
**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 and Oxford

Copyright ©2009 by Princeton University Press

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

In the United Kingdom: Princeton University Press, 6 Oxford Street, Woodstock, Oxfordshire
OX20 1TW

All Rights Reserved

Library of Congress Cataloging-in-Publication Data

Gintis, Herbert

The Bounds of Reason: Game Theory and the Unification of the Behavioral Sciences /
Herbert Gintis

p. cm.

Includes bibliographical references and index.

ISBN 978-0-691-14052-0 (hardcover: alk. paper) 1. Game theory. 2. Practical reason

3. Human Behavior. 4. Social Sciences—Methodology. 5. Psychology. I. Title.

HB144.G55.2009

519.4—dc22

2008036522

British Library Cataloging-in-Publication Data is available

The publisher would like to acknowledge the author of this volume for providing the
camera-ready copy from which this book was printed

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

Printed on acid-free paper.

pupress.princeton.edu

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

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	17
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	27
2 Game Theory: Basic Concepts	31
2.1 The Extensive Form	31
2.2 The Normal Form	34
2.3 Mixed Strategies	35
2.4 Nash Equilibrium	35
2.5 The Fundamental Theorem of Game Theory	36
2.6 Solving for Mixed-Strategy Nash Equilibria	37
2.7 Throwing Fingers	38
2.8 The Battle of the Sexes	39
2.9 The Hawk-Dove Game	41
2.10 The Prisoner's Dilemma	43
2.11 Alice, Bob, and the Choreographer	44
2.12 An Efficiency-Enhancing Choreographer	46
2.13 The Correlated Equilibrium Solution Concept	47
3 Game Theory and Human Behavior	48
3.1 Self- and Other-Regarding Preferences	49

3.2	Methodological Issues in Behavioral Game Theory	52
3.3	An Anonymous Market Exchange	55
3.4	The Rationality of Altruistic Giving	57
3.5	Conditional Altruistic Cooperation	59
3.6	Altruistic Punishment	60
3.7	Strong Reciprocity in the Labor Market	62
3.8	Altruistic Third-Party Punishment	64
3.9	Altruism and Cooperation in Groups	67
3.10	Inequality Aversion	71
3.11	The Trust Game	74
3.12	Character Virtues	76
3.13	The Situational Character of Preferences	78
3.14	The Dark Side of Altruistic Cooperation	80
3.15	Norms of Cooperation: Cross-Cultural Variation	81
4	Rationalizability and Common Knowledge of Rationality	86
4.1	Epistemic Games	86
4.2	A Simple Epistemic Game	89
4.3	An Epistemic Battle of the Sexes	91
4.4	Dominated and Iteratedly Dominated Strategies	92
4.5	Eliminating Weakly Dominated Strategies	93
4.6	Rationalizable Strategies	94
4.7	Eliminating Strongly Dominated Strategies	96
4.8	Common Knowledge of Rationality	97
4.9	Rationalizability and Common Knowledge of Rationality	98
4.10	The Beauty Contest	98
4.11	The Traveler's Dilemma	99
4.12	The Modified Traveler's Dilemma	100
4.13	Global Games	102
4.14	CKR Is an Event, Not a Premise	104
5	Extensive Form Rationalizability	106
5.1	Backward Induction and Dominated Strategies	106
5.2	Subgame Perfection	108
5.3	Subgame Perfection and Incredible Threats	109
5.4	The Surprise Examination	109
5.5	The Common Knowledge of Logicality Paradox	110
5.6	The Repeated Prisoner's Dilemma	111

