

## Acknowledgements to Volume I

My first inkling that there was something amiss with the Humean conception of the self came before I knew enough Western philosophy to call it that. I am grateful to Allen Ginsberg, Timothy Leary, Edward Sullivan and Swami Vishnudevananda for urging me to read the *Upanishads*, *Bhagavad Gita* and *Yoga Sutras* in 1965. I am grateful most of all to Phillip Zohn for his willingness to argue with me at length about the import of these texts, and for introducing me to Kant's *Critique of Pure Reason* in 1969, after reading an art text of mine on space and time ("Hypothesis") that inadvertently echoed its doctrine of transcendental idealism. The influence of all of these works on my thinking has informed my (you will pardon the pun) critical and skeptical approach to the Humean conception from the beginning.

This project has been in production for a very long time. The ancestor of the concept of pseudorationality introduced in Chapter VII of Volume II was my undergraduate Social Sciences Phi Beta Kappa Medal Honors Thesis, "Deception and Self-Deception" (City College of New York, 1974). I am grateful to Martin Tamny, Arthur Collins and David Weissman for their guidance and input at that stage. The ancestor of the analysis of cyclical and genuine preference in Chapter IV of Volume I and Chapter III of Volume II was Chapter II of my Second-Year Paper, "A Theory of Rational Agency" (Harvard University, 1976), for advice and comments on which I am indebted to John Rawls. Both ancestors liased in revised form in my dissertation, "A New Model of Rationality" (Harvard University, 1981) under John Rawls and Roderick Firth, in whose debt I permanently remain. Professor Firth provided the sounding board, the detailed and rigorous criticism, and the personal encouragement that has helped preserve my faith in the value of this project. I am deeply grateful for his involvement with it, and to have known him as a teacher and colleague.

My animated discussions with Professor John Rawls, both about my work and about the role of the utility-maximizing model in his work, were absolutely crucial to my conviction that I was on to something. His example as a scholar and teacher, the breadth and depth of his learning, and his magisterial achievement in *A Theory of Justice* have remained an inspiration to me in all of my work. I rank Rawls' achievement as a *theory-builder* – a philosopher who constructs substantive theories – with those of the middle and late Plato, Aristotle, Hobbes, Kant, and Habermas. A *critic*, by contrast, is a philosopher who mostly criticizes, improves upon, or demolishes theory-builders' theories. The quintessential critic would be the Slice-'em-and-dice-'em Socrates of the early Platonic dialogues. But some might also count St. Thomas Aquinas, Sidgwick, the later Wittgenstein, and Ryle among the philosophical critics, for different reasons. Philosophers may reasonably disagree about how some of these examples are to be classified, and most philosophers evince both theory-building and critical inclinations to varying degrees. But the distinction is nevertheless useful, because training in analytic philosophy is by default training in how to be a critic: We study the views of famous philosophers, learn how to detect areas of inconsistency or fault or lack, and then learn how to correct, supplement or level them.

There is no way to teach theory-building, except by encouraging students to have confidence in their intuitions. So if we happen to incline toward theory-building, we are pretty much on our own, because there are no ground rules about how to proceed. In developing the theory defended in this project, I was fortunate from the very beginning to receive good advice about how to proceed, from another theory-builder who had already been there and done that. The ground rules Rawls taught me were three:

- (1) Anchor your theory in relation to identifiable current problem(s) or controversies. Describe the problems, analyze some recent arguments that purport to solve them, and explain the ways in which these arguments fail. Then briefly sketch how your theory avoids these failures, so that your readers will be able to locate your theory on their own map of philosophical issues in a way that confers meaning and importance on it for them.
- (2) Anchor your theory relative to the views, with which you disagree, of other philosophers who have worked on the problem and have received attention for their efforts. Discuss those views, explain what's wrong with them, and describe how your theory avoids the criticisms you make of their views. Refer to these opposing views in developing your own, in order to bring your theory into connection with a larger, ongoing philosophical discussion among your peers.
- (3) Avoid cooking up a straw man to attack. Show that you take your opponents' views seriously, by making the best and most sympathetic case for them you possibly can, before showing how they disappoint despite your best efforts. The worst that can happen is that really understanding your opponents' views will convince you to modify your own.

