Five Principles for the Unification of the Behavioral Sciences

Herbert Gintis[*]

May 13, 2008

1 Introduction

The behavioral sciences include economics, anthropology, sociology, psychology, and political science, as well as biology insofar as it deals with animal and human behavior. These disciplines have distinct research foci, but they include four conflicting models of decision-making and strategic interaction, as determined by what is taught in the graduate curriculum, and what is accepted in journal articles without reviewer objection. The four are the psychological, the sociological, the biological, and the economic.

These four models are not only different, which is to be expected given their distinct explanatory goals, but *incompatible*. This means, of course, that at least three of the four are certainly incorrect, and I will argue that in fact all four are flawed, but can be modified to produce a unified framework for modeling choice and strategic interaction for all of the behavioral sciences. Such a framework would then be enriched in different ways to meet the particular needs of each discipline.

In the past, cross-disciplinary incoherence was tolerated because distinct disciplines dealt largely with distinct phenomena. Economics dealt with market exchange. Sociology dealt with stratification and social deviance. Psychology dealt with brain functioning. Biology avoided dealing with human behavior altogether. In recent years, however, the value of transdisciplinary research in addressing questions of social theory has become clear, and sociobiology has become a major arena of behavioral research. Moreover, contemporary social policy involves issues that fall squarely in the interstices of the behavioral disciplines, including substance abuse, crime, corruption, tax compliance, social inequality, poverty, discrimina-

---

[*] Santa Fe Institute and Central European University. I would like to thank the European Science Foundation for financial support, and Rosaria Conte for helpful comments on an earlier version.

tion, and the moral foundations of market economies. Incoherence is now an impediment to future progress.

My framework for unification includes five conceptual units: (a) gene-culture coevolution; (b) the socio-psychological theory of norms; (c) game theory, (d) the rational actor model; and (e) complexity theory. Gene-culture coevolution comes from the biological theory of social organization (sociobiology), and is foundational because *Homo sapiens* is an evolved, biological, highly social, species. The socio-psychological theory of norms includes fundamental insights from sociology that apply to all forms of human social organization, from hunter-gatherer to advanced technological societies. These societies are the product of gene-culture coevolution, but have emergent properties, including social norms and their psychological correlates, that cannot be derived analytically from lower-level constructs.

Game theory includes four related disciplines: classical, behavioral, epistemic, and evolutionary game theory. Classical game theory is the analytical elaboration of the ideas of John von Neumann and John Nash, widely used in economics. Behavioral game theory is the application of classical game theory to the social psychology of human strategic interaction, used to show that rational behavior involves deep psychological and sociological principles, including other-regarding behavior and a social epistemology. Epistemic game theory is the application of the modal logic of knowledge and belief to classical game theory, used to integrate the rational actor model with the socio-psychological theory of norms (Gintis 2009a). Evolutionary game theory is a macro-level analytical apparatus allowing the insights of biological and cultural evolution to be mathematically modeled.

The rational actor model, developed in economic and decision theory, is the single most important analytical construct in the behavioral sciences operating at the level of the individual. While gene-culture coevolutionary theory is a form of "ultimate" explanation that does not predict, the rational actor model provides a "proximate" description of behavior that can be tested in the laboratory and real life, and is the basis of the explanatory success of economic theory. Classical, epistemic, and behavioral game theory make no sense without the rational actor model, and behavioral disciplines, like sociology and psychology, that have abandoned this model have fallen into theoretical disarray.

Complexity theory is needed because human society is a complex adaptive system with *emergent properties* that cannot now, and perhaps never will be, explained starting with more basic units of analysis. The hypothetico-deductive methods of game theory and the rational actor model, and even gene-culture coevolutionary theory, must therefore be complemented by the work of behavioral scientists who deal with society in more macro-level, interpretive terms, and develop insightful schemas that shed light where analytical models cannot penetrate. Anthropological and historical studies fall into this category, as well as macroeconomic policy and comparative economic systems. Agent-based modeling of complex dynamical systems is also useful in dealing with emergent properties of complex adaptive systems.

The above principles are not meant to revolutionize research in any discipline. Indeed, they apply only in the areas of overlap among disciplines. For instance, a psychologist working on visual processing, or an economist working on futures markets, or an anthropologist documenting food sharing practices, or a sociologist gauging the effect of dual parenting on children's educational attainment, might gain little from knowing that a unified model of decision making underlay all the behavioral disciplines. On the other hand, a unified model of human choice and strategic interaction might foster innovations that come to pervade the discipline.

2 Gene-culture Coevolution

The centrality of culture and complex social organization to the evolutionary success of *Homo sapiens* implies that individual fitness in humans depends on the structure of social life. Since culture is limited and facilitated by human genetic propensities, it follows that human cognitive, affective, and moral capacities are the product of an evolutionary dynamic involving the interaction of genes and culture. This dynamic is known as *gene-culture coevolution* (Cavalli-Sforza and Feldman 1982, Boyd and Richerson 1985, Dunbar 1993, Richerson and Boyd 2004). This coevolutionary process has endowed us with preferences that go beyond the self-regarding concerns emphasized in traditional economic and biological theory, and embrace a social epistemology facilitating the sharing of intentionality across minds, as well as such non-self-regarding values as a taste for cooperation, fairness, and retribution, the capacity to empathize, and the ability to value honesty, hard work, piety, toleration of diversity, and loyalty to one's reference group.

Gene-culture coevolution is the application of *sociobiology*, the general theory of the social organization of biological species, to species that transmit culture without informational loss across generations. An intermediate category is *niche construction*, which applies to species that transform their natural environment to facilitate social interaction and collective behavior (Odling-Smee, Laland and Feldman 2003).

The genome encodes information that is used both to construct a new organism and to endow it with instructions for transforming sensory inputs into decision outputs. Because learning is costly and error-prone, efficient information transmission will ensure that the genome encode all aspects of the organism's environment that are constant, or that change only slowly through time and space. By contrast, environmental conditions that vary rapidly can be dealt with by providing the organism with the capacity to *learn*.

There is an intermediate case, however, that is efficiently handled neither by genetic encoding nor learning. When environmental conditions are positively but imperfectly correlated across generations, each generation acquires valuable information through learning that it cannot transmit genetically to the succeeding generation, because such information is not encoded in the germ line. In the context of such environments, there is a fitness benefit to the transmission of *epigenetic* information concerning the current state of the environment. Such epigenetic information is quite common (Jablonka and Lamb 1995), but achieves its highest and most flexible form in *cultural transmission* in humans and to a considerably lesser extent in other primates (Bonner 1984, Richerson and Boyd 1998). Cultural transmission takes the form of vertical (parents to children) horizontal (peer to peer), and oblique (elder to younger), as in Cavalli-Sforza and Feldman (1981), prestige (higher influencing lower status), as in Henrich and Gil-White (2001), popularity-related as in Newman, Barabasi and Watts (2006), and even random population-dynamic transmission, as in Shennan (1997) and Skibo and Bentley (2003).

The parallel between cultural and biological evolution goes back to Huxley (1955), Popper (1979), and James (1880)—see Mesoudi, Whiten and Laland (2006) for details. The idea of treating culture as a form of epigenetic transmission was pioneered by Richard Dawkins, who coined the term "meme" in *The Selfish Gene* (1976) to represent an integral unit of information that could be transmitted phenotypically. There quickly followed several major contributions to a biological approach to culture, all based on the notion that culture, like genes, could evolve through replication (intergenerational transmission), mutation, and selection.

Cultural elements reproduce themselves from brain to brain and across time, mutate, and are subject to selection according to their effects on the fitness of their carriers (Parsons 1964, Cavalli-Sforza and Feldman 1982). Moreover, there are strong interactions between genetic and epigenetic elements in human evolution, ranging from basic physiology (e.g., the transformation of the organs of speech with the evolution of language) to sophisticated social emotions, including empathy, shame, guilt, and revenge-seeking (Zajonc 1980, 1984).

Because of their common informational and evolutionary character, there are strong parallels between genetic and cultural modeling (Mesoudi et al. 2006). Like biological transmission, culture is transmitted from parents to offspring, and like cultural transmission, which is transmitted horizontally to unrelated individuals, so in microbes and many plant species, genes are regularly transferred across lineage boundaries (Jablonka and Lamb 1995, Rivera and Lake 2004, Abbott, James, Milne and Gillies 2003). Moreover, anthropologists reconstruct the history of social groups by analyzing homologous and analogous cultural traits, much as biologists reconstruct the evolution of species by the analysis of shared characters and homologous DNA (Mace and Pagel 1994). Indeed, the same computer programs developed by biological systematists are used by cultural anthropologists (Holden 2002, Holden and Mace 2003). In addition, archeologists who study cultural evolution have a similar *modus operandi* as paleobiologists who study genetic evolution (Mesoudi et al. 2006). Both attempt to reconstruct lineages of artifacts and their carriers. Like paleobiology, archaeology assumes that when analogy can be ruled out, similarity implies causal connection by inheritance (O'Brian and Lyman 2000). Like biogeography's study of the spatial distribution of organisms (Brown and Lomolino 1998), behavioral ecology studies the interaction of ecological, historical, and geographical factors that determine distribution of cultural forms across space and time (Smith and Winterhalder 1992).

Perhaps the most common critique of the analogy between genetic and cultural evolution is that the gene is a well-defined, discrete, independently reproducing and mutating entity, whereas the boundaries of the unit of culture are ill-defined and overlapping. In fact, however, this view of the gene is simply outdated. Overlapping, nested, and movable genes discovered in the past 35 years, have some of the fluidity of cultural units, whereas quite often the boundaries of a cultural unit (a belief, icon, word, technique, stylistic convention) are quite delimited and specific. Similarly, alternative splicing, nuclear and messenger RNA editing, cellular protein modification and genomic imprinting, which are quite common quite undermine the standard view of the insular gene producing a single protein, and support the notion of genes having variable boundaries and having strongly context-dependent effects.

Dawkins added a second fundamental mechanism of epigenetic information transmission in *The Extended Phenotype* (1982), noting that organisms can directly transmit environmental artifacts to the next generation, in the form of such constructs as beaver dams, bee hives, and even social structures (e.g., mating and hunting practices). The phenomenon of a species creating an important aspect of its environment and stably transmitting this environment across generations, known as *niche construction*, it a widespread form of epigenetic transmission (Odling-Smee et al. 2003). Moreover, niche construction gives rise to what might be called a *gene-environment coevolutionary process*, since a genetically induced environmental regularity becomes the basis for genetic selection, and genetic mutations that give rise to mutant niches will survive if they are fitness enhancing for their constructors.

An excellent example of gene-environment coevolution is the honey bee, in which the origin of its eusociality likely lay in the high degree of relatedness fostered by haplodiploidy, but which persists in modern species despite the fact that relatedness in the hive is generally quite low, due to multiple queen matings, multiple queens, queen deaths, and the like (Gadagkar 1991, Seeley 1997, Wilson and Holldobler 2005). The social structure of the hive is transmitted epigenetically across generations, and the honey bee genome is an adaptation to the social structure laid down in the distant past.