5.7	The Centipede Game	112
5.8	CKR Fails Off the Backward Induction Path	114
5.9	How to Play the Repeated Prisoner's Dilemma	116
5.10	The Modal Logic of Knowledge	118
5.11	Backward Induction and Extensive Form CKR	119
5.12	Rationality and Extensive Form CKR	122
5.13	On the Nonexistence of CKR	123
6	The Mixing Problem: Purification and Conjectures	125
6.1	Why Play Mixed Strategies?	125
6.2	Harsanyi's Purification Theorem	127
6.3	A Reputational Model of Honesty and Corruption	129
6.4	Purifying Honesty and Corruption	132
6.5	Epistemic Games: Mixed Strategies as Conjectures	132
6.6	Resurrecting the Conjecture Approach to Purification	133
7	Bayesian Rationality and Social Epistemology	136
7.1	The Sexes: From Battle to Ballet	137
7.2	The Choreographer Trumps Backward Induction	138
7.3	Property Rights and Correlated Equilibrium	139
7.4	Convention as Correlated Equilibrium	140
7.5	Correlated Strategies and Correlated Equilibria	140
7.6	Correlated Equilibrium and Bayesian Rationality	142
7.7	The Social Epistemology of Common Priors	143
7.8	The Social Epistemology of Common Knowledge	145
7.9	Social Norms	147
7.10	Game Theory and the Evolution of Norms	147
7.11	The Merchants' Wares	148
8	Common Knowledge and Nash Equilibrium	150
8.1	Conditions for a Nash Equilibrium in Two-Player Games	150
8.2	A Three-Player Counterexample	151
8.3	The Modal Logic of Common Knowledge	153
8.4	The Commonality of Knowledge	156
8.5	The Tactful Ladies	157
8.6	The Tactful Ladies and the Commonality of Knowledge	160
8.7	Agreeing to Disagree	162
8.8	The Demise of Methodological Individualism	165

9	Reflective Reason and Equilibrium Refinements	168
9.1	Perfect, Perfect Bayesian, and Sequential Equilibria	170
9.2	Incredible Threats	171
9.3	Unreasonable Perfect Bayesian Equilibria	174
9.4	The LBR criterion picks out the sequential equilibrium	175
9.5	Selten's Horse: Sequentiality vs. the LBR criterion	175
9.6	The Spence Signaling Model	177
9.7	Irrelevant Node Additions	178
9.8	Improper Sequential Equilibria	179
9.9	Second-order forward induction	180
9.10	Beer and Quiche Without the Intuitive Criterion	181
9.11	An Unreasonable Perfect Equilibrium	182
9.12	The Principle of Insufficient Reason	183
9.13	The Principle of Honest Communication	183
9.14	Induction: Forward is Robust, Backward is Fragile	184
10	The Analytics of Human Sociality	185
10.1	Explaining Cooperation: An Overview	185
10.2	Bob and Alice Redux	187
10.3	The Folk Theorem	189
10.4	The Folk Theorem with Imperfect Public Information	192
10.5	Cooperation with Private Signaling	197
10.6	One Cheer For the Folk Theorem	199
10.7	Altruistic Punishing in the Public Goods Game	201
10.8	The Failure of Models of Self-Regarding Cooperation	204
11	The Evolution of Property Rights	205
11.1	The Endowment Effect	205
11.2	Territoriality	208
11.3	Property Rights in Young Children	211
11.4	Respect for Possession in Nonhuman Animals	211
11.5	Conditions for a Property Equilibrium	214
11.6	Property and Antiproperty Equilibria	217
11.7	An Antiproperty Equilibrium	221
11.8	Property Rights as Choreographer	224
12	The Unification of the Behavioral Sciences	225
12.1	Gene-Culture Coevolution: The Biological Model	227

12.2	Biological and Cultural Dynamics	232
12.3	The Theory of Norms: The Sociological Model	233
12.4	Socialization and the Internalization of Norms	235
12.5	Rational Choice: The Economic Model	237
12.6	Deliberative Choice: The Psychological Model	239
12.7	Application: Addictive Behavior	241
12.8	Game Theory: The Universal Lexicon of Life	242
12.9	Epistemic Game Theory and Social Norms	243
12.10	Society as a Complex Adaptive System	245
12.11	Counterpoint: Biology	246
12.12	Counterpoint: Economics	247
12.13	Counterpoint: Psychology	248
12.14	The Behavioral Disciplines Can Be Unified	249
13	Summary	250
14	Table of Symbols	252
	References	255
	Index	284

