake of argument, let us assume that:

$(P \& L)$ is the case, and for each action, A , that S can perform, if S were to perform A , it would still be the case that $(P \& L)$.

If determinism is true, there will be some true proposition Q that describes some intentional action such that:

$((P \& L) \supset Q)$ is the case, and for each action, A , that S can perform, if S were to perform A , it would still be the case that $((P \& L) \supset Q)$.

Yet, Q may describe an action performed by S and, on any version of traditional compatibilism, S may have been able to do otherwise even if determinism is true—even if, that is, $((P \& L) \supset Q)$ is true. That is, it may well be *false* that:

Q is the case, and for each action, A , that S can perform, if S were to perform A , it would still be the case that Q .

This means that traditional compatibilism allows us to generate as many counterexamples to β as we wish, even under Huemer's interpretation of the N operator.¹¹ To be sure, those counterexamples beg the question insofar as they presuppose compatibilism. But compatibilists can reply in the same vein: to reject an account of the ability to do otherwise solely on the ground that it is compatible with determinism begs the question just as well. More importantly, incompatibilists have always appealed, at some point, to the counterexamples and the other standard objections to TCA in order to defend their position in the debate on the consequence argument. CACD avoids the counterexamples and the other standard objections. This means that incompatibilists have *no independent reason* to reject a reply to the consequence argument that is based on CACD. The consequence argument results, at best, in a genuine impasse against CACD.¹² This is not to say, however, that no dialectical progress has been made. As mentioned at the outset, my account does overcome the impasse concerning the question of whether or not general abilities are sufficient. This impasse has been overcome insofar as I have agreed with the opposition that having the general ability to do otherwise is not sufficient for having the ability that is required for free will.

Incompatibilists and libertarians, I claimed, do not want to commit themselves to the radical view that objects and agents cannot have dispositional properties, unactualized capacities, or unexercised general abilities if determinism is true. This means, in my terminology, that they do not want to deny that intending and doing otherwise can be in an agent's repertoire if determinism is true. One may object that this holds only insofar as *most* incompatibilists and libertarians are concerned. Neither incompatibilism nor libertarianism entails that there are unactualized capacities and unexercised general abilities if determinism is true. So, why should some incompatibilists and libertarians not simply deny that intending and doing otherwise can be in an agent's repertoire if determinism is true?

It is true that incompatibilists and libertarians could deny this. But, doing so would amount to making two rather substantial concessions. First, they would deny this only if they thought that they *had* to deny this in order to reject compatibilism. They would thereby concede that incompatibilism about free will entails a commitment to incompatibilism about unactualized capacities and unexercised general abilities. This is not only a price that incompatibilists and

libertarians do not want to pay, but it is a price that they think they do not have to pay. Second, they would thereby also concede that incompatibilism about free will *follows* from incompatibilism about unactualized capacities and unexercised general abilities. For, if determinism rules out the latter, then it follows, trivially, that it rules out free will as well. Again, incompatibilists and libertarians do not want to be committed to this, and they think that they are not committed to this. They think that their arguments for incompatibilism are, *specifically*, arguments for the incompatibility of determinism and free will. They do not think, that is, that the incompatibility of free will simply follows from the general metaphysical thesis that determinism is incompatible with unactualized capacities and unexercised general abilities. Given this, we can see now that we would make progress in the overall dialectic *even if* opponents denied the premise that intending and doing otherwise can be in an agent's repertoire if determinism is true. We would make progress, because opponents would be forced to make concessions they do not want to make and they thought they did not have to make.

VI. CACD COMPARED AND DISTINGUISHED

Most contemporary compatibilists do not hold that the ability to choose and do otherwise is compatible with determinism, and they do not hold that this ability is required for moral responsibility. There are, though, some notable exceptions to this recent trend. In this section, we turn to three recent compatibilist views that preserve important features of traditional compatibilism, and that are in various respects close to CACD. My aim is not to prove them wrong—they are, after all, on my side as far as the overall dialectic is concerned. My aim, rather, is to highlight the differences and the advantages of CACD.

Following Wolf (1990), Nelkin (2011: 66) offered the following account of an agent's ability to act:

S is able to *A* if (1) *S* possesses the capacities, skills, talents, knowledge and so on which are necessary for doing *A*, and (2) nothing interferes with or prevents the exercise of the relevant capacities, skills, and talents and so on.

