

CAN GAME(S) THEORY EXPLAIN CULTURE?¹

Jenna Bednar

and

Scott Page

University of Michigan

August 2000

updated 29 July 02

Culture is described as behavioral consistency and blamed for suboptimal strategy choice. That is, it is often fingered as the excuse for otherwise unexplainable “irrational” behavior. Here, we build a formal model (both agent-based and game theoretic) that *derives* behavioral consistency across games and can help explain the underlying rationality of suboptimal choice. We do so by introducing “Games Theory,” a method to analyze game *ensembles*, or sets of games played concurrently. We prove that cognitively-constrained agents will choose context specific strategies and may play common games differently. Moreover, even when cognitive constraints do not bind, agents may evolve ensemble dependent (but payoff equivalent) strategies. Our results provide game theoretic foundations for cultural diversity. We show how the model can be extended to a dynamic context to provide a theoretical mechanism for institutional path dependence. We also comment on the clash of cultures as it relates to our framework.

¹ jbednar@umich.edu and spage@umich.edu. A previous version of this paper was presented at the annual meeting of the American Political Science Association, Washington, D.C., September 2000 and the Midwest Political Science Association meeting, Chicago, April 2001. For comments on earlier drafts we thank Bob Axelrod, Rob Axtell, Ken Kollman, Paul Mahoney, and Ken Shotts, as well as seminar audiences at Columbia University, Washington University, and the Santa Fe Institute. Professor Page acknowledges funding and support from the MacArthur Foundation and the McDonnell Foundation.

1 INTRODUCTION

We provide a game theoretic foundation for the emergence of culture---a pattern of behavior--- and demonstrate the rationality of diverse cultural behaviors that can exhibit suboptimal behavior in isolated situations. We do so by introducing the modeling innovation of simultaneous multiple game analysis in which agents can play different strategies in each game. This enables us to capture the influence of context on strategic behavior. We show how agents choose behavioral routines in response to their full ensemble of games and not on a game by game basis. Ostensibly, cultural differences---the rich fabric of religions, languages, art, law, morals, customs, and beliefs that diversifies societies---and their impact would seem to be at odds with traditional game theoretic assumptions of optimizing behavior. Political scientists tend to invoke the term culture to explain regional exceptionalism. More generally, social scientists use culture to explain divergences from the rational behavior assumed by game theorists, such as when political institutions perform poorly, firms fail to profit maximize, or societies do not respond to information or reject an obvious advancement. In light of this, the idea of game theoretic foundations deriving culture sounds inherently contradictory. Yet we show that culture may be an optimal response to a multifaceted strategic environment.

That culture matters is indisputable. There are numerous empirical studies demonstrating culture's impact on the choices made by individuals and communities. But these studies do not coalesce into a tractable, general theory of how cultures emerge, making it difficult to answer even the most basic theoretical questions: What causes within-group behavioral consistency? Why do cultures have identifiable characteristics such as hot or cold? Why does culture change so slowly? When does culture lead someone to act suboptimally? How do institutions affect culture? How does culture affect institutional performance?

In this paper, we propose a specific endogenous model of culture that responds to some of these questions and provides a road map of how to analyze the others. In the model, a community-level behavioral pattern---culture---emerges collectively as agents evolve in strategic environments. The paper relies on standard game theoretic modeling techniques, but we make two empirically plausible but theoretically novel assumptions: agents play multiple distinct games and are cognitively constrained. Given these assumptions, agents' behavior in any particular game depends upon their entire strategic environment, namely the ensemble of games that they play. Although these agents do play multiple games simultaneously, they can differentiate between the games in their ensemble and play different strategies in each. This approach is a "tool kit" model of culture: agents select collections of strategies to fit the circumstance.² Culture emerges not in spite of optimizing motivations, but because of how those motivations are affected by incentives, cognitive constraints, and institutional precedents.³

² Swidler (1986:273) writes: Culture influences behavior by shaping a person's "repertoire or 'tool kit' of habits, skills, and styles from which people construct 'strategies of action.'"

Increasingly, social scientists are coming to view individuals as possessing problem solving and strategic tool kits - collections of perspectives, heuristics, and strategies (Hong and Page 2001, Gigerenzer and Selten 2001).

³ Our agents optimize relative to constraints. The unconstrained optimization assumption does not stand up well empirically unless people are confronting rather simple problems (Rabin 1998, Conlisk 1996). Even among scholars who believe that optimizing behavior is not meant to be descriptively accurate, but instead is intended as a benchmark, there is support for looking for richer alternatives or amendments. The three most promising approaches all build upon

Our model opens up the possibility of a more rigorous explanation of the remarkable and stable cultural differences that have been identified. Here, we give just a sampling from both the survey and experimental literatures. In surveys of 43 societies, Inglehart (1977, 1990, 1997) finds significant difference between societies in his respondents' values, from the way they view authority, from religion to government, and tolerate diversity, including gender and lifestyle, as well as the variety of their goals, from number of children to importance of leisure to material consumption and concern for the environment. Cultural differences correlate with diversity in important activities, including political participation (Almond & Verba 1963; Inglehart 1997). In controlled environments from both primitive and advanced societies, experimentalists have found that cultural differences affect the way individuals behave when faced with identical problems. In bargaining experiments in Israel, Japan, the United States, and Yugoslavia, Roth et al (1991) found significant differences in their subjects' behavior that they concluded could not be due to language—which would have shown up in their market experiments as well—but to cultural norms about negotiation. More recently, experimental results from fifteen small-scale societies provide evidence of cultural determinants of behavior in games (Henrich, et al. 2001). Societies that must cooperate in their daily economic activity, such as in the hunting of whales, are more likely to play cooperatively in experiments. Similarly, there is a substantial literature

theoretical foundations but are based on empirical regularities: (a) that people tend to learn good strategies gradually; they do not immediately intuit optimal behavior (Fudenberg and Levine 1997, Kalai and Lehrer 1988), (b) that people suffer from cognitive biases (Thaler 1994), and (c) that people possess limited cognitive capacity (Radner 1993, Rubinstein 1986, Kalai and Stanford 1988, Banks and Sundaram 1990). Our paper belongs to this third research agenda.

on corporate culture that demonstrates behavioral similarities of workers within firms (Ouchi and Wilkins 1985).

Evidence from psychology also points to the relevance of culture to behavior. Social psychologist Richard Nisbett and colleagues (2001) have found empirical evidence that thought processes differ between people of eastern and western cultures. In experiments, as well as reviews of the historical and social records, Nisbett and colleagues find strong support that East Asians tend to have a holistic interpretation of the world, while westerners tend to be more atomistic in their thinking. They hypothesize that the differences arise due to different social systems. Perhaps most compelling is the recent neurological research suggests that human cultures influence the formation of neural architecture. Our genes may encode a scaffolding of how neuronal connections will form but the particulars will be partially determined by our experiences. In other words, culture plays a role in brain formation (Quartz and Sejnowski 2002).

These previous studies emphasize how within a culture individuals behave and think similarly but how across cultures they differ. A second body of research demonstrates that these differences have profound impacts on institutional performance. In his seminal work explaining the divergence in democratic governance between northern and southern Italy, Putnam (1993) finds that in northern Italy, social capital developed a climate of trust that fed and legitimized democratic institutions. In southern Italy, this trust did not develop. The weight of the evidence supporting this cultural explanation for the success and failure of democratic institutions is incontrovertible. Parallel studies back culture's effect on political institutions (Chong 2000, Elster et al 1998, Huntington 1981, Putnam 1973, Schaffer 1998), while others show how it

affects the performance of economic institutions (Greif 1994, Litwack 1991, North 1990) and attempts at legal reform (Hay & Shleifer 1998).

