Individuality and Entanglement: The Moral and Material Basis of Social Life

Herbert Gintis
Dedicated to Gary Becker, Jack Hirshliefer and Edward O. Wilson for their ability to explore deep connections across the various regions of human social life. I love these guys.

Play is a uniquely adaptive act, not subordinate to some other adaptive act, but with a special function of its own in human experience.

Johan Huizinga

How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it except the pleasure of seeing it.

Adam Smith
# Contents

## Overview

1 Gene-culture Coevolution
   1.1 Culture Determines Biological Fitness  
   1.2 Modeling Gene-Culture Coevolution  
   1.3 The Physiology of Human Communication  

2 Zoon Politikon: The Evolutionary Origins of Human Socio-political Systems  
   2.1 Accounting for Human Exceptionalism  
   2.2 Models of Political Power  
   2.3 The Moral Basis of Modern Political Systems  
   2.4 The Socio-political Structure of Primate Societies  
   2.5 The Evolutionary History of Primate Societies  
   2.6 Fire and Social Sharing  
   2.7 From Gatherer to Scavenger  
   2.8 Primitive Lethal Weapons  
   2.9 Warfare  
   2.10 Dominance and Reverse Dominance Hierarchies  
   2.11 Are there Egalitarian Non-human Primates?  
   2.12 Governance by Consent  
   2.13 Cooperative Mothering: The Evolution of Prosociality  
   2.14 Lethal Weapons and Egalitarianism  
   2.15 The Long-term Evolution of Human Sociality  

3 Distributed Effectivity: Political Theory and Rational Choice  
   3.1 Public and Private Spheres  
   3.2 Private and Public *Persona*  
   3.3 Social Rationality  
   3.4 The Social Rationality of Voter Turnout  
   3.5 The Logic of Distributed Effectivity
3.6 Other Explanations of Voter Behavior 61

4 Power and Trust in Competitive Markets 64
  4.1 The Short-side Power Principle 65
  4.2 Power in Competitive Markets 70
  4.3 Trust and Integrity 71
  4.4 Reputational Equilibrium 75
  4.5 Contingent Renewal Labor Markets 77
  4.6 I’d Rather Fight than Switch 81
  4.7 Regulating Market Power 84

5 Rational Choice 85
  5.1 The Axioms of Rational Choice 88
  5.2 Choice Under Uncertainty 91
  5.3 Limitations of the Rational Actor Model 93
  5.4 Bayesian Updating with Radical Uncertainty 96
  5.5 State-Dependent Preferences 97

6 An Analytical Core for Sociology 100
  6.1 The Social Division of Labor 102
  6.2 The Socio-Psychological Theory of Norms 106
  6.3 Socialization and the Internalization of Norms 108
  6.4 A Model of Norm Internalization 109
  6.5 The Evolution of Social Conventions 111
  6.6 The Omniscient Choreographer and Moral Preferences 114
  6.7 The Evolution of Norm Internalization 117
  6.8 Social Norms as Correlated Equilibria 119
  6.9 Networked Minds and Distributed Cognition 120
  6.10 General Social Equilibrium: An Example 121
  6.11 Class Structure in General Social Equilibrium 124
  6.12 Resurrecting Sociological Theory 128

7 The Theory of Action Reclaimed 130
  7.1 The Moral and Material Bases of Choice 135
  7.2 Carving an Academic Niche for Sociology 137
  7.3 The Parsonian Synthesis 140
  7.4 The Attempt to Separate Morality and Rationality 142
  7.5 Why Did Parsons Fail? 147
7.6 The Flourishing of Middle-Range Theory 149
7.7 High Theory as Interpretation 152

8 The Evolution of Property 154
8.1 The Endowment Effect 155
8.2 Territoriality 157
8.3 Property Rights in Young Children 160
8.4 Respect for Possession in Nonhuman Animals 160
8.5 Conditions for a Property Equilibrium 163
8.6 Property and Antiproperty Equilibria 166
8.7 An Antiproperty Equilibrium 171
8.8 Property Rights as Choreographer 173

