



## I Could not have Done Otherwise--So What?

Daniel C. Dennett

*The Journal of Philosophy*, Volume 81, Issue 10, Eighty-First Annual Meeting American Philosophical Association, Eastern Division (Oct., 1984), 553-565.

Stable URL:

<http://links.jstor.org/sici?sici=0022-362X%28198410%2981%3A10%3C553%3AICNHDO%3E2.0.CO%3B2-D>

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*The Journal of Philosophy* is published by Journal of Philosophy, Inc.. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/jphil.html>.

---

*The Journal of Philosophy*  
©1984 Journal of Philosophy, Inc.

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact [jstor-info@umich.edu](mailto:jstor-info@umich.edu).

©2003 JSTOR

## I COULD NOT HAVE DONE OTHERWISE—SO WHAT?\*

WHETHER progress is stalled on a philosophical problem, a tactic worth trying is to find some shared (and hence largely unexamined) assumption and deny it. The problem of free will is such a problem, and, as Peter van Inwagen notes:

... almost all philosophers agree that a necessary condition for holding an agent responsible for an act is believing that the agent *could have* refrained from performing that act.<sup>1</sup>

Perhaps van Inwagen is right; perhaps most philosophers agree on this. If so, this shared assumption, which I will call CDO (for "could have done otherwise"), is a good candidate for denial, especially since there turns out to be so little to be said in support of it, once it is called in question. I will argue that, just like those people who are famous only for being famous, this assumption owes its traditional high regard to nothing more than its traditional high regard. It is almost never questioned. And the tradition itself, I will claim, is initially motivated by little more than inattentive extrapolation from familiar cases.

To engage the issue, I assert that it simply does not matter at all to moral responsibility whether the agent in question could have done otherwise in the circumstances. Now how does a friend of CDO set about showing that I am obviously wrong? Not by reminding me, unnecessarily, of the broad consensus in philosophy in support of the CDO principle, or by repeating it, firmly and knowingly. The inertia of a tradition is by itself scant recommendation, and if it is claimed that the assumption is not questioned because it is obvious or self-evident, I can at least ask for some supporting illustration of the self-evidence of the assumption in application to familiar cases. Can anyone give me an example of someone withholding a judgment of responsibility until he has determined (to his own satisfaction) whether the agent could have done otherwise?

It will perhaps appear that I must be extraordinarily inattentive to the topics of daily conversation if I can ask that question with a straight face. A prominent feature of many actual inquiries into the

\*To be presented in an APA symposium on Freedom and Determinism, December 30, 1984. Peter van Inwagen will comment; see this JOURNAL, this issue, 565–567.

<sup>1</sup>"The Incompatibility of Free Will and Determinism," *Philosophical Studies*, xxvii, 3 (March 1975): 185–99, p. 188, reprinted in Gary Watson, ed., *Free Will* (New York: Oxford, 1982): 46–58, p. 50.

responsibility of particular agents is the asking of the question "Could he have done otherwise?" The question is raised in trials, both civil and criminal, and much more frequently in the retrospective discussions between individuals concerning blame or excuse for particular regretted acts of omission or commission.

Before turning to a closer examination of those cases, it is worth noting that the question plays almost no role in discussions of praise or reward for felicitous, unregretted acts—except in the formulaic gracious demurrer of the one singled out for gratitude or praise: "What else could I do?" ("Anyone else would have done the same." "Shucks; 'twarn't nothin'.") And in these instances we do not take the agent to be disavowing responsibility at all, but just declaring that being responsible under those conditions was not difficult.

Perhaps one reason we do not ask "Could he have done otherwise?" when trying to assess responsibility for good deeds and triumphs is that (thanks to our generosity of spirit) we give agents the benefit of the doubt when they have done well by us, rather than delving too scrupulously into facts of ultimate authorship. Such a charitable impulse may play a role, but there are better reasons, as we shall see. And we certainly do ask the question when an act is up for censure. But when we do, we never use the familiar question to inaugurate the sort of investigation that would actually shed light on the traditional philosophical issue the question has been presumed to raise. Instead we proceed to look around for evidence of what I call a pocket of *local fatalism*: a particular circumstance in the relevant portion of the past which ensured that the agent would not have done otherwise (during the stretch of local fatalism) *no matter what he had tried, or wanted, to do*. A standard example of local fatalism is being locked in a room.<sup>2</sup>

