
**The Bounds of Reason:
Game Theory and the Unification
of the Behavioral Sciences**

**The Bounds of Reason:
Game Theory and the Unification
of the Behavioral Sciences**

Herbert Gintis

Princeton University Press
Princeton, New Jersey

Copyright ©2008 by Princeton University Press

Published by Princeton University Press, 41 William Street, Princeton, New Jersey 08540

In the United Kingdom: Princeton University Press, 3 Market Place, Woodstock, Oxfordshire
OX20 1SY

All Rights Reserved

Library of Congress Cataloging-in-Publication Data

Gintis, Herbert

The Bounds of Reason

Herbert Gintis

p. cm.

Includes bibliographical references and index.

ISBN 0-691-00942-2 (cloth: alk. paper)—ISBN 0-691-00943-0 (pbk.: alk. paper)

1. Game theory. 2. Economics, Mathematical. I. Title.

HB144.G562008

330'.01'5193—dc21 99-054923

This book has been composed in Times and Mathtime by the author

The paper used in this publication meets the minimum requirements of ANSI/NISO Z39.48-1992 (R1997) (Permanence of Paper)

www.pup.princeton.edu

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

10 9 8 7 6 5 4 3 2 1

(pbk.)

This book is dedicated to Morris and Henrietta
Malena, Gerson Gintis, and Flora and Melvin Greisler.
We get along without you very well. Of course we do.

There is no sorrow so great that it
does not find its background in joy.

Neils Bohr

Contents

Preface	xiii
1 Decision Theory and Human Behavior	2
1.1 Beliefs, Preferences, and Constraints	5
1.2 The Meaning of Rational Action	7
1.3 Why Are Preferences Consistent?	8
1.4 Time Inconsistency	9
1.5 Bayesian Rationality and Subjective Priors	12
1.6 The Biological Basis for Expected Utility	16
1.7 The Allais and Ellsberg Paradoxes	17
1.8 Risk and the Shape of the Utility Function	19
1.9 Prospect Theory	22
1.10 Heuristics and Biases in Decision-making	26
2 Game Theory: Basic Concepts	30
2.1 The Extensive Form	30
2.2 The Normal Form	33
2.3 Mixed Strategies	34
2.4 Nash Equilibrium	34
2.5 The Fundamental Theorem of Game Theory	35
2.6 Solving for Mixed Strategy Nash Equilibria	36
2.7 Throwing Fingers	37
2.8 Battle of the Sexes	38
2.9 The Hawk-Dove Game	40
2.10 The Prisoner's Dilemma	42
2.11 Alice, Bob, and the Choreographer	43
2.12 An Efficiency-enhancing Choreographer	45
2.13 The Correlated Equilibrium Solution Concept	45
3 Game Theory and Human Behavior	47
3.1 Self- and Other-regarding Preferences	48

3.2	Methodological Issues in Behavioral Game Theory	51
3.3	An Anonymous Market Exchange	54
3.4	The Rationality of Altruistic Giving	56
3.5	Conditional Altruistic Cooperation	58
3.6	Altruistic Punishment	59
3.7	Strong Reciprocity in the Labor Market	61
3.8	Altruistic Third Party Punishment	63
3.9	Altruism and Cooperation in Groups	66
3.10	Inequality Aversion	69
3.11	The Trust Game	72
3.12	Character Virtues	75
3.13	The Situational Character of Preferences	77
3.14	The Dark Side of Altruistic Cooperation	78
3.15	The Legitimacy of Cooperation: Cultural Variation	80
4	Rationalizability and Common Knowledge of Rationality	85
4.1	Epistemic Games	85
4.2	A Simple Epistemic Game	88
4.3	An Epistemic Battle of the Sexes	89
4.4	Dominated and Iteratedly Dominated Strategies	90
4.5	Eliminating Weakly Dominated Strategies	91
4.6	Rationalizable Strategies	92
4.7	Eliminating Strongly Dominated Strategies	94
4.8	Common Knowledge of Rationality	94
4.9	Rationalizability and Common Knowledge of Rationality	95
4.10	The Beauty Contest	96
4.11	The Traveler's Dilemma	97
4.12	The Modified Traveler's Dilemma	98
4.13	Global Games	99
4.14	CKR is a Condition, not a Premise	101
5	Extensive Form Rationalizability	104
5.1	Backward Induction and Dominated Strategies	104
5.2	Subgame Perfection	106
5.3	Subgame Perfection and Incredible Threats	107
5.4	The Surprise Examination	107
5.5	The Common Knowledge of Logicality Paradox	108
5.6	The Repeated Prisoner's Dilemma	109

