

ROBERT  
NOZICK

SOCRATIC  
PUZZLES

**SOCRATIC**

**PUZZLES**

**Robert Nozick**

Harvard University Press  
Cambridge, Massachusetts  
London, England • 1997

Copyright © 1997 by Robert Nozick  
All rights reserved  
Printed in the United States of America

Designed by Marianne Perlak

*Library of Congress Cataloging-in-Publication Data*

Nozick, Robert.

Socratic puzzles / Robert Nozick.

p. cm.

Includes bibliographical references and index.

ISBN 0-674-81653-6 (alk. paper)

1. Methodology. 2. Socrates. 3. Philosophy. 4. Political  
science—Philosophy. 5. Ethics. I. Title.

B945.N683S63 1997

191—dc21

96-39221

→ → → **SOCRATIC PUZZLES**

*To the memory of my parents,*

*Max Nozick (1906–1990)*

*and*

*Sophie Cohen Nozick (1908–1975)*

---

---

## CONTENTS

Introduction 1

### CHOICE AND UTILITY

- 1 Coercion 15
- 2 Newcomb's Problem and Two Principles of Choice 45
- 3 Reflections on Newcomb's Problem 74
- 4 Interpersonal Utility Theory 85
- 5 On Austrian Methodology 110

### PHILOSOPHY AND METHODOLOGY

- 6 Socratic Puzzles 145
- 7 Experience, Theory, and Language 156
- 8 Simplicity as Fall-Out 182
- 9 Invisible-Hand Explanations 191

### ETHICS AND POLITICS

- 10 Moral Complications and Moral Structures 201
- 11 On the Randian Argument 249
- 12 Weighted Voting and "One-Man, One-Vote" 265

**DISCUSSIONS AND REVIEWS**

- 13 Goodman, Nelson, on Merit, Aesthetic 273
- 14 Who Would Choose Socialism? 277
- 15 Why Do Intellectuals Oppose Capitalism? 280
- 16 The Characteristic Features of Extremism 296
- 17 War, Terrorism, Reprisals—Drawing Some Moral Lines 300
- 18 Do Animals Have Rights? 305

**PHILOSOPHICAL FICTIONS**

- 19 Fiction 313
- 20 R.S.V.P.—A Story 317
- 21 Testament 324
- 22 Teleology 329

Notes and References 333

Credits 391

Name Index 393

Subject Index 397

---

---

## INTRODUCTION

It is disconcerting to be known primarily for an early work. Others have identified me as a “political philosopher,” but I have never thought of myself in those terms. The vast majority of my writing and attention has focused on other subjects.

*Anarchy, State, and Utopia* was an accident. It was written during 1971–72, a year spent at the Center for Advanced Study in the Behavioral Sciences, adjacent to Stanford University, and my plan for that year was to write on the topic of free will. Social and political philosophy was a serious but not a predominant interest of mine. After several months of getting nowhere on the topic of free will, I was invited in early December to give a talk to a Stanford University student group, and I presented some thoughts on how a state would arise out of (individualist) anarchy. Those thoughts seemed worth writing up, and that was done by the beginning of January, when the mail from Cambridge brought John Rawls’s long-awaited book, which I had read in manuscript and discussed extensively with Rawls. In lectures at Harvard, I had been developing the entitlement theory of justice, within a libertarian framework, and when I read the final version of *A Theory of Justice* that January, I was moved to write down the entitlement theory along with my critical thoughts on Rawls’s theory, which had developed significantly since our last discussions. By the end of February, I had on my desk an essay on state-of-nature theory, an essay on distributive justice, and also an essay on utopia that I had delivered at a meeting of the American Philosophical Association and brought along with me to Stanford. These three pieces fit together and it seemed that with some integration and some additional material they might constitute a book. So, rather than returning to the intractable problem of free will, I set to intensive work to complete that book before my stay at Stanford would be up in July.



I have not responded to the sizable literature on *Anarchy, State, and Utopia*, or followed it closely. I did not want to spend my life writing “The Son of Anarchy, State, and Utopia,” “The Return of the Son of . . .,” etc. I had other philosophical questions to think about: knowledge, the self, why there is something rather than nothing, and—of course—free will. Also, I believed that responding to criticisms of that book or reading them closely would not be a good way for me to gain any critical distance from its positions. My natural tendency would be to defend strongly those positions under attack—and *attack* it was!—and how could I learn that my views were mistaken if I thought about them always with defensive juices flowing? (I returned briefly to express reservations about some of *Anarchy, State, and Utopia*’s views in *The Examined Life*, pp. 30–32, 286–296.) I have not responded to the literature on my other writing either. I could say that any intellectual contribution I could make was likely to be greater in broaching new ideas on new subjects than in elaborating or defending my previous ideas, and that would be true, but also it is true that what pleases and excites me is to think new thoughts about new topics. I have learned much from philosophers who stay with one area or one theme, gaining deeper insight, but my bent is elsewhere.