Gene-culture coevolution in humans is a special case of gene-environment coevolution in which the environment is culturally constituted and transmitted (Feldman and Zhivotovsky 1992). The key to the success of our species in the framework of the hunter-gatherer social structure in which we evolved is the capacity of unrelated, or only loosely related, individuals to cooperate in relatively large egalitarian groups in hunting and territorial acquisition and defense (Boehm 2000, Richerson and Boyd 2004). While contemporary biological and economic theory have attempted to show that such cooperation can be effected by self-regarding rational agents (Trivers 1971, Alexander 1987, Fudenberg, Levine and Maskin 1994), the conditions under which this is the case are highly implausible even for small groups (Boyd and Richerson 1988, Gintis 2005). Rather, the social environment of early humans was conducive to the development of prosocial traits, such as empathy, shame, pride, embarrassment, and reciprocity, without which social cooperation would be impossible.

Neuroscientific studies exhibit clearly the genetic basis for moral behavior. Brain regions involved in moral judgments and behavior include the prefrontal cortex, the orbitalfrontal cortex, and the superior temporal sulcus (Moll, Zahn, di Oliveira-Souza, Krueger and Grafman 2005). These brain structures are virtually unique to, or most highly developed in humans and are doubtless evolutionary adaptations (Schulkin 2000). The evolution of the human prefrontal cortex is closely tied to the emergence of human morality (Allman, Hakeem and Watson 2002). Patients with focal damage to one or more of these areas exhibit a variety of antisocial behaviors, including the absence of embarrassment, pride and regret (Beer, Heerey, Keltner, Skabini and Knight 2003, Camille 2004), and sociopathic behavior (Miller, Darby, Benson, Cummings and Miller 1997). There is a likely genetic predisposition underlying sociopathy, and sociopaths comprise 3–4% of the male population, but they account for between 33% and 80% of the population of chronic criminal offenders in the United States (Mednick, Kirkegaard-Sorenson, Hutchings, Knop, Rosenberg and Schulsinger 1977).

It is clear from this body of empirical information that culture is directly encoded into the human brain, which of course is the central claim of gene-culture coevolutionary theory.

3 Evolutionary Game Theory

The analysis of living systems includes one concept that is not analytically represented in the natural sciences: that of a *strategic interaction*, in which the behavior of agents is derived by assuming that each is choosing a *best response* to the actions of other agents. The study of systems in which agents choose best responses and in which such responses evolve dynamically, is called *evolutionary game theory*.

A *replicator* is a physical system capable of drawing energy and chemical building blocks from its environment to make copies of itself. Chemical crystals, such as salt, have this property, but biological replicators have the additional ability to assume a myriad of physical forms based on the highly variable sequencing of its chemical building blocks. Biology studies the dynamics of such complex replicators using the evolutionary concepts of replication, variation, mutation, and selection (Lewontin 1974).

Biology plays a role in the behavioral sciences much like that of physics in the natural sciences. Just as physics studies the elementary processes that underlie all natural systems, so biology studies the general characteristics of survivors of the process of natural selection. In particular, genetic replicators, the epigenetic environments to which they give rise, and the effect of these environments on gene frequencies, account for the characteristics of species, including the development of individual traits and the nature of intraspecific interaction. This does not mean, of course, that behavioral science in any sense *reduces* to biological laws. Just as one cannot deduce the character of natural systems (e.g., the principles of inorganic and organic chemistry, the structure and history of the universe, robotics, plate tectonics) from the basic laws of physics, similarly one cannot deduce the structure and dynamics of complex life forms from basic biological principles. But, just as physical principles inform model creation in the natural sciences, so must biological principles inform all the behavioral sciences.

Within population biology, evolutionary game theory has become a fundamental tool. Indeed, evolutionary game theory is basically population biology with frequency dependent fitnesses. Throughout much of the Twentieth century, classical population biology did not employ a game-theoretic framework (Fisher 1930, Haldane 1932, Wright 1931). However, Moran (1964) showed that Fisher's Fundamental Theorem, which states that as long as there is positive genetic variance in a population, fitness increases over time, is false when more than one genetic locus is involved. Eshel and Feldman (1984) identified the problem with the population genetic model in its abstraction from mutation. But how do we attach a fitness value to a mutant? Eshel and Feldman (1984) suggested that payoffs be modeled game-theoretically on the phenotypic level, and a mutant gene be associated with a strategy in the resulting game. With this assumption, they showed that under some restrictive conditions, Fisher's Fundamental Theorem could be restored. Their results were generalized by Liberman (1988), Hammerstein and Selten (1994), Hammerstein (1996), Eshel, Feldman and Bergman (1998) and others.

The most natural setting for genetic and cultural dynamics is game theoretic. Replicators (genetic and/or cultural) endow copies of themselves with a repertoire of strategic responses to environmental conditions, including information concerning the conditions under which each is to be deployed in response to the character and density of competing replicators. Genetic replicators have been well understood since the rediscovery of Mendel's laws in the early 20th century. Cultural transmission also apparently occurs at the neuronal level in the brain, in part through the action of *mirror neurons* (Williams, Whiten, Suddendorf and Perrett 2001, Rizzolatti, Fadiga, Fogassi and Gallese 2002, Meltzhoff and Decety 2003). Mutations include replacement of strategies by modified strategies, and the "survival of the fittest" dynamic (formally called a *replicator dynamic*) ensures that replicators with more successful strategies replace those with less successful (Taylor and Jonker 1978).

Cultural dynamics, however, do not reduce to replicator dynamics. For one thing, the process of switching from lower- to higher-payoff cultural norms is subject to error, and with some positive frequency, lower-payoff forms can displace higher-payoff forms (Edgerton 1992). Moreover, cultural evolution can involve conformist bias (Henrich and Boyd 1998, Henrich and Boyd 2001, Guzman, Sickert and Rowthorn 2007), as well as oblique and horizontal transmission (Cavalli-Sforza and Feldman 1981, Gintis 2003b).

4 Socio-psychological Theory of Norms

Complex social systems generally have a division of labor, with distinct social positions occupied by individuals specially prepared for their roles. For instance, a bee hive has workers, drones, queens, and workers can be nurses, foragers, or scouts. Preparation for roles is by gender and larval nutrition. Modern human society has a division of labor characterized by dozens of specialized *roles*, appropriate behavior within which is given by *social norms*, and individuals are *actors* who are motivated to fulfill these roles through a combination of *material incentives* and *normative commitments*.

The centrality of culture in the social division of labor was clearly expressed by Emile Durkheim (1933[1902]), who stressed that the great multiplicity of roles (which he called *organic solidarity*) required a commonality of beliefs (which he called *collective consciousness*) that would permit the smooth coordination of actions by distinct individuals. This theme was developed by Talcott Parsons (1937), who used his knowledge of economics to articulate a sophisticated model of the interaction between the situation (role) and its inhabitant (actor). The actor/role approach to social norms was filled out by Erving Goffman (1959), among others.

The social role has both normative and positive aspects. On the positive side, the payoffs—rewards and penalties—associated with a social role must provide the appropriate incentives for actors to carry out the duties associated with the role. This requirement is most easily satisfied when these payoffs are independent from the behavior of agents occupying other roles. However, this is rarely the case. In general, as developed in section 9, social roles are deeply interdependent, and can be modeled as the strategy sets of players in an epistemic game, the payoffs to which are precisely these rewards and penalties, the choices of actors then forming a *correlated equilibrium*, for which the required commonality of beliefs is provided by a society's common culture (Gintis 2009a). This argument provides an analytical link uniting the actor/role framework in sociological theory with game theoretic models of cooperation in economic theory.

Appropriate behavior in a social role is given by a *social norm* that specifies the duties, privileges, and normal behavior associated with the role. In the first instance, the complex of social norms has a instrumental character devoid of normative content, serving merely as an informational device that coordinates the behavior of rational agents (Lewis 1969, Gauthier 1986, Binmore 2005, Bicchieri 2006). However, in most cases, high level performance in a social role requires that the actor have a *personal commitment* to role performance that cannot be captured by the self-regarding "public" payoffs associated with the role (Conte and Castelfranchi 1999, Gintis 2009a). This is because (a) actors may have private payoffs that conflict with the public payoffs, inducing them to behave counter to proper role-performance (e.g., corruption, favoritism, aversion to specific tasks); (b) the signal used to determine the public payoffs may be inaccurate and unreliable (e.g., the performance of a teacher or physician); and (c) the public payoffs required to gain compliance by self-regarding actors may be higher than those required when there is at least partial reliance upon the personal commitment of role incumbents (e.g., it may be less costly to use personally committed rather than purely materially motivated physicians and teachers). In such cases, self-regarding actors who treat social norms purely instrumentally will behave in a socially inefficient manner.

The normative aspect of social roles flows from these considerations. First, to the extent that social roles are considered legitimate by incumbents, they will place an intrinsic positive ethical value on role-performance. We may call this the *normative bias* associated with role-occupancy (Gintis 2009a). Second, human ethical predispositions include *character virtues*, such as honesty, trustworthiness, promise-keeping, and obedience, that may increase the value of conforming to the duties associated with role-incumbency (Gneezy 2005). Third, humans are also predisposed to care about the esteem of others even when there can be no future reputational repercussions (Masclet, Noussair, Tucker and Villeval 2003), and take pleasure in punishing others who have violated social norms (Fehr and Fischbacher 2004). These normative traits by no means contradict rationality (section 6), because individuals trade off these values against material reward, and against each other, just as described in the economic theory of the rational actor (Andreoni and Miller 2002, Gneezy and Rustichini 2000).

The socio-psychological model of norms can thus resolve the contradictions between the sociological and economic approaches to social cooperation, retaining the analytical clarity of game theory and the rational actor model, while incorporating the normative and cultural considerations stressed in psycho-social models of norm compliance.

5 Socialization and the Internalization of Norms

Society is held together by *moral values* that are transmitted from generation to generation by the process of *socialization*. These values are instantiated through the *internalization of norms* (Parsons 1967, Grusec and Kuczynski 1997, Nisbett and Cohen 1996, Rozin, Lowery, Imada and Haidt 1999), a process in which the initiated instill values into the uninitiated (usually the younger generation) through an extended series of personal interactions, relying on a complex interplay of affect and authority. Through the internalization of norms, initiates are supplied with moral values that induce them to conform to the duties and obligations of the role-positions they expect to occupy. The internalization of norms of course presupposes a genetic predisposition to moral cognition that can be explained only by gene-culture coevolution.

Internalized norms are accepted not as instruments towards achieving other ends, but rather as *arguments in the preference function that the individual maximizes*. For instance, an individual who has internalized the value of "speaking truthfully" will do so even in cases where the net payoff to speaking truthfully would otherwise be negative. Such fundamental human emotions as shame, guilt, pride, and empathy are deployed by the well-socialized individual to reinforce these prosocial values when tempted by the immediate pleasures of such deadly sins as anger, avarice, gluttony, and lust.[1]

The human openness to socialization is perhaps the most powerful form of

---

[1] It is tempting to treat some norms as constraints rather than objectives, but virtually all norms are violated by individuals under some conditions, indicating that there are tradeoffs that could not exist were norms merely constraints on action.

epigenetic transmission found in nature. This epigenetic flexibility in considerable part accounts for the stunning success of the species *Homo sapiens*, because when individuals internalize a norm, the frequency of the desired behavior will be higher than if people follow the norm only instrumentally—i.e., when they perceive it to be in their to do so on other grounds. The increased incidence of prosocial behaviors are precisely what permits humans to cooperate effectively in groups (Gintis, Bowles, Boyd and Fehr 2005).