If the agent was locked in a room (or in some other way had his will rendered impotent), then independently of the truth or falsity of determinism and no matter what sort of causation reigns within the agent's brain (or Cartesian soul, for that matter), we agree that "he could not have done otherwise." The readily determinable empirical fact that an agent was a victim of local fatalism terminates the inquiry into causation. (It does not always settle the issue of responsibility, however, as Harry Frankfurt shows<sup>3</sup>; under special

<sup>2</sup>The misuse of this standard example—e.g., in the extrapolation to the theme that, if determinism is true, the whole world's a prison—is described in my *Elbow Room: The Varieties of Free Will Worth Wanting* (Cambridge, Mass.: Bradford/MIT Press, 1984), from which portions of the present argument are drawn.

<sup>3</sup>"Alternate Possibilities and Moral Responsibility," this JOURNAL, LXV, 23 (Dec. 4, 1969): 829–833.

circumstances an agent may still be held responsible.) And if our investigation fails to uncover any evidence of such local fatalism, this also terminates the inquiry. We consider the matter settled: the agent was responsible after all; "he could have done otherwise." Proving a negative existential is not generally regarded as a problem here. We can typically count on the agent himself to draw our attention to any evidence of local fatalism he is aware of (for if he can show that "his hands were tied" this will tend to exculpate him), so his failures to come forward with any such evidence is taken to be a reliable (but not foolproof) sign that the case is closed.

The first point I wish to make is that if the friends of CDO look to everyday practice for evidence for the contention that *ordinary people* "agree that a necessary condition for holding an agent responsible for an act is believing that the agent *could have* refrained from performing that act," they in fact will find no such support. When the act in question is up for praise, people manifestly ignore the question and would seem bizarre if they didn't. And when assessing an act for blame, although people do indeed ask "Could he have done otherwise?", they show no interest in pursuing that question beyond the point where they have satisfied their curiosity about the existence or absence of local fatalism—a phenomenon that is entirely neutral between determinism or indeterminism. For instance, people never withhold judgment about responsibility until after they have consulted physicists (or metaphysicians or neuroscientists) for their opinions about the ultimate status—deterministic or indeterministic—of the neural or mental events that governed the agent's behavior. And so far as I know, no defense attorney has ever gone into court to mount a defense based on an effort to establish, by expert testimony, that the accused was determined to make the decision that led to the dreadful act, and hence could not have done otherwise, and hence ("obviously") is not to be held responsible for it.

So the CDO principle is not something "everybody knows" even if most philosophers agree on it. The principle requires supporting argument. My second point is that any such supporting argument must challenge an abundance of utterly familiar evidence suggesting that often, when we seem to be interested in the question of whether the agent could have done otherwise, it is because we wish to draw the opposite conclusion about responsibility from that which the philosophical tradition endorses.

"Here I stand," Luther said. "I can do no other." Luther claimed that he could do no other, that his conscience made it *impossible* for him to recant. He might, of course, have been wrong, or have been deliberately overstating the truth, but even if he was—perhaps

especially if he was—his declaration is testimony to the fact that we simply do not exempt someone from blame or praise for an act because we think he could do no other. Whatever Luther was doing, he was not trying to duck responsibility.

There are cases where the claim “I can do no other” is an avowal of frailty: suppose what I ought to do is get on the plane and fly to safety, but I stand rooted on the ground and confess I can do no other—because of my irrational and debilitating fear of flying. In such a case I can do no other, I claim, because my rational control faculty is impaired. This is indeed an excusing condition. But in other cases, like Luther’s, when I say I cannot do otherwise I mean that I cannot because I see so clearly what the situation is and because my rational control faculty is *not* impaired. It is too obvious what to do; reason dictates it; I would have to be mad to do otherwise, and, since I happen not to be mad, I cannot do otherwise.