The impact of culture would be less important if cultures could be changed to fit circumstances. (Perhaps) unfortunately, culture proves remarkably persistent. Putnam states that patterns that emerged from the “mists of the Dark Ages” (1993:180) doom or buoy democracy today. Change occurs very slowly: decades for institutional patterns, centuries for norms of social capital. Individual values are formed largely by the time one reaches adulthood and for the most part are stable thereafter (Jennings & Niemi 1981; Jennings & Stoker unpub; Jennings, Stoker & Bowers unpub), so to account for cultural shifts, the work of Inglehart (1997) and Nisbett et al (2001) indicate that cultural shifts are attributable to generational replacement.

Studying the pattern of change, Inglehart (1997) finds that economic, political, and cultural change track one another. This interconnectivity is fascinating and perplexing, as activity in one realm holds up or drags along activity in another. Environmental and institutional changes often have unforeseen consequences on human behavior. Political and economic alterations spill over into other facets of social activity, or are held up by seemingly unrelated dimensions. The U.S. Constitution, with its system of presidential democracy, successfully balances powers, but when transplanted to the southern hemisphere, presidentialism has led to authoritarianism (e.g. Linz & Valenzuela 1994). Development, as politicians and economists have painfully learned, is too complex for piecemeal analysis and isolated institutional fixes. Post-communist transitions have shown that a market economy and democratic openness are intertwined, but a society without experience in either lacks the norms necessary to carry out the transformation (Litwack 1991). Innovative suggestions abound for “growing” these norms, including local elections to foster party development (Ordeshook & Shvetsova 1996) or micro

credit lending banks to stimulate individual entrepreneurship and collective responsibility (Pitt & Khandker 1998). In short, context seems preeminent: new institutions succeed by building upon behavioral patterns induced by extant institutions.

To model the reality suggested by this remarkable empirical work on culture, game theorists must shift their perspective from a more standard vertical time dependence (repeated game theory) to a horizontal context dependence, with an emphasis on how the historical, political, and economic dimensions of people's lives affect the decisions people make in isolated situations. With this horizontal emphasis, history matters because it shapes today's portfolio of activities and decisions. Since we aim to explain the emergence of culture, we cannot incorporate culture into the utility function or assume cultural differences and use them to "type" the agents (Greif 1994, Fearon & Laitin 1996).⁴ To do so is to formalize culture as the residual category, and not as something that emerges from the environment, entangles behavior importantly, persists beyond its apparent adaptive usefulness, and adjusts to change slowly. Nor can we make an appeal to culture as an external factor that refines equilibria. To do so would be to assume culture's existence and importance.

Our simultaneous-games approach is a cousin of Tsebelis's (1990) work on "nested games": his agents face multiple games with variable payoffs; suboptimality is a judgment made by observers who do not see the full ensemble of games played by the agent. Conceptually, we share a common spirit with Bowles and Gintis (1986), who discuss the importance of

⁴ Important exceptions exist, such as those models that incorporate a changing payoff matrix. Taylor (1987, pp. 107-108) makes this suggestion explicitly. Calvert and Johnson (1997) add uncertainty about what game the agent is playing.

overlapping games---e.g. family, state, and economy---to understanding individual behavior. Our methodology is related to Samuelson (2000), where agents play two games, the more important being the Red Queen game, in which payoffs increase in cognitive effort.⁵ This construction allows him to avoid placing the constraint on cognitive capacity that we do here. Samuelson emphasizes how cognitive attention varies with the frequency of games, a result of our model as well.

The approach taken in both our paper and Samuelson's shares a common spirit with the work of Gilboa and Schmeidler on Case-Based Decision Theory (1995) and the recent book by Chong (2000). In a Case-Based model, agents employ similarity functions to classify problems relative to those they have encountered in the past. Similarity between problems increases the likelihood that an agent takes a similar action. For example, in Chong's model agents accumulate dispositions based on their life history of interactions. Thus, one could use the Gilboa-Schmeidler approach to model diverse behavior as a function of variations in life experiences. By confronting unique ensembles of situations, they may invoke different rules by virtue of having evolved distinct ideas about what will be successful actions. The Gilboa-Schmeidler approach, as with any based on hindsight, including Bayesian updating, is a vertical comparison, where agents compare their current problem with those that they have seen in the past. In contrast, in our model the vertical force of history determines the current context. We focus our analysis on the horizontal effect of context---agents playing multiple games

⁵ The Red Queen game derives from the character in Alice in Wonderland. The faster she went, the further behind she fell. In the Red Queen game, both players must choose a time commitment. The more time a player allocates relative to his opponent, the better he does.

concurrently---on behavior.

2 A MODEL OF GAME ENSEMBLES

To show how a common culture can evolve and persist, we analyze communities of agents that interact frequently and in multiple strategic situations. All agents within one community share a common *ensemble*, a collection of distinct games that all agents play. A two game ensemble might consist of the Prisoner's Dilemma game and a variant of the Battle of the Sexes. Agents have limited cognitive capacity, so while agents can distinguish between games, in some cases they will not be capable of playing every game optimally. Instead, they can simplify their strategies in some games, they can use a common strategy in multiple games, or they can borrow portions of a strategy from one game and apply it to another. They also learn from other members of their community. Through mathematical theorems and the results of simulations, we learn about how patterns of behavior develop and the effect that those patterns have on isolated interactions.

2.1 The Games

We consider ensembles where each game belongs to a class of two-person two-action games that include incentives for being selfish (S) and benefits from cooperating (C). In the first four games, cooperation lowers own payoff and raises the payoff of the other and being selfish does the opposite, so both players being selfish is the unique pure strategy equilibrium. In the final two games, these conditions hold only if both players were cooperative. This creates two pure strategy equilibria in which one player cooperates and one is selfish. The ensemble were chosen with great care so that the following outcomes were optimal in at least one of the games:

cooperation, selfishness, alternation, deference to women, deference to men, and indifference between two equilibria.⁶

insert Table 1 here

The first game is a standard Prisoner's Dilemma and needs no explanation. The second game, the Alternation game (AL), bears some resemblance to the Battle of the Sexes. Like the Battle of the Sexes, in the repeated game the players can both be made happy by alternating which player gets the high payoff, but unlike in the Battle of the Sexes, each player in Alternation has a dominant strategy. In the third game, which we call Knife Edge (KE), the player gets the same average payoff when it alternates between being selfish and cooperative as it does if it is always cooperative. In the fourth game, being selfish is not only dominant, but the strategy pair (S, S) generates the Pareto dominant equilibrium, so we call this the Self-Interest (SI) game. In the final two games, Top Right (TR) and Bottom Left (BL), the upper right and lower left cells are the respective unique equilibrium outcomes. Each of these last two games is biased in favor of one of the two players. Asymmetric coordination games are particularly interesting as a means to study the effect of domination within a community, such as male dominance, and the effect that this domination has in choice over institutions.

2.2 Automata Strategies

⁶ Apart from including other types of indifferences and this spans the set of all possible equilibria in these games. Alternation between (C,C) and (S,S) is not reasonable given our interpretations of C and S.

In this section, we highlight some of the strategies that evolve in response to ensembles comprised of combinations of the above games. We encode the strategies that the agents use to play these games using finite state automata. Automata are among the most basic classes of models used to represent human behavior (Rubinstein 1986, Kalai and Stanford 1988, Miller 1996), and unlike more elaborate representations such as neural networks or genetic programming, they are mathematically tractable and computationally transparent. Finite state automata consist of three parts: mental states M , transition mappings T , and initial states I . The mental states are numbered and each one prescribes an action, such as “be selfish” or “be cooperative.” The transition mappings tell the agent how to update her mental state in reaction to the opponent’s action. For example, the transition mapping from mental state 0 might say “go to state 1 if the opponent is selfish and stay in state 0 if she is cooperative.” The initial state denotes the agent’s starting mental state for a particular game. Automata that differ only in their initial states can represent both Tit For Tat and Suspicious Tit For Tat.