There are, of course, limits to socialization (Tooby and Cosmides 1992, Pinker 2002), and it is imperative to understand the dynamics of emergence and abandonment of particular values, which in fact depend on their contribution to fitness and well-being, as economic and biological theory would suggest (Gintis 2003a,b). Moreover, there are often swift society-wide value changes that cannot be accounted for by socialization theory (Wrong 1961, Gintis 1975). However, socialization theory has an important place in the general theory of culture, strategic learning, and moral development.

6 The Rational Actor Model

General evolutionary principles suggest that individual decision making for members of a species can be modeled as optimizing a preference function. Natural selection leads the content of preferences to reflect biological fitness. The principle of expected utility extends this optimization to stochastic outcomes. The resulting model is called the *rational actor model* in economics, although there is some value to renaming it the *beliefs, preferences, and constraints* (BPC) model, thus avoiding the often misleading connotations attached to the term “rational.”

For every constellation of sensory inputs, each decision taken by an organism generates a probability distribution over outcomes, the expected value of which is the *fitness* associated with that decision. Since fitness is a scalar variable, for each constellation of sensory inputs, each possible action the organism might take has a specific fitness value, and organisms whose decision mechanisms are optimized for this environment will choose the available action that maximizes this value. This argument was presented verbally by Darwin (1872) and is implicit in the standard notion of “survival of the fittest,” but formal proof is recent (Grafen 1999, 2000, 2002). The case with frequency-dependent (non-additive genetic) fitness has yet to be formally demonstrated, but the informal arguments are compelling.

Given the state of its sensory inputs, if an orgasm with an optimized brain chooses action A over action B when both are available, and chooses action B over action C when both are available, then it will also choose action A over action C when both are available. Thus choice consistency follows from basic evolutionary dynamics.

The so-called *rational actor model* was developed in the Twentieth century by John von Neumann, Leonard Savage and many others. The model is often presented as though it applies only when actors possess extremely strong information processing capacities. In fact, the model depends only on choice consistency (Gintis 2009a). When preferences are consistent, they can be represented by a numerical function, often called a *utility function*, which the individual maximizes subject to his subjective beliefs.

Four *caveats* are in order. First, individuals do not consciously maximize something called "utility," or anything else. Second, individual choices, even if they are self-regarding (e.g., personal consumption) are not necessarily welfare-enhancing. Third, preferences must have some stability across time to be theoretically useful, but preferences are ineluctably a function of an individual's *current state*, and beliefs can change dramatically in response to immediate sensory experience. Finally, beliefs need not be correct nor need they be updated correctly in the face of new evidence, although Bayesian assumptions concerning updating can be made part of consistency in elegant and compelling ways (Jaynes 2003).

The rational actor model is the cornerstone of contemporary economic theory, and in the past few decades has become the heart of the biological modeling of animal behavior (Real 1991, Alcock 1993, Real and Caraco 1986). Economic and biological theory thus have a natural affinity: the choice consistency on which the rational actor model of economic theory depends is rendered plausible by evolutionary theory, and the optimization techniques pioneered in economics are routinely applied and extended by biologists in modeling the behavior of nonhuman organisms. I suggest in section 13 that this is due to the "routine choice" paradigm that applies in economics and biology, as opposed to the "deliberative choice" paradigm that applies in cognitive psychology.

In a stochastic environment, natural selection will ensure that the brain make choices that, at least roughly, maximize expected fitness, and hence satisfy the expected utility principle Cooper (1987). To see this, suppose an organism must choose from action set X, where each $x \in X$ determines a lottery that pays i offspring with probability $p_i(x)$, for $i = 0, 1, \ldots, n$. Then the expected number of offspring from this lottery is $\psi(x) = \sum_{j=1}^{n} jp_j(x)$. Let L be a lottery on X that delivers $x_i \in X$ with probability q_i for $i = 1, \ldots, k$. The probability of j offspring given L is then $\sum_{i=1}^{k} q_i p_j(x_i)$, so the expected number of offspring given L is

$$\sum_{j=1}^{n} j \sum_{i=1}^{k} q_i p_j(x_i) = \sum_{i=1}^{k} q_i \sum_{i=1}^{k} jp_j(x_i) = \sum_{i=1}^{k} q_i \psi(x_i), \tag{1}$$

which is the expected value theorem with utility function $\psi(\cdot)$.

Evidence from contemporary neuroscience suggests that expected utility maximization is not simply an "as if" story. In fact, the brain's neural circuitry actually makes choices by internally representing the payoffs of various alternatives as neural firing rates, and choosing a maximal such rate (Glimcher 2003, Dorris and Glimcher 2003, Glimcher, Dorris and Bayer 2005). Neuroscientists increasingly find that an aggregate decision making process in the brain synthesizes all available information into a single, unitary value (Parker and Newsome 1998, Schall and Thompson 1999, Glimcher 2003). Indeed, when animals are tested in a repeated trial setting with variable reward, dopamine neurons appear to encode the difference between the reward that an animal expected to receive and the reward that an animal actually received on a particular trial (Schultz, Dayan and Montague 1997, Sutton and Barto 2000), an evaluation mechanism that enhances the environmental sensitivity of the animal's decision making system. This error-prediction mechanism has the drawback of only seeking local optima (Sugrue, Corrado and Newsome 2005). Montague and Berns (2002) address this problem, showing that the obitofrontal cortex and striatum contains a mechanism for more global predictions that include risk assessment and discounting of future rewards. Their data suggest a decision making model that is analogous to the famous Black-Scholes options pricing equation (Black and Scholes 1973).

Perhaps the most pervasive critique of the BPC model is that put forward by Herbert Simon (1982), holding that because information processing is costly and humans have finite information processing capacity, individuals *satisfice* rather than *maximize*, and hence are only *boundedly rational*. There is much substance to this view, including the importance of including information processing costs and limited information in modeling choice behavior and recognizing that the decision on how much information to collect depends on unanalyzed subjective priors at some level (Winter 1971, Heiner 1983). Indeed, from basic information theory and quantum mechanics it follows that *all rationality is bounded*. However, the popular message taken from Simon's work is that we should reject the BPC model. For instance, the mathematical psychologist D. H. Krantz (1991) asserts, "The normative assumption that individuals *should* maximize *some* quantity may be wrong... People do and should act as *problem solvers*, not *maximizers*." This is incorrect. In fact, as long as individuals are involved in routine choice (see section 13), and hence have consistent preferences, they can be modeled as maximizing an objective function subject to constraints.

If there is a single objective (e.g., solve the problem with a given degree of acceptability), then preference consistency can be ignored. But, once the degree of acceptability is treated as endogenous, multiple objectives compete (e.g., cost and accuracy), and the BPC model cannot be ignored. This point is lost on even such capable researchers as Gigerenzer and Selten (2001), who reject the "optimization subject to constraints" method on the grounds that individuals do not in fact solve optimization problems. However, just as the billiards players do not solve differential equations in choosing their shots, so decision-makers do not solve Lagrangian equations, even though in both cases we may use such optimization models to describe their behavior.

7 Application: Addictive Behavior

Substance abuse appears to be irrational. Abusers are time inconsistent and their behavior is welfare-reducing. Moreover, even draconian increases in the penalties for illicit substance use lead to the swelling of prison populations rather than the abandonment of the sanctioned activity. Because rational actors generally trade off among desired goals, this curious phenomenon has led some researchers to reject the BPC model out of hand.

However, the BPC model remains the most potent tool for analyzing substance abuse on a society-wide level. The most salient target of the critics has been the "rational addiction" model of Becker and Murphy (1988). Empirical research, however, has supported the contention that illicit drugs respond normally to market forces. For instance Saffer and Chaloupka (1999) estimated the price elasticities of heroin and cocaine using a sample of 49,802 individuals from the National Household Survey of Drug Abuse to be 1.70 and 0.96, respectively. These elasticities are in fact quite high. Using these estimates, the authors judge that the lower prices flowing from the legalization of these drugs would lead to an increase of about 100% and 50% in the quantities of heroin and cocaine consumed, respectively.

How does this square with the observation that draconian punishments do not squelch the demand altogether? Gruber and Koszegi (2001) explain this by showing that drug users exhibit the commitment and self-control problems that are typical of time-inconsistent agents, for whom the possible future penalties have highly attenuated deterrent value in the present. This behavior may be welfare-reducing but, the rational actor model does not presume that preferred outcomes are necessarily welfare-improving.

8 Game Theory: The Universal Lexicon of Life

In the BPC model, choices give rise to probability distributions over outcomes, the expected values of which are the payoffs to the choice from which they arose. Game theory extends this analysis to cases where there are multiple decision makers. In the language of game theory, *players* (or *agents*) are endowed with a set of available *strategies*, and have certain *information* concerning the rules of the game, the nature of the other players and their available strategies, as well as the structure of payoffs. Finally, for each combination of strategy choices by the players, the game specifies a distribution of *payoffs* to the players. Game theory predicts the behavior of the players by assuming each maximizes its preference function subject to its information, beliefs, and constraints (Kreps 1990).

Game theory is a logical extension of evolutionary theory. To see this, suppose there is only one replicator, deriving its nutrients and energy from non-living sources. The replicator population will then grow at a geometric rate, until it presses upon its environmental inputs. At that point, mutants that exploit the environment more efficiently will out-compete their less efficient conspecifics, and with input scarcity, mutants will emerge that "steal" from conspecifics who have amassed valuable resources. With the rapid growth of such predators, mutant prey will devise means of avoiding predation, and predators will counter with their own novel predatory capacities. In this manner, strategic interaction is born from elemental evolutionary forces. It is only a conceptual short step from this point to cooperation and competition among cells in a multi-cellular body, among conspecifics who cooperate in social production, between males and females in a sexual species, between parents and offspring, and among groups competing for territorial control.

Historically, game theory did not emerge from biological considerations, but rather from strategic concerns in World War II (Von Neumann and Morgenstern 1944, Poundstone 1992). This led to the widespread caricature of game theory as applicable only to static confrontations of rational self-regarding agents possessed of formidable reasoning and information processing capacity. Developments within game theory in recent years, however, render this caricature inaccurate.

Game theory has become the basic framework for modeling animal behavior (Maynard Smith 1982, Alcock 1993, Krebs and Davies 1997), and thus has shed its static and hyperrationalistic character, in the form of evolutionary game theory (Gintis 2009b). Evolutionary and behavioral game theory do not require the formidable information processing capacities of classical game theory, so disciplines that recognize that cognition is scarce and costly can make use of game-theoretic models (Young 1998, Gintis 2009b, Gigerenzer and Selten 2001). Thus, agents may consider only a restricted subset of strategies (Winter 1971, Simon 1972), and they may use by rule-of-thumb heuristics rather than maximization techniques (Gigerenzer and Selten 2001). Game theory is thus a generalized schema that permits the precise framing of meaningful empirical assertions, but imposes no particular structure on the predicted behavior.

