

JOSEPH KEIM CAMPBELL

A COMPATIBILIST THEORY OF ALTERNATIVE  
POSSIBILITIES<sup>1</sup>

(Received 15 January 1996)

1. INTRODUCTION

In recent discussions of freedom and determinism, two views of compatibilism have emerged. One, which I call *strong compatibilism*,<sup>2</sup> assumes the following:

1. Alternative possibilities of action are necessary for both free will and moral responsibility.
2. Both free will and moral responsibility are compatible with causal determinism.

Strong compatibilism has long been thought to be incoherent by its incompatibilists, who deny (2),<sup>3</sup> but recently it has also received criticism from other compatibilists, who deny (1). In a celebrated essay, for instance, Harry Frankfurt has provided examples in which it seems that a person is morally responsible for some action even though she could not have done otherwise.<sup>4</sup> This result has delighted many philosophers who find that attempts to establish (2) in light of (1) have generally led to failure,<sup>5</sup> and has given rise to what I call *weak compatibilism*:

- 1'. Moral responsibility does not require alternative possibilities of action, so neither does any freedom that is necessary for moral responsibility.
- 2'. Moral responsibility – and any freedom that is essential to it – is compatible with causal determinism.<sup>6</sup>

The strength and uniqueness of weak compatibilism rests with (1') and, thus, the supposed counterexamples put forth by Frankfurt.

*Philosophical Studies* **88**: 319–330, 1997.

© 1997 Kluwer Academic Publishers. Printed in the Netherlands.

In this paper I prove that (1') is false. The *Frankfurt examples*,<sup>7</sup> as I call them, not only fail to establish (1') but they lend support to strong compatibilism instead. The paper is a defense of the *traditional theory of the will* which states that a person has free will provided that she has certain powers of agency and cognition which are not in any way impeded. That one has such powers entails the existence of possibilities which can be used to construct a compatibilist account of 'could have done otherwise'. Since agents in the Frankfurt examples have the appropriate powers they must also have alternative possibilities of action. At most, the agents lack categorical possibilities that are incompatible with determinism. Therefore, the Frankfurt examples promote strong, not weak, compatibilism.

## 2. FRANKFURT ON ALTERNATIVE POSSIBILITIES OF ACTION

According to the strong compatibilist, a person has free will only if she has alternative possibilities of action, that is, only if she could have, in some relevant sense, done otherwise.<sup>8</sup> In addition, the strong compatibilist will also endorse the following:

**The Principle of Alternative Possibilities [PAP]:** A person is morally responsible for what she has done only if she could have done otherwise.<sup>9</sup>

Given PAP and a standard incompatibilist argument,<sup>10</sup> one may contend that moral responsibility is also incompatible with determinism. Since many previous compatibilist accounts of alternative possibilities are flawed<sup>11</sup> some compatibilists have jettisoned the notion of alternative possibilities altogether and provided some other basis for a theory of freedom and moral responsibility. Thus, we are led to weak compatibilism, which is primarily driven by Frankfurt's apparent counterexamples to PAP.

To illustrate such an example, suppose that a woman, Eleanor, and her father, Roscoe, decide to rob a bank since both are desperately in need of money. Despite her claims to the contrary, Roscoe fears that Eleanor may change her mind about the robbery at the last minute. As a fall-back, he has a device implanted in Eleanor's brain that, when activated, will render her unable to do anything other than

follow through with the robbery as planned. As it happens, Eleanor is a willing subject and she performs the crime on her own, without the activation of the device. Eleanor is morally responsible for her action but, it seems, she could not have done otherwise. Hence, PAP appears to be false.

