

FROM:

**Philosophical Problems
of the Internal and External Worlds**

**Essays on the Philosophy
of Adolf Grünbaum**

EDITED BY John Earman, Allen I. Janis,
Gerald J. Massey, and
Nicholas Rescher

University of Pittsburgh Press / Universitätsverlag Konstanz

1993

21

Indeterminism and the Freedom of the Will

Arthur Fine

Department of Philosophy, Northwestern University

I formed them free, . . . they themselves ordain'd thir fall.
—Milton, *Paradise Lost*, Book III

Adolf Grünbaum's writings on the free will problem, although not extensive, have been widely reproduced and influential (see Grünbaum 1953, 1967, and 1972). Characteristically, at the center of Grünbaum's work are clear and forceful arguments—in this instance, for the compatibility of determinism and free will. This compatibilism, I suspect, derives in good measure from Grünbaum's passionate concern to protect the possibility of an adequate human science, which is to say (as he sees it) a causally based science of individual and social behavior. Insofar as libertarian incompatibilism seems to draw boundaries around causal analysis, exempting human action from its reach,¹ incompatibilism may seem to stand in the way of true human science. Grünbaum has been concerned with opening the path. In the course of his writings, at least to a limited extent, Grünbaum considers the bearing of indeterminism and (inevitably) of the quantum theory on the free will issue. Those limited considerations are my starting point here.

I will lay out the first half of a two-part argument intended to deconstruct the metaphysical concept of human freedom. In the first half, which constitutes this essay, I show how the libertarian conception of human freedom self-destructs. In the second half I would take up the antilibertarian (or compatibilist) conception with the same end in view (see also Earman's 1986 discussion of the difficulties with free will). I will not pursue the second half here, however, be-

cause in fact I am not yet sure just how the argument goes. It may be useful, nevertheless, to sketch the picture that motivates the whole project. In the usual picture there is some antecedent concept of human freedom and agency, which our moral, legal or social concepts of responsibility track and which they are bound to respect. In my view this has the tail wagging the dog, for it seems to me that things are just the other way around. It is we who manufacture conceptions of responsibility in order to meet the changing conditions of our social lives. We may then try to construct metaphysical concepts of freedom and agency in order to ground our attributions of responsibility. This foundational enterprise, however, (like others) is an unproductive fancy. From the fact that our metaphysical constructs are incoherent, we can see that our conceptions of responsibility actually require no such grounding. This, at any rate, is the important moral I would draw from the argument—if only I had both parts in place!

Freedom and Indeterministic Laws

I begin with an argument that turns on considerations concerning indeterministic laws; that is, laws whose statistical character, as Grünbaum puts it, “is not removeable by the possession of more complete information” (1972, 306).² He suggests the statistical laws of the quantum theory as a paradigm. Grünbaum argues that such laws are of no comfort to the libertarian. He argues, that is, that if determinism were truly incompatible with free will, as the libertarian conceives it, then indeterminism would be incompatible as well. Grünbaum draws the libertarian conception of human freedom from C. A. Campbell, who holds that one’s act is free only if one could have acted differently under the very same circumstances. In line with a long compatibilist tradition, Grünbaum rejects this conception of the freedom of the will as inadequate. However, he argues for something much stronger than its inadequacy; namely, Grünbaum argues that we would not have this libertarian kind of freedom even if human behavior were governed by indeterministic laws.

Why not? The argument proceeds as follows (Grünbaum 1972). Suppose a community is subject to an indeterministic law according to which, in the long run, 80 percent of the population will commit a certain kind of crime. Can we hold those members of the community who commit the crime morally responsible for their behavior?

Only if, as the libertarian standard would have it, they could have done otherwise—that is, refrained from committing the crime. But, argues Grünbaum, the statistical law does not entitle us to say of an individual *A* who commits the crime that *A* could have done otherwise. To be sure, on the basis of the law we cannot tell which particular individuals in the community will commit the crime. The law does not specify that. But this limitation (epistemic or otherwise) does not entail that, in the very circumstances in which an individual *A* does commit the crime, *A* could have refrained from so doing.³ This is the consequence that the libertarian needs, and thus Grünbaum concludes that irreducibly statistical laws would not ground human freedom in the special libertarian sense.

If we deal with a statistical law, then—as in the quantum case—we can restate the statistical content in purely probabilistic terms. So, for the example at hand, we can say that each member of the community has a probability of 0.80 to commit the crime. This probabilistic statement can be given what it seems appropriate to call a libertarian interpretation or an antilibertarian one. According to the *libertarian interpretation*, probability 0.80 to commit the crime for an individual *A* entails that in the very same circumstances in which *A* does in fact commit the crime *A* might not have done so. In this sense, *A*’s act is undetermined (or uncaused). According to the *antilibertarian interpretation*, probability 0.80 to commit the crime for an individual *A* entails that if the circumstance were exactly the same, then *A* would do the same again. In the antilibertarian sense *A*’s act, although governed by probabilities, is not fully undetermined. Insofar as the circumstances would entail a repetition of the act, causes seem to operate, at least partially.

These distinctions, in terms of what would or might happen in *exactly* the same circumstances, may seem to be idle and perhaps there is a hint of this attitude in Grünbaum’s evident irritation with the libertarian conception of freedom. But, as we have come to learn in connection with foundational studies in the quantum theory, just such counterfactual distinctions may turn out to have unexpected and testable consequences. I think this is precisely the case here. For if we subject the statistical laws of the quantum theory to the antilibertarian interpretation then, given other reasonable assumptions, I think we have all the machinery in place to derive the Bell inequalities, which is to say to contradict the quantum statistics. It follows that,

given certain reasonable assumptions, the indeterministic laws of the quantum theory cannot be subjected to an antilibertarian interpretation. They call for a libertarian one. If indeed quantum theory is the paradigm for indeterministic laws more generally, then the centerpiece in the argument against the libertarian fails, for an indeterministic law would entail the libertarian conclusion that A could have done otherwise. Thus it seems that one could indeed appeal to indeterministic laws to ground the libertarian conception of human freedom. My agenda will be to carry the discussion up to this last point, and then to see whether an appeal to indeterministic laws really does help the libertarian cause. (To anticipate, the answer is no.)

