

DENNETT ON THE BASIC ARGUMENT

JOHN MARTIN FISCHER

Abstract: I argue that Dennett does not adequately support his rejection of the “Basic Argument” for the incompatibility of causal determinism and the sort of free will that involves genuine access to alternative possibilities (sometimes referred to as the “Consequence Argument”). In addition, I seek to highlight the plausibility and importance of the incompatibilist’s interpretation of this sort of free will.

Keywords: Basic Argument for Incompatibilism, could have done otherwise, Dennett, Taylor, van Inwagen.

In recent work, Daniel Dennett has offered a critique of what might be called the “Basic Argument” for the incompatibility of causal determinism and the sort of free will that involves genuine access to alternative possibilities (Taylor and Dennett 2001; Dennett 2003). In previous work, Dennett has declined directly to confront this argument, but now he says:

Christopher Taylor has greatly clarified my thinking on this topic and shown me how to launch a deeper and more radical campaign in support of my earlier claims to this effect, and our coauthored paper (Taylor and Dennett 2001) provides more technical detail than is needed here. Here I will attempt a gentler version of our argument, highlighting the main points so that non-philosophers can at least see what the points of contention are, and how we propose to settle them, while leaving out almost all the logical formulae. Philosophers should consult the full-dress version, of course, to see if we have actually tied off the loose ends, and closed the loopholes that are passed by without mention in this telling. (Dennett 2003, 64)

In this article, I take up Dennett’s challenge to look more carefully at the argument. Since the (marginally) more technical formulation of the critique appears in the jointly written paper, I take as my target the views of Dennett and Taylor.

Although there have been many different formulations of what is essentially the same argument for incompatibilism, Taylor and Dennett focus on van Inwagen’s “Consequence Argument.” Van Inwagen himself gives various different formulations of the Consequence Argument (van Inwagen 1983). Taylor and Dennett do not consider the arguments as

presented in van Inwagen's book, but they discuss only their own reformulation of what they take to be the argument in van Inwagen 1975. Here is the Taylor and Dennett version:

1. Let ϕ be some event that actually occurs in agent A's life (missing a putt, say). Also let σ_0 be a comprehensive description of the universe's state at some time in the remote past, and let λ be a statement of the laws of nature.
2. Then, assuming determinism, $\lambda \wedge \sigma_0 \Rightarrow \phi$ applies in every possible world. Equivalently, $\sim \phi \Rightarrow \sim (\lambda \wedge \sigma_0)$.
3. If A has the power to cause α and $\alpha \Rightarrow \beta$ obtains in every possible world, then A has the power to cause β .
4. So if A has the power to cause $\sim \phi$, then A has the power to cause the falsity of either λ or σ_0 , which is absurd.
5. Therefore A lacks the power to cause $\sim \phi$. (Taylor and Dennett 2001, 273–74)

Taylor and Dennett contend that premise 3 should be rejected. They first contend that causation should be analyzed in terms of counterfactual necessity, rather than sufficiency. That is, when it is true that some event $e1$ causes $e2$, then the primary element in the analysis is that if $e1$ had not occurred, then $e2$ would not have occurred. The causal claim does not require that $e1$'s occurring (together with the relevant background conditions and laws) is sufficient for $e2$'s occurring.

Building on this idea, Taylor and Dennett continue:

... we would recommend that van Inwagen's 'power to cause α ' be rendered as follows:

A has the *power to cause* α iff for some sentence γ describing an action of A and a world f close to actuality, $\gamma \wedge \alpha$ holds in f and $\alpha \Rightarrow \gamma$ in every world similar to f .

In other words, within some cluster of nearby worlds, there is a possible action of A (called γ) that is a necessary condition for α to occur. But under this definition, line 3 has no warrant whatever. Line 3 hypothesizes that $\alpha \Rightarrow \gamma$ in a cluster of nearby worlds, and that $\alpha \Rightarrow \beta$ in every world; if we could deduce that $\beta \Rightarrow \gamma$ in this cluster, we would be home free. But of course in Logic 101 we learn that $\alpha \Rightarrow \gamma$ and $\alpha \Rightarrow \beta$ do not entail $\beta \Rightarrow \gamma$, and so line 3 fails, and van Inwagen's argument with it. (Taylor and Dennett 2001, 274)

I want to begin a critical examination of Taylor and Dennett by simply noting that it would be surprising if the proponents of the Basic Argument for incompatibilism were guilty of some fairly simple logical error—the sort one would learn about in Logic 101. Furthermore, since the Basic Argument can be formulated in various different ways, it would

not follow that it is unsound from the mere fact that one particular version of it is unsound. But let us first consider the particular version of the argument presented by Taylor and Dennett.