Two State Automata

A two state automaton has just two mental states, denoted by 0 and 1. These minimal automata permit a range of strategic behavior that includes Grim Trigger and Tit For Tat. We adopt the convention of writing two state automata as a list of numbers and characters beginning with an initial state (in this case state 0), and then an action and set of transitions for each state. The three-tuple (C, 0, 1) says “play C, stay in state 0 if the opponent plays C but move to state 1 if the opponent plays S.” We can then write Tit For Tat (TFT) as: {0, (C, 0, 1), (S, 0, 1)}.

In addition to TFT many familiar strategies can be encoded as two state automata. In Table 2, we identify and describe several strategies that emerge from our computational model.

Many strategies have a “suspicious” analog where the agent chooses S rather than C in the first period. We will denote those strategies by placing an S in front of the acronym, so STFT means Suspicious Tit For Tat and SALT means Alternation starting selfishly, rather than cooperatively.

The structure of several of the games makes some strategies prone to exploitation by other strategies. For example, the alternation strategy (ALT) can be exploited in the Alternation game by a strategy that chooses the player’s one-shot dominant strategy because ALT does not take into account the action of the other player. In contrast, the strategy Switch After C (SAC) cannot be exploited in the Alternation game. It only plays C after the opposing strategy has played C while SAC played S.

Three-state Automata

With more states, automata may evolve more sophisticated strategies. For example, using the notation from the previous section, we can write the strategy Tit-For-Every-Two-Tats, TFE2T, as: $\{0, (C,0,1), (C,1,2), (S,0,2)\}$. In the initial state, state 0, the agent plays C. If the other agent is selfish, then the agent moves to state 1, where the agent still plays C. However, whenever the other agent plays S again, the TFE2T agent will move to state 2 and play S. In state 2, the TFE2T will continue to play S as long as the other agent does. A similar strategy would be tit for two consecutive tats, TF2CT: $\{0, (C,0,1), (C,0,2), (S,0,2)\}$. In this strategy, the agent stops playing C if the other agent plays S twice in a row.⁷

⁷ State 1 is a memory of the other agent having been selfish; its transition mapping is the only technical difference between the strategies TFE2T and TF2CT. In TFE2T, the agent remains in state 1 if the other agent plays C. In TF2CT, the agent does not stay in state 1 for more than one period; if the opponent follows a play of S with another S immediately, the opponent is punished,

These examples demonstrate that simple automata can carry out sophisticated strategies, but they also reveal a weakness with measuring sophistication by the number of states. If history matters, automata need a minimum of $m + 1$ mental states and transition mapping combinations to describe m periods of memory or m particular actions. This problem notwithstanding, we will use the number of mental states as our measure of cognitive capacity.

2.3 From Automaton to Agent

Agents play ensembles of games with other members of their community. We represent the agents in our model much like the automata we just described with one key difference: when an agent plays more than one game, the agent must have initial states for each game. This allows them to differentiate between games. For example, suppose that an agent plays PD and SI. The automata would include an initial state for each game. An agent with only two states of capacity that plays GRIM in the PD and ALL S in the SI game would be written: {PD =0, SI = 1, (C,0,1), (S,1,1)}. In state 1 the agent plays S. The transition mapping creates an absorbing state; once the agent enters mental state 1, it does not leave it. State 1 is used as the starting point in SI, where S is the best strategy. State 1 is also used as a threat for the PD game. The strategy for the PD game begins in state 0 and plays C until the other agent plays S, at which point the

but if it plays C, all is forgotten. While the notational difference is slight, the implication regarding the agent's memory is enormous. TF2CT-playing agents only maintain a memory of the most recent two periods, while TFE2T maintain a potentially infinite memory of a single past deviation, although no record of when that injury occurred.

strategy moves to state 1 and the agent plays S. Thus, the strategy GRIM and the strategy ALL S can both be played with the same two-state automata.

We differentiate between row agents and column agents and number each type from 1 to N . Each row (column) agent plays with five neighboring column (row) agents. For example, row agent 21 will play with column agents 19, 20, 21, 22, and 23. To maintain symmetry, we place the agents on a circle, i.e. if there are 100 agents, then row agent 50 will play with column agents 48, 49, 50, 1, and 2.

Agents play a subset of the possible agents for empirical, theoretical, and practical reasons. Empirically, the phenomena we wish to study, differences in behavior as a function of ensemble, occurs through local interactions. The theoretical and practical reasons have to do with the speed that new strategies spread. Ellison (1993) and others have shown that movement to new equilibria is more likely when interactions are among subsets of agents. In practice, we found that the program ran more quickly without qualitative changes in outcomes, allowing us to perform more robustness checks.

3 RESULTS

We emphasize three phenomena that emerge in our model that are broadly consistent with the literature on culture. First, we see two types of *behavioral consistency*: both within an agent's toolkit, and across agents in a community. Agents choose or evolve strategies similar to others that they use, and like others in their community. Second, our model generates *contextual effects*: agents from different communities may play a game common to both ensembles differently. The strategic differences are predictable based upon the context each faces in its distinct community. Third, our model generates *isolated underperformance*: while agents evolve

strategies that are optimal across all games in their ensemble, when one game is examined independently, the strategy used to play it may be suboptimal.

We also demonstrate three implications related to these phenomena that align with the empirical, experimental, and theoretical literatures on culture: *frequency dependence*: how often a game is played influences how likely the agents evolve optimal strategies; *behavioral rigidity*: agents may not immediately respond to changes in their ensemble and this might lead to variance in institutional performance; and *institutional path dependence*: the selection and efficiency of political and economic institutions may depend upon how past choices have shaped the culture.

We support these results using both mathematical and agent-based models. We use mathematics to prove the inevitability of cultural effects such as behavioral consistency or bias and the agent-based models to generate particular examples that enable a more nuanced analysis of these cultural effects. We also rely on the agent-based model to counter the obvious criticism of our mathematical formulation. Our mathematical results only prove the existence of efficient equilibria. Were this the extent of our analysis, we would need to assume that our agents both find and coordinate on these efficient equilibrium strategies, an assumption that would raise eyebrows because of our agents' cognitive constraints. We use agent-based models to show that a simple learning dynamic enables our communities of agents to experiment and imitate, collectively evolving strategies that are optimal within their constraints.

3.1 Behavioral Consistency

We begin by showing that our agents develop consistent strategies to play the games within an ensemble. We do this by comparing strategies when games are played independently with the strategies that evolve when a game is part of an ensemble. To compare equilibrium outcomes,

we use a utilitarian social welfare function: the sum of the players' utilities. We begin by stating a claim that there exists a utilitarian social welfare maximizing equilibrium in two state automata for each of the six individual games. This claim is important because it shows that our agents are not overly constrained: they could always choose to play any one game optimally.

Claim 1: *For each of the six individual games there exists an equilibrium in two state automata that maximizes a utilitarian social welfare function.*⁸

In Table 3, we report results from our agent-based model⁹ that show that our agents can locate these equilibria. As it is equally important how they do so, we include the strategies. Our simple learning dynamic *evolves* intuitively and empirically plausible strategies using automata that are nearly always efficient. These models tend to evolve reasonable strategies quickly and robustly,

⁸ We prove all non-trivial claims in Appendix 3.