9 Epistemic Game Theory and Social Norms

Economics and sociology have highly contrasting models of human interaction. Economics traditionally considers individuals to be rational, self-regarding, payoff maximizers while sociology considers individuals to be highly socialized, other-regarding, moral agents who strive to fill social roles and whose self-esteem depends on the approbation of others. The project of unifying the behavioral sciences must include a resolution of these inconsistencies in a manner that preserves the key insights of each.

Behavioral game theory helps us adjudicate these disciplinary differences, providing experimental data supporting the sociological stress on moral values and other-regarding preferences, but also supports the economic stress on rational payoff maximization. For instance, most individuals care about reciprocity and fairness as well as personal gain (Gintis et al. 2005), value such character virtues as honesty for their own sake (Gneezy 2005), care about the esteem of others even when there can be no future reputational repercussions (Masclet et al. 2003), and take pleasure in punishing others who have hurt them (deQuervain et al. 2004). Moreover, as suggested by socialization theory, individuals have consistent values, based on their particular socio-cultural situation, that they apply in the laboratory even in one-shot games under conditions of anonymity (Henrich, Boyd, Bowles, Camerer, Fehr and Gintis 2004, Henrich et al. 2006). This body of evidence suggests that sociologists would benefit from reincorporating the rational actor model into sociological theory, and economists broaden their concept of human preferences.

A second discrepancy between economics and sociology concerns the contrasting claims of game theory and the socio-psychological theory of norms in explaining social cooperation. My research in this area Gintis (2009a) can be summarized as follows.

The basic model of the division of labor in economic theory is the Walrasian general equilibrium model, according to which a system of flexible prices induces firms and individuals to supply and demand goods and services in such amount that all markets clear in equilibrium (Arrow and Debreu 1954). However, this model assumes that all contracts among individuals can be costlessly enforced by a third party, such as the judicial system. In fact, however, many critical forms of social cooperation are not mediated by third-party enforceable contract, but rather take the form of repeated interactions in which the informal, but very real, threat of rupturing the relationship is used to induce mutual cooperation (Fudenberg et al. 1994, Ely and Välimäki 2002). For instance, an employer hires a worker, who works hard under the threat of dismissal, not the threat of an employer lawsuit.

Repeated game theory thus steps in for economists to explain forms of face-to-face cooperation that do not reduce to simple price-mediated market exchanges. Repeated game theory shows that in many cases the activity of many individuals can be coordinated, in the sense that there is a Nash equilibrium ensuring that no self-regarding player can gain by deviating from the strategy assigned to him by the equilibrium, assuming other players also use the strategies assigned to them. If this theory were adequate, which most economists believe is the case, then there would be no role for the socio-psychological theory of norms, and sociological theory would be no more that a thick description of a social mechanism analytically accounted for by repeated game theory.

However, repeated game theory with self-regarding agents does not solve the problem of social cooperation. When the group consists of more than two individuals and the signal indicating how well a player is performing his part is imperfect and private (i.e., players receive imperfectly correlated signals of another player's behavior), the efficiency of cooperation may be quite low, and the roles assigned to each player will be extremely complex mixed strategies that players have no incentive to use. As we suggested in section 4, the socio-psychology of norms can step in at this point to provide mechanisms that induce individuals to play their assigned parts. A social norm may provide the rules for each individual in the division of labor, players may have general predilection for honesty that allows them to consolidate their private signals concerning another player's behavior into a public signal that can be the bases for coordinated collective punishment and reward, and players may have a personal normative bias in favor of following the social roles assigned to them. The sociological and economic forces thus complement rather than contradict one another.

A central analytical contribution to this harmonization of economics and sociology was provided by Robert Aumann (1987), who showed that the natural concept of equilibrium in game theory for rational actors who share common beliefs is not the Nash equilibrium, but the correlated equilibrium. A correlated equilibrium is the Nash equilibrium in the game formed by adding to the original game a new players, whom I call the *choreographer* (Aumann calls this simply a "correlating device") that samples the probability distribution given by the players (common) beliefs, and then instructs each player what action to take. The actions recommended by the choreographer are all best responses to one another, conditional on their having been simultaneously ordered by the choreographer, so self-regarding players can do no better than following the choreographer's advice.

Sociology then comes in not only by supplying the choreographer, in the form of a complex of social norms, but also supplies cultural theory to explain why players might have a common set of beliefs, without which the correlated equilibrium would not exist, and cognitive psychology explains the normative bias that induces players to take the advice of the choreographer (i.e., to follow the social norm) when in fact there might be many other actions with equal, or even higher, payoff that the player might have an inclination to choose.

10 Society as Complex Adaptive System

The behavioral sciences advance not only by developing analytical and quantitative models, but by accumulating historical, descriptive and ethnographic evidence that pays heed to the detailed complexities of life in the sweeping array of wondrous forms that nature reveals to us. Historical contingency is a primary focus for many students of sociology, anthropology, ecology, biology, politics, and even economics. By contrast, the natural sciences have found little use for narrative along side analytical modeling.

The reason for this contrast between the natural and the behavioral sciences is that *living systems are generally complex, dynamic adaptive systems* with emergent properties that cannot be fully captured in analytical models that attend only to the local interactions. The hypothetico-deductive methods of game theory, the BPC model, and even gene-culture coevolutionary theory must therefore be complemented by the work of behavioral scientists who adhere to a more historical and interpretive traditions, as well as researchers who use agent-based programming techniques to explore the dynamical behavior of approximations to real-world complex adaptive systems.

A *complex system* consists of a large population of similar entities (in our case, human individuals) who interact through regularized channels (e.g., networks, markets, social institutions) with significant stochastic elements, without a system of centralized organization and control (i.e., if there a state, it controls only a fraction of all social interactions, and itself is a complex system). A complex system is *adaptive* if undergoes an evolutionary (genetic, cultural, agent-based, or other) process of reproduction, mutation, and selection (Holland 1975). To characterize a system as complex adaptive does not explain its operation, and does not solve any problems. However, it suggests that certain modeling tools are likely to be effective that have little use in a non-complex system. In particular, the traditional mathematical methods of physics and chemistry must be supplemented by other modeling tools, such as agent-based simulation and network theory.

The stunning success of modern physics and chemistry lies in their ability to avoid or control emergence. The experimental method in natural science is to create highly simplified laboratory conditions, under which modeling becomes analytically tractable. Physics is no more effective than economics or biology in analyzing complex real-world phenomena *in situ*. The various branches of engineering (electrical, chemical, mechanical) are effective because they recreate in everyday life artificially controlled, non-complex, non-adaptive, environments in which the discoveries of physics and chemistry can be directly applied. This option is generally not open to most behavioral scientists, who rarely have the opportunity of "engineering" social institutions and cultures.

11 Counterpoint: Biology

Biologists are generally comfortable with three of the five principles laid out in the introduction to this paper. Only gene-culture coevolution and the socio-psychology of norms have generated significant opposition.

Gene-culture coevolutionary theory is only two decades old and applies only to one species—*Homo sapiens*. Not surprisingly, many sociobiologists have been slow to adopt it, and have deployed a formidable array of population biology concepts towards explaining human sociality in more familiar terms—especially kin selection (Hamilton 1964) and reciprocal altruism (Trivers 1971). The explanatory power of these models convinced a generation of researchers that what appears to be altruism— personal sacrifice on behalf of others—is really just long-run self-interest, and elaborate theories drawn from anthropology, sociology, and economics are unnecessary to explain human cooperation and conflict.

Richard Dawkins, for instance, in *The Selfish Gene* (1989[1976]), asserts "We are survival machines—robot vehicles blindly programmed to preserve the selfish molecules known as genes.. . . This gene selfishness will usually give rise to selfishness in individual behavior." Similarly, in his *The Biology of Moral Systems* (New York: Aldine, 1987), R. D. Alexander asserts, "ethics, morality, human conduct, and the human psyche are to be understood only if societies are seen as collections of individuals seeking their own self-interest.. . . " (p. 3). In a similar vein, Michael Ghiselin (1974) wrote "No hint of genuine charity ameliorates our vision of society, once sentimentalism has been laid aside. What passes for cooperation turns out to be a mixture of opportunism and exploitation. . . Scratch an altruist, and watch a hypocrite bleed" (p. 247).

Evolutionary psychology, which has been a major contributor to human sociobiology, has incorporated the kin selection/reciprocal altruism perspective into a broadside critique of the role of culture in society (Barkow, Cosmides and Tooby 1992), and of the forms of group dynamics upon which gene-culture coevolution depends (Price, Cosmides and Tooby 2002). I believe these claims have been effectively refuted (Richerson and Boyd 2004, Gintis, Henrich, Bowles, Boyd and Fehr 2009), although the highly interesting debate in population biology concerning group selection has been clarified but not completely resolved (Lehmann and Keller 2006, Lehmann, Rousset, Roze and Keller 2007, Wilson and Wilson 2007).

12 Counterpoint: Economics

Economists generally believe in methodological individualism, a doctrine claiming that all social behavior can be explained by strategic interactions among rational agents. Were this correct, gene-culture coevolution would be unnecessary, complexity theory irrelevant, and the socio-psychological theory of norms could be derived from game theory. I have tried to show in section 9 that this is not the case.

Economists also generally reject the idea of society as a complex adaptive system, on grounds that we may yet be able to tweak the Walrasian general equilibrium framework, suitably fortified by sophisticated mathematical methods, so as to explain macroeconomic activity. In fact, there has been virtually no progress on this front since the mid-Twentieth century existence proofs (Arrow and Debreu 1954). Particularly noteworthy has been the absence of any credible stability model (Fisher 1983). Indeed, the standard models predict price instability and chaos (Saari 1985, Bala and Majumdar 1992). Moreover, analysis of excess demand functions suggests that restrictions on preferences are unlikely to entail the stability of Walrasian price dynamics (Sonnenschein 1972,1973; Debreu 1974; Kirman and Koch 1986).

My response to this sad state of affairs has been to show that agent-based models of generalized exchange, based on the notion that the economy is a complex nonlinear system, exhibit a high degree of stability and efficiency (Gintis 2006, Gintis 2007). There does not appear to be any serious doctrinal impediment to the use of agent-based modeling in economics.

13 Counterpoint: Psychology

Psychology has contributed strongly to our understanding of human choice behavior in recent years. Psychologists have used these contributions, improperly, I believe, to mount a sustained attack on the rational actor model. I believe this attack is based on the fact that cognitive psychology and economic theory fail to recognize that they have distinct and non-overlapping research projects that each calls "decision theory." When this confusion is cleared up, the contradictions between the psychological model of decision-making and our five principles, including the rational actor model, disappear. The psychological and economic models are in fact complementary, mutually supportive, and both stand to gain from synthesizing their research agendas.

The psychological paradigm in decision theory applies to ambiguous choice situations in which the characteristics of the various alternatives are multi-faceted and their value imperfectly known, the probabilities of various outcomes must be inferred from imperfect data, and deliberative processes required to balance the costs and benefits of various options are mentally challenging and subject to imprecision. Instances where this paradigm seems applicable include deciding whether, when, and whom to marry, choosing allies and enemies, deciding on a career plan, assessing candidates for public office, and opting for a medical intervention.