Libertarianism and the Quantum Theory

I turn now to the conflict between antilibertarianism and the quantum theory (for more of the quantum details see Cushing and McMullin 1989 and Redhead 1987). Consider a typical EPR-type experiment that involves the measurement of four variables in two spatially separate locations: two variables in the "A-wing" of the experimental apparatus (say, A and \bar{A}) and two in a distant "B-wing" (say, B and \bar{B}). The reasonable assumption that we need is that the experiment involves no action-at-a-distance. Specifically, we assume a locality principle according to which the circumstances affecting a measurement outcome in one wing of the experiment do not depend on which variable is being measured in the other wing. To confront antilibertarianism consider two experimental runs each consisting of n pairs of measurements. In the first ("AB") run we measure A in one wing simultaneously with (or, at spacelike separation from) B in the other wing. For concreteness, let α be the sequence of outcomes of the n A -measurements and β the sequence of outcomes of the n measurements of B . In the second ("AB") run we measure \bar{A} with \bar{B} and obtain outcome sequences $\bar{\alpha}$ and $\bar{\beta}$. Suppose, hypothetically, that in the first run we had measured A with \bar{B} instead of with B . According to the locality assumption, in this hypothetical $A\bar{B}$ run, the circumstances affecting the measurement outcome in the A-wing would have been exactly as they were in the actual AB run. Hence in this hypothetical case the antilibertarian interpretation of the quantum statistical laws entails that the A -outcomes would be just as before, namely, the sequence α . Since the circumstances affecting the measurement outcome in the B -wing do not depend on which variable is

measured in the A -wing, the results of the \bar{B} measurement in the hypothetical $A\bar{B}$ run might have been the same as they were in the actual AB run; that is, β . Assuming that no special bias attaches to an outcome sequence due to the fact that one variable rather than another is measured in the opposite wing, the statistics of the experiment for the AB and $A\bar{B}$ runs (that is, the probability distributions for the variables A , B and \bar{B} and the correlations or joint distributions for the pairs $\langle A, B \rangle$ and $\langle A, \bar{B} \rangle$) can be computed from the three outcome sequences α , β and $\bar{\beta}$. Suppose, again hypothetically, that in the second ($A\bar{B}$) run we had measured \bar{A} with B instead of with \bar{B} . Then, just as before, locality combined with antilibertarianism entails that in this hypothetical $A\bar{B}$ run the \bar{A} -outcomes would be just as in the actual $A\bar{B}$ run, namely, $\bar{\alpha}$. Again, as in the first hypothetical case, the results of the B measurement in the $A\bar{B}$ run might have been the same as they were in the actual AB run, that is, β . Once again, assuming our sample of outcomes is fair, we conclude that statistics of the experiment, this time for the $A\bar{B}$ and AB runs, can be computed from three outcome sequences, namely, from $\bar{\alpha}$, β and $\bar{\beta}$. Thus, taken together, the four outcome sequences that occur in our two actual runs carry statistics for all of the four possible runs of the experiment, that is, the four single distributions and the four A -wing, B -wing pairs. However, it is well known (e.g., Fine 1982) that the single and joint distributions that can be computed from four fixed outcome sequences satisfy the Bell inequalities, which the quantum statistics violate for certain experimental configurations. Hence the antilibertarian reading of the statistical laws of the quantum theory, together with the principle of locality, conflicts with the quantum statistics. Assuming that the statistical laws of the quantum theory are correct, we can conclude that the antilibertarian reading of those laws entails action-at-a-distance; that is, it entails that, at least in certain experiments, measuring one variable rather than another would immediately alter the circumstances affecting some measurement outcome in a distant region of space. If we assume that the quantum theory is correct in its statistical predictions and we hold to the reasonable no action-at-a-distance condition involved in the stated locality principle, then it follows that the statistical laws of the quantum theory cannot be given an antilibertarian interpretation.

The quantum theory, we have argued, requires a libertarian reading of its probabilistic assertions, on pain of action-at-a-distance. In view of that argument, it seems reasonable to conclude that when the

quantum theory says that an individual *A* has an 80 percent chance of doing something (e.g., a radioactive atom of decaying within the hour) and then *A* does it, other things being equal, *A* could have done otherwise. Thus the antilibertarian position finds little room to breathe in a statistical world if we take laws of the quantum theory as the exemplars of the statistical laws in such a world. So, it appears that, contrary to what Grünbaum claims, the libertarian's "could have done otherwise" does indeed find support from indeterminism if we take the indeterministic laws to be of the sort found in the quantum theory. It remains to be seen, however, whether such an indeterminism provides refuge for the libertarian position on the freedom of the will from a more general perspective, for the conclusion that Grünbaum draws may turn out to be more robust than the particular argument he gives for it.

Freedom and Indeterminism

The conclusion that Grünbaum set us after is a conditional one: If free will is incompatible with determinism, then it is also incompatible with indeterminism. I will call this *Grünbaum's conditional*. It is a strong conditional assertion and a very important one since, if it holds, then there is no free will in the libertarian sense. For, assuming that determinism and indeterminism are jointly exhaustive, if the will were free, both would fail.⁴ Hence, the assumption that the will is free leads to a contradiction if Grünbaum's conditional holds.