It is interesting that, in developing their account of causation, Taylor and Dennett tend to discuss causal sentences that straightforwardly refer to two events (where one of the events is an action), putatively causally related: “Bill’s tripping Arthur caused him to fall,” “The sharpshooter’s actions caused the victim’s death,” “Susie’s throwing the rock causes the bottle to shatter,” and so forth. But philosophers who have developed versions of the Basic Argument have sometimes introduced special locutions, such as “*S* makes it the case that *p*,” or “*S* renders a certain proposition *r* true (or false).” Indeed, it is precisely the latter locution that Peter van Inwagen (1975) employs in the article to which Taylor and Dennett refer. And yet Taylor and Dennett reformulate the argument employing variables that sometimes range over actions (“ α ” and “ β ”), and sometimes range over “descriptions” or “statements” (such as “ σ ” and “ λ ”).¹ This tends to conflate issues about causation of events and causation of states of affairs or propositions or the truth of sentences, that is, causing it to be the case that a certain state of affairs obtains (or a certain proposition or sentence is true).

Now some of the complexities of giving a proper account of “making it the case that *p*” or “rendering it true (or false) that *r*” have been explored in the literature (Fischer 1983, 1988, and 1994; Ekstrom 2002, 1–55). I shall simply point out here that states of affairs that involve conjunctions or disjunctions of events, of conditional relationships among events, or events that are in the past relative to the time at which the “causing event” is alleged to occur, are complex and require delicate analysis. I do not believe that Taylor and Dennett provide an account sufficiently nuanced to handle such complex states of affairs, and their conflation of causation of events with making it the case that some state of affairs obtains (or rendering some proposition true) leads them astray here.

What exactly is it to render some proposition *p* true? Well, suppose that *p* is the “simple” proposition that event *e* occurs. On one account, to render *p* true here would be to perform some act the occurrence of which causes *e* to occur. Call this the “strongly causal” account of rendering a proposition true. Adopting this strongly causal account of rendering a proposition true, and employing this notion in the Taylor and Dennett reformulation of the argument, the argument does indeed appear to be invalid, and it is precisely line 3 that is problematic (as pointed out by Taylor and Dennett). But, on this strongly causal reading, it is just *obvious* that line 3 is false, and we do not need any complicated reasoning to see this. For example, line 3 would imply (on the relevant reading) that

¹ Note that in his 2003 Dennett employs “ γ ” to refer to both a sentence and a possible action.

if *A* has the power to walk to school, and if he were to walk to school the sun would not flicker out in every possible world, then *A* has it in his power to cause the sun's not flickering out (or the sun's continuing to shine). But this is absurd. (Of course, the possible worlds referred to in the second conjunct of the antecedent of line 3 must have the same causal laws as the world relative to which the first conjunct is analyzed—otherwise line 3 is obviously false in another way.)

So it is not news that on this strongly causal reading of “rendering proposition *p* true” line 3 is false and the argument unsound. But I would contend that there is a weakly causal reading, accepting which makes line 3 defensible against the criticisms of Taylor and Dennett, and arguably defensible tout court. So consider the weakly causal reading: *S* renders *p* true (false) iff *S* performs some action (that is, causes some event *e* to occur) which is such that, if *e* were to occur, *p* would obtain (fail to obtain). On this account, *S*'s causing some event is counterfactually (or, better, subjunctively) connected to *p*'s occurring. Note that this account is consistent with *p*'s being a simple proposition of the sort discussed above or a more complex proposition involving various events (perhaps related in complex ways), and it is completely consistent with the Taylor and Dennett point about causal necessity in the relationship between *S*'s action and its effect. Note also that his sort of account seems particularly pertinent when *p* is about some event *e** which would have occurred *prior* to *e*. When (if!) *S* has it in his power to cause some event *e* which is such that if *e* were to occur, then *e** would have occurred, then *S* is said to have “subjunctive or counterfactual power over the past” (see Saunders 1968 and Plantinga 1986). When an agent has such power, he does not have the power to initiate a backward-flowing causal chain; rather, he has the power to cause some event that is subjunctively related to a past event.