⁹ We present output from a computational model using 50 row agents and 50 column agents who played 25 rounds of each game. We assume no discounting; the utility-maximizing equilibrium generates the highest joint utility for the row and column agents totaled over the 25 rounds. We refer to one round of mutations and imitation as a cycle. We evolved strategies over 1000 cycles and performed at least twenty runs for each of the games. Typically, systems settled into equilibrium in 100 to 200 cycles. In the results that we describe in the next section, we ran the model for 500 cycles. We wrote two independent programs, one in C and one in Objective C. Most of the situations that we analyze were run in both platforms to ensure robustness.

that is, for a wide range of parameters. In the Prisoner's Dilemma, our agents evolved the cooperative but provokable strategies of TFT and GRIM. In the Alternation Game (AL), if the row agents play TFT or SAC and the column agents play STFT or SSAC, then the outcomes will alternate between the top right and the upper left. It can be shown that these pairs of strategies are an equilibrium and are utility maximizing. Unsurprisingly, in the Self-Interest game (SI) where Selfish is the dominant strategy, agents evolved automata generating selfish ALL S behavior.

insert Table 3 here

The Knife Edge game (KE), a hybrid between a coordination game and a Prisoner's Dilemma, generated mixed results. In most runs, agents evolve pairs of strategies that alternate between the upper right and lower left such as TFT for the row agents and SSAC for the column agents. All of the pairs generated are equilibria and are utility maximizing. In about 5% of the runs, all agents evolve TFT and GRIM, with the outcome being cooperative behavior. In two trial runs, we found hybrids in which some players play ALL C and others alternate. To see how this can happen suppose that row agents all play TFT and that column agents 8 through 10 play STFT while that the rest of the column agents play TFT. Any row agent playing a column agent numbered 8 through 10 will generate outcomes that alternate in the upper right and lower left. All other pairs of agents will always play C.

To show evidence of behavioral consistency, we compare the results from one game ensembles to two game ensembles. The six games generate fifteen ensemble pairs. We ran computational experiments on each. For each pair of games we performed approximately 20 runs, each of 1000 iterations. These experiments demonstrate whether the agents can attain the

utilitarian social welfare maximizing equilibria (when they exist), whether the strategies used to achieve those equilibria are ensemble dependent, and provide insight into the six cases where the automata lack the capacity to reach the utilitarian social welfare maximizing equilibria. We summarize the findings in Table 4.

We highlight two findings. First, when the Self-Interest game (SI) is paired with the Prisoner's Dilemma (PD) game, the agents that learn to cooperate in the PD use GRIM, while in the Self-Interest game they play ALL S. This combination of strategies, GRIM and ALL S, is behaviorally consistent in its unwillingness to forgive. This combination evolves because SI is an easier game. Agents tend to evolve an absorbing state in which they take action S to exploit the Self-Interest game. From there it takes fewer mutations to get to GRIM than it does to get to TFT. Mathematically, it is easy to prove that the agents could have also used a two state automata that played TFT in PD and STFT in SI. This strategy would also have been behaviorally consistent in that both strategies would have been forgiving. Second, consider the agents playing the PD/AL ensemble. The Alternation game requires coordination by type among the agents to alternate between outcomes. GRIM will not allow them to alternate in AL. In this ensemble the unique efficient equilibrium requires one type of agent (row or column) to play TFT in both games and the other type to play TFT in the Prisoner's Dilemma and STFT in the Alternation game. The agent-based model evolves this equilibrium behavior. As Table 4 reports, all use TFT rather than GRIM in the Prisoner's Dilemma when they do manage to maximize social welfare.

insert Table 4 here

We also ran experiments for each of the 20 possible three game ensembles. Here, we found

one interesting case where even though the outcome was identical to what occurred in the individual games, the behavior was different. In the ensemble of PD, AL, and SI, the row agents play TFT in the PD and AL games, and STFT in the Self-Interest game. The column agents play TFT in PD and STFT in AL and SI. The agents in our model evolved this strategy in 16 of the 20 runs on that ensemble of three games. With these strategies, the agents play (C, C) in PD, they alternate between (S, C) and (C, S) in AL, and they play (S, S) in SI. Comparing these to the single game strategies, we see that they evolve almost completely different strategies than when there is no cognitive pressure to behave consistently across games.

3.2 Contextual Behavior

One phenomenon often attributed to culture is when two people react differently to the same situation. When we invoke the adage “When in Rome....” we mean that we adjust our behavior to suit the context. One example of contextual behavior from the two game ensembles appears in the Knife Edge (KE) game. When played alone, the cooperative outcome (C, C) rarely emerged. However, when paired with the PD, the cooperative outcome in the KE game was supported over 35% of the time. But when paired with AL, (C, C) never arose. It is straightforward to prove that there is no mathematical constraint on playing either (C, C) or alternating in KE when paired with either of these games. Therefore, the contextual behavior must result from the learning dynamic. One possible explanation is that once the agents have found the efficient equilibrium strategy for PD (or AL), then with a single mutation in the initial state, that same strategy will be efficient in the Knife Edge game.

We found the most compelling evidence for emergent contextual effects on strategies and outcomes in the three game ensembles. In the PD/KE/AL ensemble, the agents played KE and

AL identically in all 20 runs. Similarly, when KE and AL are paired with SI, in all but one run the agents played KE and AL identically. But when PD and KE are paired with SI, the agents played KE and PD identically (i.e. (C, C)) and they did so using GRIM. Recall our discussion in Section 3.1: when PD is paired with SI, agents tend to play GRIM in the Prisoner's Dilemma. The fact that strategies in KE depend on the ensemble follows. The following observation provides a partial explanation.

Observation: *No two state automaton agent is capable of playing both GRIM and TFT.*

Note that the observation does not imply that there must be a contextual effect. The agents could evolve TFT and STFT so that they play (S, S) in SI, (C, C) in PD, and alternate in KE. We find, however, that they do not.

3.3 Isolated Underperformance

To demonstrate the inevitability of suboptimal behavior, we begin with some formal results and then show how underperformance manifests itself in particular ensembles using our agent-based model. Our first two lemmas provide the foundation for our main result that suboptimal behavior must exist in our model.

Lemma 1: *There exists an equilibrium in two state automata that maximizes a utilitarian social welfare function for the following nine two game ensembles: PD/AL, PD/KE, PD/SI, AL/KE, AL/SI, KE/SI, SI/TR, SI/BI, and TR/BL, but not the other six two game ensembles.*

The proof of the lemma is straightforward and contained in the Appendix. The combination of TR or BL with PD, AL, or KE creates too complex an environment for a two state automaton. To maintain (S, C) as an equilibrium in BL, the column agents must have a state in which they continue to play C even though the row agent is playing S. To play TFT or to play GRIM in the PD, they would need two other states for a total of three.

The next lemma states a similar result for the three game ensembles.

Lemma 2: *There exists an equilibrium in the space of two state automata that maximizes a utilitarian social welfare function for the following five three game ensembles: PD/AL/KE, PD/AL/SI, PD/KE/SI, AL/KE/SI, and TR/BL/SI, but not for the other 15 three game ensembles.*

Our final lemma shows that even with three state automata, the cognitive constraint binds and affects behavior.

Lemma 3: *There does not exist an equilibrium in the space of three state automata that maximizes a utilitarian social welfare function when the agents play BL and TR and at least one game from the set {PD, AL, KE}.*

These three lemmas enable us to state a claim that follows from them directly.

Claim 2: *If agents use automata with three or fewer states, then there exist game ensembles for which they must exhibit cultural bias in their behavior.*

We should note that the three state restriction on the size of the automata binds. Four state automata are sufficient.

Corollary: *For any ensemble comprised from our set of games, there exists an equilibrium in the space of four state automata that maximizes a utilitarian social welfare function in each game separately.*

We now explore the suboptimalities that we see in our agent-based model. For the two game ensembles, we focus on the six pairings described in Lemma 1 for which the agents cannot evolve utilitarian social welfare maximizing strategies. We see that cognitive simplicity dominates. In each game in the ensemble the agents play the one-shot Nash Equilibrium of the game. No sophisticated strategies evolve. As a result, the agents play (S, S) in PD when it belongs to an ensemble with either TR or BL. Furthermore, the Self-Interest game is unaffected in any of the ensembles. In the Self-Interest game, not only is being selfish dominant, but also the selfish equilibrium is Pareto efficient, rendering the game less susceptible to ensemble dependence. As long as (S, S) is an attracting state in the automata for each player, we get the efficient equilibrium in the SI game. In fact, when we increase the number of games, we still find that (S, S) is always played in SI.

