

---

**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.G552009

519.3—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.

[press.princeton.edu](http://press.princeton.edu)

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

---

# Contents

Preface	xiii
<b>1 Decision Theory and Human Behavior</b>	<b>1</b>
1.1 Beliefs, Preferences, and Constraints	4
1.2 The Meaning of Rational Action	6
1.3 Why Are Preferences Consistent?	7
1.4 Time Inconsistency	8
1.5 Bayesian Rationality and Subjective Priors	11
1.6 The Biological Basis for Expected Utility	16
1.7 The Allais and Ellsberg Paradoxes	16
1.8 Risk and the Shape of the Utility Function	18
1.9 Prospect Theory	21
1.10 Heuristics and Biases in Decision Making	26
<b>2 Game Theory: Basic Concepts</b>	<b>30</b>
2.1 The Extensive Form	30
2.2 The Normal Form	33
2.3 Mixed Strategies	34
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	38
2.9 The Hawk-Dove Game	39
2.10 The Prisoner's Dilemma	40
2.11 Alice, Bob, and the Choreographer	41
2.12 An Efficiency-Enhancing Choreographer	43
2.13 The Correlated Equilibrium Solution Concept	44
<b>3 Game Theory and Human Behavior</b>	<b>45</b>
3.1 Self- and Other-Regarding Preferences	46

3.2	Methodological Issues in Behavioral Game Theory	49
3.3	An Anonymous Market Exchange	52
3.4	The Rationality of Altruistic Giving	54
3.5	Conditional Altruistic Cooperation	56
3.6	Altruistic Punishment	57
3.7	Strong Reciprocity in the Labor Market	59
3.8	Altruistic Third-Party Punishment	61
3.9	Altruism and Cooperation in Groups	64
3.10	Inequality Aversion	68
3.11	The Trust Game	71
3.12	Character Virtues	73
3.13	The Situational Character of Preferences	75
3.14	The Dark Side of Altruistic Cooperation	77
3.15	Norms of Cooperation: Cross-Cultural Variation	78
<b>4</b>	<b>Rationalizability and Common Knowledge of Rationality</b>	<b>83</b>
4.1	Epistemic Games	83
4.2	A Simple Epistemic Game	86
4.3	An Epistemic Battle of the Sexes	87
4.4	Dominated and Iteratedly Dominated Strategies	88
4.5	Eliminating Weakly Dominated Strategies	89
4.6	Rationalizable Strategies	90
4.7	Eliminating Strongly Dominated Strategies	92
4.8	Common Knowledge of Rationality	93
4.9	Rationalizability and Common Knowledge of Rationality	94
4.10	The Beauty Contest	94
4.11	The Traveler's Dilemma	95
4.12	The Modified Traveler's Dilemma	96
4.13	Global Games	98
4.14	CKR Is an Event, Not a Premise	100
<b>5</b>	<b>Extensive Form Rationalizability</b>	<b>102</b>
5.1	Backward Induction and Dominated Strategies	102
5.2	Subgame Perfection	104
5.3	Subgame Perfection and Incredible Threats	105
5.4	The Surprise Examination	105
5.5	The Common Knowledge of Logicality Paradox	106
5.6	The Repeated Prisoner's Dilemma	107

