

all to you: the doubts and worries and uncertainties as well as the beliefs, convictions, and arguments.

At those particular points in my arguments, transitions, assumptions, and so forth, where I feel the strain, I try to comment or at least to draw the reader's attention to what makes me uneasy. In advance, it is possible to voice some general theoretical worries. This book does not present a precise theory of the moral basis of individual rights; it does not contain a precise statement and justification of a theory of retributive punishment; or a precise statement of the principles of the tripartite theory of distributive justice it presents. Much of what I say rests upon or uses general features that I believe such theories would have were they worked out. I would like to write on these topics in the future. If I do, no doubt the resulting theory will differ from what I now expect it to be, and this would require some modifications in the superstructure erected here. It would be foolish to expect that I shall complete these fundamental tasks satisfactorily; as it would be to remain silent until they are done. Perhaps this essay will stimulate others to help.

ACKNOWLEDGMENTS

THE first nine chapters of this essay were written during 1971–1972, while I was a Fellow at the Center for Advanced Study in the Behavioral Sciences at Palo Alto, a minimally structured academic institution bordering on individualist anarchy. I am very grateful to the Center and its staff for providing an environment so conducive to getting things done. Chapter 10 was presented in a symposium on “Utopia and Utopianism” at a meeting of the Eastern Division of the American Philosophical Association in 1969; some points from that delivered address appear scattered in the other chapters. The whole manuscript was rewritten during the summer of 1973.

Barbara Nozick’s objections to some of the positions defended here helped me to sharpen my views; in addition she helped enormously in innumerable other ways. Over several years, I have benefited from Michael Walzer’s comments, questions, and counter-arguments as I tried out on him ideas on some topics of this essay. I have received detailed and very helpful written comments on the whole manuscript written at the Center from W. V. Quine, Derek Parfit, and Gilbert Harman, on Chapter 7 from John Rawls and Frank Michelman, and on an earlier draft of Part I from Alan Dershowitz. I also have benefited from a discussion with Ronald Dworkin on how competing protective agencies would(n’t) work, and from suggestions by Burton Dreben. Various stages of various portions of this manuscript were read and discussed, over the years, at meetings of the Society for Ethical and Legal Philosophy (SELF); the regular discussions with its members have been a source of intellectual stimulation and pleasure. It was a long conversation about six years ago with Murray Rothbard that stimulated my interest in individualist anarchist theory. Even

longer ago, arguments with Bruce Goldberg led me to take libertarian views seriously enough to want to refute them, and so to pursue the subject further. The result is before the reader.

PART

I

*State-of-Nature Theory,
or How to Back into a State
without Really Trying*

CHAPTER

1

Why State-of-Nature Theory?

IF the state did not exist would it be necessary to invent it? Would one be *needed*, and would it have to be *invented*? These questions arise for political philosophy and for a theory explaining political phenomena and are answered by investigating the “state of nature,” to use the terminology of traditional political theory. The justification for resuscitating this archaic notion would have to be the fruitfulness, interest, and far-reaching implications of the theory that results. For the (less trusting) readers who desire some assurance in advance, this chapter discusses reasons why it is important to pursue state-of-nature theory, reasons for thinking that theory would be a fruitful one. These reasons necessarily are somewhat abstract and metatheoretical. The best reason is the developed theory itself.

POLITICAL PHILOSOPHY

The fundamental question of political philosophy, one that precedes questions about how the state should be organized, is whether there should be any state at all. Why not have anarchy? Since anarchist theory, if tenable, undercuts the whole subject of *political* philosophy, it is appropriate to begin political philosophy with an examination of its major theoretical alternative. Those who consider anarchism not an unattractive doctrine will think it possible that political philosophy *ends* here as well. Others impatiently will await what is to come afterwards. Yet, as we shall see, anarchists and anarchists alike, those who spring gingerly from the starting point as well as those reluctantly argued away from it, can agree that beginning the subject of political philosophy with state-of-nature theory has an *explanatory* purpose. (Such a purpose is absent when epistemology is begun with an attempt to refute the skeptic.)

Which anarchic situation should we investigate to answer the question of why not anarchy? Perhaps the one that would exist if the actual political situation didn't, while no other possible political one did. But apart from the gratuitous assumption that everyone everywhere would be in the same nonstate boat and the enormous unmanageability of pursuing that counterfactual to arrive at a particular situation, that situation would lack fundamental theoretical interest. To be sure, if that nonstate situation were sufficiently awful, there would be a reason to refrain from dismantling or destroying a particular state and replacing it with none, now.