In the three game ensembles in which the cognitive constraint binds (Lemma 2), we found a large influence of TR and BL. When two games from the set {PD, AL, KE} join either TR or BL in an ensemble, in over 90% of the cases the agents played equilibrium strategies in TR or BL and played (S, S) in the other two games. This appears to occur because TR and BL are simpler games, so the equilibria are easier to find. Since BL and TR are games mimicking social

structures biased toward one type, we see how the presence of these games can create a behavioral residue across other games by stifling cooperation. Within the context of our ensembles, the effect of being able to win big in one game by locking into a certain behavior appears to squelch any incentive to abandon that behavior for a norm of cooperation.

3.4 Frequency Dependence

We next analyze whether changing the frequency of a game's appearance can lead agents to evolve different strategies. To investigate this hypothesis, we constructed several ten game ensembles. The first ensemble contained five PD games, one SI game, two AL games, and one TR and BL game. In the second ensemble, we switched two of the PD games to SI games. The results were striking. In the first (PD-laden) ensemble, agents evolved cooperative behavior: in over 70% of the runs, they reached the cooperative outcome over 70% of the time. They failed to cooperate (less than 5% playing (C, C)) in less than 20% of the runs. When agents played the second ensemble, with more Self-Interest games, the agents played the cooperative equilibria over 50% of the time in only three of the 50 runs. They failed to cooperate at all in over 60% of the runs. The agents tend to play TR and BL optimally in this ensemble. The effect of the self-interest games ripple through the ensemble, and payoffs in the PD game suffer.

One implication of this finding is that if institutions can intervene to alter the ensemble structure, then they can change behavior. Imagine two societies. The first contains institutions that reward cooperation, like the Prisoner's Dilemma. Suppose that the second society's institutions offer relative fewer opportunities to cooperate. This second society's institutions, through ensemble effects, could breed selfish behavior, even in the few games that do reward cooperation. This intuition is supported by the recent ultimatum game experiments in fifteen

small-scale societies (Henrich et al 2001). The societies that had more opportunities for cooperative behavior in their daily lives, such as through collective hunting or with informal markets, tended to cooperate more in the experiments. The authors note the consistency between the behavior in the experiments and the economic behavior in these societies. Our model provides a theory for their experimental findings. People who play lots of games that demand cooperation, like the PD, will be more likely to cooperate in all games. One interpretation of these findings is that cooperative behavior can be made more likely through the creation of opportunities to be cooperative. Thus, institutional tinkering can have consequences for behavior that spill over from one game to another.

3.5 Behavioral Stickiness

One conflict in the political culture literature is empirical evidence of culture's persistence as well as change. Among those who notice its stasis, its persistence is puzzling because it resists change despite incentives to shift. Among those who see it shifting, it does so slowly but completely, in patterns that permeate the full fabric of people's lives. In our agent-based model, we experimented with changing the ensemble to see if the behavior responded immediately or exhibited stickiness. The natural experiment to run is to change the payoffs so that an alternative set of strategies has higher payoff than the current configuration and to see if the agents quickly adjust. We found that movements will be particularly unlikely if the new strategies conflict with existing patterns of behavior.

For example, in a society where cooperation (in the form of alternation schemes) frequently occurs, the switch to self interested behavior may not occur until the benefits of selfishness are sufficiently pronounced to effect a change in culture. We find just such a result in our model.

Consider a large ensemble containing multiple copies of AL. If we interchange copies of AL with TR and BL, we find that the agents often continue to alternate in all three games, even in ensembles where the agents would evolve the efficient equilibrium in BL and TR, rather than alternate, were they to evolve strategies from scratch.

This finding is particularly relevant in light of Ferejohn's (1991) paper analyzing the selection of candidates for Parliamentary elections in seventeenth-century England. In the early Stuart years, the majority of elections were not contested, but by the latter part of the period, potential candidates switched from the gentlemanly alternation of candidacy to a form of contestation more recognizable today, characterized by competitive campaigns. Ferejohn demonstrates how game theory could predict either outcome; a closer examination of history is needed to rate one as more plausible than the other. Ferejohn argues that the gentlemanly alternation generated a higher expected utility than probabilistic entry when there were few players and elections were held often. Thus, rational players should choose to alternate. Ferejohn then explains that the rise of competitive elections corresponded with a perceived decrease in the frequency of elections and a diminishment of the social hierarchy, which increased the number of potential candidates. Both changes made the cooperative agreements, necessary for the alternating equilibrium, less likely.

Notice that while Ferejohn's argument could not be made without a rich interpretation of history, it rests solely on efficiency. The alternating equilibrium is abandoned when it is no longer optimal. But did the transition occur at the moment when the expected utility comparison reversed? Calculations of expected utility based upon historical interpretation are fuzzy math. It is quite possible that the transition lagged by a generation or more; evidence indicates as much, with 15 contests in 1604 to 91 in the Long Parliament, starting in 1629 (Ferejohn 1991:292).

The delayed evolution is consistent with survey evidence of intergenerational change. Our model would explain that lag: the time taken to adjust behavior is a product of the influence of other strategic situations that the agent faces, and of the effect of routine. Adjustments are most natural between generations, when a new stock of potential candidates can evaluate their decision in ways consistent with other transformations in their lives and without the experiences that corset their elders' strategies.

Along similar lines, our model suggests an alternative explanation for the seeming inevitability of clashes between cultures. Suppose that a new trade route opens between two communities who did not interact before, and now they must share access to a common resource, such as a harbor. One society may play PD games with cooperative strategies, while the other plays a selfish strategy, perhaps because it rarely plays the PD. The second society's strategy of ALL S might lead the players in the first society to switch to ALL S and to think poorly of the second society. Alternatively, the second society might switch and play TFT or GRIM, but doing so might mean that it changes behavior in other contexts as well. Given the subtle interplay between strategies and ensembles, it does not follow immediately which strategy will win out, or if in fact the clash will continue unabated. One might anticipate the emergence of tagging strategies---selective cooperation based upon identifiable characteristics (Axelrod, Cohen, and Riolo 2001)---or in-group policing (Fearon & Laitin 1996).

3.6 Institutional Path Dependence

Our model also has implications for the historical development of institutions. Early institutional choices can influence later choices and the success of a society. Greif (1994) makes this point in his formal analysis of the 11th century Maghribi and the 12th century Genovese. Greif shows

how economic outcomes and institutional performance depend on characteristics of the societies: the trust-based, segregated economic relations of the Maghribi worked well as long as the trading circle was small, but the individualistic Genovese had institutions in place to enforce contracts, giving them the advantage in long-distance trading. In our framework, the societies choose different economic and political institutions as a function of previous decisions. So like Greif, we can explain why an institution might perform differently in two societies. In addition, we provide a framework within which we can explain the emergence of the cultural consistencies that drive these differences.

To demonstrate how path dependence arises, we construct a simple example.¹⁰ Consider a world with many distinct communities. Within each society agents play only one game, but they can choose which game to play. Next, assume that over time, population growth, changes in technology, or shocks to the environment result in new types of interactions, such as the need to trade an additional good, to create a primitive court system, or to manage a common resource. Suppose as well that different institutional approaches can structure these interactions, defining a game form. We define the addition of a new institution---a new game---as the start of an era.

We propose through an example the possibility of institutional path dependence. Suppose

¹⁰ The concept of path dependence is widely misunderstood (Pierson 2000). A process is state dependent if the future path depends upon the current state, in our case the set of institutional choices. One way that a process can be path dependent is if the order that previous choices have been made matters for future decisions. In the example that we describe, we only show evidence of set dependence – the past set of choices matters – but in light of the previous section on behavioral stickiness, path dependence seems plausible, so we use that term.

that there are five eras and that at the start of each era, the community chooses a new institution. Suppose further that the institutional choices are limited; in fact, each choice is from a set of only two. The implicit assumption here is that the strategic environment is such that the games that can be created take one of only two forms. Suppose that institutional choices appear in the following ordering:

Era 1: PD or TR

Era 2: PD or SI

Era 3: TR or AL

Era 4: TR or AL

Era 5: AL or SI

We assume that agents make optimal choices over institutions in each era but that they do not take into account future choices over institutions. If we were assuming fully rational behavior, we would have to make explicit assumptions about distributions of game pairs to substantiate this assumption, but we are assuming computationally constrained agents who do not take future eras into account when making current decisions.