The rational actor paradigm avoids these difficulties, assuming that the decision maker has a probability distribution over states of nature that resolves all ambiguities, so the decision itself is reduced to the mechanical process of choosing an alternative from a fixed set of possibilities that has the highest expected payoff. Instances where this paradigm appears applicable include choosing a basket of items at the supermarket, making income/leisure choices, trading off the costs and benefits of various commuting alternatives, and deciding on a pattern of charitable contributions.

The paradigms are sufficiently distinct that they deserve different names. I will call them *deliberative* vs. *routine* choice paradigms. Despite its having been an active area of psychological research for several decades, there is no general theory of deliberative choice. By contrast, the theory of routine choice is well developed and lies at the heart of the standard economic approach to modeling market behavior.

Because there is no general theory of deliberative decision-making, cognitive psychology peripheralizes decision-making in the introductory textbooks. For instance, a widely used text of graduate-level readings in cognitive psychology, (Sternberg and Wagner 1999) devotes the *ninth* of eleven chapters to “Reasoning, Judgment, and Decision Making,” offering two papers, the first of which shows that human subjects generally fail simple logical inference tasks, and the second shows that human subjects are irrationally swayed by the way a problem is verbally “framed” by the experimenter. A leading undergraduate cognitive psychology text (Goldstein 2005) placed “Reasoning and Decision Making” the *last* of twelve chapters. This includes one paragraph describing the rational actor model, followed by many pages purporting to explain why it is wrong.

A similar situation obtains for behavioral psychology, which generally avoids positing internal states, of which preferences and beliefs, and even some constraints (e.g. such character virtues as keeping promises), are examples. When the rational actor model is mentioned, it is summarily rejected (Herrnstein, Laibson and Rachlin 1997). Not surprisingly, in a leading behavioral psychology text (Mazur 2002), choice is covered in the *last* of fourteen chapters, and is limited to a review of the literature on choice between concurrent reinforcement schedules and the capacity to defer gratification.

Summing up a quarter century of psychological research in 1995, Paul Slovic asserted, accurately I believe, that “it is now generally recognized among psychologists that utility maximization provides only limited insight into the processes by which decisions are made." (Slovic 1995):365 "People are not logical," psychologists are fond of saying, "they are *psycho*logical."

My suggestion for resolving the conflict between psychological and economic models of decision-making has four points. First, the two disciplines should recognize the distinction between deliberative and routine decision-making. Second, psychology should introduce the evolution of routine decision-making into its core framework, based on the principle that the brain is a fitness-enhancing adaptation. Third, deliberative decision-making is an adaptation to the increased social complexity of primates and hominid groups. Finally, routine decision-making shades into deliberative decision-making under conditions that are only imperfectly known, but are of great potential important for understanding human choice. The remainder of this section is an elaboration on these points.

The fitness of an organism depends on the effectiveness of its decision-making in a stochastic environment. Effective choice is a function of the organism's state of knowledge, which consists of the information supplied by the sensory inputs that monitor the organism's internal states and its external environment. In relatively simple organisms, the choice environment is primitive and distributed in a decentralized manner over sensory inputs. But, in three separate groups of animals, the craniates (vertebrates and related creatures), arthropods (including insects, spiders, and crustaceans) and cephalopods (squid, octopuses, and other mollusks) a central nervous system with a brain (a centrally located decision making and control apparatus) evolved. The phylogenetic tree of vertebrates exhibits increasing complexity through time, and increasing metabolic and morphological costs of maintaining brain activity. The brain thus evolved because larger and more complex brains, despite their costs, enhanced the fitness of their carriers. Brains therefore are ineluctably structured to make on balance fitness-enhancing decisions in the face of the various constellations of sensory inputs their bearers commonly experience. The idea that human choice is on balance illogical and irrational, according to this reasoning, is highly implausible.

The human brain shares most of its functions with that of other vertebrates, including the coordination of movement, maintenance of homeostatic bodily functions, memory, attention, processing of sensory inputs, and elementary learning mechanisms. The distinguishing characteristic of the human brain, however, lies in its power in deliberative decision making. Human beings, with the largest brains and the highest proportion of caloric resources devoted to supporting brain activity, are the most competent and effective decision-makers in the biosphere. However, this brain power is probably not needed for routine decision-making. This may be why the treatment of choice in economics and biology, both of which deal with routine choice, is comfortable with the rational actor model, while the psychology of human choice, which focuses on the *differentia specifica* of human choice, which involve deliberative choice, is not.

Nevertheless, the widespread claims by experimental psychologists that humans are illogical and irrational when dealing with routine choice is simply incorrect, because there are almost always more plausible explanations of the observed behavior than failure of logic or reason. In some case, the alternative explanation suggests that humans are highly effective decision-makers. More often, however, the alternative is compatible with the axioms of rational choice over an appropriate (often non-obvious) choice space, but involve imperfect decision-making.

Consider, for instance, Linda the Bank Teller, a well-known experiment performed by Tversky and Kanheman (1983) that falls into the former category. A young woman Linda is described as politically active in college and highly intelligent, and the subject is asked which of the following two statements is more likely: "Linda is a bank teller" or "Linda is a bank teller and is active in the feminist movement." Many subjects rate the second statement more likely, despite the fact that elementary probability theory asserts that if p implies q, then p cannot be more likely than q. Since the second statement implies the first, it cannot be more likely than the first.

However, there are several plausible alternatives to faulty reasoning in explaining this behavior. One plausible alternative is based on the notion that in normal conversation, a listener assumes that any information provided by the speaker is relevant to the speaker's message (Grice 1975). Applied to this case, the norms of conversation lead the subject to believe that the experimenter wants Linda's politically active past to be taken adequately into account (Hilton 1995, Wetherick 1995). Moreover, the meaning of such terms as "more likely" or "higher probability" are vigorously disputed even in the professional literature, and may well have a different meaning for the average subject and for the expert. For instance, if the subject were given two piles of identity folders and ask to search through them to find the one belonging to Linda, and one of the piles was "all bank tellers" while the other was "all bank tellers who are active in the feminist movement," the subject might reasonable look through the second (doubtless much smaller) pile first, even though being well aware that there is a "higher probability" that the folder is in the first (much larger) pile rather than the second.

More generally, subjects may appear irrational because basic terms have different meanings in propositional logic and in everyday logical inference. For instance, "if p then q" is true in formal logic except when p is true and q is false. In everyday usage "if p then q" may be interpreted as a material implication, in which there is something about p that cause q to be the case. In particular, in material logic "p implies q" means "p is true and this situation causes q to be true." Thus, "if France is in Africa, then Paris is in Europe" is true in propositional logic, but false as a material implication. Part of the problem is also that individuals without extensive academic training simply lack the expertise to follow complex chains of logic, so psychology experiments often exhibit a high level of *performance error* (Cohen 1981). For instance, suppose Pat and Kim live in a certain town where all men have beards and all women wear dresses. Then the following can be shown to be true in propositional logic: "Either if Pat is a man then Kim wears a dress or if Kim is a woman, then Pat has a beard." It is quite hard to see why this is formally, true, and it is not true if the implications are material. Finally, the logical meaning of "if p then q" can be context dependent. For instance, "if you eat dinner (p), you may go out to play (q)" formally means "you may go out to play (q) only if you eat dinner (p)."

We may apply this insight to an important strand of experimental psychology that purports to have shown that subjects systematically deviate from simple principles of logical reasoning. In a widely replicated study, Wason (1966) showed subjects cards each of which had a "1" or "2" on one side and "A" or "B" on the other, and stated the following rule: a card with a vowel on one side must have an odd number on the other. The experimenter then showed each subject four cards, one showing "1", one showing "2", one showing "A", and one showing "B", and asked the subject which cards must be turned over to check whether the rule was followed. Typically, only about 15% of college students point out the correct cards ("A" and "2"). Subsequent research showed that when the problem is posed in more concrete terms, such as "any person drinking beer must over eighteen," the correct response rate increases considerably (Stanovich 1999, Shafir and LeBoeuf 2002). This accords with the observation that most individuals do not appear to have difficulty making and understanding logical arguments in everyday life.

Just as the rational actor model began to take hold in the mid-Twentieth century, vigorous empirical objections began to surface. The first was Allais (1953), who discovered cases where subjects exhibited clear choice inconsistency in choosing among simple lotteries. This behavioral entails a loss of decision-making efficiency, but it has been shown that Allais' examples can be explained by regret theory (Bell 1982, Loomes and Sugden 1982), which can be represented by consistent choices over pairs of lotteries (Sugden 1993), and satisfies the conditions for rational choice over this extended choice set.

Close behind Allais came the famous Ellsberg Paradox (Ellsberg 1961), which can be shown to violate the most basic axioms of choice under uncertainty. Consider two urns. Urn A has 51 red balls and 49 white balls. Urn B also has 100 red and white balls, but the fraction of red balls is unknown. One ball is chosen from each urn but remains hidden from sight. Subjects are asked to choose in two situations. First, a subject can choose the ball from urn A or urn B, and if the ball is red, the subject wins $10. In the second situation, two new balls are drawn from the urns, with replacement, the subject can choose the ball from urn A or urn B, and if the ball is white, the subject wins $10. Many subjects choose the ball from urn A in both cases. This obviously violates the expected utility principle, no matter what probability the subject places on the probability the ball from urn B is white.

It is easy to see why unsophisticated subjects make this error: urn B seems to be *riskier* than urn A, because we know the probabilities in A but not B. It takes a relatively sophisticated probabilistic argument—one that no human being ever made or could have made (to our knowledge) prior to the modern era—to see that in fact in this case uncertainty does not lead to increased risk. Indeed, most intelligent subjects who make the Ellsberg error will be convinced, when presented with the logical analysis, to modify their choices without modifying their preferences. In cases like this, we speak of *performance error*, whereas in cases such as the Allais Paradox, even the highly sophisticated subjects will not change their choice.

Numerous experiments document that many people have beliefs concerning probabilistic events that are without scientific foundation, and which will most likely lead them to sustain losses if acted upon. For instance, virtually every enthusiast believes that athletes in competitive sports run "hot and cold," although this has never been substantiated empirically. I have not met a basketball fan who does not believe in the hot hand. Yet, Gilovich, Vallone and Tversky (1985) have shown, using professional basketball data, that the hot hand does not exist. This is but one instance of the general rule that our brains often lead us to perceive a pattern when faced with purely random data. In the same vein, I have talked to professional stock traders who believe, on the basis of direct observation of stock volatility, that stocks follow certain laws of inertia and elasticity that simply cannot be found through a statistical analysis of the data. Another example of this type is the "gambler's fallacy," which is that in a fair game, after a run of exceptionally bad luck (e.g., a series of ten incorrect guesses in a row in a coin flip game), a run of good luck is likely to follow. Those who believe this cannot be dissuaded by scientific evidence. Many who believe in the "law of small numbers," which says that a small sample from a large population will have the same distribution of characteristics as the population (Tversky and Kahneman 1971), simply cannot be dissuaded either by logical reasoning or presentation of empirical evidence.