I hope it is true—and think it very likely is true—that it would be impossible to induce me to torture an innocent person by offering me a thousand dollars. “Ah”—comes the objection—“but what if some evil space pirates were holding the whole world ransom, and promised not to destroy the world if only you would torture an innocent person? Would that be something you would find impossible to do?” Probably not, but so what? That is a vastly different case. If what one is interested in is whether *under the specified circumstances* I could have done otherwise, then the other case mentioned is utterly irrelevant. I claimed it would not be possible to induce me to torture someone *for a thousand dollars*. Those who hold the CDO principle dear are always insisting that we should look at whether one could have done otherwise in *exactly* the same circumstances. I claim something stronger; I claim that I could not do otherwise even in any roughly similar case. I would *never* agree to torture an innocent person for a thousand dollars. It would make no difference, I claim, what tone of voice the briber used, or whether I was tired and hungry, or whether the proposed victim was well illuminated or partially concealed in shadow. I am, I hope, immune to all such offers.

Now why would anyone’s intuitions suggest that, if I am right, then if and when I ever have occasion to refuse such an offer, my refusal would not count as a responsible act? Perhaps this is what some people think: they think that if I were right when I claimed I could not do otherwise in such cases, I would be some sort of zombie, “programmed” always to refuse thousand-dollar bribes. A genuinely free agent, they think, must be more volatile somehow. If I am to be able to listen to reason, if I am to be flexible in the right way, they think, I mustn’t be too dogmatic. Even in the most pre-

posterous cases, then, I must be able to see that "there are two sides to every question." I must be able to pause, and weigh up the pros and cons of this suggested bit of lucrative torture. But the only way I could be constituted so that I can always "see both sides"—no matter how preposterous one side is—is by being constituted so that *in any particular case* "I could have done otherwise."

That would be fallacious reasoning. Seeing both sides of the question does not require that one not be overwhelmingly persuaded, in the end, by one side. The flexibility we want a responsible agent to have is the flexibility to recognize the one-in-a-zillion case in which, thanks to that thousand dollars, not otherwise obtainable, the world can be saved (or whatever). But the general capacity to respond flexibly in such cases does not at all require that one could have done otherwise in the particular case, or in any particular case, but only that under some variations in the circumstances—the variations that matter—one would do otherwise. Philosophers have often noted, uneasily, that the difficult moral problem cases, the decisions that "might go either way", are not the only, or even the most frequent, sorts of decisions for which we hold people responsible. They have seldom taken the hint to heart, however, and asked whether the CDO principle was simply wrong.

If our responsibility really did hinge, as this major philosophical tradition insists, on the question of whether we ever could do otherwise than we in fact do *in exactly those circumstances*, we would be faced with a most peculiar problem of ignorance: it would be unlikely in the extreme, given what now seems to be the case in physics, that anyone would ever know whether anyone has ever been responsible. For today's orthodoxy is that indeterminism reigns at the subatomic level of quantum mechanics; so, in the absence of any general and accepted argument for universal determinism, it is possible for all we know that our decisions and actions truly are the magnified, macroscopic effects of quantum-level indeterminacies occurring in our brains. But it is also possible for all we know that, even though indeterminism reigns in our brains at the subatomic quantum-mechanical level, our macroscopic decisions and acts are all themselves determined; the quantum effects could just as well be self-canceling, not amplified (as if by organic Geiger counters in the neurons). And it is extremely unlikely, given the complexity of the brain at even the molecular level (a complexity for which the word 'astronomical' is a vast understatement), that we could ever develop good evidence that any particular act was such a large-scale effect of a critical subatomic indeterminacy. So if someone's responsibility for an act did hinge on whether, at the moment of decision, that decision was (already) determined by a

prior state of the world, then barring a triumphant return of universal determinism in microphysics (which would rule out all responsibility on this view), the odds are very heavy that we will never have *any* reason to believe of any particular act that it was or was not responsible. The critical difference would be utterly inscrutable from every macroscopic vantage point and practically inscrutable from the most sophisticated microphysical vantage point imaginable.