The suggestion of an incompatibility between free will and indeterminism *tout court*, however, which is the consequent of that conditional, has a respectable history of its own. In the *Treatise* ([1739] 1902, book 2, part 3, chaps. 1 and 2), Hume questioned whether indeterminism (or "chance") provided a sound basis for the idea of human agency. In Hume ([1777] 1902), he summarized it this way, "Actions are by their very nature, temporary and perishing; and where they proceed not from some *cause* in the character and disposition of the person who performed them, they can neither redound to his honour, if good; nor infamy, if evil. . . . According to the principle, therefore, which denies necessity, and consequently causes, a man is as pure and untainted after having committed the most horrid crime, as at the first moment of his birth" (sec. 8, part 2, sec. 75). As I understand it, the line of thought goes something like this.

If after due deliberation and under circumstances free of any coercion (in the ordinary sense) my choice to do a certain deed, and my doing of it, are at best under the rule of a probabilistic law, then in what sense can I be said to be the author of my deed, and responsible for it? To say that I determined my action would seem to require that under the same circumstances (via my choice, perhaps) *I* made the difference between doing or not doing. But this would only be the case if, under the *very* same circumstances, things were bound to come out the same way. If the determining laws, however, are statistical and if we give them a libertarian reading, then even were circumstances exactly the same (including everything that pertains to me) the outcome might have been different; that is; I might not have done the deed. The libertarian reading leaves room for chance and change. According to that kind of indeterminism, the kind of indeterminism the quantum theory requires, literally nothing determines the outcome. But if nothing determines the outcome, literally, then *I* do not determine the outcome. A freedom of the will that can support attributions of responsibility (that can "redound to [one's] honour, if good; . . . infamy, if evil"), therefore, seems incompatible with indeterminism, under the libertarian conception.⁵

Like the arguments for the incompatibility of free will and determinism, this incompatibilist line too contains a number of moves that can be questioned. Moreover, one might think that the stochastic conception of causality can rescue the idea of agency in an indeterminist setting. One might think, that is, that although I do not determine the outcome of my behavior in the sense of strict causality (no probability or chance), still in the stochastic sense of "cause" I am a causal factor (maybe even the most significant causal factor) in my behaving as I do, and that this is enough for agency. The relevant sense of causal factor here is, roughly, this (see Cartwright 1989 and Humphreys 1989 for the problems with this rough account of stochastic causality, and for different proposals for how to polish and trim it). To say that I cause *B* to happen (or that my willing it does) is to say that my willing *B* makes *B* more likely to happen than would be the case if I did not will it. *I* make a difference. To be sure I do, but concerning what do I make the difference? Unfortunately for the application to agency, it is not with regard to *B*'s happening or not happening. On the libertarian reading, *nothing* makes a difference concerning that. What I (or my will) influences are the odds that *B*

happens. My will raises the odds. Is this enough for agency with regard to the act itself? The basic rule of thumb for causal efficacy in the context of human agency (for example in criminal law and the law of torts)⁶ is the “but for” test: But for my willing it *B* would not have happened. Of course this is the test whose failure is built into stochastic cases of the kind we are supposing here, since in *no* stochastic case is it true that *B* would not have happened without my willing it. In every case, *B* might have happened anyway. Thus the customary underpinning for agency has no grounding when chance enters in. We can see the difficulty clearly by considering stochastic cases where I will the act (thus upping the odds) but nevertheless the act does not happen. In these cases I may get credit for trying, but I cannot be blamed when the act does not transpire. That result, we might say, is simply the luck of the draw. But then, we need to ask, why are we supposed to be accountable for the act if, in identical circumstances, it does transpire? That is, in the usual case, where is agency supposed to get its grip? Since I play the same part whether the act occurs or not, I do not see how I may be counted out when the act does not occur, but held to account for the act when it does.

The conditions for agency do not fit the stochastic conception of causality. That conception is an attempt to extend the idea of a cause in circumstances where some of the usual concomitants of causality are absent. It shifts the effects of a cause from outcomes to the odds (or probability) for outcomes. This may be a perfectly good extension, which is to say one that is useful in some circumstances for certain purposes. We should not presume its universal applicability, however. Extending concepts is a context-sensitive business. Like making puns, whether it succeeds depends on whether the particular circumstances are just right. When it comes to the circumstances of human agency in an indeterminist setting, the stochastic move from outcomes to odds does not work well enough to rescue the freedom of the will.

The line that traces an incompatibility between indeterminism and free will has been attacked directly. Daniel Dennett, for example, advertises himself as having refuted it, “It has often been claimed that responsibility and indeterminism are incompatible. The argument typically offered is fallacious, as I show in. . . .” (1984, 142, n. 8) Dennett (1978, chap. 15) attempts to show that random elements in the decision process down the line from the act itself can still leave

room for the efficacy of judgement and choice. Following this argument Dennett pleads. “The libertarian could not have wanted to place the indeterminism at the end of the agent’s assessment and deliberation. . . . It would be insane to hope that after all rational deliberation had terminated with an assessment of the best available course of action, indeterminism would then intervene to flip the coin before action” (ibid., 298). Dennett goes on to urge the standard compatibilist gloss on “could have done otherwise,” a gloss he could find in Grünbaum, namely, that the circumstances in which the agent could have done otherwise are *not* exactly the same—in particular, that they do not involve the same beliefs and desires with which rational deliberation ended. If the incompatibility of agency (or responsibility) and indeterminism (in the libertarian sense) rests on a fallacy, Dennett has certainly not displayed it. Instead he concedes the incompatibility and calls it “insane,” hoping, no doubt, to draw our attention away from his failure to come to terms with it.