### 3. FREE WILL AND MORAL RESPONSIBILITY<sup>12</sup>

Traditionally, philosophers have viewed the will as a *faculty*, “a set of interconnected powers” specifically related to action and thus practical reasoning.<sup>13</sup> Some regard the term ‘faculty’ with suspicion<sup>14</sup> but a faculty is nothing more than a set of powers and is, therefore, no more mysterious than the powers themselves. Keeping with tradition, I identify one’s will with those *active powers* and *cognitive capacities* relevant to moral responsibility.<sup>15</sup> As I see it, having a *free will* is nothing more than having a will that is in good working order, one that is unhampered. When one’s will is not free it means that a particular power or capacity is in some way either lacking or impeded.<sup>16</sup>

Far from being clouded in mystery, the traditional approach allows one to construct a naturalistic account of freedom. Psychology can explain the relevant powers and capacities and once this is done there is nothing more to providing an account of free will. A free will is merely a will that is unobstructed, one in which the appropriate powers and capacities remain intact. In this section, I explain the active powers and the cognitive capacities in more detail and discuss some cases which motivate this traditional theory of the will. This will help to pave the way for my compatibilist account of alternative possibilities.

An *active power* is any power essential to human action whereas a *cognitive capacity* is a capacity which pertains to an epistemic or psychological state of the agent. Often these will overlap but I think that in combination they are exhaustive of the powers and capacities that comprise the will. Examples of cognitive capacities include those relevant to understanding, reasoning and deliberation.<sup>17</sup> As indicated, some of these, particularly those concerned with practical reasoning and choice, may also be classified as active powers but this latter category will include other capacities, as well.

An examination of specific types of compulsion is helpful in establishing the traditional theory of the will. First, there are *internal compulsions* where one's behavior is driven by some psychological disorder such as kleptomania. Second, there are *external compulsions*. Examples include behavior influenced by hypnotic suggestion or even the activation of a physical device like the one in the Frankfurt examples. The difference here might be subtle. By classifying the second group as external I am merely suggesting that the person's behavior has a more immediate external cause. Both groups are similar in that they involve a disruption of one's cognitive capacities. One might suggest that they diminish a person's agency, as well. Either way, they undermine one's free will by frustrating important powers and capacities.

Other types of mitigating circumstances involve elements of *faulty reasoning*. Like compulsions they too can be divided into *internal* and *external* groupings. Cases of the former kind include schizophrenia or other psychological ailments whereas the latter types include behavior resulting from hallucinogenic drugs such as LSD or PCP. These examples differ from those of compulsion in that the resulting act is not in any way forced. Rather, one's decision making processes are impeded. Still, they violate one's will by hampering one's cognitive capacities.

When analyzing examples like those above we usually assume that a person has performed some wrong action for which, we presume, she is not responsible, and then we look for an explanation that accounts for our intuitions. My suggestion is that with mitigating circumstances pertaining to a loss of free will there is an excuse which either involves the absence or hindrance of a cognitive capacity or persuades us to reevaluate our original judgment that the agent is the true cause. Clearly, certain possibilities are relevant to agency but in order to establish that an agent had alternate possibilities of action we must also consider those possibilities relevant to cognition. Once this is done a response to Frankfurt can be given.

#### 4. A COMPATIBILIST ACCOUNT OF ALTERNATIVE POSSIBILITIES

'Could' is a modal concept and it is fashionable to explain these concepts in terms of possible worlds.<sup>18</sup> When talking about alternative

possibilities of action, though, not all possible worlds are equal since some are not relevant to attributions of freedom and moral responsibility. There needs to be a way of restricting possible worlds, saving the relevant ones while eliminating those that are irrelevant. The remaining worlds are ones which are *accessible* to the agent. Hence, a person has alternative possibilities of action if and only if there is some world which is accessible to that agent. In this section, I establish criteria proving that there are worlds still accessible to Eleanor in the Frankfurt examples. Thus, these examples are not genuine counterexamples to PAP. To motivate this strategy, I begin with cases of omission and the corresponding version of PAP.

**The Principle of Possible Action [PPA]:** A person is morally responsible for failing to perform a given act only if she could have performed that act.<sup>19</sup>

PPA suggests that abilities to do otherwise are necessary for moral responsibility in circumstances where an action is omitted. It is generally recognized that this principle is unaffected by the Frankfurt examples.<sup>20</sup> As a first case, suppose that Eleanor is gardening on the roof of a four-story building next to a yard where a young kitten has accidentally fallen into a pool. Since she cannot jump down four stories, she cannot save the cat; should the cat die Eleanor would not be negligent, according to PPA.