It would be more promising to focus upon a fundamental abstract description that would encompass all situations of interest, including "where we would now be if." Were this description awful enough, the state would come out as a preferred alternative, viewed as affectionately as a trip to the dentist. Such awful descriptions rarely convince, and not merely because they fail to cheer. The subjects of psychology and sociology are far too feeble to support generalizing so pessimistically across all societies and persons, especially since the argument depends upon *not* making *such* pessimistic assumptions about how the *state* operates. Of

course, people know something of how actual states have operated, and they differ in their views. Given the enormous importance of the choice between the state and anarchy, caution might suggest one use the “minimax” criterion, and focus upon a pessimistic estimate of the nonstate situation: the state would be compared with the most pessimistically described Hobbesian state of nature. But in using the minimax criterion, this Hobbesian situation should be compared with the most pessimistically described possible state, including *future* ones. Such a comparison, surely, the worst state of nature would win. Those who view the state as an abomination will not find minimax very compelling, especially since it seems one could always bring back the state if that came to seem desirable. The “maximax” criterion, on the other hand, would proceed on the most optimistic assumptions about how things would work out—Godwin, if you like that sort of thing. But imprudent optimism also lacks conviction. Indeed, no proposed decision criterion for choice under uncertainty carries conviction here, nor does maximizing expected utility on the basis of such frail probabilities.

More to the point, especially for deciding what goals one should try to achieve, would be to focus upon a nonstate situation in which people generally satisfy moral constraints and generally act as they ought. Such an assumption is not wildly optimistic; it does not assume that all people act exactly as they should. Yet this state-of-nature situation is the best anarchic situation one reasonably could hope for. Hence investigating its nature and defects is of crucial importance to deciding whether there should be a state rather than anarchy. If one could show that the state would be superior even to this most favored situation of anarchy, the best that realistically can be hoped for, or would arise by a process involving no morally impermissible steps, or would be an improvement if it arose, this would provide a rationale for the state's existence; it would justify the state.*

This investigation will raise the question of whether all the ac-

* This contrasts with a theory that presents a state's arising from a state of nature by a natural and inevitable process of *deterioration*, rather as medical theory presents aging or dying. Such a theory would not “justify” the state, though it might resign us to its existence.

tions persons must do to set up and operate a state are themselves morally permissible. Some anarchists have claimed not merely that we would be better off without a state, but that any state necessarily violates people's moral rights and hence is intrinsically immoral. Our starting point then, though nonpolitical, is by intention far from nonmoral. Moral philosophy sets the background for, and boundaries of, political philosophy. What persons may and may not do to one another limits what they may do through the apparatus of a state, or do to establish such an apparatus. The moral prohibitions it is permissible to enforce are the source of whatever legitimacy the state's fundamental coercive power has. (Fundamental coercive power is power not resting upon any consent of the person to whom it is applied.) This provides a primary arena of state activity, perhaps the only legitimate arena. Furthermore, to the extent moral philosophy is unclear and gives rise to disagreements in people's moral judgments, it also sets problems which one might think could be appropriately handled in the political arena.

EXPLANATORY POLITICAL THEORY

In addition to its importance for political philosophy, the investigation of this state of nature also will serve explanatory purposes. The possible ways of understanding the political realm are as follows: (1) to fully explain it in terms of the nonpolitical; (2) to view it as emerging from the nonpolitical but irreducible to it, a mode of organization of nonpolitical factors understandable only in terms of novel political principles; or (3) to view it as a completely autonomous realm. Since only the first promises full understanding of the whole political realm,¹ it stands as the most desirable theoretical alternative, to be abandoned only if known to be impossible. Let us call this most desirable and complete kind of explanation of a realm a *fundamental* explanation of the realm.

To explain fundamentally the political in terms of the nonpolitical, one might start either with a nonpolitical situation, showing how and why a political one later would arise out of it, or with a

political situation that is described nonpolitically, deriving its political features from its nonpolitical description. This latter derivation either will identify the political features with those features nonpolitically described, or will use scientific laws to connect distinct features. Except perhaps for this last mode, the illumination of the explanation will vary directly with the independent glow of the nonpolitical starting point (be it situation or description) and with the distance, real or apparent, of the starting point from its political result. The more fundamental the starting point (the more it picks out basic, important, and inescapable features of the human situation) and the less close it is or seems to its result (the less political or statelike it looks), the better. It would not increase understanding to reach the state from an arbitrary and otherwise unimportant starting point, obviously adjacent to it from the start. Whereas discovering that political features and relations were reducible to, or identical with, ostensibly very different nonpolitical ones would be an exciting result. Were these features fundamental, the political realm would be firmly and deeply based. So far are we from such a major theoretical advance that prudence alone would recommend that we pursue the alternative of showing how a political situation would arise out of a nonpolitical one; that is, that we begin a fundamental *explanatory* account with what is familiar within political philosophy as state-of-nature theory.