We have already seen that ordinary people, when interested in assigning responsibility, do not in fact pursue inquiries into whether their fellows could have done otherwise. Now we see a reason why they would be unwise to try: the sheer impossibility of conducting any meaningful investigation into the question—except in cases where macroscopic local fatalism is discovered. What then can people think they are doing when they ask the CDO question in particular cases? Are they really asking the philosophers' metaphysical question about whether the agent was determined to do what he did, but just giving up as soon as the investigation gets difficult and the prospects get dim of striking a lucky negative answer (with the discovery of some local fatalism)? No, for there is a better question they can have been asking all along, a question that stops for *principled reasons* with the conclusion that local fatalism is ruled out (or in). It is better for two reasons: it is usually empirically answerable, and its answer matters. For not only is the traditional metaphysical question unanswerable; its answer, even if you knew it, would be useless.

What good would it do to know, about a particular agent, that on some occasion (or on every occasion) he could have done otherwise than he did? Or that he could not have done otherwise than he did? Let us take the latter case first. Suppose you knew (because God told you, presumably) that when Jones pulled the trigger and murdered his wife at time *t*, he could *not* have done otherwise. That is, given Jones's microstate at *t* and the complete microstate of Jones's environment (including the gravitational effects of distant stars, etc.) at *t*, no other Jones-trajectory was possible than the trajectory he took. If Jones were ever put back into exactly that state again, in exactly that circumstance, he would pull the trigger again—and if he were put in that state a million times, he would pull the trigger a million times.

Now if you learned this, would you have learned anything important about Jones? Would you have learned anything about his character, for instance, or his likely behavior on merely similar occasions? No. Although people are physical objects which, like

atoms or ball bearings or bridges, obey the laws of physics, they are not only more complicated than anything else we know in the universe; they are also designed to be so sensitive to the passing show that they never can be in the same microstate twice. One doesn't even have to descend to the atomic level to establish this. People learn, and remember, and get bored, and shift their attention, and change their interests so incessantly, that it is as good as infinitely unlikely that any person is ever in the same (gross) *psychological* or *cognitive* state on two occasions. And this would be true even if we engineered the surrounding environment to be "utterly the same" on different occasions—if only because the second time around the agent would no doubt think something that went unthought the first time, like "Oh my, this all seems so utterly familiar; now what did I do last time?"

Learning (from God, again) that a particular agent was *not* thus determined to act would be learning something equally idle, from the point of view of character assessment or planning for the future. A genuinely undetermined agent is no more flexible, versatile, sensitive to nuances, or reformable than a deterministic near-duplicate would be.<sup>4</sup>

So if anyone at all is interested in the question of whether one could have done otherwise in *exactly* the same circumstances (and internal state) this will have to be a particularly pure metaphysical curiosity—that is to say, a curiosity so pure as to be utterly lacking in any ulterior motive, since the answer could not conceivably make any noticeable difference to the way the world went.

If it is unlikely that it matters whether a person could have done otherwise—when we look microscopically closely at the causation involved—what is the other question that we are (and should be) interested in when we ask "But could he have done otherwise?"? Consider a similar question that might arise about a robot, destined (by hypothesis) to live its entire life as a deterministic machine on a deterministic planet. Even though this robot is, by hypothesis, completely deterministic, it can be controlled by "heuristic" programs that invoke "random" selection—of strategies, policies, weights, or whatever—at various points. All it needs is a pseudo-random number generator, either a preselected list or table of pseudo-random numbers to consult deterministically when the occasion demands or an algorithm that generates a pseudo-random sequence of digits. Either way it can have a sort of bingo-parlor machine for providing it with a patternless and arbitrary series of digits on which to pivot some of its activities.

<sup>4</sup> This is shown in "Designing the Perfect Deliberator," in *Elbow Room*, *op. cit.*



Whatever this robot does, it could not have done otherwise, if we mean that in the strict and metaphysical sense of those words that philosophers have concentrated on. Suppose then that one fine Martian day it makes a regrettable mistake: it concocts and executes some scheme that destroys something valuable—another robot, perhaps. I am not supposing, for the moment, that *it* can regret anything, but just that its designers, back on Earth, regret what it has done and find themselves wondering a wonder that might naturally be expressed: *Could it have done otherwise?* Let us suppose that they first satisfy themselves that no obvious local fatalism (locked room, dead battery) has afflicted their robot. But still they press their question: Could it have done otherwise? They know it is a deterministic system, of course; so they know better than to ask the metaphysical question. Their question concerns the design of the robot; for in the wake of this regrettable event they may wish to redesign it slightly, to make this *sort* of event less likely in the future.<sup>5</sup> What they want to know, of course, is what information the robot was relying on, what reasoning or planning it did, and whether it did “enough” of the right sort of reasoning or planning.