Philosophers seek structures telling how things are related and constituted. They find complexity in what appears simple, and simplicity in what appears complex. Philosophy’s understanding is structural, its method is clarity of thought and reasoning.

Socrates, at western philosophy’s beginning, made critical thinking an explicit, methodical activity. The aim was to bring everything to conscious, explicit, and critical awareness: the definition of important concepts, the articulation of principles of conduct, the methods of reasoning themselves. Philosophers since then have been indebted to Socrates not only for his intellectual daring and incisive trains of thought, but for the vivid and inspiring way he embodies what a philosopher is. Medieval thinkers paid tribute to the scope and depth of Aristotle’s theories by referring to him as “the philosopher.” But when I think of “the” philosopher, I think of Socrates. Scholars have disagreed over central questions about this sharp and clear thinker. What did he mean when he said he did not know the answers to his questions, what does his method of critical discussion presuppose, and what is his concern in im-

proving the souls of those he encounters? In “Socratic Puzzles” (Essay 6), I propose some answers.

The intellectual problems and puzzles in the other essays of this volume\* are not all of them “Socratic,” if that means focusing on the particular questions Socrates investigated. However, the writing here grows out of what Socrates did and is continuous with it. Sometimes an interesting question motivates analyzing a concept, sometimes a central notion leads to a new and puzzling question, and sometimes these links iterate to form a longer chain of reasoning. That is true of much of philosophy, of course, so while the title of this book marks it as philosophy, it does not specifically describe its contents. Rather, it pays Socrates homage.

One of philosophy’s central tasks has been to formulate and justify rules, norms, and principles to help guide us through the welter of possibilities we face, different possibilities of action, of belief, of ways we could be. Let me indicate how the essays here fit within this philosophical task, and how they connect to my other writing. Normative principles of action will make actions depend upon our beliefs (about what is the case, and what will happen if we act a certain way) and upon our desires or evaluations (what we want to happen or value happening). But what normative rules or principles should our beliefs and desires satisfy? And, given (appropriate) beliefs and desires, what specific principle or rule should then select the action?

The most widely discussed rule mandates that our actions maximize expected utility. Along with normative conditions on the structure of preferences, this rule is part of a normative theory of decision that was formulated by economists and mathematicians in the mid-twentieth century. Philosophers, with their interest in explicit normative or rational rules, were not slow in taking up this theory. I was drawn seriously into the study of philosophy in Sidney Morgenbesser’s philosophy classes at Columbia College, and there I first encountered these decision theories. I was led to Morgenbesser’s classes in particular by my earlier experience in a section that he taught of a required general education course at Columbia on twentieth-century social and political thought. Every time I said something in that course, Morgenbesser raised an objection—I was overlooking a distinction or ignoring an objection or failing to consider how some structural feature of my point might generalize—and

\*I am very grateful to Tamar Gendler for her persuasiveness in urging me to bring together a volume of essays and for her help in assembling this work for press.

his objection would convince me that I wasn't thinking clearly enough, and that there was a clearer way to think. Thus I was lured to take more courses in the hope of learning to think that clearly—soon I had amassed enough courses to “major in Morgenbesser.” It was in graduate school at Princeton, studying with Carl Hempel, that I took a really close look at decision theory, and a projected dissertation (spurred initially by Hempel's writing) on the confirmation and acceptance of scientific hypotheses, after some months and some work, was transformed into “The Normative Theory of Individual Choice” (1963; published in 1990 by Garland Press). Here, I discussed normative conditions on preference, the theory of choice under certainty, risk and uncertainty, and some issues about game theory.

It was in the dissertation that I first discussed Newcomb's Problem, a problem in decision theory thought up by a physicist in California, William Newcomb, and told to me at a party by a mutual friend and physicist at Princeton, David Martin Kruskal. (It was, for me, the most consequential party I have attended.) This intriguing problem was first published in “Newcomb's Problem and Two Principles of Choice” (Essay 2); the treatment there differed from that in the dissertation. Martin Gardner wrote a “Mathematical Games” column on the problem in *Scientific American*, and he asked me to do a guest column (Essay 3) responding to the vast amount of mail he had received. I returned to the problem in *The Nature of Rationality* (pp. 40–50), where I propose the maximization of decision-value (a weighted sum of causally expected, evidentially expected, and symbolic utility). Although three distinct positions on Newcomb's Problem are a small fraction of the positions in the by-now vast literature, they are more than enough for one writer!