A theory of a state of nature that begins with fundamental general descriptions of morally permissible and impermissible actions, and of deeply based reasons why some persons in any society would violate these moral constraints, and goes on to describe how a state would arise from that state of nature will serve our explanatory purposes, *even if no actual state ever arose that way*. Hempel has discussed the notion of a potential explanation, which intuitively (and roughly) is what would be the correct explanation if everything mentioned in it were true and operated.² Let us say that a *law-defective* potential explanation is a potential explanation with a false lawlike statement and that a *fact-defective* potential explanation is a potential explanation with a false antecedent condition. A potential explanation that explains a phenomenon as the result of a process *P* will be defective (even though it is neither law-defective nor fact-defective) if some process *Q* other than *P* produced the

phenomenon, though P was capable of doing it. Had this other process Q not produced it, then P would have.* Let us call a potential explanation that fails in this way actually to explain the phenomenon a *process-defective* potential explanation.

A *fundamental* potential explanation (an explanation that would explain the whole realm under consideration were it the actual explanation) carries important explanatory illumination even if it is *not* the correct explanation. To see how, in principle, a *whole realm* could fundamentally be explained greatly increases our understanding of the realm.† It is difficult to say more without examining types of cases; indeed, without examining particular cases, but this we cannot do here. Fact-defective fundamental potential explanations, if their false initial conditions “could have been true,” will carry great illumination; even wildly false initial conditions will illuminate, sometimes very greatly. Law-defective fundamental potential explanations may illuminate the nature of a realm almost as well as the correct explanations, especially if the “laws” together form an interesting and integrated theory. And process-defective fundamental potential explanations (which are neither law-defective nor fact-defective) fit our explanatory bill and purposes almost perfectly. These things could not be said as strongly, if at all, about nonfundamental explanation.

State-of-nature explanations of the political realm *are* fundamental potential explanations of this realm and pack explanatory

* Or, perhaps yet *another* process R would have if Q hadn't, though had R not produced the phenomenon, then P would have, or. . . . So the footnoted sentence should read: P would have produced the phenomenon had no member of [Q , R , . . .] done so. We ignore here the complication that what would prevent Q from producing the phenomenon might also prevent P from doing so.

† This claim needs to be qualified. It will not increase our understanding of a realm to be told as a potential explanation what we know to be false: that by doing a certain dance, ghosts or witches or goblins made the realm that way. It is plausible to think that an explanation of a realm must present an underlying mechanism yielding the realm. (Or do something else equally productive of understanding.) But to say this is not to state precisely the deep conditions an underlying mechanism must satisfy to explain a realm. The precise qualification of the claim in the text awaits advances in the theory of explanation. Yet other difficulties call for such advances; see Jaegwon Kim, “Causation, Nomic Subsumption, and the Concept of Event,” *The Journal of Philosophy*, 70, no. 8 (April 26, 1973), 217–236.

punch and illumination, even if incorrect. We learn much by seeing how the state could have arisen, even if it didn't arise that way. If it didn't arise that way, we also would learn much by determining why it didn't; by trying to explain why the particular bit of the real world that diverges from the state-of-nature model is as it is.

Since considerations both of political philosophy and of explanatory political theory converge upon Locke's state of nature, we shall begin with that. More accurately, we shall begin with individuals in something sufficiently similar to Locke's state of nature so that many of the otherwise important differences may be ignored here. Only when some divergence between our conception and Locke's is relevant to *political* philosophy, to our argument about the state, will it be mentioned. The completely accurate statement of the moral background, including the precise statement of the moral theory and its underlying basis, would require a full-scale presentation and is a task for another time. (A lifetime?) That task is so crucial, the gap left without its accomplishment so yawning, that it is only a minor comfort to note that we here are following the respectable tradition of Locke, who does not provide anything remotely resembling a satisfactory explanation of the status and basis of the law of nature in his *Second Treatise*.