Given these choices there would appear to be 32 possible institutional paths, but some of these would never occur if agents choose institutions that maximize payoffs at the start of each era. Suppose that a community chose a PD in Era 1. In Era 2, they might choose either PD or SI and still play both games optimally. If another community chose TR in Era 1, it must choose SI in Era 2. Using similar logic, it can be shown that if agents are restricted to using two-state automata and at the start of each era choose the game that gives the higher total payoff, then that there are only five possible institutional paths, as follows:

ERA 1

ERA 2

ERA 3

ERA 4

ERA 5

PD	PD	AL	AL	AL or SI
PD	SI	AL	AL	AL or SI
TR	SI	TR	TR	SI

This example demonstrates multiple equilibria in the “meta game” of choosing which games to play. And given that distinct sets of games can breed unique cultures, this example shows how institutional path dependence and cultural diversity go hand in hand. If we interpret the asymmetry of the Top-Right game as favoring the males in a community, then we see how that community later chooses institutions that also favor men because such institutions require less cognitive capacity in their operation.

There are many examples of this sort of path dependence in the corporate culture literature (Cohen and Sastry 2000). Organizational theorists refer to the importance of early stages in corporate culture formation as imprinting. Empirical evidence of imprinting has been found in craft unions, department stores, banks, newspapers, and high-tech firms (Stinchcombe 1965, Swaminathan 1996, and Boeker, 1989). Cohen and Bacdayan (1994) have found evidence of cultural imprinting, where early learning influence later strategies. For example, if subjects are sufficiently imprinted with a right-turn heuristic, they may take “left” turns by executing three successive right turns. While this is inefficient, it is simpler given the cognitive tools at their disposal. Imprinting may occur at the societal level as well. For example, Ebay’s method of auctioning off goods has become prevalent because so many people have evolved strategies for playing in that game. It is even used in settings where other types of auctions could (at least in theory) work better because people are less comfortable with the other auctions.

4 CONCLUSION

In this paper, we have shown in a game theoretic framework that when agents have cognitive constraints and face an ensemble of games, the ensemble's composition may influence behavior in individual games in ways that suggest the emergence of culture. In other words, if people living in different climates or followers of different religions or citizens of various political institutions daily face different mixes of strategic situations, they may be led to evolve different practices in common games, hence distinct behavioral cultures. We have then applied this model to demonstrate behavioral rigidity, frequency dependence, and institutional path dependence.

Our model also suggests a role for institutions in helping to change culture. To the extent that culture is a soup-pot descriptor of behavior---which we argue it is---institutional tinkering can alter behavior, although it may take some time. The introduction of market mechanisms may not immediately create a culture that will make those mechanisms successful. More broadly, we would argue that culture, a set of behaviors, is not primordial, but adapts to fit the ensemble of problems which itself is only partially institutionally derived, a conclusion that agrees with the findings of Nisbett et al., discussed above.

As presented, our model is incomplete in that it considers only behavioral repertoires that apply to strategic situations. We have omitted those aspects of culture that relate to art, religion, physical mannerisms, belief systems, and norms. In this discussion we provide a framework for extending the model to include beliefs and norms. We also suggest how the model's claims may be tested empirically and experimentally.

Some authors conceive of norms as reactions to single problems. Ensminger and Knight (1997) define a norm as the strategic equilibrium employed to overcome collective action problems. We widen the perspective: we expect norms to develop contextually (Picker 1997). A

society's solution to a collective action problem should depend on the ensemble of problems that the society faces. Norms that may be cognitively difficult for some societies may be easier for others. Alternatively, Bendor and Swistak (1997) define norms as strategies that are collectively enforced. Although our simple model did not allow for collective enforcement, there is no reason that our model could not be emended to include it. And, further, we see no reason why we would expect norms, in the Bendor and Swistak sense, to violate context dependence. If anything, the increased cognitive costs of collective enforcement mechanisms would seem to argue for similarity in their construction across social situations.

It is also possible to adapt our model to analyze the evolution of belief systems within and across societies. This can be accomplished without abandoning a Bayesian foundation. Agents can evolve encodings or representations of the world that are coarse partitionings of the "real" world. Constraints on the fineness of the partition will play a role akin to our restriction of the number of states in the automata. We expect that a community of agents would develop representations of the world, and therefore belief systems, that are context dependent and different from those of other communities.

We expect that our framework of agents evolving constrained optimal behaviors---whether they are strategies, norms, decision rules, or representations---can be applied in both the laboratory and the field. Laboratory experiments with human agents playing ensembles of games as opposed to individual games could help us understand how in the short term humans develop cognitive links between environments. By linking these experiments to more advanced computational agent-based models using more descriptively accurate models of human behavior such as neural nets or classifier systems, we might more accurately identify those game forms

(institutions) that create cognitive complementarities. We might also better understand which institutions create more rigid behaviors.

The application of our model to historical cases will be difficult but the potential payoff to researchers is large. If the strategic environment of people can be mapped into some common game forms and if the frequencies of the various games can be crudely approximated, then there is the possibility of explaining why some cultures are more altruistic or more vicious using game theory as opposed to appealing to Putnam's "mist". Scholars might also link experimental work on behavior rigidity and ensemble composition with empirical investigations of failed and successful transitions to democracy and markets.

Cultural similarity can then be compared formally by analyzing the number of like strategies players in different ensembles adopt. Axelrod (1997) has investigated how cultures may evolve over time if we begin with a set of diverse cultures. Our model provides an underpinning for his investigation by explaining the initial differences and by enabling a metric on behaviors.

Finally, the specific model presented in this paper can be modified in many directions. The possibility of noise could be included. Alternative methods of modeling boundedly rational agents could be considered (see Rubinstein 1998). Other definitions of cognitive costs could be employed, such as counting the number of games that are played differently. And, most interestingly, different games ensembles, especially those that are drawn from empirical research, could be considered.

APPENDIX: Proof of Claims

Recall that players use finite state automata to pick strategies and play each game in their ensemble each period. In the mathematical results that we describe, we assume that there are at least twenty periods and that there is no discounting. Given an ensemble of games, an *equilibrium in k-state automata* exists when each agent chooses a utility maximizing k-state automata given the automata choice of the other player.

Proof of Claim 1: We will describe the equilibrium strategies and where necessary explain why neither type of player could choose a two-state automata that would yield a strictly higher payoff. The outcome (C, S) will denote the row player choosing C and the column player choosing S.

PD: Row and column players choose GRIM. If either player deviates, given that they must use two state automata, the players would have to deviate in either period one or period two. The result follows immediately.

ALT: Row players choose TFT. Column players choose STFT. To get an average payoff above four the column player needs to get the (C, S) outcome more than once. After (C, S), the row player plays S. The only way the column player can get the (C, S) outcome is to play C. If the column player waits one period to play C, then the column player's average payoff equals $(10+2-2)/3 = 10/3 < 4$.

SI: Row and Column players choose ALL S

KE: Row and column players choose GRIM. Proof identical to PD. Alternatively, row players choose TFT, and column players choose STFT and the proof is then similar to ALT

proof.

TR: Row players choose C, Column players choose S. This is the unique Nash to the one shot game.

BL: Row players choose S, Column players choose C. This is the unique Nash to the one shot game.

Proof of Lemma 1: In all cases except the PD/AL game, the proofs will just describe the equilibrium strategies because these strategies are identical to the utilitarian social welfare maximizing strategies for the games being played individually.