Now it is *logically* possible for Eleanor to safely jump four stories so there is *some* possible world in which this happens. One may argue that in some sense Eleanor could have saved the kitten. But those worlds where Eleanor saves the cat are not accessible to her. Any world in which Eleanor can safely jump four stories is a world where Eleanor has quite different abilities and capacities than she has in the actual world. Or perhaps it is a world which does not share all of our laws of nature. In any event, it is not possible for Eleanor to save the cat in a sense that is relevant to attributions of freedom and moral responsibility.

This example suggests at least two important constraints on those worlds that are genuinely accessible to an agent.

- A1. Accessible worlds must have the same laws of nature as the actual world.<sup>21</sup>
- A2. In accessible worlds, agents cannot have any abilities or capacities that they lack in the actual world.

(A2) is supported by another example, as well. Suppose Eleanor is closer to the cat but, unfortunately, Eleanor has not learned how to swim. Since she cannot swim, it is natural to say that she could not have saved the cat and PPA would suggest that she is not responsible for its death. Clearly, there are possible worlds in which Eleanor is able to swim but none of them are accessible to her in this situation.

Thus far we have the beginnings of a possible worlds analysis of ‘could have done otherwise,’ similar to one pioneered by Keith Lehrer.<sup>22</sup> Lehrer is also well known for his criticisms of the *hypothetical analysis* of alternative possibilities, supported by many compatibilists. According to this view “a person could have done otherwise” simply means that the person would have done otherwise if she had so chosen. To see why this view is faulty, let’s suppose that Eleanor is standing beside the pool after the cat falls in. In order to save it, she need only bend down and pick it up. However, due to an unfortunate childhood incident, resulting in a pathological fear of cats, she cannot choose to pick the cat up and the cat drowns. Intuitively, she could not have saved the cat because of her neurosis. The hypothetical account, though, would have us believe that Eleanor could have saved the cat since it is still true that she would have saved the cat if she had chosen to do so. What is wrong with the hypothetical account is that it suggests worlds in which Eleanor lacks her psychological affliction are still accessible to her though they should not be.<sup>23</sup>

Lehrer recognized that worlds which are accessible to agents “must not bestow . . . any advantage for performing” actions that the person lacks in the actual world.<sup>24</sup> Worlds where a person lacks psychological aversions that she has in the actual world bestow such an advantage. This result leads to another criterion for accessibility.

- A3. Any advantage that a person has in an accessible world which she lacks in the actual world must result from something that she does in that possible world.<sup>25</sup>

(A3) allows that persons may acquire advantages even though they cannot be unfairly granted. Suppose, for instance, that Eleanor has a form of kleptomania which is treatable and can be easily overcome through therapy. The first few times she steals we might excuse her but after awhile we would not. Even if she never seeks treatment, there are still accessible worlds where she lacks the affliction.

One might think that (A3) confirms Frankfurt's arguments. Recall from the above examples that any world where Eleanor has the device implanted in her brain is one where she robs the bank. She can only avoid robbing the bank if the device is not implanted. In such worlds, though, Eleanor seems to have an advantage that she lacks in the actual world. If this line of reasoning is correct, then the Frankfurt examples would indeed require the rejection of PAP.

However, there is an important difference between the Frankfurt examples and cases of psychological compulsion. In the latter, the affliction is a causal factor in the performance of the wrong action whereas in the former it is not. The implanted device insures that Eleanor robs the bank but it does not play any causal role. As such, worlds in which the device is absent do not give Eleanor any advantage. Contrast this with the case in which the device is activated. This leads to one final criterion of accessibility.

- A4. Accessible worlds need not include factors which are causally irrelevant to the performance of actions.<sup>26</sup>

Examples of compulsion and faulty reasoning are causally efficacious whereas non-activated devices like the one in the Frankfurt examples are not. Thus, worlds which exclude the latter do not bestow relevant advantages unlike worlds which exclude the former.