9 The Sociology of the Genome 175
9.1 The Core Genome 179
9.2 Inclusive Fitness and Hamilton’s Rule 184
9.3 Kin Selection and Inclusive Fitness 190
9.4 A Generalized Hamilton’s Rule 194
9.5 Harmony and Disharmony Principles 196
9.6 The Utterly Selfish Nature of the Gene 197
9.7 Prosocial Genes Maximize Inclusive Fitness 199
9.8 The Boundaries of Inclusive Fitness Maximization 200
9.9 The One Mutation at a Time Principle 201
9.10 The Phenotypic Gambit 202
9.11 The Anatomy of the Core Genome 203
9.12 Explaining Social Structure 206
A1 Hamilton’s Rule with General Social Interaction 207

10 Gene-Culture Coevolution and the Internalization of Norms 215
10.1 Socialization and Fitness-Enhancing Norms 217
10.2 Altruism 221
10.3 Copying Phenotypes: The Replicator Dynamic 225
10.4 Why is Altruism Predominantly Prosocial? 227
10.5 The Power of Altruistic Punishment 230
10.6 Final Considerations 231

11 The Economy as Complex Dynamical System 234
11.1 The Walrasian Economy 236
Contents

11.2 Exchange Processes with Private Prices 238
11.3 Strict Nash Equilibria and Stability 238
11.4 The Characterization of Stable Exchange Processes 239
11.5 A Markov Implementation of Walrasian Dynamics 244
11.6 Complex Dynamics 248

12 The Future of the Behavioral Sciences 249
12.1 What are Analytical Foundations? 253
12.2 The Unification of the Behavioral Sciences 256
12.3 Gene-Culture Coevolution 259
12.4 The Rational Actor Model 261
12.5 Game Theory 266
12.6 Complexity 269

Acknowledgements 274
References 275
Subject Index 328
Author Index 333
Overview

The eternal mystery of the world is its comprehensibility.

Albert Einstein

There are more things in Heaven and Earth, Horatio, than are dreamt of in your philosophy.

William Shakespeare

HG: Who are you?
C: I am the Choreographer
HG: What are you?
C: An emergent property of social systems.
HG: What are you here for?
C: To ask you some questions.

Choreographer interview

I develop several related themes in this book. The first theme is that society is a game with rules, people are players in this game, and politics is the arena in which we affirm and change these rules. Unlike the rules in standard game theory, however, social rules are continually contested by players allying to scrap old rules and create new rules to serve their purposes.

Rule contestation is conflictual. People do not always agree on what rules should govern their lives. Societies usually have rules concerning how the rules themselves get changed, but people do not always play by the rules when they want to change the rules. Individuality is central to our species because the rules do not change through inexorable macrosocial forces or through a biochemical dynamic of gene substitution. Rather, individuals band together to change the rules.

Everything distinctive about human social life flows from the fact that we construct and then play social games. Other animals are playful, but they do not make up the games they play. They do not change the rules of the game to suit their purposes. Similarly, other animals live in societies. But the rules of the game for nonhuman societies are inscribed in the genome of the species, while ours is not. Other animals do not change these rules to suit their purposes. We do.
Second theme: playing games with socially constructed rules requires a *moral sense*. Humans treat some rules a purely *conventional*, such as what side of the road to drive on. But many social rules are treated as *morally binding*. Especially when there is general agreement about the rules, people gain satisfaction by playing by the rules, ashamed when they break the rules, and are offended when others break the rules. Indeed, individuals often reward others who play by the rules and punish others who break the rules. They often do this even at considerable personal cost, without any prospect of some personal material return in the future.

Even societies that lack government, judges, juries, and jails effectively reward and punish behavior. This is why Thomas Hobbes was profoundly incorrect in saying that before the emergence of modern institutions, life was “solitary, poor, nasty, brutish, and short” (Hobbes 1968[1651]). Our moral sense was developed long before there were courts, jails, and teachers to lecture us on morality. Despots would love to be able to determine the moral sense of their subjects, but they can never rest securely in their beds knowing that they have done so. Morality has a dynamic that, at least up to now, cannot be controlled by states. New technologies may change this. Hopefully not.