PD/AL: Row players choose TFT for both games. Column players choose TFT for PD and STFT for AL. TFT and STFT was an equilibrium for AL played alone. Any two state automata that plays C in every period is a best response to TFT.

PD/KE: Row and column players choose GRIM for both games.

PD/SI: Row and column players choose GRIM for the PD and ALL S for SI. ALL S can be created from GRIM by starting in the S state. These were both equilibrium strategies for the individual games.

AL/KE: Row players choose TFT for both games. Column players choose STFT for both games.

AL/SI: Row players choose TFT for AL and STFT for SI. Column players choose STFT for both games.

KE/SI: Same as the PD/SI proof. Row and column players choose GRIM for KE and ALL S for SI.

SI/TR: Row and column players each have two one-state automatas. The row players use an All S automaton for the SI game and an All C automaton for the TR game. The column

players use an All S automaton for the SI game and an All S automaton for the TR game.

SI/BL: Row and column players each have two one-state automatons. The row players use an All S automaton for the SI game and an All S automaton for the BL game. The column players use an All S automaton for the SI game and an All C automaton for the BL game.

TR/BL: Row and column players each have two one state automatons. The row players use an All C automaton for the TR game and an All S automaton for the BL game. The column players use an All S automaton for the TR game and an All C automaton for the BL game.

For the second part of the lemma, we prove the results for the PD/TR and the AL/TR ensembles. The PD/BL and AL/BL cases follow by switching the roles of the row and column players. The KE/TR and KE/BL ensembles follow from the previous cases.

PD/TR: In order to play (C, C) in the PD, the row agents must each have one state that prescribes action C. If the transition mappings from that state remain in the same state if the column agent chooses S, then this strategy can be exploited by a column agent who plays All S. Therefore, the transition mapping causes them to move to a state that plays S. However, this means that there is not a state in which the row agent plays C and continues to play C even though the column player chooses S. This is required for an equilibrium of (C, S) in TR.

AL/TR: In order to alternate playing (C, S) and (S, C) in AL, the row agents must (1) each have one state that prescribes C and one state that prescribes S and (2) if the column agent plays S when the row agent plays C, then the row agent must play S in the next period. Thus, the agent cannot play (C, S) in every period in the game TR.

Proof of Lemma 2: The proofs will just describe the equilibrium strategies because these strategies are identical to the utilitarian social welfare maximizing strategies for the games being played individually or in pairs. The second half of the lemma follows from lemma 1.

PD/AL/KE: Row agents play TFT in all three games. Column agents play TFT in PD and STFT in AL and KE.

PD/AL/SI: Row agents play TFT in PD and AL, and STFT in SI. Column agents play TFT in PD and STFT in AL and SI.

PD/KE/SI: Row agents play TFT in PD and KE, and STFT in SI. Column agents play TFT in PD and STFT in KE and SI.

AL/KE/SI: Row agents play TFT in AL and KE, and STFT in SI. Column agents play STFT in all three games.

TR/BL/SI: Row agents play ALL C in TR and ALL S in BL and SI. Column agents play ALL C in BL and ALL S in BL and SI.

Proof of Lemma 3: The proof covers only the game AL played with TR and BL. The proofs for the other ensembles are similar. In order to alternate playing (C, S) and (S, C) in AL, (1) the row agents must each have one state that prescribes C and one state that prescribes S and (2) if the column agent plays S when the row agent plays C, then the row agent must play S in the next period. Thus, if the agents play (C, S) in every period in the game TR, the row agent must devote his third state to that strategy, leaving no state for the row agent to use to play BL.

References

- Almond, Gabriel and Sidney Verba. 1963. *The Civic Culture: Political Attitudes and Democracy in Five Nations*. Princeton, NJ: Princeton University Press.
- Axelrod, Robert M. 1997. *The Complexity of Cooperation: Agent-based Models of Competition and Collaboration*. Princeton, NJ: Princeton University Press.
- Banks, Jeffrey and Ragu Sundaram. 1990. "Repeated Games, Finite Automata, and Complexity." *Games and Economic Behavior* 2:97-117.
- Bendor, Jonathon and Piotr Swistak. 1997. "Evolutionary Stability of Cooperation." *American Political Science Review* 91(2): 290-307.
- Boeker, W. 1989. "Strategic Change: The effects of founding and history." *Academy of Management Journal* 32(3):484-515.
- Bowles, Samuel and Herbert Gintis. 1986. *Democracy and Capitalism*. New York: Basic Books.
- Calvert, Randall and James Johnson. 1997. "Interpretation and Coordination in Constitutional Politics." University of Rochester manuscript.
- Chong, Dennis. 2000. *Rational Lives: Norms and Values in Politics and Society*. Chicago: University of Chicago Press.
- Cohen, Michael D., Rick Riolo and Robert Axelrod. 2001. "The Role of Social Structure in the Maintenance of a Cooperative Regime." *Rationality and Society* 13(1):5-32.
- Cohen, Michael D. and Paul Bacdayan. 1994. "Organizational Routines are Stored as Procedural Memory: Evidence from a Laboratory Study." *Organization Science*. December 1994.
- Conlisk, John. 1996. "Why Bounded Rationality?" *Journal of Economic Literature* 34(2):669-700.
- Ellison, Glenn. 1993. "Learning, Local Interaction, and Coordination." *Econometrica* 61(5):1047-

1072.

- Elster, Jon, Claus Offe, and Ulrich K. Preuss. 1998. *Institutional Design in Post-communist Societies : Rebuilding the Ship at Sea*. New York: Cambridge University Press.
- Ensminger, Jean and Jack Knight. 1997. "Changing Social Norms." *Current Anthropology* 38(1):1-24.
- Ferejohn, John. 1991. "Rationality and Interpretation: Parliamentary Elections in Early Stuart England," in K.R. Monroe, ed., *The Economic Approach to Politics: A Critical Reassessment of the Theory of Rational Action*. New York: HarperCollins.
- Fudenberg, Drew, and David K. Levine. 1999. *The Theory of Learning in Games*. Cambridge: MA: MIT Press.
- Gigerenzer, Gerd and R. Selten, eds. 2001. *Bounded Rationality: The Adaptive Toolbox*. Cambridge, MA: MIT Press.
- Gilboa, Itzhak, and David Schmeidler. 1995. "Case-Based Decision Theory." *Quarterly Journal of Economics* 110(3):605-39.
- Greif, Avner. 1994. "Cultural Beliefs and the Organization of Society: A Historical and Theoretical Reflection on Collectivist and Individualist Societies." *Journal of Political Economy* 102(5):912-950.
- Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath. 2001. "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies." *American Economic Review* 91(2):73-78.
- Hong, Lu and Scott E. Page. 2001. "Problem Solving by Heterogeneous Firms." *Journal of Economic Theory* 97:123-163.
- Huntington, Samuel P. 1981. *American Politics: the Promise of Disharmony*. Cambridge, MA:

- Belknap Press.
- Inglehart, Ronald. 1977. *The Silent Revolution: Changing Values and Political Styles*. Princeton, NJ: Princeton University Press.
- Inglehart, Ronald. 1990. *Culture Shift in Advanced Industrial Society*. Princeton, NJ: Princeton University Press.
- Inglehart, Ronald. 1997. *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*. Princeton, NJ: Princeton University Press.
- Jennings, M. Kent and Richard G. Niemi. 1981. *Generations and Politics*. Princeton, NJ: Princeton University Press.
- Jennings, M. Kent and Laura Stoker. n.d. "The Persistence of the Past: The Class of 1965 Turns Fifty." manuscript.
- Jennings, M. Kent, Stoker, Laura, and Jake Bowers. n.d. "Politics Across Generations: Family Transmission Reexamined." manuscript.
- Kalai, Ehud and Ehud Lehrer. 1993. "Rational Learning Leads to Nash Equilibrium." *Econometrica* 61(5):1019-1045.
- Kalai, Ehud and William Stanford. 1988. "Finite Rationality and Interpersonal Complexity in Repeated Games." *Econometrica* 56(2):397-410.
- Linton, Ralph. 1936. *The Study of Man*. New York: Appleton-Century.
- Linz, Juan J. and Arturo Valenzuela. 1994. *The Failure of Presidential Democracy*. Baltimore: The Johns Hopkins University Press.
- Litwack, John M. 1991. "Legality and Market Reform in Soviet-Type Economies." *Journal of Economic Perspectives* 5(4):77-89.
- Miller, John H. 1996. "The Coevolution of Automata in the Repeated Prisoner's Dilemma."