(A4) is also supported by examples of causal overdetermination. Suppose that Eleanor is a kleptomaniac but that she is not always burdened by her ailment and that she decides to steal something not because she is compelled to but for the fun of it. The fact that she has kleptomania is not, then, an excuse. (A3) cannot suggest that every world accessible to Eleanor must be one in which her kleptomania is efficacious or it would entail that she could not do otherwise. Therefore, accessible worlds should not include overdetermining causes. Since Eleanor's kleptomania is not a causal factor in her theft,

worlds in which she does not have kleptomania become relevant in assessing her responsibility.

Likewise, since the device in the Frankfurt examples is not ever activated, worlds without the device may still be accessible to Eleanor. For this reason, I claim that she has alternative possibilities of action in the examples noted above. First, consider the following counterfactual:

3. If Eleanor had chosen not to rob the bank and the device had not been implanted, then Eleanor would not have robbed the bank.

Counterfactuals like (3) are essential to confirming that Eleanor was causally responsible for the bank robbery. Hence, if (3) is either false or vacuous, then Eleanor cannot be held morally responsible for her action.<sup>27</sup> Since it is not vacuous, the antecedent of (3) entails the existence of a possible world in which Eleanor does otherwise. Let's call this world *W*. Is *W* accessible to Eleanor? There is no reason to suggest that *W* lacks our laws of nature or that Eleanor has different abilities or capacities in *W* than she has in the actual world. Also, since the device implanted by Roscoe is not operative, the fact that it does not exist in *W* does not bestow on Eleanor any advantage that she lacks in the actual world. Given our criteria we can conclude that *W* is accessible to Eleanor and that she indeed could have done otherwise.

## 5. CONCLUDING REMARKS

When we ask if a person has alternate possibilities of action we are really asking whether or not she has certain powers and capacities that are relevant to moral responsibility. More specifically, we are inquiring about the status of those active powers and cognitive capacities which constitute her will. In order to give a full account of these powers and capacities one needs to consider possible worlds other than the actual one. If the person's will is intact, then certain possible worlds exist which are accessible to the agent. For this reason, any person who freely performs an action 'could have done otherwise.' Since persons in the Frankfurt examples perform actions and have

cognitive capacities which are not impeded in any way, they must also have alternate possibilities of action.

I admit that in the Frankfurt examples there is a sense in which Eleanor could not do otherwise but, since Eleanor is free and morally responsible for her actions, this sense is not relevant to either free will or moral responsibility. Eleanor could not have done otherwise in the *categorical sense*: each world which shares the past and laws of nature of the actual world is one in which Eleanor performs the same action.<sup>28</sup> Hence, categorical possibilities are not relevant to either free will or moral responsibility though alternate possibilities of some kind clearly are relevant. Therefore, the Frankfurt examples support strong, not weak compatibilism.

In summary, free will is a collection of related powers and capacities, a faculty. These include cognitive capacities, relevant to practical reasoning, deliberation and choice, as well as active powers. All of these powers and capacities are necessary for moral responsibility and merely having them entails the existence of alternate possibilities of action. In spite of this, there is no reason to think that such possibilities are incompatible with determinism, and there is much reason to think that they are not.

#### NOTES

<sup>1</sup> I have benefited from many helpful comments on previous drafts of this paper. I especially thank Henry Byerly, Tom Christiano, Randolph Clarke, Robert Cummins, Hud Hudson, Robert Kane, Keith Lehrer, Kenneth Lucey, George Rainbolt, Russ Shafer-Landau, Jim Stone, Kadri Vibvelin, Jeff Wattles, Michael J. Zimmerman and a host of referees from assorted journals.

<sup>2</sup> See especially Moore 1912, Lehrer 1966b, 1976 and 1980, and Lewis 1981.

<sup>3</sup> See Campbell 1957, Taylor 1963, Chisholm 1966b, van Inwagen 1983 and Kane 1985.