Of course in one sense of ‘enough’ they know the robot did not do enough of the right sort of reasoning; if it had, it would have done the right thing. But it may be that the robot’s design in this case could not really be improved. For it may be that it was making optimal use of optimally designed heuristic procedures—but this time, unluckily, the heuristic chances it took didn’t pay off. Put the robot in a *similar* situation in the future, and, thanks to no more than the fact that its pseudo-random number generator is in a different state, it will do something different; in fact it will usually do the right thing. It is tempting to add: it *could* have done the right thing on this occasion—meaning by this that it was well enough designed, at the time, to have done the right thing (its “character” is not impugned); its failure depended on nothing but the fact that something *undesigned* (and unanticipatable) happened to intervene in the process in a way that made an unfortunate difference.

A heuristic program is not guaranteed to yield the “right” or sought-after result. Some heuristic programs are better than others; when one fails, it may be possible to diagnose the failure as assignable to some characteristic weakness in its design, but even the best are not foolproof, and when they fail, as they sometimes must,

<sup>5</sup> “We are scarcely ever interested in the performance of a communication-engineering machine for a single input. To function adequately it must give a satisfactory performance for a whole class of inputs, and this means a statistically satisfactory performance for the class of inputs which it is statistically expected to receive.” Norbert Wiener, *Cybernetics* (Cambridge, Mass: Technology Press; New York: Wiley, 1948), p. 55.

there may be no *reason* at all for the failure; as Cole Porter would say, it was just one of those things.

Such failures are not the only cases of failures that will "count" for the designers as cases where the system "could have done otherwise." If they discover that the robot's failure, on this occasion, was due to a "freak" bit of dust that somehow drifted into a place where it could disrupt the system, they may decide that this was such an unlikely event that there is no call to redesign the system to guard against its recurrence.<sup>6</sup> They will note that, in the microparticulate case (as always) their robot could not have done otherwise; moreover, if (by remotest possibility) it ever found itself in exactly the same circumstances again, it would fail again. But the designers will realize that they have no rational interest in doing anything to improve the design of the robot. It failed on the occasion, but its design is nevertheless above reproach. There is a difference between being optimally designed and being infallible.

Consider yet another sort of case. The robot has a ray gun that it fires with 99.9% accuracy. That is to say, sometimes, over long distances, it fails to hit the target it was aiming at. Whenever it misses, the engineers want to know something about the miss: was it due to some *systematic* error in the controls, some foible or flaw that will keep coming up, or was it just one of those things—one of those "acts of God" in which, in spite of an irreproachable execution of an optimally designed aiming routine, the thing just narrowly missed? There will always be such cases; the goal is to keep them to a minimum—consistent with cost-effectiveness, of course. Beyond a certain point it isn't worth caring about errors. W. V. Quine notes that engineers have a concept of more than passing philosophical interest: the concept of "don't-cares"—the cases one is rational to ignore.<sup>7</sup> When they are satisfied that a particular miss was a don't-care, they may shrug and say: "Well, it could have been a hit."

What concerns the engineers when they encounter misperformance in their robot is whether the misperformance is a telling one: does it reveal something about a pattern of systematic weakness, likely to recur, or an inappropriate and inauspicious linking between sorts of circumstances and sorts of reactions? Is this *sort* of thing apt to happen again, or was it due to the coincidental convergence of fundamentally independent factors, highly unlikely to

<sup>6</sup> Strictly speaking, the recurrence of an event of *this general type*; there is no need to guard against the recurrence of the particular event (that is logically impossible) or against the recurrence of an event of *exactly* the same type (that is nomologically impossible).

<sup>7</sup> *Word and Object* (Cambridge, Mass.: MIT Press, 1960), pp. 182, 259.

recur? So long as their robot is *not* misperforming but rather making the "right" decisions, the point in asking whether it could have done otherwise is to satisfy themselves that the felicitous behavior was not a fluke or mere coincidence but rather the outcome of good design. They hope that their robot, like Luther, will be imperturbable in its mission.