What Dennett might claim to show (at least this is where he engages in argument rather than invective) is that if the indeterminism is limited to a certain place in the deliberative process that leads to action, then the agent might still intelligibly be thought of as the author of the act. (He uses the term “authentic.”) Dennett thinks that indeterminism might come into play, somehow, in generating considerations which rational choice then converts deterministically into action. This separatist scheme is similar to ideas that Grünbaum (and others) have entertained with respect to the impact of quantum indeterminacy on human freedom. The common idea is to keep the indeterminacy suitably confined and to hope that deterministic principles, sufficient to support a decent sense of agency, function appropriately somewhere at the level of molar behavior and choice. It seems to me, however, that this form of separatism is not stable unless the indeterminacy is taken out of the intentional stream entirely. That move, however, involves rejecting the attractive Leibnizian idea that motives (or whatever) incline but do not necessitate. It would take us too far afield to explore the cogency of that rejection. Let me just show that Dennett’s idea of placing the indeterminism somewhere downline but still within the intentional stream does not work.

Suppose we grant, with Dennett, that if the coin flips between the end of deliberation and the consequent act, then indeed agency is called into question. So we grant that the action is not authentically

mine if, having determined a rational course of action, it is a then not up to me but rather a matter of chance whether the act transpires. To be *my* act, the relation between the act's happening and what we may call its intentional determinant has to be nonprobabilistic, or so we will suppose. Well, what if the coin flips a little sooner; that is, what if chance enters between the considerations that form the background of my choice and the rational decision process itself? Then, surely, application of the same principle of ownership requires that the determination of the course of action is, likewise, not authentically mine. For how can background consideration be said to form a basis for my decision if it is not up to me but rather a matter of chance how I proceed to deliberate, given those considerations? The fact that Dennett does not consider this possibility at all, but simply assumes that the background considerations go hand in glove with a decision process, which in turn fixes the action, seems to corroborate the analysis. To be my deliberations, their determinants have to be nonprobabilistic as well, or so Dennett seems to assume.

Dennett would flip the coin just before the considerations themselves, where he thinks it is harmless. But suppose I am a lobbyist for *Birds First!* (a radical environmental organization) and have been offered a job by *Pollution, Inc.* (a large and notoriously irresponsible oil company). In Dennett's sense, considerations on which I base my decision about accepting the job might include the difference in salaries, the quality of the support staff, the workload, the location of the jobs, the amount of travel involved, and so on. Some or all of these might occur to me and let us say, with Dennett, that whether they (or other considerations) do occur involves an element of chance. In the end I do not determine which considerations occur to me and which do not. Nothing fixes that. However, I am a sincere person, committed to the welfare of birds and the environment, and generally scrupulous in keeping that commitment. I would never wittingly put my talent to work for *Pollution, Inc.*, not, that is, if it occurred to me to consider these commitments of mine to birds and the environment, or if I considered the self-image problem that would be engendered by earning my livelihood out of the ill-gained profits of *Pollution, Inc.*—and so on. But suppose that despite my most sincere efforts in making a good decision, these considerations just never occur to me. Chance intervenes to flip the coin only on the first list above. Pursuing a rational decision policy, I wind up accepting the offer. Now,

given the role of chance in that result, I do not see why I am any more the captain of my fate in these circumstances than I would be had chance intervened later down the road, say, to mediate the execution of a more balanced decision. (Indeed, would it be different had the environmental considerations occurred to me by chance and, because the coin landed the other way around, I made a different decision?) For an act to be authentically mine, we have supposed that chance must not divide it from its intentional antecedents. On that basis, it would seem that the coin may not flip to separate me and my characteristic concerns from the considerations on which my choice is based.

I hope that one can see in this the outline of a good inductive argument. It starts with Dennett's concession that agency is compromised if chance enters between the end of deliberation and the act. The argument proceeds via the inductive rule that agency excludes chance between any act and its intentional predecessors. We conclude that chance may not enter at all in any chain of intentional antecedents of the act, if the act is to be mine. This line supports Hume's intuition that responsibility is rooted in connections between character and behavior that leave no room for chance. Nothing in the separatist strategy that we have been looking at suggests that Hume was mistaken, for it appears that regardless of where in the intentional stream you put it, one chancy apple may spoil the whole pile. To get around Hume, one would have to question whether necessity is really required for agency at all. I propose to avoid that issue and return us instead to the conditional incompatibilism with which we started.

Conditional Incompatibility

Is indeterminism of any help to the libertarian? A negative answer follows from Grünbaum's conditional, that if free will is incompatible with determinism, then it is also incompatible with indeterminism. Insofar as the results of the preceding section support the consequent here they support the conditional itself. In this section I will try to provide even stronger support by arguing that if there were an adequate indeterminist account of an agent's behavior there would be an equally adequate determinist account. It follows that if indeed determinism is not compatible with an agent's free will then neither is indeterminism, which is the conditional we are after.

The argument derives from the trivial observation that if the probability of an event is 8/10, say, then one can think of that as involving ten possible cases in eight of which the event occurs. The argument simply gives a formal expansion of that idea, uniformly, for a class of probabilistic assertions. Suppose then that we consider an agent A . I will piggyback on the state-observable framework of quantum theory to suppose that we can talk meaningfully about “the state” of our agent at a given time, which I will denote by σ (suppressing a temporal index unless required). We suppose that an indeterminist theory of A 's behavior treats possible acts B from a well-defined class and yields a set of probabilistic laws of the form

$$\text{Prob}_A(B \mid \sigma) = p$$

(“The probability that agent A does act B in state σ is p .”). Indeterminism requires that the probability p be different from 0 or 1 for some acts B . Were the probabilities for all acts either 0 or 1, then (relative to the state) nothing would be left to chance and the theory would be deterministic. If the state includes enough relevant information about the intentional situation of the agent just prior to the act (beliefs, desires, or whatever) the indeterminist theory is precisely of the “insane” kind that (*pace* Dennett) interests the libertarian, and which was the starting point for the considerations in the previous section.