Decision theory (“rational choice theory”) has received extensive application in the social sciences, and it provides the structure for some of my later philosophical work. It plays a role in *Anarchy, State, and Utopia's* thoughts about how individual action would give rise to a state, and also in its game-theoretic analysis of utopia; in *The Nature of Rationality* I reformulate decision theory and apply it to issues of rational belief. Decision theory provides the implicit background structure in the essay on “Coercion” (Essay 1), and that part of decision theory concerned with the structure and measurement of preference (utility theory) is one of the models for the theory of moral structures, and the measurement of moral weight, presented in “Moral Complications and Moral Structures” (Essay 10).

Another model was Noam Chomsky's *Syntactic Structures*, and I meant “Moral Complications” to begin to uncover the structure under-

lying our moral judgments. The essay presents several possible structures (the maximization structure, the deductive structure, and the simple balancing structure) before discussing the more complicated moral balancing structure with its own special governing principles. I added a bit to the formulation and investigation of alternative moral structures in the presentation and discussion of “side-constraints” in *Anarchy, State, and Utopia* (pp. 26–42), but with rare exceptions, the task has not been pursued very extensively by philosophers since then.\* One of the new principles in “Moral Complications” (Principle VII) is put to use in a review of Michael Walzer’s *Just and Unjust Wars* (Essay 17) to sharpen his proposal about what risks should be incurred to protect civilians in warfare.

Utility theory, the component of decision theory concerning the structure of rational preferences and of their measurement, has been of interest to normative social theory, especially within the utilitarian tradition which mandates maximizing the total happiness or the aggregate sum of utility. That tradition’s use of the utility notion, and some other uses as well (hence the interest of the problem, even for critics of utilitarianism), faces the well-known problem of interpersonal comparisons of utility: how can we systematically compare the strength of different people’s preferences or desires? “Interpersonal Utility Theory” (Essay 4) presents a new strategy for attacking this problem and makes specific new proposals.

A normative concern about what conditions an individual’s choices should satisfy fits in naturally with a concern about the range of choice a person should be free to make. Given my background view that coercion is justified only under very stringent conditions, I was led to investigate the notion of coercion directly (Essay 1). Philosophers had not previously looked closely at this notion, and it presented interesting problems, not simply of formulating an adequate definition that would handle a range of complicated cases, but of understanding why the contours of the notion of coercion take the particular form they do. Why should *threats* be an important component of coercion when offers also can establish a significant difference between the utility yielded by two actions? The answer went beyond the utilities involved in the actual

\* However, the structure of “prima facie” principles, investigated in one direction in “Moral Complications,” has been developed in other intricate and fruitful directions by writers in the field of artificial intelligence studying default structures; for instance, in John Holland, Keith Holyoak, Richard Nisbett, and Paul Thagard, *Induction: Processes of Inference, Learning, and Discovery* (Cambridge, MA: MIT Press, 1986).

choice situation (after the threat or offer has been made); it lay in the hypothetical choice of entering into those threat or offer situations. A person normally would be willing to be the recipient of an offer yet be unwilling to be the recipient of a threat. The difference in willingness to move *into* the situation makes a difference in the voluntariness of the action that is done *within* the situation. A moral evaluation must look beyond the immediate situation and beyond the agent's first-level desires within that situation. (The literature on coercion has grown greatly since my essay appeared. For a discussion, see Alan Wertheimer, *Coercion* [Princeton, NJ: Princeton University Press, 1987].)

The issue of how wide a range of factors must be considered by moral or evaluative principles recurs elsewhere. The moral treatment owed to a person who is mentally retarded, I claim in a review of Tom Regan's book where he discussed this topic (Essay 18), depends not only on the traits he or she possesses but also upon the typical and normal traits possessed by members of the human species, even when the particular person being treated lacks some of these traits. (The closest-continuer theory of identity over time, in *Philosophical Explanations*, pp. 29–70, also widens a context: to determine whether two entities are identical, one must look beyond the traits and relations of those two entities themselves.)