Third theme: Our minds are socially *entangled*. Cognition (what people know and believe) does not reside in individual minds. Minds are *social networked* with cognition distributed *across* social networks. An individual in making a decision draws not only on his own information and personally ascertained beliefs, but on information and beliefs located in the minds with which he is networked as well. Entangled minds produce social behavior that is *rational*, as we shall see, although it does not always conform to the standard axioms of rational choice in contemporary decision theory. We extend rational choice theory to deal with entangled minds.

Fourth theme: Human morality has an important *nonconsequentialist* dimension: individuals generally do the right thing, when they do the right thing, not because of the *personal or social consequences* of their actions, but simply because they believe it is the right thing to do. I am honest, when I am honest, not simply because I care about the people with whom I daily interact (although I may care about them), and not simply because I am afraid to be caught cheating (although I may indeed be afraid), but because *honesty is the right way to be*. Being dishonest dirties me, like not brushing my teeth. I vote not because my vote can possibly alter the outcome of an election, but because contributing to the election of good leaders is the
right thing to do, even if one makes no difference to the outcome. I vote for the candidate I want to win because this is my contribution to a collective effort, even though I know my vote will not affect the outcome of the election. This nonconsequentialist moral sense is akin to Aristotle’s notion of virtue (Aristotle 2002[350BC]), and even more closely to Immanuel Kant’s (2012[1797]) notion of the categorical imperative.

The nonconsequentialist dimension of morality is connected to the entanglement of minds. Our minds work with a form of causal efficacy that is collectively rational as opposed to being instrumentally rational. I call this distributed effectivity, a principle analyzed in Chapter 3. An example is the belief that one has helped a candidate for political office to win the election, despite the obvious and well understood fact that the candidate would have won even if one had not voted, or even if one had voted for a different candidate.

Fifth theme: careful inspection of human behavior reveals that there are three distinct dimensions of human preferences: self-regarding, other-regarding, and universal. Self-regarding preferences deal with what we want for our personal selves, including consumption, leisure, wealth, love, health, and the respect of others. Other-regarding preferences deal with our care about other people’s well-being. Wanting to help someone for whom we feel compassion, or hurt someone who has offended or annoyed us, are other-regarding preferences. Universal morality is neither self- nor other-regarding. Universal moral principles can have consequences, as when I help a stranger in need, but include Aristotle’s nonconsequentialist character virtues, such as courage, truthfulness, and loyalty. Nonconsequentialist moral principles correspond to Kant’s notion of acting on purely moral grounds.

As a scientist rather than a philosopher, I am of course not concerned with what people should value, but rather with what they do value. I am particularly not concerned with the relative moral coherence of utilitarian, deontological, and virtue theories. Real human beings mix and match, and so will I.

Sixth theme: individuals trade off among these three sorts of values. Someone, for instance, may sacrifice personal reward (self-regard) on behalf of the well-being of others (other-regard), and/or on behalf of universal moral principles. Similarly, an individual may lie (violation of virtue) out of regard for the feelings of others (other-regard) or because he can gain materially therefrom (self-regard). Aristotle understood this idea by cham-
pioning the mean over both excess and deficiency, whereas Kant went off track by asserting a rigorous nonconsequentialism in which perfectly moral behavior does not take self-interest or the welfare of others into account. We discover, through game-theoretic experiments in the laboratory (Fehr and Gintis 2007) and in the field (Herbst and Mas 2015) not only that people trade off, but these trade-offs can be modeled in terms of rational choice theory. Indeed, unless an individual satisfies some basic principles of rational choice, it is virtually meaningless to say he is moral or immoral.

There are whole academic disciplines in which rational choice is actively rejected by most researchers. These include psychology, sociology, and anthropology. The grounds for rejection are uniformly unpersuasive and ill-informed. Even great thinkers have offered embarrassingly shoddy reasons for rejecting rational choice. Rejection of rational choice theory accounts in part for these fields’ lack of a coherent analytical core. The current vogue of behavioral economists also include many who spend considerable time arguing that people are not rational (Ariely 2010). While there are certainly systematic violations of the rational actor model (see Chapter 5), rationality is generally strongly empirically supported. There is absolutely no way to do serious social theory without recognizing that human behavior is purposive and can generally be modeled as rational choice. Denying this is like asserting that people cannot see what is really there because there are optical illusions.