Presumably, this account can avoid the consequence argument in the same way as TCA. Does it avoid the counterexamples? Nelkin does not address this, but she does make it clear that having the relevant capacities, skills, and so on, amounts to having the general ability to perform the

action—it amounts, in my terminology, to having the action in one’s act repertoire. One could then argue, as I do in section V.B, that Alan and Betty lack the capacities to act, because they lack the relevant capacities to intend, at the time in question. Likewise, one could argue, as I do in section V.C, that the possession of the relevant capacities requires the possession of the relevant intrinsic properties. Given this, Nelkin’s account can avoid the counterexamples and Lehrer’s objection in the same way as CACD.

But there are important differences. According to Nelkin, all that is required in addition to the general ability to act is the absence of interference and prevention. This is not as clear-cut as it may seem. Is interference and prevention restricted to external conditions? Or does a neurosis, a phobia, or being in a coma count as interference or prevention as well? More importantly, by requiring merely the absence of interference and prevention, the account neglects the important distinction between freedom of action and freedom of the will, and it raises the issue of regress. Suppose an agent is prevented from exercising her capacities in a certain way. Could she have chosen to exercise her capacities in another way? Does this generate a regress of mental acts? Nelkin’s account does not provide the resources to address these issues, because it does not offer an account of the freedom of choice, and because it does not offer an account of what the exercise of the relevant capacities, skills, and so on, consists in.

Vihvelin (2004 and 2013) and Fara (2008) have proposed compatibilist accounts of the ability to do otherwise in terms of dispositions. According to Fara, we can leave the notion of a disposition unanalyzed for this purpose, and he proposed the following “dispositional analysis” (with minor alterations from Fara 2008: 848):

S has the ability to *A* in circumstances *C* if and only if *S* has the disposition to *A* when, in circumstances *C*, *S* tries to *A*.

Fara shows that this account avoids the consequence argument. Does it avoid the counterexamples? Consider again E1 and E2. Is Alan disposed to lift the spider when he tries? Is Betty disposed to have breakfast when she tries? According to Fara, the answer is “surely not”, for it is “not merely odd to ascribe to an object a disposition with impossible manifestation conditions: it is wrong to do so” (2008: 851). That is because agents or “objects are not prone or inclined to do *anything* in situations that could never obtain” (852). Clarke (2009), however, has argued that this response to the counterexamples is unconvincing. Consider a rubber ball that is

nailed to the wall, so that it cannot drop to the floor. According to Fara, the ball is not disposed to bounce when dropped, because it “simply lacks any disposition to behave one way or the other when it is placed in conditions that it cannot be placed in” (852). This example is supposed to support the suggested diagnosis of the counterexamples. But, as Clarke points out, the ball does seem to have the disposition to bounce when dropped, because it is “still rubbery” after all (2009: 335). Not only does the example fail to support Fara’s diagnosis of the counterexamples, but it seems to undermine it: in some sense, the ball cannot drop, but it nevertheless has the disposition to bounce when dropped. I find myself in full agreement with Clarke’s assessment, and so I am not convinced by Fara’s diagnosis of the counterexamples.

Further, the account faces also difficult questions concerning freedom of the will and, in particular, concerning the notion of trying. It is widely agreed that trying can be a mental action. This raises the obvious question of whether the agent is, in such cases, able to try otherwise. This, in turn, gives rise to an obvious regress problem. One could avoid this, if one could show that all actions, including all mental actions, are ultimately based on instances of trying that are themselves not actions. This is not impossible, and it may not even be difficult to show this. But Fara owes us an account of trying that avoids the regress problem. Finally, Fara did not show that the account can avoid Lehrer’s objection.

According to Vihvelin, to have an ability to act is to have a disposition (or a bundle of dispositions). On the basis of Lewis’ (1997) “reformed conditional analysis” of dispositions, she proposed the following conditional analysis of the ability to act (adopted with some modifications and simplifications from Vihvelin 2004: 438 and 2013: Ch. 6):

S has the ability to *A* as the result of trying if and only if *S* has some intrinsic property *B* in virtue of which *S* would *A*, if *S* had the opportunity to *A* and tried to *A*, and if *S* retained *B*.