<sup>4</sup> Frankfurt 1969. See Davidson 1973 for a similar argument.

<sup>5</sup> Many strong compatibilists adopt the *hypothetical analysis* of 'could have done otherwise' which is discussed and criticized in section (4) above.

<sup>6</sup> Weak compatibilist accounts differ though most accept a notion of 'acting freely' which embraces (1') and (2') (See Fischer 1986b, pp. 46–47). Of course, weak compatibilists would reject any notion of freedom associated with alternative possibilities, such as 'freedom of action.' A collection of important writings is contained in Fischer 1986a, part I. John Martin Fischer and Mark Ravizza (1991) endorse a version of weak compatibilism but call it 'semicompatibilism.' Susan Wolf (1990) calls it the 'Real Self View.'

<sup>7</sup> Peter van Inwagen (1978) calls these 'Frankfurt-type counter-examples' but

since I do not think that they are genuine counterexamples I have slightly altered the name.

<sup>8</sup> Along with others, I regard the expressions ‘could have,’ ‘was able to,’ and ‘had it within her power’ as equivalent. Likewise, I treat the terms ‘capacity,’ ‘ability,’ and ‘power’ as synonyms. See van Inwagen 1983, p. 8.

<sup>9</sup> Frankfurt 1969, p. 143.

<sup>10</sup> See, for instance, van Inwagen 1983 for such an argument.

<sup>11</sup> Many compatibilists – following Hume 1975, p. 447, and Moore 1912 – endorse a hypothetical account of ‘could have done otherwise.’ See section (4) below for criticisms.

<sup>12</sup> This section was greatly influenced by similar discussions on moral responsibility in Aristotle 1985, Hart 1968 and Feinberg 1970b.

<sup>13</sup> See Kane 1985, pp. 19–20. My discussion of the will borrows significantly from this work.

<sup>14</sup> Van Inwagen 1983, p. 8.

<sup>15</sup> The name ‘active power’ comes from Thomas Reid (1983). Reid, though, adopts a libertarian conception of freedom, and I intend no such constraints on the notion. There may be other differences, as well, though I do endorse much of what Reid says. See, especially, Reid 1983, pp. 323–351.

<sup>16</sup> This is reminiscent of John Locke’s theory of the will. See Locke 1979, Book 2, Chapter XXI, ‘Of Power.’

<sup>17</sup> See Hart 1968, p. 227, and Kane 1985, p. 19.

<sup>18</sup> See, especially, the work of David Lewis (1983 and 1986).

<sup>19</sup> See van Inwagen 1978.

<sup>20</sup> In van Inwagen 1978, the author defends PPA and proves that it is not defeated by the Frankfurt examples. See also Fischer and Ravizza 1991, p. 227.

<sup>21</sup> Some philosophers might disagree with this requirement. See especially Lewis 1981.

<sup>22</sup> Lehrer 1976.

<sup>23</sup> See Lehrer 1966b, 1968 and 1976, pp. 248–250. Criticisms of the hypothetical analysis started with J.L. Austin (1956) but were articulated in their finest form in the work of Roderick M. Chisholm (1964a and 1966) and Lehrer.

<sup>24</sup> Lehrer 1976, p. 254.

<sup>25</sup> Ibid., p. 256.

<sup>26</sup> See Fischer 1982 for his *actual-sequence theory of responsibility* which led me to this criterion. Fischer’s theory is quite different from mine, however. See also Fischer and Ravizza 1991.

<sup>27</sup> See Cross 1992 for some important connections between causation and counterfactuals. Cross is especially concerned with cases of causal overdetermination.

<sup>28</sup> See Campbell 1957.

#### REFERENCES

- Aristotle (1985): *Nicomachean Ethics*, translated by Terence Irwin. Indianapolis: Hackett.
- Austin, J. L. (1956): ‘Ifs and Cans’, *Proceedings of the British Academy* 42, 109–132.