- Journal of Economic Behavior and Organization* 29(1): 87-112.
- Nisbett, Richard E., Kaiping Peng, Incheol Choi, and Ara Norenzayan. 2001. "Culture and Systems of Thought: Holistic Versus Analytic Cognition." *Psychological Review* 108(2): 291--310.
- North, Douglass C. 1990. *Institutions, Institutional Change, and Economic Performance*. New York: Cambridge University Press.
- Ordeshook, Peter C. and Olga Shvetsova. 1995. "If Madison and Hamilton Were Merely Lucky, What Hope is There for Russian Federalism?" *Constitutional Political Economy* 6(2):107-126.
- Ouchi, William G. and Alan L. Wilkins. 1985. "Organizational Culture." *Annual Review of Sociology* 11:457-483.
- Picker, Randal C. 1997. "Simple Rules in a Complex World: A Generative Approach to the Adoption of Norms." *University of Chicago Law Review* 64:1225-1288.
- Pierson, Paul. 2000. "Increasing Returns, Path Dependence, and the Study of Politics" *American Political Science Review*, 251-268.
- Putnam, Robert D. 1973. *The Beliefs of Politicians: Ideology, Conflict, and Democracy in Britain and Italy*. New Haven: Yale University Press
- Quartz, Steven R. And Terrence J Sejnowski, (2002) *Liars, Lovers, and Heroes: What The New Brain Science Reveals About How We Become Who We Are*, W.B. Morrow
- Rabin, Matthew. 1998. "Psychology and Economics." *Journal of Economic Literature* Vol. XXXVI, pp. 11-46.
- Radner, Roy. 1993. "The Organization of Decentralized Information Processing." *Econometrica* 61(5):1109-1146.

- Roth, Alvin, Vesna Prasnikar, Jasahiro Okuno-Fujiwara, and Shmeul Zamir. 1991. "Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: An Experimental Study." *American Economic Review* 81(5):1068-1095.
- Rubinstein, Ariel. 1986. "Finite Automata Play a Repeated Prisoner's Dilemma." *Journal of Economic Theory* 39(1):83-96.
- Rubinstein, Ariel. 1998. *Modeling Bounded Rationality*. New York: Zeuthen Lecture Books.
- Samuelson, Larry. 2001. "Analogies, Adaptations, and Anomalies." *Journal of Economic Theory* 97:320-367.
- Sastry, Anjali and Corrine Coen. 2001. "Beyond the Beginning: Building a Theory of Organizational Imprinting." Working paper, University of Michigan.
- Schaffer, Frederic C. 1998. *Democracy in Translation: Understanding Politics in an Unfamiliar Culture*. Ithaca, NY: Cornell University Press.
- Stinchcombe, A.L. 1965. "Social Structure and Organizations," in James G. March, ed. *Handbook of Organizations*. Chicago: Rand McNally.
- Swaminathan, A. 1996. "Environmental Conditions at Founding and Organizational Mortality: A Trial by Fire Model." *Academy of Management Journal* 39(5) 1350-1385.
- Swidler, Ann. 1986. "Culture in Action: Symbols and Strategies." *American Sociological Review* 51(2):273-286.
- Taylor, Michael. 1987. *The Possibility of Cooperation*. New York: Cambridge University Press.
- Thaler, Richard. 1994. *Quasi Rational Economics*. New York: Russell Sage Foundation.
- Tsebelis, George. 1990. *Nested Games: Rational Choice in Comparative Politics*. Berkeley: University of California Press.

Table 1: Description of Games

Prisoner's Dilemma Game (PD)

		Column	
		C	S
Row	C	(4, 4)	(-2, 6)
	S	(6, -2)	(2, 2)

Alternation Game (AL)

		Column	
		C	S
Row	C	(2, 2)	(-2, 10)
	S	(10, -2)	(2, 2)

Knife Edge Game (KE)

		Column	
		C	S
Row	C	(4, 4)	(-2, 10)
	S	(10, -2)	(2, 2)

Self-Interest Game (SI)

		Column	
		C	S
Row	C	(2, 2)	(0, 8)
	S	(8, 0)	(4, 4)

Top Right Game (TR)

		Column	
		C	S
Row	C	(2, 4)	(2, 6)
	S	(4, 0)	(0, 2)

Bottom Left Game (BL)

		Column	
		C	S
Row	C	(4, 2)	(4, 0)
	S	(6, 2)	(2, 0)

Table 2: Description of Strategies

Name of Strategy		Initial Action	Continued Play
Tit For Tat	TFT	C	Copy other's previous action
Grim Trigger	GRIM	C	C until other plays S, then S forever
All C	All C	C	Always play C
All S	All S	S	Always play S
Switch After C	SAC	C	After C, play S until other plays C
Alternate	ALT	C	Alternate between S and C
Do The Opposite	DTO	C	Do the opposite of what other plays last period

Table 3: Evolution of Strategies in One Game Ensembles		
Game	Outcome	Strategies
Prisoner's Dilemma (PD)	90% runs: All agents play C 10%: All agents play S	61% GRIM. 39% TFT for all C outcomes.
Alternation (AL)	45%: greater than 90% alternate 50%: between 70%-90% alternate	40% SAC, 35% TFT, and 25% ALT
Self Interest (SI)	All agents play S	90% All S 10% STFT
Knife Edge (KE)	40%: greater than 90% alternate 50%: between 70%-90% alternate 5%: All agents play C	68% SAC 25% TFT 7% ALT
Top Right (TR)	Top Right	Column: 100% All S, Row: 75% All C, 25% DTO
Bottom Left (BL)	Bottom Left	Row: 100% All S, Column: 70% All C, 30% DTO

Table 4: Simulation Results with Two -Game Ensembles

Ensemble Pair	Comment
PD / AL	*95% over 70% USWF; 100% TFT in PD when USWF optimized
PD / KE	**100% over 90% USWF; 35% CC in both games
PD / SI	*100% UWSF in SI, 75% USWF in PD; GRIM in every all C PD
PD / TR	† 100% (S, S) in PD; 60% of row players could reach two states.
PD / BL	† 100% (S, S) in PD; 50% of column players could reach two states
AL / KE	95% > 65% USWF, no (C, C) in KE; 70% SAC and 25%TFT
AL / SI	*100% USWF in both games; 65% TFT in AL, 35% SAC
AL / TR	†100% SS in AL; 65% of row players could reach two states
AL / BL	† 100% SS in AL; 45% of column players could reach two states
KE / SI	100% USWF in both games; 85% use an alternating strategy in KE, of this 70% SAC in KE
KE / TR	† 100% SS in KE; 65% of row players could reach two states
KE / BL	† 100% SS in KE; 60% of column players could reach two states
SI / TR	100% USWF in both games; 45% of row players had two automata
SI / BL	100% USWF in both games; 50% of column players had two automata
TR / BL	100% USWF in both games; 55% of players had two automata

Notes: * An asterisk denotes that the strategies employed in the ensemble differ from the strategies that were evolved when the games were played separately. ** Two asterisks denote that different equilibria are evolved. † A cross denotes that the strategy or strategies that evolved are not included in the strategies that evolved when the game was played separately. Strategies in **bold** denote those that maximize a utilitarian social welfare function most of the time. The comment “95% over 70% USWF” means that in 95% of the runs, at least 70% of the agents were obtaining the utilitarian social welfare maximizing payoff.