Now I do not claim that the weakly causal reading of such locutions as “*S* can render *p* true (false)” captures some ordinary, commonsense idea. It perhaps gets close to doing so but departs in certain ways from related commonsense notions. So, for example, on the weakly causal reading, I have it in my power to render it true that the sun continues to shine: I have it in my power to do something (say, type this article) which is such that if I were to do so, the sun would continue to shine. So it is evident that the weakly causal account is not intended to capture entirely some antecedently held commonsense idea. Rather, it builds on such ideas but departs slightly from them to create a theoretical notion that can then be employed in the Basic Argument. When the weakly causal reading is used, the argument as formulated by Taylor and Dennett is not obviously problematic, and certainly not problematic in the way indicated by Taylor and Dennett. That is, it is not obvious that line 3 is false.

To help to see this, consider line 3 and accept, for the sake of the argument, the Taylor and Dennett analysis of causation of the relevant sort. Now, given this analysis of causation (or the relevant kind), it

follows from the fact that A has the power to cause α that there is some sentence γ describing an action of A and a world f close to actuality such that $\gamma \wedge \alpha$ holds in f and $\alpha \Rightarrow \gamma$ in every world similar to f . Now if it is the case that $\alpha \Rightarrow \beta$ obtains in every possible world, then β will obtain in f . Thus, A can perform some act (described by γ) which is such that if he were to perform it, β would obtain. That is, A can render true β . Let's say the act performed by A corresponds to event e , and the event referred to in β is e^* . It does not follow from A 's ability to render true β , on the weakly causal reading, that A has it in his power to cause e^* . Thus, the worries invoked by Taylor and Dennett are irrelevant. (Of course, I am simply pointing out that line 3—a conditional—is not obviously problematic; this does not entail that any agent in fact has counterfactual power over the past; this is one of the major issues that divide compatibilists and incompatibilists.)

To be a bit more explicit in the diagnosis of the problem with Taylor and Dennett's critique of the Basic Argument for Incompatibilism, note that they focus on the claim that necessity (of the relevant sort) is a more important feature of causation between events than sufficiency. Of course, there are well-known and notorious problems involved with reconciling this contention with judgments that robust causation can occur in contexts of simultaneous or preemptive overdetermination. I put these worries to the side for now (although the reader will surely know that I have indicated at least a mild interest in issues pertaining to moral responsibility in contexts of preemptive overdetermination!). My point here is that the Taylor and Dennett claim is naturally taken to apply to the relationship between an agent S 's action and a specific subsequent event (its effect). One can grant the claim as regards this relationship. But the issue relevant to the Basic Argument concerns what is true when an agent S causes some sentence or statement to be true, where that sentence or statement may refer to various events (which may themselves be related in complex ways). Alternatively, the issue concerns what is true when an agent S makes it the case that some state of affairs obtains, where that state of affairs may contain various events (which may themselves be related in complex ways).

Nothing that Taylor and Dennett say about the relationship between an agent's action and its causal effect is directly relevant to the proper analysis of such locutions as " S causes it to be true that p " or " S makes it the case that p ". " S causes it to be the case that p " entails that S does something that causes some event e , which is such that if e were to occur, p would obtain. S 's action may be granted to be necessary for e to occur (for the sake of this discussion), but nothing follows about the relationship between e and the events "mentioned in" or "contained in" p .

The above discussion illustrates a point that has emerged from the literature on the Basic Argument. When some sort of "transfer principle," such as line 3, is employed in the argument, one can understand the

relevant causal locution in different ways. When it is understood in certain (strongly causal) ways, the transfer principle will be false. But when interpreted in a weakly causal way, the transfer principle will not obviously be false, and the argument is, arguably, valid. (But see Saunders 1968 and Lewis 1981.) Here the main point is that Taylor and Dennett cannot dispose of the “modal” or “transfer” version of the Basic Argument so easily; it is quite irrelevant whether or not causation between an agent’s action and its particular effect is analyzed in terms of necessity or sufficiency (or some combination of these ingredients).