We want to show that if an agent's states and behavior are governed by an indeterminist theory, they are also governed by a deterministic one. It is important that the deterministic theory we will produce treat the same states and behavior as the indeterminist one. Otherwise, one might suppose that the determinism depends on a trick of redescription, for example, shifting from an intentional idiom (waves goodbye) to a nonintentional one (arm moves 27 centimeters vertically). To be sure, what is undetermined relative to one sort of description might be determined relative to another, but that has nothing to do with the point here at issue. My trick is different. It connects, rather, with the concept of indeterminism already introduced, that is, with whether the statistics are “removeable by the possession of more complete information.” When such a reduction is possible one obtains a deterministic theory, as previously defined. We now show that, in principle, it is always possible to reduce the statistics of an indeterminist theory.

Suppose the state of the agent A is σ . To achieve the reduction we introduce a set of indices, which we take just to be the numbers between 0 and 1 (inclusive). These indices correspond to the “ten possible cases” mentioned, and there are several interpretations that one might give to them. Here I will treat the indices as marking different possible ways of being in a given state, and assume that each agent has just one way. Keeping track of the ways provides the additional information needed to get rid of the statistics. For, corresponding to each act B , we can use the indeterminist theory to introduce a function $B_A(\cdot)$ taking only 0 or 1 as values and defined as follows for any index x between 0 and 1:

$$B_A(x) = 1 \text{ if } 0 \leq x \leq \text{Prob}_A(B) \text{;} \\ \text{otherwise, } B_A(x) = 0.$$

We understand $B_A(x) = 1$ (or 0) just in case it is true (or false) that if A were in state σ in the way marked by x , A would do B . Thus our new theory supplements the indeterminist one by adding the indices and the “truth” functions defined with respect to them as just given.

The new theory involves a new class of probabilistic ascriptions of the form

$$\text{Prob}_A(B \mid \sigma, x)$$

where x is an index. This is the probability that A does B , given that A is in state σ in the way marked by index x . The assumption that every agent is in a state in just one way implies that these probabilities are either 1 or 0, depending (respectively) on whether it is true or false that the way A is in σ yields that A does B . In fact

$$\text{Prob}_A(B \mid \sigma, x) = B_A(x)$$

and, since $B_A(x) = 0$ or 1, the new theory is deterministic. The probabilities from the indeterministic theory, however, all follow from the deterministic one as averages. That is,

$$\text{Prob}_A(B \mid \sigma) = \text{AVERAGE} [\text{Prob}_A(B \mid \sigma, x)] = \text{AVERAGE} [B_A(x)],$$

where the average is taken over all x between 0 and 1. This completes the reduction.

The indeterminist theory distinguishes one agent from another only by their states, treating agents generically as randomly selected from all the ways of being in a state. The deterministic theory individuates in a more precise way, one that reduces the posited statistics. That reduction eliminates chance. No coins flip. Hence, if there is an indeterministic theory of an agent's behavior, there is a deterministic one too. Therefore, if determinism is incompatible with free will, so is indeterminism. This result gives strong support to Grünbaum's conclusion that indeterminism is of no help to the libertarian. Before we rest content with that conclusion, however, we need to study the possibility for a determinist reduction in a little more detail.

Reducibility

In claiming that indeterminism implies determinism, we show that the statistics of an indeterminist theory are always reducible to a determinist base. But we have already argued that the statistics of the quantum theory require a libertarian reading that would prevent their deterministic reduction. Attention to two different factors will help us see how to reconcile these claims. First of all, in the quantum case we deal with a more complicated set of probabilistic assertions than we have just considered. Those all concerned the probability for a single agent to do this, or that. In the quantum theory we deal with joint probabilities: the probability that *A* does *B* and *C* does *D* (to put it in agent terms). So the quantum theory gives us a family of joint (actually multiple) probability distributions, whereas before I deal only with single distributions. It turns out that difficulties with reducibility in the quantum theory (technically, with the introduction of deterministic hidden variables) always relate to the special character of the family of quantum joint distributions (see Fine 1982). The Bell inequalities, which we previously used to block reducibility, are a case in point. They are inequalities constraining joint distributions which the twin requirements of antilibertarianism and locality imply cannot be matched by the quantum joints. This points to the second issue, which is that of external conditions. To get the irreducibility of the quantum statistics we need to go beyond the formal requirements of the quantum theory itself and assume locality. (Other no-go results in the foundational literature; e.g., the Kochen-Specker theorem, or the Heywood-Redhead result, also require external assumptions, al-

though in these cases the assumptions are less well motivated and plausible than locality, at least from a physical point of view—see Redhead 1987 and Elby 1990). Nonlocal reductions of the quantum statistics along the lines sketched in the general argument just given are perfectly possible, indeed trivial from a formal point of view. (Somewhat less trivial is the nonlocal de Broglie-Bohm “pilot wave” theory that also reduces the quantum statistics and offers a challenging deterministic alternative to the usual understanding of the quantum theory—Bell 1987 contains the relevant details.) By taking account of the joint distributions and the external constraints, we can reconcile the claims of irreducibility in the quantum case with the claim of reducibility for agent indeterminism.