Seventh theme: the above aspects of human behavior can be understood only through transdisciplinary research. Many social scientists take pride in their deep knowledge of one discipline uncontaminated by ‘foreign’ ideas. Some of these researchers do fine work, but the above themes cannot be effectively handled in this way. There is therefore a pressing need for a thorough cleansing of the behavioral disciplines of the glaring incompatibilities among them. For it is precisely these incompatibilities that prevent effective transdisciplinary research. I stressed this in The Bounds of Reason (Princeton, 2009), and the theme is reflected in virtually every chapter of this book.

The central fact of the behavioral sciences is that social species live in complex social systems that must be studied through the lens of evolutionary theory. The first common background of all the behavioral disciplines is the scientific method, mathematical modeling, statistical testing, and where formal analytical methods fail, and fail they must at some point in concep-
tualizing complex systems, historical insight, ethnographic description, and agent-based modeling become critical.

Chapter 1 (Gene-culture Coevolution), explores the evolutionary dynamic that applies to any species for which epigenetic information takes the form of culture that accumulates reliably and long-term across generations. That applies strongly only to *Homo sapiens*. This chapter is based on a paper that appeared in special issue of the *Philosophical Transactions of the Royal Society B* (Biological Sciences) in 2011. The main point is that culture includes techniques and social norms that determine what genes will be evolutionarily rewarded in a given society. Therefore human genes are the product of human culture as much as the reverse (Boyd and Richerson 1985). I give as an extended example the evolution of the physiology of communication in humans (Section 1.3). This example is important because it involves the development of key human genetic capacities that are the product of social structure as opposed to simple individual adaptation. Some researchers have called this group selection, but I avoid it, being rather fed up with the arcane and futile arguments that have arisen over the use of this term. The physiology of communication in humans simply is what it is—a product of social structure.

The natural habitats for gene-cultural coevolution are the fields of sociology and anthropology. It is curious, then, that the most powerful models of cultural structure and evolution come from biology. It is equally curious that economic theory has no concept of culture whatever, and as I show in my book *The Bounds of Reason* (Princeton, 2009), deploys an implausible form of methodological individualism that attempts, quite unsuccessfully, to explain human behavior on the basis of rational choice and game theory alone. Rational choice theory and game theory are simply mathematical constructs. Standing alone, like vector spaces or semi-simple algebras, they say nothing.

Chapter 2 (Zoon Politikon: The Evolutionary Origins of Human Socio-political Systems) addresses the evolutionary roots of human sociality. The argument is a direct application of gene-culture evolutionary theory. The material taken from an interdisciplinary collaboration with coauthors Carel van Schaik and Christopher Boehm, primatologists and anthropologists. The paper appeared as a target article in *Current Anthropology* in 2015. My inspiration for this analysis was two-fold. First, studying animal behavior gives the realization that *Homo sapiens* is also *Homo ludens*—Man, the game player. This is the root of political theory, where politics is viewed as
a process of creating, maintaining, and transforming the social rules of the
game. Second, an exciting genetic study by Shultz et al. (2011) in Nature
using coalescent theory specified the most likely socio-political structure of
our last common ancestor with the primates. This allowed my coauthors
and I to chart a plausible evolutionary cascade that led from this common
ancestor, through the hominin line, to *Homo sapiens*.

Chapter 3 (*Distributed Effectivity: Political Theory and Rational Choice*)
explains the implications of the moral predispositions, both personal and
political, that have emerged from the coevolutionary process described in
Chapters 1 and 2. This chapter is based on a paper that appeared in the
*Journal of Economic Behavior and Organization* in 2015. The value of the
rational actor model becomes clear in this chapter as it allows us to explore
in some depth the interplay between self-regarding, other-regarding, and
universal preferences.

In developing the rational actor model throughout this book I argue that
the notion of the *subjective prior* that is standard in rational decision theory
(Savage 1954) must be expanded to include entanglement: *human minds are
networked* and *cognition is distributed across minds*. Among the themes ex-
plored in this chapter is that in dealing with political participation, individ-
uals reason not using standard consequentialist logic, but rather a logic of
distributed *effectivity*. From this understanding flows the insight that such
phenomena as voting in mass elections and participating in large collective
actions is both rational and has a deeply moral dimension. For instance,
each individual supporter believes he helped elect the winner even if the
outcome would have been no different had he not voted. And they are of
course right. The alternative, which is that no one helped the winner get
elected, is just silly.