In this project I have tried to honor Rawls' ground rules as best I can, in order to honor him as my teacher and their author, and also all of those others from whom I have learned so much by disputing their views in the following pages.

I have also benefited by teaching and discussing extensive portions of both volumes of this project with several generations of graduate students at the University of Michigan, Stanford, Georgetown and USCD – particularly Richard Dees, Jeffrey Kahn, Brian Leiter, Alan Madry, Minerva San Juan McGraw, David Reed-Maxfield, Joel Richeimer, Laura Shanner, Cristel Steinworth, and Sigrun Svavarsdottir; and fifteen years' worth of brilliant and feisty undergraduates at Wellesley College.

Chapter I of both volumes, "General Introduction to the Project: The Enterprise of Socratic Metaethics," was drafted during an unpaid leave of absence from Wellesley College during early 1998 and funded by an NEH College Teachers' Research Fellowship. The NEH support came at a crucial moment and I am deeply grateful for it. This chapter incorporates and modifies some passages and sections of my "Two Conceptions of the Self," published in *Philosophical Studies* 48, 2 (September 1985), 173-197 and reprinted in *The Philosopher's Annual VIII* (1985), 222-246. The discussion of Anglo-American philosophical practice that appears in Sections I.2 and I.3 benefited from comments by Anita Allen, Houston Baker, Paul Boghossian, Ann

Congleton, Joyce Carol Oates, Ruth Anna Putnam and Kenneth Winkler, as well as by members of the audience to the 1994 Greater Philadelphia Philosophy Consortium symposium, "Philosophy as Performance" at which these remarks were originally presented. The chapter received its near-final form during my tenure as a Research Scholar at the Getty Research Institute during the academic years 1998-1999. For providing me with all of the conditions I requested – some very idiosyncratic – as necessary for me to make substantial progress on this and many other parts of this project, my gratitude to the Institute knows no bounds. My debt of thanks to Brian Davis, Larry Hertzberg, Karen Joseph, Michael Roth, and Sabine Schlosser is particularly great. While there I also benefited a great deal from discussion of these and related topics with Reinhart Meyer-Kalkus. I would also like to thank Naomi Zack for her interest and willingness to publish an earlier version of this chapter, despite its length, in her edited collection, *Women of Color and Philosophy* (New York: Blackwell, 2000).

Chapter II, "The Belief-Desire Model of Motivation," was first drafted in 1981, while I was an Assistant Professor at the University of Michigan. I learned much from discussing the issues with Richard Brandt, William Frankena, Allan Gibbard, Jaegwon Kim, David Velleman, Nicholas White, and Stephen White, however much we in the end agreed to disagree. The chapter was redrafted in 1985, after having spent two wonderful and productive years at the Stanford University Philosophy Department on an Andrew Mellon Post-Doctoral Fellowship from 1982 to 1984. While there I benefited from discussing action theory with Michael Bratman and philosophy of science with John Duprés. Not until my year at the Getty Research Institute in 1998 was I able to return to this part of the project. The enthusiasm and dedication of the Getty staff in putting at my disposal all of the research and administrative assistance I needed, and more, to update and revise it in light of more recent discussions helped me to believe in the importance of doing so.