Furthermore, it is very important to note that there are *other* versions of the Basic Argument that do *not* employ modal or transfer principles (such as line 3) and such locutions as “*S* can make it the case that some state of affairs obtains,” and “*S* can render some proposition *p* true (or false).” The Basic Argument employs certain very powerful and plausible intuitions, such as the view that the past is “fixed” and out of our control and, similarly, that the natural laws are out of our control (or not within our power to affect). As I mentioned above, van Inwagen himself gives two other formulations in *An Essay on Free Will*.²

I shall now present a version that does not appear to employ any premise such as line 3.³ It employs the extremely plausible and intuitively attractive Principle of the Fixity of the Past and Laws: an agent has it within his power to do *A* only if his doing *A* can be an extension of the actual past, holding the natural laws fixed. This captures Carl Ginet’s point that our freedom is the freedom to add to the given past (holding the natural laws fixed). If one accepts this fundamental idea, then one can state an argument for incompatibilism that is, so far as I can see, completely invulnerable to the sorts of worries raised by Taylor and Dennett.

I and my coauthor have presented the argument as follows:

... start with the assumption that causal determinism obtains. Suppose further ... that someone *S* does *A* at time *T3*. It follows from the truth of causal determinism that the state of the world at *T1* together with the natural laws *entails* that *S* does *A* at *T3*; that is, in all possible scenarios in which that state of the world obtains at *T1* conjoined with the actual natural laws, *S* does *A* at *T3*. But the Principle of the Fixity of the Past and Laws states that an agent can perform an act only if his performing that act could be an extension of the actual past, holding the natural laws fixed. Given the entailment just described, *S*’s refraining from doing *A* at *T3 cannot* be an extension of the actual past, holding the laws of nature fixed. Given the truth of causal determinism, one

² van Inwagen 1983, 55–105. Van Inwagen has subsequently argued that all three versions depend, at a deep level, on a modal or transfer principle: But I disagree. For my views on this issue, see Fischer and Ravizza 1992 and 1996.

³ I and my coauthor presented this version in Fischer and Ravizza 1998, 21–23. The argument is based on a version presented in Ginet 1990, 90–123; also, it is similar to a version presented in van Inwagen 1983, 55–105.

could not even in principle trace out a path along which the [actual] natural laws obtain from the actual past to *S*'s refraining from doing *A* at *T3*. Thus, *S* cannot at *T2* refrain from doing *A* at *T3* (where, as earlier, we assume that *T2* is prior to or contemporaneous with *T3*). That is to say, given the truth of causal determinism, it follows that *S* cannot do other than he actually does—he lacks the sort of control that involves alternative possibilities. (Fischer and Ravizza 1998, 22)

The above argument employs simple ingredients from common sense, and it seems to be a sound argument. It does not appear to depend in any way on causal locutions that are contentious or unclear or vulnerable to the considerations invoked by Taylor and Dennett. I would challenge Taylor and Dennett to say exactly where the problem is in the above argument. How exactly can one deny that our freedom is to add to the given past, holding the natural laws fixed? How could I actualize a state of affairs (say) with a different past from the actual past—how could I get there from here? It seems to me that the genuinely accessible pathways into the future branch off a fixed past—to deny this would be to substitute a complicated, radically unintuitive picture in which various pathways into the future come with separate pasts of their own (or different natural laws).

So, even if Taylor and Dennett are correct about their particular formulation of the Basic Argument (employing the strongly causal interpretation), it does *not* follow that the fundamental intuitive ingredients that drive the argument cannot be put together in slightly different ways to reach the incompatibilistic conclusion. One can formulate the argument employing the weakly causal interpretation of “rendering a proposition true (false)” or “making (or causing) it to be the case that *p*.” Also, one can offer various different formulations of the argument that crystallize the fundamental ideas of the fixity of the past and natural laws in different ways. These versions appear to be entirely invulnerable to the sorts of considerations adduced by Taylor and Dennett. And this should not be surprising, since outstanding and sophisticated philosophers have been troubled by the Basic Argument for thousands of years. Whereas good philosophers can certainly make mistakes, it would be surprising if the Basic Argument could be dismissed abruptly as involving a simple mistake about the analysis of causation, or a simple logical blunder.