Another principle of normative evaluation is subject to examination in "Goodman, Nelson on Merit, Aesthetic" (Essay 13), where I criticize Goodman's view that a work has aesthetic merit if it is a work of art and changes the way we view the world. That essay contains my only published poem; readers will discover why I want them to think it is a poor poem. "The Characteristic Features of Extremism" (Essay 16) was written for a 1987 conference in Jerusalem (organized by the American Jewish Congress) on political extremism in the United States and Israel, and it was presented in a panel with Geula Cohen, a member of the Israeli Knesset. (Afterwards, some members of the audience asked how, without having met her beforehand, I could so accurately describe the other panelist!) The phenomenon of extremism has grown even more widespread and serious since then. A United States Supreme Court decision (*Baker v. Carr*) formulates the principle that every person's vote is to count equally, and Essay 12 applies formal apparatus to that decision (and in passing presents a counter-example to the Shapley-Shubik power index).

Preference and decision fall under normative rules, and so does belief. Rules encourage or mandate some beliefs, and they discourage

or exclude others. Such rules also might be used to criticize the beliefs of other people—always a satisfying activity. (Surely that was part of the appeal of the verifiability criterion of meaning—now known to be defective—put forth by the logical positivists. Allied with the rule not to believe or pay attention to anything (cognitively) meaningless, it allowed one, without further thought, to dismiss whatever was not verifiable, and to criticize others for indulging in meaningless talk.) Rules of rational belief, and the justification of these rules, has been a—perhaps *the*—longtime interest of philosophers. (I investigated such rules recently in *The Nature of Rationality*.) One rule often proposed is to believe the simplest hypothesis compatible with the evidence. It has been a puzzle why the simplest hypothesis is most likely to be the true one. “Simplicity as Fall-Out” (Essay 8) shows how a procedure for fixing belief, without seeking simplicity, might lead one nevertheless to end up believing the simplest hypothesis compatible with the evidence. Thus it proposes an invisible-hand explanation of simplicity.

The reviews of Walzer and of Regan are the only two book reviews I have written. Perhaps this surprising paucity of book reviews is explained by my first attempt at one in the late 1960s. I agreed to review H. L. A. Hart’s book of essays, *Punishment and Responsibility*, for the *Harvard Law Review*, and after reading this admirable work I set out with what seemed to me a reasonable plan: first, to develop the correct theory of punishment, and then, after doing so, to stand on that platform to state what was right and what was wrong with Hart’s views. The review never got written. (The thinking done then, however, found its way into the discussions of punishment in *Anarchy, State, and Utopia* and *Philosophical Explanations*.)

My interest in the minimization of coercion led me to investigate “libertarian” views and arguments. “Who Would Choose Socialism?” (Essay 14) formulates one such argument, looking at the percentage of Israelis who choose life in a kibbutz. “Why Do Intellectuals Oppose Capitalism?” (Essay 15) considers why intellectuals are especially resistant to capitalism and to arguments in its favor, and it hypothesizes a new sociological law. The essay on Ayn Rand, “On the Randian Argument” (Essay 11), examines her moral arguments, which academic philosophers previously had neglected, but since it treated them with the standard critical tools of the trade, it brought me vituperation from some of Rand’s followers. It is interesting to notice the themes of part I of *Anarchy, State, and Utopia* that are first adumbrated in this essay:

side-constraints and rights, whether the night-watchman state of classical liberal theory is redistributive, whether what matters is only people's *experiences*.

This aspect of early work holds true of other philosophers, as I discovered once while teaching a graduate seminar where the work of contemporary philosophers was looked at with the aim of understanding how they generated their ideas and theories. In addition to mature work, the seminar looked also at some early papers, where (I assumed) working methods of thought would be closer to the surface. What surprised me—in the early writing of Quine and Goodman, for example—was how many of these philosophers' later striking ideas were present already in these very early works, for those who had eyes to see.

The essay on Ayn Rand was written as a sidenote to my own independent thinking about topics in social philosophy and the moral foundations of capitalism. The essay on Quine, "Experience, Theory, and Language" (Essay 7), however, grew not from thinking independently about topics in the philosophy of language but from thinking about the ideas of Quine himself. Despite its particular points and its discussion of some new questions (such as why Duhem's thesis holds), I was struck, in rereading it soon after its writing, by the difference between thinking independently about a subject and thinking that is focused upon someone else's ideas. When you approach a topic through the route of someone's theories, that person's mode of structuring the issues limits how far you can stray and how much you can discover. You think within their "problematic."