We are indebted to Daniel Kahneman, Amos Tversky, and their colleagues for a series of brilliant papers documenting the various errors intelligent subjects commit in dealing with probabilistic decision making. Subjects systematically underweigh base rate information in favor of salient and personal examples, they reverse lottery choices when the same lottery is described emphasizing probabilities rather than monetary payoffs, when described in term of losses from a high baseline as opposed to gains from a low baseline, they treat proactive decisions differently from passive decisions even when the outcomes are exactly the same, and when outcomes are described in terms of probabilities as opposed to frequencies

(Kahneman, Slovic and Tversky 1982, Kahneman and Tversky 2000).

These findings are important for understanding human decision-making but they are not threat to the BPC model (Gigerenzer and Selten 2001). They are simply performance errors in the form of incorrect beliefs as to how payoffs can be maximized. Decision theory did not exist until this century. Before the contributions of Bernoulli, Savage, von Neumann and other experts, no creature on Earth knew how to value a lottery. It takes years of study to feel at home with the laws of probability. Moreover, it is costly, in terms of time and effort, to apply these laws even if we know them. Of course, if the stakes are high enough, it is worthwhile to go to the effort, or engage an expert who will do it for you. But generally, as Kahneman and Tversky suggest, we apply a set of heuristics that more or less get the job done (Gigerenzer and Selten 2001). Among the most prominent heuristics is simply *imitation*: decide what class of phenomenon is involved, find out what people "normally do" in that situation, and do it. If there is some mechanism leading to the survival and growth of relatively successful behaviors and if the problem in question recurs with sufficient regularity, the choice-theoretic solution will describe the winner of a dynamic social process of trial, error, and replication through imitation.

The BPC model assumes that individuals have stable preferences and beliefs that are functions of the individual's personality and current needs. Yet, in many cases laboratory experiments show that individuals can be induced to make choices over payoffs based on subtle or obvious cues that ostensibly do not affect the value of the payoffs to the decision maker. For instance, if a subjects' partner in an experimental game is described as a "competitor" or an "opponent," or the game itself is described as a "bargaining game," subjects may make very different choices from a situation where the partner is described as a "teammate", or the game is described as a community participation game. Similarly, a subject in a bargaining game may reject an offer if made by his bargaining partner, but accept the same offer if made by the random draw of a computer (Blount 1995) on behalf of the proposer.

Sensitive to this critique, experimenters in the early years of behavioral game theory attempted to minimize the possibility of framing effects by rendering as abstract and unemotive as possible the language in which a decision problem or strategic interaction was described. It is now widely recognize that it is in fact impossible to avoid framing effects, because abstraction and lack of real-world reference is itself a frame rather than an absence thereof. A more productive way to deal with framing is to make the frame a part of the specification of the experiment itself, and vary the frame systematically to discover the effect of the frame on the choices of the subjects, and by inference, on their beliefs and preferences.

We do not have a complete model of framing, but we do know enough to know that its existence does not undermine the BPC model. Indeed, we can model this explicitly based on the fact that a frame can lead players to share particular beliefs, while a distinct frame for the same game can elicit different shared beliefs. There is surely nothing "irrational" about such a phenomenon. Moreover, when confronted with a novel social setting in the laboratory, subjects must first decide what moral values to apply to the situation by *mapping the game onto some sphere of everyday life* to which they are accustomed. The verbal and other cues provided by experimenters are the clues that subjects use to "locate" the interaction in their social space, so that moral principles can be properly applied to the novel situation. Moreover, such framing instruments as calling subjects "partners" rather than "opponents" in describing the game can increase cooperation because strong reciprocators (Gintis 2000), who prefer to cooperate if others do the same, may increase their prior as to the probability that others will cooperate, given the "partner" as opposed to the "opponent" cue. In sum, framing is in fact an ineluctable part of the BPC model, properly construed.

It is widely argued that in many situations of extreme importance choice consistency fails, so preferences are not maximized. These cases include time inconsistency, in which individuals have very high short-term discount rates and much lower long-term discount rates (Herrnstein 1961, Ainslie 1975, Laibson 1997). As a result, people lack the will-power to sacrifice present pleasures for future wellbeing. This leads to such well-known behavioral problems as unsafe sex, crime, substance abuse, procrastination, under-saving, and obesity. It is thus held that these phenomena of great public policy importance are irrational and cannot be treated with the BPC model. Yet, adding a reference point to the objective function allows such behavior to be modeled insightfully within the rational actor framework (Gintis 2009a).

There remains perhaps the most widely recognized example of inconsistency, that of preference reversal in the choice of lotteries. Lichtenstein and Slovic (1971) were the first to find that in many cases, individuals who prefer lottery A to lottery B are nevertheless willing to take less money for A than for B. Reporting this to economists several years later, Grether and Plott (1979) assert "A body of data and theory has been developed. . . [that] are simply inconsistent with preference theory. . . (p. 623). These preference reversals were explained several years later by Tversky, Slovic and Kahneman (1990) as a bias toward the higher probability of winning in lottery choice and toward the higher the maximum amount of winnings in monetary valuation. If this were true for lotteries in general it might compromise the BPC model.[2] However, the phenomenon has been documented only when the

---

[2] I say "might" because in real life individuals generally do not choose among lotteries by observing or contemplating probabilities and their associated payoffs, but by imitating the behavior of

lottery pairs A and B are so close in expected value that one needs a calculator (or a quick mind) to determine which would be preferred by an expected value maximizer. For instance, in Grether and Plott (1979) the average difference between expected values of comparison pairs was 2.51% (calculated from Table 2, p. 629). The corresponding figure for Tversky et al. (1990) was 13.01%. When the choices are so close to indifference, it is not surprising that inappropriate cues are relied upon to determine choice.

14 Conclusion

I have proposed five analytical tools that together serve to provide a common basis for the behavioral sciences. These are gene-culture coevolution, the socio-psychological theory of norms, game theory, the rational actor model and complexity theory. While there are doubtless formidable scientific issues involved in providing the precise articulations between these tools and the major conceptual tools of the various disciplines, as exhibited, for instance, in harmonizing the socio-psychological theory of norms and repeated game theory (Gintis 2009a), these intellectual issues are likely to be dwarfed by the sociological issues surrounding the semi-feudal nature of modern the behavioral disciplines.

REFERENCES

Abbott, R. J., J. K. James, R. I Milne, and A. C. M Gillies, "Plant Introductions, Hybridization and Gene Flow," *Philosophical Transactions of the Royal Society of London B* 358 (2003):1123–1132.

Ainslie, George, "Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control," *Psychological Bulletin* 82 (July 1975):463–496.

Alcock, John, *Animal Behavior: An Evolutionary Approach* (Sunderland, MA: Sinauer, 1993).

Alexander, R. D., *The Biology of Moral Systems* (New York: Aldine, 1987).

Allais, Maurice, "Le comportement de l'homme rationnel devant le risque, critique des postulats et axiomes de l'école Américaine," *Econometrica* 21 (1953):503–546.

Allman, J., A. Hakeem, and K. Watson, "Two Phylogenetic Specializations in the Human Brain," *Neuroscientist* 8 (2002):335–346.

---

others who appear to be successful in their daily pursuits. In frequently repeated lotteries, the law of large numbers ensures that the higher expected value lottery will increase in popularity by imitation without any calculation by participants.

Andreoni, James and John H. Miller, "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism," *Econometrica* 70,2 (2002):737–753.

Arrow, Kenneth J. and Gerard Debreu, "Existence of an Equilibrium for a Competitive Economy," *Econometrica* 22,3 (1954):265–290.

Aumann, Robert J., "Correlated Equilibrium and an Expression of Bayesian Rationality," *Econometrica* 55 (1987):1–18.

Bala, V. and M. Majumdar, "Chaotic Tatonnement," *Economic Theory* 2 (1992):437–445.

Barkow, Jerome H., Leda Cosmides, and John Tooby, *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (New York: Oxford University Press, 1992).

Becker, Gary S. and Kevin M. Murphy, "A Theory of Rational Addiction," *Journal of Political Economy* 96,4 (August 1988):675–700.

Beer, J. S., E. A. Heerey, D. Keltner, D. Skabini, and R. T. Knight, "The Regulatory Function of Self-conscious Emotion: Insights from Patients with Orbitofrontal Damage," *Journal of Personality and Social Psychology* 65 (2003):594–604.

Bell, D. E., "Regret in Decision Making under Uncertainty," *Operations Research* 30 (1982):961–981.

Bicchieri, Cristina, *The The Grammar of Society: The Nature and Dynamics of Social Norms* (Cambridge: Cambridge University Press, 2006).

Binmore, Kenneth G., *Natural Justice* (Oxford: Oxford University Press, 2005).

Black, Fisher and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* 81 (1973):637–654.

Blount, Sally, "When Social Outcomes Aren't Fair: The Effect of Causal Attributions on Preferences," *Organizational Behavior & Human Decision Processes* 63,2 (August 1995):131–144.

Boehm, Christopher, *Hierarchy in the Forest: The Evolution of Egalitarian Behavior* (Cambridge, MA: Harvard University Press, 2000).

Bonner, John Tyler, *The Evolution of Culture in Animals* (Princeton, NJ: Princeton University Press, 1984).

Boyd, Robert and Peter J. Richerson, *Culture and the Evolutionary Process* (Chicago: University of Chicago Press, 1985).

— and — , "The Evolution of Reciprocity in Sizable Groups," *Journal of Theoretical Biology* 132 (1988):337–356.

Brown, J. H. and M. V. Lomolino, *Biogeography* (Sunderland, MA: Sinauer, 1998).

Camille, N., "The Involvement of the Orbitofrontal Cortex in the Experience of Regret," *Science* 304 (2004):1167–1170.

Cavalli-Sforza, Luca L. and Marcus W. Feldman, "Theory and Observation in Cultural Transmission," *Science* 218 (1982):19–27.

Cavalli-Sforza, Luigi L. and Marcus W. Feldman, *Cultural Transmission and Evolution* (Princeton, NJ: Princeton University Press, 1981).

Cohen, L. Jonathan, "Can Human Irrationality be Experimentally Demonstrated?," *Behavioral and Brain Sciences* 4 (1981):317–331.

Conte, Rosaria and Cristiano Castelfranchi, "From Conventions to Prescriptions. Towards an Integrated View of Norms," *Artificial Intelligence and Law* 7 (1999):323–340.

Cooper, W. S., "Decision Theory as a Branch of Evolutionary Theory," *Psychological Review* 4 (1987):395–411.

Darwin, Charles, *The Origin of Species by Means of Natural Selection* (London: John Murray, 1872). 6th Edition.

Dawkins, Richard, *The Selfish Gene* (Oxford: Oxford University Press, 1976).

— , *The Extended Phenotype: The Gene as the Unit of Selection* (Oxford: Freeman, 1982).

— , *The Selfish Gene, 2nd Edition* (Oxford: Oxford University Press, 1989).

Debreu, Gerard, "Excess Demand Function," *Journal of Mathematical Economics* 1 (1974):15–23.

deQuervain et al., Dominique J.-F., "The Neural Basis of Altruistic Punishment," *Science* 305 (27 August 2004):1254–1258.

Dorris, M. C. and P. W. Glimcher, "Monkeys as an Animal Model of Human Decision Making during Strategic Interactions," 2003. Under submission.

Dunbar, R. I. M., "Coevolution of Neocortical Size, Group Size and Language in Humans," *Behavioral and Brain Sciences* 16,4 (1993):681–735.