To get evidence about this they ignore the micro-details, which will never be the same again in any case, and just average over them, analyzing the robot into a finite array of *macroscopically* defined states, organized in such a way that there are links between the various degrees of freedom of the system. The question they can then ask is this: Are the links the right links for the task?

This rationale for ignoring micro-determinism (wherever it may "in principle" exist) and squinting just enough to blur such fine distinctions into probabilistically related states and regions that can be *treated as* homogeneous is clear, secure, and unproblematic in science, particularly in engineering and biology. That does not mean, of course, that this is also just the right way to think of people, when we are wondering whether they have acted responsibly. But there is a lot to be said for it.

Why do we ask "Could he have done otherwise?"? We ask it because something has happened that we wish to interpret. An act has been performed, and we wish to understand how the act came about, why it came about, and what meaning we should attach to it. That is, we want to know what conclusions to draw from it about the future. Does it tell us anything about the agent's character, for instance? Does it suggest a criticism of the agent that might, if presented properly, lead the agent to improve his ways in some regard? Can we learn from this incident that this is or is not an agent who can be trusted to behave *similarly* on *similar* occasions in the future? If one held his character constant, but changed the circumstances in minor—even major—ways, would he almost always do the same lamentable sort of thing? Was what we observed a fluke, or was it a manifestation of a robust trend—a trend that persists, or is constant, over an interestingly wide variety of conditions?<sup>8</sup>

When the agent in question is oneself, this rationale is even more plainly visible. Suppose I find I have done something dreadful. *Who cares* whether, in exactly the circumstances and state of mind I found myself, I could have done something else? I didn't do some-

<sup>8</sup> We are interested in trends and flukes in both directions (praiseworthy and regretted); if we had evidence that Luther was just kidding himself, that his apparently staunch stand was a sort of comic-opera coincidence, our sense of his moral strength would be severely diminished. "He's not so stalwart," we might say. "He could well have done otherwise."

thing else, and it's too late to undo what I did. But when I go to interpret what I did, what do I learn about myself? Ought I to practice the sort of maneuver I botched, in hopes of making it more reliable, less vulnerable to perturbation, or would that be wasted effort? Would it be a good thing, so far as I can tell, for me to try to adjust my habits of thought in such sorts of cases in the future? Knowing that I will always be somewhat at the mercy of the considerations that merely happen to occur to me as time rushes on, knowing that I cannot entirely control this process of deliberation, I may take steps to bias the likelihood of certain sorts of considerations routinely "coming to mind" in certain critical situations. For instance, I might try to cultivate the habit of counting to 10 in my mind before saying anything at all about Ronald Reagan, having learned that the deliberation time thus gained pays off handsomely in cutting down regrettable outbursts of intemperate commentary. Or I might decide that, no matter how engrossed in conversation I am, I must learn to ask myself how many glasses of wine I have had every time I see someone hovering hospitably near my glass with a bottle. This time I made a fool of myself; if the situation had been quite different, I certainly would have done otherwise; if the situation had been virtually the same, I might have done otherwise and I might not. The main thing is to see to it that I will jolly well do otherwise in (merely) similar situations in the future.

That, certainly, is the healthy attitude to take toward the regrettable parts of one's recent past. It is the self-applied version of the engineers' attitude toward the persisting weaknesses in the design of the robot. Of course, if I would rather find excuses than improve myself, I may dwell on the fact that I don't *have* to "take" responsibility for my action, since I can always imagine a more fine-grained standpoint from which my predicament looms larger than I do. (If you make yourself really small, you can externalize virtually everything.) But we wisely discourage this refuge in finer-grained visions of our embedding in the world, for much the same reason it is shunned by the engineers: what we might learn from such an investigation is never of any consequence. It simply does not matter whether one could have done otherwise.