Psychologists have investigated a phenomenon they call "anchoring and adjustment." For example, a subject is asked to estimate a person's height by estimating how far that height deviates from a fixed benchmark—from, say, five feet tall. If he thinks the person is five foot seven, he says "plus seven inches." The interesting fact is that the expressed estimates of a person's height will differ, depending upon the fixed benchmark. In theory, that particular benchmark should make no difference. If it is six feet instead of five feet, then the height of someone who is five foot seven can be said to deviate by "minus five inches" from that taller benchmark. Any given height can equally well be located by its distance from either benchmark. Nevertheless, the estimate of a given person's height by a group of judges who start with the five-foot benchmark will be less than the estimate by a group of judges who start with the benchmark of six feet. It is as though the benchmark exerts a gravitational attraction on the estimate, pulling that estimate

---

---

toward itself. It is similar, I think, when you approach a topic through the thought of another. Even when your own conclusions do deviate, they are “gravitationally” pulled toward those of your source. I cannot say, however, that I have consciously refrained from writing extensively on historical figures in order to avoid being trapped within their mode of thinking. I simply have not felt the urge to do so—except in the case of Socrates.

I do not work within the framework of any one philosopher’s thought but I do draw upon a kitbag of intellectual structures (in addition to the tools of mathematical logic which all contemporary philosophers use). Decision theory (and the associated material of utility theory, game theory, and social choice theory) is one structure I keep in mind, and use, and tinker with developing. There also is evolutionary theory, micro-economic theory, and inductive logic. I find such intellectual structures attractive and lovely, and I tend to scout other fields for already developed structures that might find philosophical application or might inspire the development of an analogous new philosophical structure.

Another structure I draw upon is presented in models of explanation, formulated by philosophers of science. Because of the presence of Carl Hempel at Princeton, the topic of explanation was central in the intellectual life of its philosophy graduate students. (And it was an admiration for Hempel’s work, promulgated in classes at Columbia with Sidney Morgenbesser and with Ernest Nagel, that was the major factor leading me to do graduate study at Princeton.) Hempel’s models of deductive-nomological and statistical explanation, their elaborations, difficulties, etc., was the *lingua franca* of study there. Everyone, no matter what his interests—unfortunately, there were no “hers” there then—felt obliged to keep current on that subject. It is no accident that my second book is entitled *Philosophical Explanations*, and that it focuses upon that notion, even though in ways different from Hempel. (*Anarchy, State, and Utopia* also was concerned with explanation: with fundamental explanation and with invisible-hand explanations of the state.)

Essays in the present volume also are concerned with explanation. “Invisible-Hand Explanations” (Essay 9) investigates that notion further, including new instances beyond those discussed and presented in *Anarchy, State, and Utopia*. “On Austrian Methodology” (Essay 5) sharpens and then examines the claim of “methodological individualism,” made by economists in the Austrian tradition, that all of the institutional structures of a society can and must be explained in terms of the actions of individuals. It also contains a critical examination of their



thesis that social science should be based upon an *a priori* theory of human choice. Thus the interests in theories of choice and in explanation come together.

A philosophical interest in explanation also is normative in part. We want to understand understanding, but we also seek explanations and therefore do not want to believe what could not be an adequate explanation. An account of what structure an explanation must possess thereby specifies our aim and guides us in our beliefs. “Experience, Theory, and Language” (Essay 7) pursues the theme of explanation in several directions: What explains why we cannot test isolated explanatory hypotheses but only large conjunctions of them? What is linguistics needed to explain?

The common intellectual interest in explanation was a fruitful stimulus for philosophical interchange at Princeton. At Columbia in the late 1950s, C. Wright Mills had just published *The Power Elite*, and no matter what their subject of study, students had to know what they thought about its thesis of a ruling elite in America. (In Essay 9, I speculate about an invisible-hand explanation of a ruling class.) Thomas Kuhn’s *Structure of Scientific Revolutions* held this focal position at many universities in the 1970s, and perhaps sociobiology did for a while afterwards. I sometimes ask students whether there is any serious work of non-fiction which all (or most) of their friends have read or feel they have to read. For the past fifteen years or so, the answer has been *no*. I do not know why there now is no such common intellectual idea or book that students debate with interest and excitement—have films displaced books as students’ common focus?—but it is unfortunate for students (and also for their teachers, who cannot take for granted the students’ passion over any recent ideas).

Philosophers seek intellectual understanding but philosophy, and what propels it, is not solely intellectual. This volume closes with several philosophical fictions (selections 19–22) that bring forth the emotional content of philosophical problems and ideas. These different aspects of philosophy are explored more fully, and more personally, in *The Examined Life*.

Philosophy is endlessly interesting and ample; *anything* can be thought about philosophically. The philosopher’s deepest urge, though, is to articulate and understand the ultimate basis and nature of things. What is the basis of our beliefs and our ethical principles, our standards of rea-