Taking these factors into account, however, seems to point to serious shortcomings in the argument of the preceding section, for the claim that agent indeterminism leads to agent determinism appears to be misleading. That result is purely formal and, as we learn by comparison with the situation in the quantum theory, it seems to depend on attending only to a restricted class of statistical theories (namely, ones without joint probabilities) and on ignoring plausible external constraints on the proposed scheme of reduction—or so one might object. Despite the lesson of the quantum theory, however, I am not so sure that these objections are well founded.

The problem concerning joint distributions can be put this way. How does one deal with correlations between the behavior of different agents? I think this question has two good answers. The first is simply to deny that we have to deal with them at all; that is, that we have to consider the behavior of more than one agent at a time. After all, we are not speculating about the reducibility of some general science of human behavior. We are only addressing the theory of agency for a single individual, any *one* of us. At any rate, that is the traditional philosophical setting for the discussion of free will, and it is hard to see how possible difficulties in reducibility that might arise in theories of group behavior would affect that discussion. Even if one is inclined to a social theory of the mind I would have thought that, over the freedom of the will, a person is not to be treated as part of a herd. So the first good answer is to assert that, in the absence of a specific limitation on reducibility that is derived from theories of group behavior, our single distribution framework is adequate (at least *prima facie*). The second good response to the question of how

to treat correlations between agents is to offer the conventional wisdom, which is this. If the correlations are not spurious then either they derive from a direct causal connection between the agents (perhaps a chain of causes) or they can be explained in terms of common background causes that affect the behavior of both agents. Let us look at these alternatives in turn.

Direct causal connections can give rise to the full range of possible correlations in the behavior of agents. To see that, just consider a class of students taking an essay exam. If they are in communication with one another (the causal connection) then they can arrange for any correlation at all, from all their answers being exactly the same (word for word) on every question, to none of them being the same on any question. However, whatever they arrange can be explained deterministically, if it can be explained stochastically, for a stochastic explanation must yield a probability distribution for all the possible answers the class might give to the examination questions. Any such (multiple) distribution, however, can be obtained by averaging over a large number of determinate answers, just as in our given simple proof of reduction. The field of determinate answers, arranged as how the students would have answered (if), provides the deterministic reduction that we were looking for. Thus direct causal connections do not produce any challenge to reduction that goes beyond what we have already treated. What then of common causes?

To explain correlations by means of common causes is to derive the correlations by averaging over the effects of background causes that uncouple the correlated variables. To take a quantitative example, the crime rate in Chicago is inversely correlated with the attendance rate at the movies during mid-week for theaters located in suburban shopping centers. The more suburbanites who attend local movies during the week, the lower the crime rate in the city. This is not because all the criminals live in the suburbs. Nor is there a direct chain of causes linking inner city crime and suburban movie attendance. (No one is planning a mass suburban movie-in to stop crime in the city.) But the conditions that make for less crime in the city (for example, high employment and bad weather) also make for movie-going in the suburbs. High city employment means business is good and that contributes to the size and affluence of the families of suburban professionals and businesspeople (which, in turn, contributes to their mobility and restlessness), and (within limits) bad weather is

good for the movie business. However, for fixed levels of employment and weather conditions, there is no significant correlation between urban crime and suburban movie-going. It is only by averaging over various levels that the correlation appears. If we adapt this example to the behavior of agents, then we would explain correlations in behavior by averaging over independent factors, that is, factors that contribute to the behavior of each agent independently of each other. These factors can be stochastic in the sense discussed in an earlier section; that is, they may merely change the odds ("incline") without necessitating the behavior. The question is whether this kind of explanatory framework for treating joint agent behavior is also subject to a deterministic reduction in the sense previously defined. The answer is "yes." If there is a common cause explanation for correlated behavior, even if the causal factors are only stochastic, there is also a deterministic explanation for the behavior. I refer the reader to Fine (1982) for the details of the proof, but the idea is simple enough. We start with the fact exploited in the discussion of direct causal links, namely, that any multiple probability distribution has a deterministic reduction. In the case of common cause explanations, corresponding to each of the causal factors there is a single probability distribution for each agent. Because these distributions are independent (in the stochastic sense) their product is a well-defined multiple distribution that can be averaged over all the causal factors. That average has a deterministic reduction, which is the deterministic theory that we were seeking.

The second answer to the question of how one deals with correlations between the behavior of different agents, then, is this. If the connections between the behavior of several agents arise from causal links or common causes, we can deal with them by effecting exactly the same sort of deterministic reduction that we gave for the single agent theory. To back up the charge that this framework is too limited, one would have to produce correlations in behavior that could not be treated this way. The quantum theory contains just such correlations, that is, ones that defy direct causal or common causal explanations. In Fine (1989), I have urged that the natural way to understand them is to acknowledge the irreducibility of the quantum statistical laws, in which case the correlations emerge as irreducible as well (see also the Appendix to van Fraassen 1989). Whatever is the right setting in the quantum case, however, we have yet to see any

correspondingly “anomalous” correlations in the behavior of human agents. There is a good reason why not. Such correlations can only arise if there is no possible way to integrate the behavior of a number of agents into a single probabilistic framework, or, more specifically, if there is no joint probability function for them all. This happens in the quantum theory where noncommuting observables do not have joint distributions. But it is difficult to see how this could occur in the arena of human behavior, and we certainly have no instances at hand. A simple sort of case would be that of three individuals *X*, *Y* and *Z* where the behavior of *X* and *Y* is jointly describable, and that of *X* and *Z* is also jointly describable, but there is no possible description for *Y* and *Z* taken together. This may remind one of intransitivity in preference rankings. But such intransitivity still allows for a composite preference ranking for all the individuals, which we obtain by simply conjoining all the separate rankings. A much deeper incoherence would have to obtain in order to forestall any conjoint probabilistic description at all. Thus, barring the demonstration of such a deep incoherence in group behavior, the reducibility of indeterminist theories of agency to deterministic ones seems well founded.