- Campbell, C. A. (1957): *On Selfhood and Godhood*, London: George Allen & Unwin.
- Chisholm, R. (1964a): 'J. L. Austin's *Philosophical Papers*', *Mind* 73, 20–25.
- Chisholm, R. (1964b): 'Human Freedom and the Self', The Lindley Lecture: 3–15. Department of Philosophy, University of Kansas.
- Chisholm, R. (1966): 'Freedom of Action', in Lehrer 1966a.
- Cross, C. (1992): 'Counterfactuals and Event Causation', *Australasian Journal of Philosophy* 70, 307–323.
- Davidson, D. (1973): 'Freedom to Act', in Ted Honderich (ed.), *Essays on Freedom of Action*, London: Routledge and Kegan Paul.
- Feinberg, J. (1970a): *Doing and Deserving*, Princeton: Princeton University Press.
- Feinberg, J. (1970b): 'Action and Responsibility', in Feinberg 1970a.
- Fischer, J. M. (1982): 'Responsibility and Control', *Journal of Philosophy* 89, 24–40. Reprinted in Fischer 1986a.
- Fischer, J. M. (ed.) (1986a): *Moral Responsibility*, Ithica: Cornell University Press.
- Fischer, J. M. (1986b): 'Introduction: Responsibility and Freedom', in Fischer 1986a.
- Fischer, J. M. and M. Ravizza (1991): 'Responsibility and Inevitability', *Ethics* 101, 258–278.
- Frankfurt, H. G. (1969): 'Alternate Possibilities and Moral Responsibility', *Journal of Philosophy* 66, 828–839. Reprinted in Fischer 1986a.
- Hart, H. L. A. (1968): *Punishment and Responsibility*, Oxford: Oxford University Press.
- Hume, D. (1975): *Enquiries Concerning Human Understanding*, third edition, edited by L. A. Selby-Bigge and revised by P. M. Nidditch. Oxford: Oxford University Press.
- Kane, R. (1985): *Free Will and Values*, Albany: State University of New York Press.
- Lehrer, K. (ed.) (1966a): *Freedom and Determinism*, New York: Random House.
- Lehrer, K. (1966b): 'An Empirical Disproof of Determinism?', in Lehrer 1966a.
- Lehrer, K. (1968): 'Cans Without Ifs', *Analysis* 29, 29–32.
- Lehrer, K. (1976): '"Can" in Theory and Practice: A Possible Worlds Analysis', in M. Brand and D. Walton (eds.), *Action Theory*, Dordrecht: D. Reidel.
- Lehrer, K. (1980): 'Preferences, Conditionals and Freedom', in Peter van Inwagen (ed.), *Time and Cause: Essays Presented to Richard Taylor*, Dordrecht: D. Reidel.
- Lewis, D. (1981): 'Are We Free to Break the Laws?', *Theoria* 47, 113–121. Reprinted in Lewis 1986.
- Lewis, D. (1983): *Philosophical Papers*, Volume I, Oxford: Oxford University Press.
- Lewis, D. (1986): *Philosophical Papers*, Volume II, Oxford: Oxford University Press.
- Locke, J. (1979): *An Essay Concerning Human Understanding*, edited by Peter H. Nidditch. Oxford: Clarendon Press.
- Moore, G. E. (1912): *Ethics*, Chapter VI, New York: Oxford University Press.
- Reid, T. (1983): *Inquiry and Essays*, edited by Ronald E. Beanblossom and Keith Lehrer. Indianapolis: Hackett Publishing Company.
- Taylor, R. (1963): *Metaphysics*, first edition, Englewood Cliffs: Prentice-Hall.
- Wolf, S. (1990): *Freedom Within Reason*, New York: Oxford University Press.

van Inwagen, P. (1978): 'Ability and Responsibility', *Philosophical Review* 87, 201–224. Reprinted in Fischer 1986a.

van Inwagen, P. (1983): *An Essay on Free Will*, Oxford: Clarendon Press.

*Department of Philosophy*  
*Washington State University*  
*Pullman, WA 99164–5130*  
*USA*

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.