5.7	The Centipede Game	108
5.8	CKR Fails Off the Backward Induction Path	110
5.9	How to Play the Repeated Prisoner's Dilemma	112
5.10	The Modal Logic of Knowledge	114
5.11	Backward Induction and Extensive Form CKR	115
5.12	Rationality and Extensive Form CKR	118
5.13	On the Nonexistence of CKR	119
<b>6</b>	<b>The Mixing Problem: Purification and Conjectures</b>	<b>121</b>
6.1	Why Play Mixed Strategies?	121
6.2	Harsanyi's Purification Theorem	123
6.3	A Reputational Model of Honesty and Corruption	125
6.4	Purifying Honesty and Corruption	128
6.5	Epistemic Games: Mixed Strategies as Conjectures	128
6.6	Resurrecting the Conjecture Approach to Purification	129
<b>7</b>	<b>Bayesian Rationality and Social Epistemology</b>	<b>132</b>
7.1	The Sexes: From Battle to Ballet	133
7.2	The Choreographer Trumps Backward Induction	134
7.3	Property Rights and Correlated Equilibrium	135
7.4	Convention as Correlated Equilibrium	136
7.5	Correlated Strategies and Correlated Equilibria	136
7.6	Correlated Equilibrium and Bayesian Rationality	138
7.7	The Social Epistemology of Common Priors	139
7.8	The Social Epistemology of Common Knowledge	141
7.9	Social Norms	143
7.10	Game Theory and the Evolution of Norms	143
7.11	The Merchants' Wares	144
<b>8</b>	<b>Common Knowledge and Nash Equilibrium</b>	<b>146</b>
8.1	Conditions for a Nash Equilibrium in Two-Player Games	146
8.2	A Three-Player Counterexample	147
8.3	The Modal Logic of Common Knowledge	149
8.4	The Commonality of Knowledge	152
8.5	The Tactful Ladies	153
8.6	The Tactful Ladies and the Commonality of Knowledge	156
8.7	Agreeing to Disagree	158
8.8	The Demise of Methodological Individualism	161

<b>9</b>	<b>Reflective Reason and Equilibrium Refinements</b>	164
9.1	Perfect, Perfect Bayesian, and Sequential Equilibria	166
9.2	Incredible Threats	167
9.3	Unreasonable Perfect Bayesian Equilibria	170
9.4	The LBR criterion picks out the sequential equilibrium	171
9.5	Selten's Horse: Sequentiality vs. the LBR criterion	171
9.6	The Spence Signaling Model	173
9.7	Irrelevant Node Additions	174
9.8	Improper Sequential Equilibria	175
9.9	Second-Order Forward Induction	176
9.10	Beer and Quiche Without the Intuitive Criterion	177
9.11	An Unreasonable Perfect Equilibrium	178
9.12	The Principle of Insufficient Reason	179
9.13	The Principle of Honest Communication	179
9.14	Induction: Forward is Robust, Backward is Fragile	180
<b>10</b>	<b>The Analytics of Human Sociality</b>	181
10.1	Explaining Cooperation: An Overview	181
10.2	Bob and Alice Redux	183
10.3	The Folk Theorem	185
10.4	The Folk Theorem with Imperfect Public Information	188
10.5	Cooperation with Private Signaling	193
10.6	One Cheer For the Folk Theorem	195
10.7	Altruistic Punishing in the Public Goods Game	197
10.8	The Failure of Models of Self-Regarding Cooperation	200
<b>11</b>	<b>The Evolution of Property Rights</b>	201
11.1	The Endowment Effect	201
11.2	Territoriality	204
11.3	Property Rights in Young Children	207
11.4	Respect for Possession in Nonhuman Animals	207
11.5	Conditions for a Property Equilibrium	210
11.6	Property and Antiproperty Equilibria	213
11.7	An Antiproperty Equilibrium	217
11.8	Property Rights as Choreographer	220
<b>12</b>	<b>The Unification of the Behavioral Sciences</b>	221
12.1	Gene-Culture Coevolution: The Biological Model	223

12.2	Culture and Physiology of Human Communication	228
12.3	Biological and Cultural Dynamics	229
12.4	The Theory of Norms: The Sociological Model	231
12.5	Socialization and the Internalization of Norms	233
12.6	Rational Choice: 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 a Complex Adaptive System	242
12.12	Counterpoint: Biology	244
12.13	Counterpoint: Economics	245
12.14	Counterpoint: Psychology	245
12.15	The Behavioral Disciplines Can Be Unified	247
<b>13</b>	<b>Summary</b>	248
<b>14</b>	<b>Table of Symbols</b>	250
	References	253
	Index	283