The most important feature of this account is the reference to the causal basis: some intrinsic property *B* in virtue of which *S* would *A*. This is also the main feature of Lewis’ reformed conditional analysis of dispositions, which enables a straightforward response to Lehrer’s objection (see section V.C). Vihvelin shows that this account avoids the consequence argument, and so the question is, once again, whether it can avoid the counterexamples. Vihvelin (2013: 201–208) argues that examples of type E1 and E2 are not counterexamples at all. She considers

Petra, who has a pathological fear of wide open spaces, so that she is even unable to try to walk outdoors. In such cases, we are asked to suppose that the agent is unable to perform the action, that she is unable to try, and that she would perform the action, if she tried. Why, Vihvelin asks, do we accept all this? We presume, it seems, that in a world in which the agent is able to try, she is not afflicted by the pathology. And we presume that in a world in which the agent is not afflicted by the pathology, she would perform the action, if she tried. But according to Vihvelin, it is a mistake to use such a “backtracking argument” in the assessment of a *causal* counterfactual (2013: 207–208). In her example, the relevant causal question is this: if Petra tried, would her trying cause the action? In order to assess this correctly we must, according to Vihvelin, consider only those worlds “where the present, at the time of her trying, is as similar to the way things actually are as can be, consistent with her trying” (208). At these worlds, Petra is still afflicted by the pathology, but she is “a tiny bit braver”. And at these worlds, Vihvelin claims, Petra fails if she tries.

This is a credible line of response, but it is not as straightforward as the one provided by CACD. First, some incompatibilists reject the causal theory of action. They might simply deny that the relevant counterfactuals are *causal* counterfactuals, and so they might deny that the counterfactuals are to be evaluated in the way Vihvelin suggests. Second, Vihvelin’s response appeals to particular views on backtracking and on the evaluation of causal counterfactuals, which some may reject. The response that I proposed is free from further commitments of this kind. Third, is it really obvious that Petra would be able to try, but not able to act, if she were “a tiny bit braver”? This raises the question of what trying consists in. Elsewhere, Vihvelin says that all that is required for trying is that the agent acquires an “*effective desire or intention*”, by which she means a “desire or intention that is causally effective” (2013: 176). But if Petra’s being “a tiny bit braver” enables her to acquire an effective desire or intention, by virtue of trying, why should it not also enable her to perform the action? An *effective* desire or intention is, after all, a mental state that tends to cause the relevant action.

Finally, there are once again issues concerning trying and freedom of the will. Vihvelin (2013: 175–181) argues that an account of the ability to act in terms of trying does not require an account of the ability to try, because the ability to act does not require the *additional* ability to try. She considers an infant who successfully grasps a desired object. Here, trying consists simply in the acquisition and pursuit of a desire. In such cases, Vihvelin argues, there is no regress of

abilities or of trying to try. I agree. But the response shows only that the ability to act does not *require* an additional ability to try, and it shows only that there are *some* cases that do not generate a regress. In other cases, trying consists in the acquisition of an effective intention, and the acquisition of an intention is itself, *sometimes*, the mental action of making a decision. Such cases do raise the question of whether or not the agent has the additional ability to try, because they raise the question of whether or not the agent has the ability to choose to do otherwise. They raise thereby a regress problem that cannot be solved in the same way as in cases where the agent's trying consists in the passive (or nonactional) acquisition of a desire or intention. Again, I do not want to claim that it is impossible or even difficult to solve those problems. But Vihvelin owes us an account of the ability to choose (intend or try) to do otherwise. Without this, the regress problem cannot be fully put to rest.

VII. CONCLUDING REMARKS

Two main features distinguish the proposed account from other versions of traditional compatibilism. First, it gives an account of both the ability to do otherwise and the ability to choose to do otherwise, and it grounds the former in the latter. Second, it concedes that the possession of a general (or repertoire) ability to do otherwise is insufficient for having the ability that is required for free will. We have seen that these two features enable a straightforward response to the counterexamples and to other standard objections. And this, in turn, enabled me to fend off the challenge from the strongest version of the consequence argument for incompatibilism.

Let me close, then, with some brief remarks on two important issues that have not been addressed. All versions of compatibilism face manipulation arguments which purport to show that the causal determination of one's actions is indistinguishable from manipulation (Pereboom 2001, for instance). Further, some philosophers have argued that the ability to do otherwise is not sufficient for free will, because free will also requires that one is the 'source' or 'ultimate origin' of one's actions (Kane 1996, for instance). And some have argued that having the ability to do otherwise is not necessary either, whereas being the source or ultimate origin of one's actions is (Pereboom 2001, for instance). A proper discussion of those points is beyond the scope of this paper. Let me add, however, a few remarks that are partly based on Bernard Berofsky's recent book on free will.