Durkheim, Emile, *The Division of Labor in Society* (New York: The Free Press, 1933[1902]).

Edgerton, Robert B., *Sick Societies: Challenging the Myth of Primitive Harmony* (New York: The Free Press, 1992).

Ellsberg, Daniel, "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics* 75 (1961):643–649.

Ely, Jeffrey C. and Juuso Välimäki, "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory* 102 (2002):84–105.

Eshel, Ilan and Marcus W. Feldman, "Initial Increase of New Mutants and Some Continuity Properties of ESS in two Locus Systems," *American Naturalist* 124 (1984):631–640.

— , — , and Aviv Bergman, "Long-term Evolution, Short-term Evolution, and Population Genetic Theory," *Journal of Theoretical Biology* 191 (1998):391–396.

Fehr, Ernst and Urs Fischbacher, "Third Party Punishment and Social Norms," *Evolution & Human Behavior* 25 (2004):63–87.

Feldman, Marcus W. and Lev A. Zhivotovsky, "Gene-Culture Coevolution: Toward a General Theory of Vertical Transmission," *Proceedings of the National Academy of Sciences* 89 (December 1992):11935–11938.

Fisher, Franklin M., *Disequilibrium Foundations of Equilibrium Economics* (Cambridge, UK: Cambridge University Press, 1983).

Fisher, Ronald A., *The Genetical Theory of Natural Selection* (Oxford: Clarendon Press, 1930).

Fudenberg, Drew, David K. Levine, and Eric Maskin, "The Folk Theorem with Imperfect Public Information," *Econometrica* 62 (1994):997–1039.

Gadagkar, Raghavendra, "On Testing the Role of Genetic Asymmetries Created by Haplodiploidy in the Evolution of Eusociality in the Hymenoptera," *Journal of Genetics* 70,1 (April 1991):1–31.

Gauthier, David, *Morals by Agreement* (Oxford: Clarendon Press, 1986).

Ghiselin, Michael T., *The Economy of Nature and the Evolution of Sex* (Berkeley: University of California Press, 1974).

Gigerenzer, Gerd and Reinhard Selten, *Bounded Rationality* (Cambridge, MA: MIT Press, 2001).

Gilovich, T., R. Vallone, and A. Tversky, "The Hot Hand in Basketball: On the Misperception of Random Sequences," *Journal of Personality and Social Psychology* 17 (1985):295–314.

Gintis, Herbert, "Welfare Economics and Individual Development: A Reply to Talcott Parsons," *Quarterly Journal of Economics* 89,2 (February 1975):291–302.

— , "Strong Reciprocity and Human Sociality," *Journal of Theoretical Biology* 206 (2000):169–179.

— , "The Hitchhiker's Guide to Altruism: Genes, Culture, and the Internalization of Norms," *Journal of Theoretical Biology* 220,4 (2003):407–418.

— , "Solving the Puzzle of Human Prosociality," *Rationality and Society* 15,2 (May 2003):155–187.

— , "Behavioral Game Theory and Contemporary Economic Theory," *Analyze & Kritik* 27,1 (2005):48–72.

— , "The Emergence of a Price System from Decentralized Bilateral Exchange," *Contributions to Theoretical Economics* 6,1,13 (2006). Available at www.bepress.com/bejte/contributions/vol6/iss1/art13.

— , "The Dynamics of General Equilibrium," *The Economic Journal* 117 (October 2007):1289–1309.

— , *The Bounds of Reason: Game Theory for the Behavioral Sciences* (Princeton, NJ: Princeton University Press, 2009).

— , *Game Theory Evolving* (Princeton, NJ: Princeton University Press, 2009). Second Edition.

— , Joseph Henrich, Samuel Bowles, Robert Boyd, and Ernst Fehr, "Strong Reciprocity and the Roots of Human Morality," *Social Justice Research (*2009).

— , Samuel Bowles, Robert Boyd, and Ernst Fehr, *Moral Sentiments and Material Interests: On the Foundations of Cooperation in Economic Life* (Cambridge: The MIT Press, 2005).

Glimcher, Paul W., *Decisions, Uncertainty, and the Brain: The Science of Neuroeconomics* (Cambridge, MA: MIT Press, 2003).

— , Michael C. Dorris, and Hannah M. Bayer, "Physiological Utility Theory and the Neuroeconomics of Choice," 2005. Center for Neural Science, New York University.

Gneezy, Uri, "Deception: The Role of Consequences," *American Economic Review* 95,1 (March 2005):384–394.

— and Aldo Rustichini, "A Fine is a Price," *Journal of Legal Studies* 29 (2000):1–17.

Goffman, Erving, *The Presentation of Self in Everyday Life* (New York: Anchor, 1959).

Goldstein, E. Bruce, *Cognitive Psychology: Connecting Mind, Research, and Everyday Experience* (New York: Wadsworth, 2005).

Grafen, Alan, "Formal Darwinism, the Individual-as-maximizing-agent Analogy, and Bet-hedging," *Proceedings of the Royal Society B* 266 (1999):799–803.

— , "Developments of Price's Equation and Natural Selection Under Uncertainty," *Proceedings of the Royal Society B* 267 (2000):1223–1227.

— , "A First Formal Link between the Price Equation and an Optimization Program," *Journal of Theoretical Biology* 217 (2002):75–91.

Grether, David and Charles Plott, "Economic Theory of Choice and the Preference Reversal Phenomenon," *American Economic Review* 69,4 (September 1979):623–638.

Grice, H. P., "Logic and Conversation," in Donald Davidson and Gilbert Harman (eds.) *The Logic of Grammar* (Encino, CA: Dickenson, 1975) pp. 64–75.

Gruber, J. and B. Koszegi, "Is Addiction Rational? Theory and Evidence," *Quarterly Journal of Economics* 116,4 (2001):1261–1305.

Grusec, Joan E. and Leon Kuczynski, *Parenting and Children's Internalization of Values: A Handbook of Contemporary Theory* (New York: John Wily & Sons, 1997).

Guzman, R. A., Carlos Rodriguez Sickert, and Robert Rowthorn, "When in Rome do as the Romans do: the Coevolution of Altruistic Punishment, Conformist Learning, and Cooperation," *Evolution and Human Behavior* 28 (2007):112–117.

Haldane, J. B. S., *The Causes of Evolution* (London: Longmans, Green & Co., 1932).

Hamilton, William D., "The Genetical Evolution of Social Behavior I & II," *Journal of Theoretical Biology* 7 (1964):1–16,17–52.

Hammerstein, Peter, "Darwinian Adaptation, Population Genetics and the Streetcar Theory of Evolution," *Journal of Mathematical Biology* 34 (1996):511–532.

— and Reinhard Selten, "Game Theory and Evolutionary Biology," in Robert J. Aumann and Sergiu Hart (eds.) *Handbook of Game Theory with Economic Applications* (Amsterdam: Elsevier, 1994) pp. 929–993.

Heiner, Ronald A., "The Origin of Predictable Behavior," *American Economic Review* 73,4 (1983):560–595.

Henrich et al., Joseph, "Costly Punishment Across Human Societies," *Science* 312 (2006):1767–1770.

Henrich, Joseph and Francisco Gil-White, "The Evolution of Prestige: Freely Conferred Status as a Mechanism for Enhancing the Benefits of Cultural Transmission," *Evolution and Human Behavior* 22 (2001):165–196.

— and Robert Boyd, "The Evolution of Conformist Transmission and the Emergence of Between-Group Differences," *Evolution and Human Behavior* 19 (1998):215–242.

— and — , "Why People Punish Defectors: Weak Conformist Transmission can Stabilize Costly Enforcement of Norms in Cooperative Dilemmas," *Journal of Theoretical Biology* 208 (2001):79–89.

— , — , Samuel Bowles, Colin Camerer, Ernst Fehr, and Herbert Gintis, *Foundations of Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-scale Societies* (Oxford: Oxford University Press, 2004).

Herrnstein, Richard J., "Relative and Absolute Strengths of Responses as a Function of Frequency of Reinforcement," *Journal of Experimental Analysis of Animal Behavior* 4 (1961):267–272.

— , David Laibson, and Howard Rachlin, *The Matching Law: Papers on Psychology and Economics* (Cambridge, MA: Harvard University Press, 1997).

Hilton, Denis J., "The Social Context of Reasoning: Conversational Inference and Rational Judgment," *Psychological Bulletin* 118,2 (1995):248–271.

Holden, C. J., "Bantu Language Trees Reflect the Spread of Farming Across Sub-Saharan Africa: A Maximum-parsimony Analysis," *Proc. Roy. Soc. B* 269

(2002):793–799.

— and Ruth Mace, “Spread of Cattle Led to the Loss of Matrilineal Descent in Africa: A Coevolutionary Analysis,” *Proc. Roy. Soc. B* 270 (2003):2425–2433.

Holland, John H., *Adaptation in Natural and Artificial Systems* (Ann Arbor: University of Michigan Press, 1975).

Huxley, Julian S., “Evolution, Cultural and Biological,” *Yearbook of Anthropology (*1955):2–25.

Jablonka, Eva and Marion J. Lamb, *Epigenetic Inheritance and Evolution: The Lamarckian Case* (Oxford: Oxford University Press, 1995).

James, William, “Great Men, Great Thoughts, and the Environment,” *Atlantic Monthly* 46 (1880):441–459.

Jaynes, E. T., *Probability Theory: The Logic of Science* (Cambridge: Cambridge University Press, 2003).

Kahneman, Daniel and Amos Tversky, *Choices, Values, and Frames* (Cambridge: Cambridge University Press, 2000).

— , Paul Slovic, and Amos Tversky, *Judgment under Uncertainty: Heuristics and Biases* (Cambridge, UK: Cambridge University Press, 1982).

Kirman, Alan P. and K. J. Koch, “Market Exchess Demand in Exchange Economies with Identical Preferences and Collinear Endowments,” *Review of Economic Studies* LIII (1986):457–463.

Krantz, D. H., “From Indices to Mappings: The Representational Approach to Measurement,” in D. Brown and J. Smith (eds.) *Frontiers of Mathematical Psychology* (Cambridge: Cambridge University Press, 1991) pp. 1–52.

Krebs, J. R. and N. B. Davies, *Behavioral Ecology: An Evolutionary Approach* fourth ed. (Oxford: Blackwell Science, 1997).

Kreps, David M., *A Course in Microeconomic Theory* (Princeton, NJ: Princeton University Press, 1990).

Laibson, David, “Golden Eggs and Hyperbolic Discounting,” *Quarterly Journal of Economics* 112,2 (May 1997):443–477.

Lehmann, Laurent and Laurent Keller, “The Evolution of Cooperation and Altruism—A General Framework and a Classification of Models,” *Journal of Evolutionary Biology* 19 (2006):1365–1376.

— , F. Rousset, D. Roze, and Laurent Keller, “Strong reciprocity or Strong Ferocity? A Population Genetic View of the Evolution of Altruistic Punishment,” *The American Naturalist* 170,1 (July 2007):21–36.

Lewis, David, *Conventions: A Philosophical Study* (Cambridge, MA: Harvard University Press, 1969).

Lewontin, Richard C., *The Genetic Basis of Evolutionary Change* (New York: Columbia University Press, 1974).