5.7	The Centipede Game	110
5.8	CKR Fails off the Backward Induction Path	112
5.9	How to Play the Repeated Prisoner’s Dilemma	114
5.10	The Modal Logic of Knowledge	116
5.11	Backward Induction and Extensive Form CKR	117
5.12	Rationality and Extensive Form CKR	120
5.13	On the Nonexistence of CKR	121
6	The Mixing Problem: Purification and Conjectures	123
6.1	Why Play Mixed Strategies?	123
6.2	Harsanyi’s Purification Theorem	125
6.3	A Reputational Model of Honesty and Corruption	127
6.4	Purifying Honesty and Corruption	129
6.5	Epistemic Games: Mixed Strategies as Conjectures	130
6.6	Resurrecting the Conjecture Approach to Purification	131
7	Bayesian Rationality and Social Epistemology	133
7.1	From Battle to Ballet of the Sexes	134
7.2	The Choreographer Trumps Backward Induction	135
7.3	The Bourgeois Equilibrium of the Hawk-Dove Game	136
7.4	Convention as Correlated Equilibrium	136
7.5	Correlated Strategies and Correlated Equilibria	137
7.6	Correlated Equilibrium and Bayesian Rationality	139
7.7	The Social Epistemology of Common Priors	140
7.8	The Social Epistemology of Common Knowledge	142
7.9	Social Norms	144
7.10	Game Theory and the Evolution of Norms	144
7.11	The Merchants’ Wares	145
8	Common Knowledge and Nash Equilibrium	147
8.1	Conditions for Nash Equilibrium in Two-Player Games	147
8.2	A Three-player Counterexample	148
8.3	The Modal Logic of Common Knowledge	150
8.4	The Commonality of Knowledge	153
8.5	The Tactful Ladies	154
8.6	The Tactful Ladies and the Commonality of Knowledge	157
8.7	Agreeing to Disagree	158
8.8	The Demise of Methodological Individualism	162

9	Reflective Reason and Equilibrium Refinements	165
9.1	Perfect, Perfect Bayesian, and Sequential Equilibria	167
9.2	Incredible Threats	168
9.3	Unreasonable Perfect Bayesian Equilibria	171
9.4	The LBR Picks Out the Sequential Equilibrium	172
9.5	Selten's Horse: Sequentiality and LBR Disagree	172
9.6	The Spence Signaling Model	174
9.7	Irrelevant Node Additions	175
9.8	Improper Sequential Equilibria	176
9.9	Second-order Forward Induction	177
9.10	Beer and Quiche without the Intuitive Criterion	178
9.11	An Unreasonable Perfect Equilibrium	179
9.12	The Principle of Insufficient Reason	180
9.13	The Principle of Honest Communication	180
9.14	Discussion	181
10	The Analytics of Human Sociality	182
10.1	Explaining Cooperation: An Overview	182
10.2	Bob and Alice Redux	184
10.3	The Folk Theorem	186
10.4	The Folk Theorem with Imperfect Public Information	189
10.5	Cooperation with Private Signaling	194
10.6	One Cheer for the Folk Theorem	196
10.7	Altruistic Punishing in the Public Goods Game	197
10.8	The Failure of Models of Self-regarding Cooperation	201
11	The Evolution of Private Property	202
11.1	The Endowment Effect	202
11.2	The Territoriality	205
11.3	Property Rights in Young Children	208
11.4	Respect for Possession in Nonhuman Animals	208
11.5	Conditions for Private Property Equilibrium	211
11.6	Property and Anti-Property Equilibria	214
11.7	An Anti-Private Property Equilibrium	218
11.8	Conclusion	220
12	The Unification of the Behavioral Sciences	222
12.1	Introduction	222

12.2	Gene-culture Coevolution: The Biological Model	224
12.3	Biological and Cultural Dynamics	229
12.4	The Theory of Norms: The Sociological Model	230
12.5	Socialization and the Internalization of Norms	232
12.6	The Rational Actor: The Economic Model	234
12.7	Deliberative Choice: The Psychological Model	236
12.8	Application: Addictive Behavior	238
12.9	Game Theory: The Universal Lexicon of Life	239
12.10	Epistemic Game Theory and Social Norms	240
12.11	Society as Complex Adaptive System	242
12.12	Counterpoint: Biology	243
12.13	Counterpoint: Economics	244
12.14	Counterpoint: Psychology	245
12.15	Conclusion	246
13	Summary	248
14	Table of Symbols	250
	References	253
	Index	281