Chapter 4 (*Power and Trust in Competitive Markets*) forges links among
economic, political, and sociological theory by showing how incomplete
contracts, which are ubiquitous and unavoidable in market economies, lead
to the power relations studied in political theory and the moral behaviors
characteristic of highly productive advanced economic systems, including
legitimacy, integrity, trust, and trustworthiness. The notion that economic
theory can provide the conditions for economic success without specifying
the distribution of political power in the economy, or without specifying
the moral commitments of market participants, is simply and dramatically
incorrect. *Sic transit gloria* neoclassical economics. The material in this
Chapter draws on several years of research with Samuel Bowles (Bowles and Gintis 1988, 1990, 1993, 1999).

Chapter 5 (Rational Choice) is an explication and a defense of the rational actor model as foundational in analyzing human behavior. Rational decision theory is de rigeur in economics and biology but rather strongly spurned by the majority of researchers in the other behavioral disciplines. This chapter explains why this is a serious error. The critics of rational choice invariably—and I mean invariably—misrepresent the theory. In particular, it does not imply that rational actors are egoists, or that they maximize pleasure, or in fact, that they maximize anything. It is useful to keep in mind at all times that the rational choice model is a key tool of animal behavior theory (Maynard Smith 1982; Alcock 1993). It is difficult to consider a creature lacking nociceptors (e.g., most insects) as a happiness maximizer, and yet the rational actor model is very illuminating even for such creatures. They maximize fitness. This chapter also explores extensions of the rational actor model that take account of social rationality and the character of beliefs as products of social networking. The most important extension replaces the subjective prior assumption with the notion that minds are socially networked and cognition is distributed across networks.

Chapter 6 (An Analytical Core for Sociology) is based on a target article that appeared in the Review of Behavioral Economics in 2015, coauthored with Dirk Helbing. I have long been a fan of sociological theory and cannot imagine that a serious behavioral scientist could get away with ignoring this immense body of social knowledge. Yet, it is not taken seriously by most researchers outside sociology. The reason is that this material is not properly theoretically grounded, so its results are difficult to interpret and integrate with the other behavioral disciplines. I think of this chapter as an extension of Talcott Parsons’ abandoned quest for an analytical core for sociology. The most ambitious aspect of this chapter is the suggestion that the general equilibrium model of economic theory be extended to a general social equilibrium model based on the actor-role model of social participation. Also very important is the idea of treating the game-theoretic notion of a social norm as choreographer, developed at greater length in Gintis (2009a).

This argument will be called economic imperialism by some because it suggests that all the behavioral disciplines, including sociology, social and cognitive psychology, and anthropology, be organized theoretically using the analytical tools of rational choice and game theory. My retort is that
these fields are in such serious need of a unifying theoretical framework that a little imperialism from more successful fields should be welcome. In fact, however, I am equally happy to apply sociological and sociobiological imperialism to the reform of economic theory. This book thoroughly rejects neoclassical economic theory’s methodological individualism. By embracing an evolutionary perspective, we can understand why many previous applications of economic theory outside its traditional domain have an outlandish and bizarre character.

The great sociologist Talcott Parsons attempted as a young man to develop an analytical core for sociological theory (Parsons 1937), but he abandoned the attempt for reasons discussed in Chapter 7 (The Theory of Action Reclaimed). The most important reasons, I believe, were his lack of mathematical training, the fact that game theory was at the time in an infant stage, and the strong fear that sociologists had of being swallowed up by economic theory.

I never met Talcott Parsons, though he had an office nearby when I began teaching at Harvard. I dedicated my economics Ph.D. dissertation to Karl Marx and Talcott Parsons. Perhaps Parsons was uncomfortable being approved of by a raving anti-Vietnam-war dissident. At any rate, he wrote an article in the Quarterly Journal of Economics (Parsons 1975) criticizing some of my work (his criticisms were remarkably off-track). I took the occasion of a reply to return the favor (Gintis 1975). Looking back, I find my criticisms of Parsons valid but unconstructive. This chapter says what I should have said forty years ago in the Quarterly Journal of Economics.