Work on Chapters III and IV was partially supported by the Mellon Post-Doctoral Fellowship, a Georgetown University Faculty Research Grant in 1988, and a Woodrow Wilson International Scholars' Fellowship in 1988-1989. At Georgetown I profited from discussions with Wayne Davis, Terry Pinkard and Henry Richardson. I also spent many, very fruitful hours discussing this chapter with colleagues at the University of Michigan, all Humeans to a man. To them I am most grateful of all for pulling no punches in their attempts to dissuade me from my views, from this chapter, and not least of all from this project. Had they not put those views to the test by resorting to every possible tactic of dissuasion, I would have had no proof that my views could withstand them. To have that proof – to know that my philosophical position was able to survive the gauntlets devised by some of the very best minds in the field – is the invaluable gift that I owe to them. I should particularly like to thank Allan Gibbard and David Velleman for conversation. Many other individuals have helped me in the writing of these two chapters, including Glenn Loury, Michael Slote, Robert Audi, David Levy, and especially Ned McClennen for extensive comments on earlier drafts. I was honored by the opportunity to present both to a group of trained economists at the Economics and Rhetoric Seminar, held at the

Academia Vitae in Deventer, The Netherlands, in June 2006. I am particularly grateful to Arjo Klamer, Dierdre McCloskey and P. W. Zuidhof for beneficial discussion that has improved their final form.

I was helped by discussion of the first draft of Chapter V at a University of Michigan Faculty Seminar, and particularly by comments from Richard Brandt, Arthur Burks, Allan Gibbard, Louis Loeb, Peter Railton, Nicholas White, and Stephen White. An earlier version was published under the title, "A Distinction Without a Difference," in *Midwest Studies in Philosophy VII: Social and Political Philosophy* (1982), 403-435. Chapter VI was completed and delivered to the Scholars' Seminar at the Getty Research Institute in November 1998, and to the Philosophy Department at the University of Minnesota in October 1999. I am grateful to both audiences for constructive comments and suggestions for improvement. Chapter VII is the outcome of over two decades of intense and satisfying – and, aside from Paul Coppock's insightful comments, largely solitary – labor on Thomas Nagel's *The Possibility of Altruism*. This part of the project taught me much more about patience and persistence than I ever could have expected when, after completing the chapter to my satisfaction a first time, I then allowed it to be irretrievably misplaced and had to reconstitute it from scratch (a scholar's worst nightmare in the pre-computer era). I am deeply grateful to the Woodrow Wilson International Research Center of the Smithsonian Institution for extending my International Scholars' Fellowship for a second year, 1989-90, so that I could do this.

Earlier versions of parts of Chapter VIII were published under the following titles: "Two Conceptions of the Self," *Philosophical Studies* 48, 2 (September 1985), 173-197, reprinted in *The Philosopher's Annual VIII* (1985), 222-246; "Moral Theory and Moral Alienation," *The Journal of Philosophy* LXXXIV, 2 (February 1987), 102-118; and "Michael Slote's *Goods and Virtues*," reviewed for *The Journal of Philosophy* LXXXIII, 8 (August 1986), 468-73. Work on this chapter was supported by the Mellon Fellowship. Earlier versions of the discussions of Frankfurt and Watson were presented to the Philosophy and Anthropology Group and the Department of Philosophy, both at the University of Michigan; and the Departments of Philosophy at Stanford, U. C. Berkeley, the University of Minnesota, and the University of Pennsylvania. I learned much from comments received on those occasions, and from detailed criticism and feedback by Michael Bratman, Jeffrey Evans and Allan Gibbard. I am equally grateful to Akeel Bilgrami, Jeffrey Evans and members of the Philosophy Department audiences at Wayne State University, Penn State, Georgetown, the University of California at San Diego, North Carolina State, Wesleyan, Memphis State, and the University of Minnesota for comments and criticism of my discussion of Williams.

Section 1 of Chapter IX was delivered in a slightly different form to the American Philosophical Association Pacific Division Convention in March 1995 in an Author Meets Critics session on Elizabeth Anderson's *Value in Ethics and Economics*; and later published under the title, "Making Sense of Value," in *Ethics* 106, 2 (April 1996), 525-537. An earlier version of Section 4 was delivered to the Moral Philosophy Colloquium at the American Philosophical Association Pacific Division Convention, Los Angeles, California in March 1986; and published under the title,

"Instrumentalism, Objectivity, and Moral Justification," in *American Philosophical Quarterly* 23, 4 (October 1986), 373-381.