It does not matter for the robot, someone may retort, because a robot could not *deserve* punishment or blame for its moments of malfeasance. For us it matters because we are candidates for blame and punishment, not mere redesign. You can't *blame* someone for something he did, if he could not have done otherwise. This, however, is just a reassertion of the CDO principle, not a new consideration, and I am denying that principle from the outset. Why indeed shouldn't you blame someone for doing something he could not

have refrained from doing? After all, if he did it, what difference does it make that he was determined to do it?

"The difference is that if he was determined to do it, then he *had no chance not to do it*." But this is simply a *non sequitur*, unless one espouses an extremely superstitious view of what a chance is. Compare the following two lotteries for fairness. In Lottery A, after all the tickets are sold, their stubs are placed in a suitable mixer, and, after suitable mixing (involving some genuinely—quantum-mechanically—random mixing if you like), the winning ticket is blindly drawn. In Lottery B, this mixing and drawing takes place *before* the tickets are sold, but otherwise the lotteries are conducted the same. Many people think the second lottery is unfair. It is unfair, they think, because the winning ticket is determined before people even buy their tickets; one of those tickets is *already* the winner; the other tickets are so much worthless paper, and selling them to unsuspecting people is a sort of fraud. But in fact, of course, the two lotteries are equally fair: *everyone has a chance of winning*. The timing of the selection of the winner is an utterly inessential feature. The reason the drawing in a lottery is typically postponed until after the sale of the tickets is to provide the public with first-hand eyewitness evidence that there have been no shenanigans. No sneaky agent with inside knowledge has manipulated the distribution of the tickets, because the knowledge of the winning ticket did not (and could not) exist in any agent until after the tickets were sold.

It is interesting that not all lotteries follow this practice. Publishers' Clearinghouse and Reader's Digest mail out millions of envelopes each year that say in bold letters on them "YOU MAY ALREADY HAVE WON"—a million dollars or some other prize. Surely these expensive campaigns are based on market research that shows that in general people do think lotteries with pre-selected winners are fair so long as they are honestly conducted. But perhaps people go along with these lotteries uncomplainingly because they get their tickets for free. Would many people *buy* a ticket in a lottery in which the winning stub, sealed in a special envelope, was known to be deposited in a bank vault from the outset? I suspect that most ordinary people would be untroubled by such an arrangement, and would consider themselves to have a real opportunity to win. I suspect, that is, that most ordinary people are less superstitious than those philosophers (going back to Democritus and Lucretius) who have convinced themselves that, without a continual supply of genuinely random *cruces* to break up the fabric of causation, there cannot be any real opportunities or chances.

If our world is determined, then we have pseudo-random number

generators in us, not Geiger counter randomizers. That is to say, if our world is determined, all our lottery tickets were drawn at once, eons ago, put in an envelope for us, and doled out as we needed them through life. "But that isn't fair!" some say, "For some people will have been dealt more winners than others." Indeed, on any particular deal, some people have more high cards than others, but one should remember that the luck averages out. "But if all the drawings take place before we are born, some people are *destined* to get more luck than others!" But that will be true even if the drawings are held not before we are born, but periodically, on demand, throughout our lives.

Once again, it makes no difference—this time to fairness and, hence, to the question of desert—whether an agent's decision has been determined for eons (via a fateful lottery ticket lodged in his brain's decision-box, waiting to be used), or was indeterministically fixed by something like a quantum effect at, or just before, the moment of ultimate decision.

It is open to friends of the CDO principle to attempt to provide other grounds for allegiance to the principle, but since at this time I see nothing supporting that allegiance but the habit of allegiance itself, I am constrained to conclude that the principle should be dismissed as nothing better than a long-lived philosophical illusion. I may be wrong to conclude this, of course, but under the circumstances I cannot do otherwise.

DANIEL C. DENNETT

Tufts University

### DENNETT ON 'COULD HAVE DONE OTHERWISE'\*

DANIEL DENNETT attacks what he describes as a "shared assumption" of writers on free will:

*A is responsible for having done X only if A could have refrained from doing X.*

If this sentence is to express a thesis that has been widely accepted, then 'could have' must be read as the past *indicative* of 'can', where

\*Abstract of a paper to be presented in an APA symposium on Freedom and Determinism, December 30, 1984, commenting on Daniel C. Dennett, "I Could Not Have Done Otherwise—So What?," this JOURNAL, this issue, 553-565.