This way of viewing the correlation problem is also responsive to the second issue raised in contrasting the reducibility of theories of agency with the irreducibility of the quantum statistics. That was the issue of external constraints. We saw that in order to get quantum irreducibility we needed to impose some extratheoretical assumption, like locality. Could there be plausible constraints of this sort that would block the reduction in the case of human agency? We have just rehearsed what the consequences would be if there were such constraints. The theory of agency could not treat all the agents in one probabilistic framework. There would be incompatibilities between agents that prevented any conjoint probabilistic description. Of course one cannot rule this possibility out a priori. However, one can say that there is nothing in our human experience (including theories of human behavior with which we are familiar) that suggests the plausibility of limitations on theorizing that would issue in this sort of incoherence. On the grounds that one need not be frightened away by skeptical possibilities that run counter to ordinary human experience, I feel comfortable in putting the idea of reduction-limiting constraints on the shelf. If some specific limiting principles were offered, then (of course) we would have to take them seriously.

This last response also addresses the final charge, bound to be brought, that this whole discussion of a deterministic reduction is merely formal. All I have shown, after all, is that nothing stands in the way of determinism from a purely formal point of view. But surely that does not mean that *in reality* we could always find an adequate deterministic reduction to a given indeterminist theory of behavior. To be sure, my proof is merely formal. But I think it shows something relevant to free will nevertheless. For “in reality” there is no indeterminist theory to reduce. As described in an earlier section, our standard practice with regard to human agency employs the “but for” test and other rules of thumb that embody deterministic presuppositions. Likewise, theories of behavior look for causes or determinants in a strict sense. The realistic issue is whether anything stands in the way of these deterministic presuppositions and practices. The example of the quantum theory shows that there can be serious obstacles. By examining the conditions required for such obstacles to arise, the preceding discussion shows that in the case of human agency there are none. This is the sense in which we may suppose that any indeterminist theory would be reducible.

Concluding Remarks

In the first part of this essay I argued that the statistical laws of the quantum theory require a libertarian interpretation. This is a reading that incorporates the libertarian idea that when something happens it could have been otherwise, and under the very same circumstances. If human behavior were governed by statistical laws like those of the quantum theory, this would seem to lend support to the libertarian (“could have done otherwise”) picture of a free agent. However, the libertarian also thinks that free agency is incompatible with a deterministic account of human behavior. In the second part of this essay I support Grünbaum’s conditional and show that if there were an indeterminist account of the behavior of an agent (that is, an account that made essential use of probability) then there would also be a determinist account of the same behavior (that is, a probability-free account). Thus indeterminism is also incompatible with free will (assuming that determinism is). It follows that on the libertarian conception there is no free will since, as I use them in these demonstrations, determinism and indeterminism are logical opposites.

This is not the conclusion for which the libertarian was hoping. The libertarian would rather have it that human behavior is governed by indeterministic laws that cannot be given a deterministic reduction. For that to be the case, however, there would have to be a deep incoherence in group behavior, one that prevented complete statistical descriptions of the behavior of several agents together. One might contrast this circumstance with the usual understanding of the libertarian position, which is nicely captured in this remark by a well-known legal scholar, "No matter how well or fully we learn the antecedent facts, we can never say of a voluntary action that it had to be the case that the person would choose to act in a certain way. In a word, every volitional actor is a wild card" (Kadish 1985, 360). Now in dealing wild cards (or, to use Dennett's metaphor, tossing coins) we give odds, and make book on the outcomes. If the laws could not be given a deterministic reduction, however, then just this sort of probabilistic description would be ruled out, at least in certain cases. Thus the burden on the libertarian is to find a way of treating human behavior that involves neither probabilistic description nor nonprobabilistic description. The only way that recommends itself is to treat people as arational, that is, as behaving outside the bounds of rational description altogether. This seems a self-defeating move for the libertarian, who wishes to uphold the dignity of humankind. Thus the libertarian conception of free will seems truly incoherent.

The upshot of these considerations leaves room for a causal and deterministic description of human behavior, the kind of theorizing that is widely practiced and familiar. This kind has its own account of free will, a strained version that fits the compatibilist tradition and according to which (although "free") we really could *not* have done otherwise. The libertarian conception of free will turns out to be incoherent. I would argue that this twisted compatibilist conception is hardly better.

NOTES

I want to dedicate this essay to Adolf Grünbaum, whose concern for clear reasoned discussion provides an admirable model for us all. Work on this essay was supported by NSF Grant DIR-8905571. My thanks to Micky Forbes, Richard Manning, and Aaron Snyder for helpful discussions, although it must be said

that they do not necessarily agree with my lines of argument. I owe a debt to Elizabeth Dipple for direction in the classics. John Earman made useful comments on an earlier draft and that has helped too, although, I fear, he will think not enough.

1. Witness the formulation by Hart and Honoré, "A deliberate human act is . . . something through which we do not trace the cause of a later event" (1959, 41).

2. Grünbaum calls "irreducibly statistical" the laws that I label *indeterministic*.

3. See Feigl et al. (1972, 614) where Grünbaum shifts from an epistemic to a nonepistemic formulation of this line of argument.

4. In this essay I employ a very minimal sense of determinism and indeterminism, depending on whether (indeterminism) the laws of nature involve an essentially probabilistic element or (determinism) not. In this sense, I take the alternatives here to be mutually exclusive and jointly exhaustive. There are other senses. See Earman (1986). In what follows, I sometimes mix in causal talk. Given the subject matter, the mix is inevitable.