The Basic Argument essentially conceives of “free will possibility” narrowly—as (among other things) the possibility to extend the given past, holding fixed the laws of nature. How does this narrow sort of possibility fit with the wider, more flexible sorts of possibility in which we are so keenly interested? Dennett points out that in our ordinary lives and scientific theorizing (more and less formal) we are interested in a wider sort of possibility. This is quite correct. Elsewhere I have argued that our interest in a wide notion of possibility stems from our pragmatic interests

in predicting the future, together with our epistemic limitations (Fischer 2003). We want to know what will happen in the future, or what might happen, and so forth. This will help us in our planning, and our interest in seeing to it that our lives go according to plan—and that our vital interests are secured to the greatest extent possible—is surely one of the primary reasons for theorizing about the future. If we knew the laws of nature (deterministic or indeterministic) and had available a complete description of the state of the universe in the past or present, then we could predict exactly what will happen, or we would know an objective probability distribution pertaining to what will happen. We do not, however, have such a description of the laws and the past, so we do the best we can, employing approximations to laws of nature, empirical regularities, and so forth. We employ a broad version of possibility, given our epistemic limitations and our interest in predicting the future so as to construct and pursue our life plans optimally.

But it is quite consistent with this to suppose that a given event can happen (in the sense relevant to free will or to our planning) only if that event's happening is an extension of the actual past, holding the laws of nature fixed. So we might want to know if an earthquake will destroy a certain sort of house in a certain area (say, of California). If we knew that determinism were true and we also knew both the natural laws and the complete description of the universe at the present (or at any point in the past), we could predict with certainty whether or not the house will be destroyed by an earthquake. But in fact we do not know the natural laws or whether they are deterministic; and we do not have available such a description. Thus we make predictions based on what we take to be possibilities, broadly construed. Suppose that we predict that the house will in fact be destroyed by an earthquake next year. In so predicting, we are picturing (explicitly or implicitly) the house's being destroyed in a future scenario that is an extension of the actual past (whatever that is), holding the natural laws (whatever they are) fixed. Similarly, in assigning say a 75 percent probability to the house's being destroyed next year, we are picturing (at least implicitly) a possible scenario that is an extension of the actual past, holding fixed the laws of nature, in which the house is destroyed next year. If we are not explicitly or implicitly picturing such scenarios, then we are at least committed to them in virtue of the claims in question.

So, for example, if Dennett wishes to predict what a chess-playing computer will do in the future, he may study regularities among such computers, without reference to the particular details of the past. But in coming to a conclusion about what this chess-playing computer will do next, Dennett is in fact presupposing that the computer's next move will be an extension of the actual past, holding the laws of nature fixed. Our pragmatic interests and epistemic situation force an interest in broad possibility; but that interest is completely compatible with the idea that

our freedom is the freedom to add to the given past, holding fixed the laws of nature. It is also completely compatible with the idea that what is possible, in the sense that is relevant to our planning, occurs along future paths that branch off the present, holding the past fixed.

Department of Philosophy
University of California
Riverside, CA 92521
 USA
John.Fischer@ucr.edu

References

- Dennett, D. *Freedom Evolves*. 2003. New York: Viking.
- Ekstrom, L. 1994. *Free Will: A Philosophical Study*. Boulder, Colo.: Westview.
- Fischer, J. M. 1983. "Incompatibilism." *Philosophical Studies* 43:127–37.
- . 1988. "Freedom and Miracles." *Noûs* 22:235–52.
- . 1994. *The Metaphysics of Free Will*. Oxford: Blackwell.
- . 2003. Review of Daniel Dennett, *Freedom Evolves*. *Journal of Philosophy* 100:632–37.
- Fischer, J. M., and M. Ravizza. 1992. "When the Will is Free." In Tomberlin 2001, 423–51.
- . 1996. "Free Will and the Modal Principle." *Philosophical Studies* 83:213–30.
- . 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Ginet, C. 1990. *On Action*. Cambridge: Cambridge University Press.
- Kane, R., ed. 2001. *Oxford Handbook on Free Will*. New York: Oxford University Press.
- Lewis, D. 1981. "Are We Free to Break the Laws?" *Theoria* 47:13–21.
- Plantinga, A. 1986. "On Ockham's Way Out." *Faith and Philosophy* 3:235–69.
- Saunders, J. T. 1968. "The Temptations of 'Powerlessness.'" *American Philosophical Quarterly* 5:100–08.
- Taylor, C., and D. Dennett. 2001. "Who's Afraid of Determinism?: Rethinking Causes and Possibilities." In Kane 2001, 257–77.
- Tomberlin, J. E., ed. 2001. *Philosophical Perspectives VI: Ethics*. Atascadero, Calif.: Ridgeview.
- van Inwagen, Peter. 1975. "The Incompatibility of Free Will and Determinism." *Philosophical Studies* 27:189–99.
- . 1983. *An Essay on Free Will*. Oxford: Clarendon Press.