Berofsky (2012: 73–92) rejects the conditional analysis and dispositional accounts. But he argues that manipulation arguments depend, ultimately, on the intuition that the manipulated agent *could not* have done otherwise (49–55). If this is correct, traditional compatibilists can argue that there is a significant difference between causal determination and manipulation, because causal determination is, on their view, compatible with having the ability to do otherwise. If, however, the manipulation of an agent merely brings it about that the agent *will not* do otherwise, traditional compatibilists can argue that the agent may well have the ability to do otherwise. In particular, they can argue that whether or not the agent is able to do otherwise depends, then, on the further details of the case, and not on whether or not the action is causally determined.

Further, Berofsky argues that the worry about ultimate origination and the source of one's actions can also be generated under the assumption that determinism is false (2012: 50–51). Suppose you find it worrying that, if determinism is true, the deterministic causation of all your actions can be traced back to factors in the distant past. You may find it equally worrying if, under indeterminism, the probabilistic causation of all your actions can be traced back to factors in the distant past—to factors that might not have but *actually did cause* all your actions. If, however, you do not find this equally worrying, it must be because you think that indeterminism opens up possibilities to do otherwise, not because it stops the tracing of explanatory chains into the distant past (55–62).¹³ Berofsky concluded that the “worry about ultimate origination reduces to the more familiar worry about alternative possibilities” (62). If this diagnosis is correct, then traditional compatibilists have again a response, because they hold that the ability to do otherwise is compatible with determinism. What I would like to emphasize, here, is that the sketched responses to manipulation and tracing arguments are available only to *traditional* compatibilists, as they depend, crucially, on the claim that the ability to do otherwise is compatible with the thesis of determinism.

NOTES

¹ Throughout, I follow the fairly common practice of leaving temporal modifiers implicit.

² Vihvelin (2013: Ch. 4) suggests that one reason for this is the influence of Frankfurt's attack on the principle of alternative possibilities. Frankfurt (1969) discusses cases in which it seems clear that the agent

is not able to perform an alternative overt action. Many commentators have argued that such cases raise only the question of whether the agent is able to *choose* to do otherwise.

³ The truth of *particular* TCA-counterfactuals does not require deterministic causal connections between choosing (wanting, intending, or trying) and acting. It may happen, in some cases, that the antecedent causes the consequent in all the closest possible worlds, even if the causal connection is merely probabilistic. But there is no good reason to generalize from such cases. If TCA-counterfactuals are to be true, *in general*, it seems that the relevant causal connections and mechanisms must, in general, be deterministic. For more on this see Bennett 2003: Ch. 16.

⁴ Throughout, I use the terms ‘disposition’, ‘capacity’, and ‘power’ interchangeably. But I consider it to be an open question whether or not the ascription of abilities can be analyzed in terms of dispositions (capacities or powers). In section VI, we will turn to a dispositional analysis of abilities.

⁵ According to some philosophers, agency *disappears* in this event-causal account (Nagel 1986; Velleman 1992). Elsewhere, I have argued that there is no such *problem* of disappearing agency. In essence, this challenge merely denies that agency can be reduced to event-causal processes, and denying a view does not amount to identifying a problem for it. For more on this see Schlosser 2010.

⁶ Few incompatibilists make this explicit, but the way in which they set up their arguments against compatibilism usually makes it clear that they do not want to commit themselves to the claim that unactualized capacities and unexercised general abilities are incompatible with determinism (see, for instance, van Inwagen 1983: 10–13; Kane 1996: 58–59; O’Connor 2000: 17). A notable exception is Clarke (2009), who is a committed incompatibilist about free will, and who explicitly acknowledged that “there is no doubt that having unexercised abilities of this [general] sort is compatible with determinism” (348).

⁷ Roughly, *basic* actions are things that we do without doing something else, whereas *non-basic* actions are things that we do *by* doing something else. For more on this see Enç 2003, for instance.

⁸ One may wonder what the truth-makers of the relevant counterfactuals are supposed to be. As mentioned, I presume that facts about what is in an agent’s act repertoire are grounded in the agent’s neurophysiological properties. I presume, further, that many of those properties are dispositional properties, which is only plausible, as they are assumed to ground facts about what the agent is able to do. Dispositional properties support counterfactuals, and so it seems reasonable to assume that the truth-makers of the relevant counterfactuals are among the intrinsic properties mentioned above: the neurophysiological properties of a human agent that explain why certain actions are in the agent’s act

repertoire (and why others are not). They are, in other words, the intrinsic properties that provide the ‘causal bases’ of the relevant dispositions. It may be helpful to note, here, that a conditional analysis of dispositions is perfectly compatible with realism about dispositional properties and intrinsic causal bases. One may hold that the ascription of dispositional properties can be analyzed in terms of counterfactuals *and* that the truth of those counterfactuals is grounded in the possession of the relevant intrinsic properties. See also section V.C.