Chapter 8 (The Evolution of Property), based on Gintis (2007b), argues that there is a biologically rooted concept of property that is prehistorical in the sense of not depending on formal social institutions, yet ubiquitous in daily life. We can understand the social norm of rights of ownership by virtue of its close relationship to territoriality in many birds and animals. A natural property right is enforced endogenously within the animal society itself: the incumbent is prepared to fight harder to keep his property than an intruder is willing to fight to take it from him, and the intruder behaves as if he knows that this is the case. Without secure property rights, many social species would simply not exist. Take nesting birds: a pair of robins would not spend two days building a nest if they were likely to be confronted with another pair, equally capable and willing to fight to take the completed nest from them.
In their boisterous attack on inclusive fitness theory (described and analyzed in Chapter 9), Nowak et al. (2010) wisely note that when one actually models the social behavior of a species, issues concerning levels of selection simply do not arise. Interpreting territoriality as a natural property right is a wonderful example of their point. It is impossible to understand territoriality—when it occurs and when it does not—in terms of either genes, or individuals. Territoriality is an emergent property of social systems. Nevertheless, all of territorial dynamics can be described using gene-level fitness parameters.

Moreover, the territoriality example violates one of the cherished tenets of group selection theory. This is the notion that group selection necessarily involves altruistic behavior. There is no altruism in the models developed in this chapter, yet respect for property emerges through biological evolution at the level of the population itself.

Proper analytical modeling of social behavior allows us to abandon the traditional levels of selection debate. In Chapter 9 (*The Sociology of the Genome*) I suggest an alternative in which a social species is characterized by its core genome, which specifies the social structure of the species. This gives us the proper balance between the gene’s-eye view that has seized the hearts of so many evolutionary biologists and the multilevel selection view that has captured the minds of so many others. Believe it or not, there is a correct view, arrived at by applying standard sociological principles to biological species, that includes the valid insights of both camps.

Chapter 10 (*Gene-Culture Coevolution and the Internalization of Norms*) forges a strong link between sociology and sociobiology by modeling the evolutionary emergence of the human capacity to internalize social norms. This chapter is based on a paper published in the *Journal of Theoretical Biology* in 2003. An internal norm is a pattern of behavior enforced in part by internal psychological self-sanctions, such as shame or guilt, as opposed to purely external sanctions, such as material rewards and punishment. The ability to internalize norms is widespread among humans, although in some so-called ‘sociopaths,’ this capacity is diminished or lacking. Our model shows that if an internal norm is fitness enhancing, then for plausible patterns of socialization, an allele that promotes the internalization of norms is evolutionarily stable.

This framework shows that social norms to which adherence is costly to the individual can ‘hitchhike’ on the general tendency of respecting social norms to be personally fitness-enhancing. An important implication of this
analysis is that the standard argument that Darwinian fitness maximization implies that organisms maximize inclusive fitness it incorrect. But of course we already knew that from Chapter 9.

Chapter 11 (*The Economy as Complex Dynamical System*) applies complexity theory and evolutionary game theory to general equilibrium theory. The Walrasian general equilibrium model of competitive market systems is in fact a *complex, dynamical, nonlinear, adaptive system* that simply fails to fit into the mold of classical dynamic systems theory. This chapter shows that when the market economy is modeled as a complex system, problems that traditional economic theory has been unable to penetrate for nearly three score years simply melt away. In particular, it becomes relatively straightforward to prove stability of general equilibrium, and using agent-based simulations of complex market dynamics, we can begin to address the question of what institutions protect the economy against extreme sensitivity to exogenous shocks.

Chapter 12 (*The Future of the Behavioral Sciences*) began as a target article in Behavioral and Brain Sciences (Gintis 2007c), and my further ideas on the subject of the unification of the behavioral sciences were presented in the final chapter of Gintis (2009a). I recapitulate here the main ideas of these earlier contributions, with a summary of some of the main themes of this book.