Liberman, Uri, "External Stability and ESS Criteria for Initial Increase of a New Mutant Allele," *Journal of Mathematical Biology* 26 (1988):477–485.

Lichtenstein, Sarah and Paul Slovic, "Reversals of Preferences Between Bids and Choices in Gambling Decisions," *Journal of Experimental Psychology* 89 (1971):46–55.

Loomes, G. and Robert Sugden, "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty," *Economic Journal* 92 (1982):805–824.

Mace, Ruth and Mark Pagel, "The Comparative Method in Anthropology," *Current Anthropology* 35 (1994):549–564.

Masclet, David, Charles Noussair, Steven Tucker, and Marie-Claire Villeval, "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism," *American Economic Review* 93,1 (March 2003):366–380.

Maynard Smith, John, *Evolution and the Theory of Games* (Cambridge, UK: Cambridge University Press, 1982).

Mazur, James E., *Learning and Behavior* (Upper Saddle River, NJ: Prentice-Hall, 2002).

Mednick, S. A., L. Kirkegaard-Sorenson, B. Hutchings, J Knop, R. Rosenberg, and F. Schulsinger, "An Example of Bio-social Interaction Research: The Interplay of Socio-environmental and Individual Factors in the Etiology of Criminal Behavior," in S. A. Mednick and K. O. Christiansen (eds.) *Biosocial Bases of Criminal Behavior* (New York: Gardner Press, 1977) pp. 9–24.

Meltzhoff, Andrew N. and J. Decety, "What Imitation Tells us About Social Cognition: A Rapprochement Between Developmental Psychology and Cognitive Neuroscience," *Philosophical Transactions of the Royal Society of London B* 358 (2003):491–500.

Mesoudi, Alex, Andrew Whiten, and Kevin N. Laland, "Towards a Unified Science of Cultural Evolution," *Behavioral and Brain Sciences (*2006).

Miller, B. L., A. Darby, D. F. Benson, J. L. Cummings, and M. H. Miller, "Aggressive, Socially Disruptive and Antisocial Behaviour Associated with Fronto-temporal Dementia," *British Journal of Psychiatry* 170 (1997):150–154.

Moll, Jorge, Roland Zahn, Ricardo di Oliveira-Souza, Frank Krueger, and Jordan Grafman, "The Neural Basis of Human Moral Cognition," *Nature Neuroscience* 6 (October 2005):799–809.

Montague, P. Read and Gregory S. Berns, "Neural Economics and the Biological Substrates of Valuation," *Neuron* 36 (2002):265–284.

Moran, P. A. P., “On the Nonexistence of Adaptive Topographies,” *Annals of Human Genetics* 27 (1964):338–343.

Newman, Mark, Albert-Laszlo Barabasi, and Duncan J. Watts, *The Structure and Dynamics of Networks* (Princeton, NJ: Princeton University Press, 2006).

Nisbett, Richard E. and Dov Cohen, *Culture of Honor: The Psychology of Violence in the South* (Boulder: Westview Press, 1996).

O’Brian, M. J. and R. L. Lyman, *Applying Evolutionary Archaeology* (New York: Kluwer Academic, 2000).

Odling-Smee, F. John, Kevin N. Laland, and Marcus W. Feldman, *Niche Construction: The Neglected Process in Evolution* (Princeton: Princeton University Press, 2003).

Parker, A. J. and W. T. Newsome, “Sense and the Single Neuron: Probing the Physiology of Perception,” *Annual Review of Neuroscience* 21 (1998):227–277.

Parsons, Talcott, *The Structure of Social Action* (New York: McGraw-Hill, 1937).

— , “Evolutionary Universals in Society,” *American Sociological Review* 29,3 (June 1964):339–357.

— , *Sociological Theory and Modern Society* (New York: Free Press, 1967).

Pinker, Steven, *The Blank Slate: The Modern Denial of Human Nature* (New York: Viking, 2002).

Popper, Karl, *Objective knowledge: An Evolutionary Approach* (Oxford: Clarendon Press, 1979).

Poundstone, William, *Prisoner’s Dilemma* (New York: Doubleday, 1992).

Price, Michael, Leda Cosmides, and John Tooby, “Punitive Sentiment as an Anti-Free Rider Psychological Device,” *Evolution & Human Behavior* 23,3 (May 2002):203–231.

Real, Leslie A., “Animal Choice Behavior and the Evolution of Cognitive Architecture,” *Science* 253 (30 August 1991):980–986.

Real, Leslie and Thomas Caraco, “Risk and Foraging in Stochastic Environments,” *Annual Review of Ecology and Systematics* 17 (1986):371–390.

Richerson, Peter J. and Robert Boyd, “The Evolution of Ultrasociality,” in I. Eibl-Eibesfeldt and F.K. Salter (eds.) *Indoctrinability, Idology and Warfare* (New York: Berghahn Books, 1998) pp. 71–96.

— and — , *Not By Genes Alone* (Chicago: University of Chicago Press, 2004).

Rivera, M. C. and J. A. Lake, “The Ring of Life Provides Evidence for a Genome Fusion Origin of Eukaryotes,” *Nature* 431 (2004):152–155.

Rizzolatti, G., L. Fadiga, L Fogassi, and V. Gallese, “From Mirror Neurons to Imitation: Facts and Speculations,” in Andrew N. Meltzhoff and Wolfgang Prinz (eds.) *The Imitative Mind: Development, Evolution and Brain Bases* (Cambridge: Cambridge University Press, 2002) pp. 247–266.

Rozin, Paul, L. Lowery, S. Imada, and Jonathan Haidt, "The CAD Triad Hypothesis: A Mapping Between Three Moral Emotions (Contempt, Anger, Disgust) and Three Moral Codes (Community, Autonomy, Divinity)," *Journal of Personality & Social Psychology* 76 (1999):574–586.

Saari, Donald G., "Iterative Price Mechanisms," *Econometrica* 53 (1985):1117–1131.

Saffer, Henry and Frank Chaloupka, "The Demand for Illicit Drugs," *Economic Inquiry* 37,3 (1999):401–11.

Schall, J. D. and K. G. Thompson, "Neural Selection and Control of Visually Guided Eye Movements," *Annual Review of Neuroscience* 22 (1999):241–259.

Schulkin, J., *Roots of Social Sensitivity and Neural Function* (Cambridge, MA: MIT Press, 2000).

Schultz, W., P. Dayan, and P. R. Montague, "A Neural Substrate of Prediction and Reward," *Science* 275 (1997):1593–1599.

Seeley, Thomas D., "Honey Bee Colonies are Group-Level Adaptive Units," *The American Naturalist* 150 (1997):S22–S41.

Shafir, Eldar and Robyn A. LeBoeuf, "Rationality," *Annual Review of Psychology* 53 (2002):491–517.

Shennan, Stephen, *Quantifying Archaeology* (Edinburgh: Edinburgh University Press, 1997).

Simon, Herbert, "Theories of Bounded Rationality," in C. B. McGuire and Roy Radner (eds.) *Decision and Organization* (New York: American Elsevier, 1972) pp. 161–176.

— , *Models of Bounded Rationality* (Cambridge, MA: MIT Press, 1982).

Skibo, James M. and R. Alexander Bentley, *Complex Systems and Archaeology* (Salt Lake City: University of Utah Press, 2003).

Slovic, Paul, "The Construction of Preference," *American Psychologist* 50,5 (1995):364–371.

Smith, Eric Alden and B. Winterhalder, *Evolutionary Ecology and Human Behavior* (New York: Aldine de Gruyter, 1992).

Sonnenschein, Hugo, "Market Excess Demand Functions," *Econometrica* 40 (1972):549–563.

— , "Do Walras' Identity and Continuity Characterize the Class of Community Excess Demand Functions?," *Journal of Economic Theory* 6 (1973):345–354.

Stanovich, Keith E., *Who is Rational? Studies in Individual Differences in Reasoning* (New York: Lawrence Erlbaum Associates, 1999).

Sternberg, Robert J. and Richard K. Wagner, *Readings in Cognitive Psychology* (Belmont, CA: Wadsworth, 1999).

Sugden, Robert, "An Axiomatic Foundation for Regret Theory," *Journal of Economic Theory* 60,1 (June 1993):159–180.

Sugrue, Leo P., Gregory S. Corrado, and William T. Newsome, "Choosing the Greater of Two Goods: Neural Currencies for Valuation and Decision Making," *Nature Reviews Neuroscience* 6 (2005):363–375.

Sutton, R. and A. G. Barto, *Reinforcement Learning* (Cambridge, MA: The MIT Press, 2000).

Taylor, Peter and Leo Jonker, "Evolutionarily Stable Strategies and Game Dynamics," *Mathematical Biosciences* 40 (1978):145–156.

Tooby, John and Leda Cosmides, "The Psychological Foundations of Culture," in Jerome H. Barkow, Leda Cosmides, and John Tooby (eds.) *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (New York: Oxford University Press, 1992) pp. 19–136.

Trivers, Robert L., "The Evolution of Reciprocal Altruism," *Quarterly Review of Biology* 46 (1971):35–57.

Tversky, Amos and Daniel Kahneman, "Belief in the Law of Small Numbers," *Psychological Bulletin* 76 (1971):105–110.

— and Daniel Kanheman, "Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment," *Psychological Review* 90 (1983):293–315.

— , Paul Slovic, and Daniel Kahneman, "The Causes of Preference Reversal," *American Economic Review* 80,1 (March 1990):204–217.

Von Neumann, John and Oskar Morgenstern, *Theory of Games and Economic Behavior* (Princeton, NJ: Princeton University Press, 1944).

Wason, P. C., "Reasoning," in B. Foss (ed.) *New Horizons in Psychology* (Harmonsworth: Penguin, 1966) pp. 135–151.

Wetherick, N. E., "Reasoning and Rationality: A Critique of Some Experimental Paradigms," *Theory & Psychology* 5,3 (1995):429–448.

Williams, J. H. G., A. Whiten, T. Suddendorf, and D. I Perrett, "Imitation, Mirror Neurons and Autism," *Neuroscience and Biobheavioral Reviews* 25 (2001):287–295.

Wilson, David Sloan and Edward O. Wilson, "Rethinking the Theoretical Foundation of Sociobiology," *The Quarterly Review of Biology* 82,4 (December 2007):327–348.

Wilson, E. O. and Bert Holldobler, "Eusociality: Origin and Consequences," *PNAS* 102,38 (2005):13367–71.

Winter, Sidney G., "Satisficing, Selection and the Innovating Remnant," *Quarterly Journal of Economics* 85 (1971):237–261.

Wright, Sewall, "Evolution in Mendelian Populations," *Genetics* 6 (1931):111–178.

Wrong, Dennis H., "The Oversocialized Conception of Man in Modern Sociology," *American Sociological Review* 26 (April 1961):183–193.

Young, H. Peyton, *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions* (Princeton, NJ: Princeton University Press, 1998).

Zajonc, R. B., "Feeling and Thinking: Preferences Need No Inferences," *American Psychologist* 35,2 (1980):151–175.

Zajonc, Robert B., "On the Primacy of Affect," *American Psychologist* 39 (1984):117–123.

c\Papers\Unity of Science\NewUnity.tex May 13, 2008