⁹ Let P be a true proposition describing the entire state of the universe at some time in the past, and let L be a true proposition describing the conjunction of all the laws of nature. Assume for the sake of argument that $N(P \& L)$. If determinism is true, there will be some true proposition, Q , that describes some intentional action performed by some agent at some time, such that it is true that $N((P \& L) \supset Q)$. According to a conditional analysis, it may well be false that NQ , because it may well be that the agent who performed that action would have done otherwise, had that agent chosen (wanted, intended, or tried) to do otherwise. So, on any such analysis, β is invalid, and we are able to do otherwise even if our actions are causally determined by the past and in accord with the laws of nature. For details and discussion see, for instance, van Inwagen 1983; Kane 1996; O’Connor 2000.

¹⁰ In brief, β entails, in conjunction with $\Box p \vdash Np$, the following agglomeration rule: $Np \& Nq \vdash N(p \& q)$. Now, suppose that Peter had a choice about whether or not to toss a coin. Peter did not have a choice about whether the coin would land ‘heads up’ or ‘tails up’. In particular, he did not have a choice about the truth value of the proposition, P , that ‘the coin did not land heads up’, and he did not have a choice about the truth value of the proposition, Q , that ‘the coin did not land tails up’. But he did have a choice about the truth value of $(P \& Q)$. This is a counterexample to the agglomeration rule, and it shows, therefore, that β is invalid. See also Huemer 2000.

¹¹ In case the reader finds it difficult to see why that is, it should help to see note 9.

¹² Essentially the same response applies, *mutatis mutandis*, to the interpretation of the N operator and the version of the consequence argument in Pruss 2013.

¹³ Agent-causal libertarians might argue that this favors their position, as it seems that only genuine and irreducible substance-causation by agents can stop the tracing and account for ultimate origination. However, elsewhere I have argued that appeal to irreducible agent-causation is of no help, as origination by agents, *qua* substances, does not by itself confer any recognizable form of self-governance or agential control (Schlosser 2008 and 2010).

REFERENCES

- Anscombe, G. Elizabeth M. 1957. *Intention*. Oxford: Basil Blackwell.
- Atmanspacher, Harald and Stefan Rotter. 2011. "On Determinacy or Its Absence in the Brain." In *Free Will and Modern Science*, ed. Richard Swinburne. Oxford: Oxford University Press, 84–101.
- Austin, John L. 1956. "Ifs and cans." *Proceedings of the British Academy* 42: 109–32.
- Ayer, Alfred J. 1954. Freedom and Necessity. In his *Philosophical Essays*. New York: St. Martin's Press, 3–20.
- Bennett, Jonathan, 2003. *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press.
- Berofsky, Bernard. 2012. *Nature's Challenge to Free Will*. Oxford: Oxford University Press.
- Chisholm, Roderick. 1964. "Human Freedom and the Self." *The Lindley lectures*, Copyright by the Department of Philosophy, University of Kansas. Reprinted in *Free Will*, ed. Gary Watson. Oxford: Oxford University Press, 1982, 24–35.
- Clarke, Randolph. 2009. "Dispositions, Abilities to Act, and Free Will: The New Dispositionalism." *Mind* 118: 323–351.
- . 2010. "Skilled Activity and the Causal Theory of Action." *Philosophy and Phenomenological Research*, 80 (3): 523–550.
- Davidson, Donald. 1963. "Actions, Reasons, and Causes." *Journal of Philosophy* 60: 685–700. Reprinted in his *Essays on Actions and Events*. Oxford: Clarendon Press, 1980, 3–20.
- . 1973. "Freedom to Act." In *Essays on Freedom of Action*, ed. Ted Honderich. London: Routledge and Kegan Paul, 137–56. Reprinted in his *Essays on Actions and Events*. Oxford: Oxford University Press, 1980, 63–81.
- Dancy, Jonathan. 2000. *Practical Reality*. Oxford: Oxford University Press.
- Dennett, Daniel C. 1984. *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge: MIT Press.
- Enç, Berent. 2003. *How We Act: Causes, Reasons, and Intentions*. Oxford: Oxford University Press.
- Ekstrom, Laura W. 2002. "Libertarianism and Frankfurt-Style Cases." In *The Oxford Handbook of Free Will*, ed. Robert Kane. Oxford: Oxford University Press, 309–322.
- Fara, Michael. 2008. "Masked Abilities and Compatibilism." *Mind* 117: 843–65.
- Fischer, John M. 1994. *The Metaphysics of Free Will*. Oxford: Blackwell.
- Fischer, John M. and Mark Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Frankfurt, Harry. 1969. "Alternate Possibilities and Moral Responsibility." *Journal of Philosophy* 66: 829–39.