5. For the purposes of the deconstructive *reductio*, I follow the tradition in discussions of the freedom of the will in supposing that responsibility tracks agency, although I do not believe any such thing. As explained in the beginning of the essay, I think agency is crafted out of the need to ground the social practices involved in assigning and judging responsibility. Like other foundational "needs," we can do better.

6. In tort law, assignments of liability require proximate cause and may use several variations of the "but for" criterion, variations like the INUS conditions familiar in philosophical discussions. In a stochastic setting, however, these variations do not fare any better than the "but for" condition itself. There is some discussion among legal scholars concerning how to apportion damages where, say, a polluter is responsible for increasing certain health risks by a determinate percentage. Increasing the risk of disease, here, corresponds to inducing certain medical conditions which would be considered harmful in themselves, and for which damages could be assessed. This is different from holding a polluter responsible for (stochastically) "causing" a disease that may or may not occur, which would be closer to the issue discussed in the text. So far, all this is legal speculation, not (I believe) supported by case law. Moreover, the awarding of damages in torts is not the most reliable guide to responsibility (just think of the strict liability).

REFERENCES

- Bell, J. S. 1987. *Speakable and Unsayable in Quantum Mechanics*. Cambridge, England: Cambridge University Press.
- Cartwright, N. 1989. *Nature's Capacities and Their Measurement*. Oxford: Clarendon Press.

- Cushing, J., and E. McMullin, eds. 1989. *Philosophical Consequences of Quantum Theory*. Notre Dame: University of Notre Dame Press.
- Dennett, D. 1978. *Brainstorms*. Cambridge, Mass.: Bradford Books.
- . 1984. *Elbow Room*. Cambridge, Mass.: MIT Press.
- Earman, J. 1986. *A Primer of Determinism*. Dordrecht: Reidel.
- Elby, A. 1990. "On the Physical Interpretation of Heywood and Redhead's Algebraic Impossibility Theorem." *Foundations of Physics Letters* 3: 239–47.
- Feigl, H.; W. Sellars; and K. Lehrer. 1972. *New Readings in Philosophical Analysis*. New York: Appleton-Century-Crofts.
- Fine, A. 1982. "Joint Distributions, Quantum Correlations and Commuting Observables." *Journal of Mathematical Physics* 23: 1306–10.
- . 1989. "Do Correlations Need to Be Explained?" In J. Cushing and E. McMullin, eds., *Philosophical Consequences of Quantum Theory*. Notre Dame: University of Notre Dame Press, pp. 175–94.
- Grünbaum, A. 1953. "Causality and the Science of Human Behavior." In H. Feigl and M. Brodbeck, eds., *Readings in the Philosophy of Science*. New York: Appleton-Century-Crofts, pp. 766–77. (Originally published in *American Scientist* 40: 665–76.)
- . 1967. "Science and Man." In M. Mandelbaum et al., eds., *Philosophic Problems*. New York: Macmillan, pp. 448–62. (Originally published in *Perspectives in Biology and Medicine* 5: 483–502.)
- . 1972. "Free Will and the Laws of Human Behavior." In H. Feigl, W. Sellars, and K. Lehrer, eds., *New Readings in Philosophical Analysis*. New York: Appleton-Century-Crofts, pp. 605–27. (Originally published in *American Philosophical Quarterly* 8: 299–317.)
- Hart, H. L. A., and A. M. Honoré. 1959. *Causation in the Law*. Oxford: Clarendon Press.
- Hume, D. [1739] 1902. *A Treatise of Human Nature*. 1st ed. Edited by L. A. Selby-Bigge. Oxford: Clarendon Press.
- . [1777] 1902. *An Enquiry Concerning Human Understanding*. 2d ed. Edited by L. A. Selby-Bigge. Oxford: Clarendon Press.
- Humphreys, P. 1989. *The Chances of Explanation*. Princeton: Princeton University Press.
- Kadish, S. 1985. "Complicity, Cause and Blame: A Study of the Interpretation of Doctrine." *California Law Review* 73: 323–410.
- Redhead, M. 1987. *Incompleteness, Nonlocality and Realism*. Oxford: Clarendon Press.
- van Fraassen, B. 1989. "The Charybdis of Realism." In J. Cushing and E. McMullin, eds., *Philosophical Consequences of Quantum Theory*. Notre Dame: University of Notre Dame Press, pp. 97–113.

22

Adolf Grünbaum and the
Compatibilist Thesis

John Watkins

Department of Philosophy, Logic and Scientific Method, London School of Economics

Compatibilism has been a long-standing thesis of Adolf Grünbaum (he first defended it in his 1952 article). His version involves three subtheses: (1) all human behavior is open to scientific explanation with the help of causal laws; (2) it makes no difference, with respect to human freedom and responsibility, whether the laws are deterministic or statistical; and (3) there is "*no incompatibility* between the existence of either causal or statistical laws of voluntary behavior, on the one hand, and the feelings of freedom which we actually do have, the meaningful assignment of responsibility, the rational infliction of punishment, and the existence of feelings of remorse or guilt on the other" (1972, 608; emphasis added). I will challenge all three subtheses. I will first set the scene by indicating the direction from which the challenges will come.

Grünbaum and I share a staunchly naturalistic outlook. Although my deductivism obliges me to call myself an agnostic rather than an atheist, since it does not allow me to conclude from our human experience that no God exists, it does allow me to conclude that human affairs are not under the (continuous and effective) supervision of a caring God. I added that parenthetical remark in view of a hypothesis of Martin Buber's, which Grünbaum (1987, 178) mentions, to the effect that there is a caring God who, however, goes into eclipse from time to time—as He did, for instance, during the Nazi holocaust. I accept unquestioningly that we are part of nature, in the sense that we