Chapter X originated in September 1976 as a paper, "Continuing Persons and the Original Position," for John Rawls' graduate seminar in Moral Psychology, and I am grateful for his comments on it. A lecture by Joshua Cohen on Social Contract Theory in the Fall of 1978 at MIT had a salutary effect on Section IX.2. I have also benefited from criticisms of an earlier draft of this chapter by Peter Dalton. Parts were published under the title, "Personal Continuity and Instrumental Rationality in Rawls' Theory of Justice," in *Social Theory and Practice* 13, 1 (Spring 1987), 49-76. Work on this chapter was supported by a University of Michigan Rackham Faculty Fellowship and the Mellon Fellowship. The final draft was completed during my year at the Getty, as was the final draft of Chapter XI. Chapter XII, originally my term paper for John Rawls' Moral and Political Philosophy course at Harvard in the Spring of 1975, was also revised and completed during my wonderful and productive year at the Getty. I am grateful to Rawls, David Auerbach and Warner Wick for helpful criticisms of earlier drafts. An earlier version was published under the title, "Utility, Publicity and Manipulation," in *Ethics* 88, 3 (April 1978), 189-206.

For criticisms of an earlier draft of Chapter XIII I would like to thank Anita Allen, Annette Baier, Margaret Carroll, John Deigh, Michael Stocker, and Judith Jarvis Thomson. A protodraft of Chapter XIV originally formed the Appendix to my dissertation. I am grateful to Rawls for persuading me of the importance of dealing with Hume straight off, and for his criticisms and encouragement throughout. I also would like to thank Marcia Baron for comments on an earlier draft of this chapter. A more recent version was published under the title, "Hume on Rational Final Ends," in *Philosophy Research Archives XIV* (1988-89), 193-228. Robert Audi's comments and criticisms, and that of an unidentified referee, improved that version immensely. Chapter XV is entirely my own doing, and I have no one to blame but myself.

There is no way for me to express my gratitude and indebtedness to the very few individuals who provided encouragement and support during the final stretch of time in which I brought this project to completion. During two years of unpaid and extremely stressful medical leave from Wellesley College from Winter 2001 to Fall 2002, Bill Cain, Joe Feagin, Terry Irwin, Mark Kaplan, James Kodera, Ruth Barcan Marcus, Julie Matthaei, Reinhart Meyer-Kalkus, Susan Neiman, Robert Rubinowitz, Stephen Schiffer, Hedwig Saxenhuber, Georg Schollhammer, Ann Stephens, and Joan Weiner extended themselves beyond the bounds of collegial or moral obligation by letting me know, each in their own way, the importance and value to them that I do so. Their encouragement was crucial. My debt to Ruth Barcan Marcus for her steadfast friendship is beyond measure. The research and administrative help provided, under less than ideal conditions and great generosity of spirit, by Robert Del Principe was invaluable. His patience, resourcefulness, persistence and good humor in obtaining the sources I needed under the most stressful conditions, and tolerating without complaint twelve years' worth of my unending incipient hysteria has manifested both heroism and martyrdom of the highest order.

My debt to him is incalculable. Without the moral support of all of these good people this project would not have been possible. The final draft was begun under conditions of extreme personal hardship, in virtually complete solitude during the long, hot summer of 2003; and received its final form in the sheltering anonymity and safety of the city of Berlin in early 2008. I am profoundly grateful that it is there, and that I am there. For the unique opportunity to live and test the values defended in this project, I would like to thank the faculty and administration of Wellesley College; I commend this work in exile to them. For the strength, the solace and the sanctuary I have been blessed to find in reading, writing and teaching philosophy I am grateful most of all.