- . 1971. “Freedom of the Will and the Concept of a Person.” *Journal of Philosophy* 68: 5–20.
- Goetz, Stewart. 2005. “Frankfurt-Style Counterexamples and Begging the Question.” *Midwest Studies in Philosophy* 29: 83–105.
- Hobbes, Thomas. 1958. *Leviathan*. Indianapolis: Bobbs-Merrill. Originally published in 1651.
- Huemer, Michael. 2000. “Van Inwagen’s Consequence Argument.” *Philosophical Review* 109: 525–544.
- Hume, David. 1955. *An Enquiry Concerning Human Understanding*, ed. Lewis Selby-Bigge. Oxford: Clarendon Press. Originally published in 1748.
- Kane, Robert. 1996. *The Significance of Free Will*. New York: Oxford University Press.
- Lehrer, Keith. 1968. “Cans without ifs.” *Analysis* 29: 29–32.
- Lewis, David. 1997. “Finkish Dispositions.” *Philosophical Quarterly* 47: 143–158.
- Locke, John. 1975. *An Essay Concerning Human Understanding*, ed. Peter Nidditch. Oxford: Oxford University Press, 1975. Originally published in 1689.
- McKay, Thomas and David Johnson. 1996. “A Reconsideration of an Argument Against Compatibilism.” *Philosophical Topics* 24: 113–122.
- Mele, Alfred R. 2003. *Motivation and Agency*. Oxford: Oxford University Press.
- Moore, George E. 1912. “Free Will.” In his *Ethics*. Oxford: Oxford University Press, 196–222.
- Nagel, Thomas. 1986. *The View from Nowhere*. New York: Oxford University Press.
- Nelkin, Dana K. 2011. *Making Sense of Freedom and Responsibility*. New York: Oxford University Press.
- O’Connor, Timothy. 2000. *Persons and Causes: The Metaphysics of Free Will*. New York: Oxford University Press.
- Pereboom, Derk. 2001. *Living without Free Will*. Cambridge: Cambridge University Press.
- Pruss, Alexander R. 2013. “Incompatibilism Proved.” *Canadian Journal of Philosophy* 43: 430–437.
- Schlosser, Markus E. 2008. “Agent-Causation and Agential Control.” *Philosophical Explorations* 11: 3–21.
- . 2010. “Agency, Ownership, and the Standard Theory.” In *New Waves in Philosophy of Action*, ed. Jesús Aguilar, Andrei A. Buckareff, and Keith Frankish. Basingstoke: Palgrave Macmillan, 13–31.
- Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford: Oxford University Press.
- Van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford: Clarendon Press.
- Vargas, Manuel. 2007. “Revisionism.” In *Four Views on Free Will*, ed. John M. Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas. Oxford: Blackwell Publishing, 126–165.
- Velleman, J. David. 1992. “What Happens When Someone Acts?” *Mind* 101: 461–481.
- Vihvelin, Kadri. 2004. “Free Will Demystified: A Dispositional Account.” *Philosophical Topics* 32: 427–50.

- . 2013. *Causes, Laws, and Free Will*. New York: Oxford University Press.
- Widerker, David. 2003. “Blameworthiness and Frankfurt’s Argument Against the Principle of Alternative Possibilities.” In *Moral Responsibility and Alternative Possibilities*, ed. Michael McKenna and David Widerker. Aldershot: Ashgate, 53–73.
- Williams, Bernard. 1979. “Internal and External Reasons.” In *Rational Action*, ed. Ross Harrison. Cambridge: Cambridge University Press, 101-113. Reprinted in *Moral Luck*. Cambridge: Cambridge University Press, 1981, 101–13.
- Wolf, Susan. 1990. *Freedom within Reason*. Oxford: Oxford University Press.