Can Game(s) Theory Explain Culture? The Emergence of Cultural Behavior within Multiple Games

Jenna Bednar
Scott Page

SFI WORKING PAPER: 2004-12-039

SANTA FE INSTITUTE

CAN GAME(S) THEORY EXPLAIN CULTURE?[1]

THE EMERGENCE OF CULTURAL BEHAVIOR WITHIN MULTIPLE GAMES

Jenna Bednar
and
Scott Page

University of Michigan

August 2000
updated 20dec04

The hallmarks of "cultural behavior" are consistency within and across individuals, variance between populations, behavioral stickiness, and suboptimal performance. In this paper, we build a formal framework within which we can *derive* each of these five behavioral attributes. Our framework rests on two primary assumptions: (i) agents play ensembles of games, not just single games as is traditionally the case in evolutionary game theory models and (ii) agents have finite cognitive capacity. Our analysis combines agent-based techniques and mathematics. The former enable us to explore dynamics and the latter allow us to prove when the behaviors produced by the agents are equilibria. Our results provide game theoretic foundations for cultural diversity and agent-based support for how cultural behavior might emerge.

---

[1] jbednar@umich.edu and spage@umich.edu. We thank the Summer Workshop in Computational Economics at the Santa Fe Institute, where Scott presented early versions of this work. We also presented this paper 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 Rob Axtell, Jake Bowers, Sam Bowles, Ken Kollman, Mika Lavaque-Manty, Paul Mahoney, Dwight Read, Ken Shotts, Cara Wong, and especially Bob Axelrod, as well as seminar audiences at Columbia University, Washington University, the Santa Fe Institute, and Harvard University. Scott is grateful for funding and support from the MacArthur Foundation and the McDonnell Foundation.

Culture can be defined as individual and community level behavioral patterns that depend upon context and are often suboptimal. Cultural behavior influences the performance of institutions. Whether looking horizontally across countries or vertically through time, one lesson is clear: the efficacy of markets, democracies, and law hinges upon behavior, particularly on the tendency for people to cooperate with and trust one another. A theory of institutions, therefore, must come to grips with culture. Game theory, the preferred formal framework for analyzing institutions, assumes isolated, context free, strategic environments and optimal behavior within them. Thus, game theory would seem to be at a loss to explain the patterned, contextual, and sometimes suboptimal behavior we think of as culture, let alone its emergence and persistence across space and time.

Yet, surprisingly, game theory is up to the task. In this paper, we provide game-theoretic foundations for cultural behavior and an agent-based implementation that demonstrates its emergence. We construct a framework with purposeful agents whose rational impulses lead them to acquire diverse cultural behaviors. Our results are surprising in light of the fact that 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 the traditional game theoretic assumption of optimizing behavior. When purposeful, incentive-sensitive agents confront multiple strategic situations rather than just one, and when cognitive effort is costly, we find that culturally distinct behavior is likely and in many cases unavoidable.

Our modeling framework expands upon the traditional single game approach: agents play multiple games simultaneously and evolve or choose separate strategies in each. The whole differs from the sum of the parts because agents can apply similar strategies to distinct games. We refer to our framework as *games theory.*[2]

In our framework, cognitive thought is costly. The more sophisticated an agent's strategies in these games the more cognitive resources those strategies require. Therefore, within any individual game, a rational agent will balance a strategy's performance and its cognitive costs. As a result, an agent's strategy in one game will be dependent upon the full ensemble of games it faces. Two agents facing different ensembles of games may choose distinct strategies on games that are common to both ensembles. Furthermore, the potential for agents to exploit common subroutines across strategies for different games leads to the emergence of intra- and inter-agent behavioral consistency. In other words, rational agents choose (for rational reasons) to act culturally.

Our approach is a "tool kit" model of culture: agents develop a variety of strategies, each tailored for specific situations.[3] Community-level behavioral patterns---cultures---emerge collectively as agents evolve behaviors in strategic environments. Diverse cultures emerge not in

---

[2] The games theory idea was informally present in Long (1985), who called it an "Ecology of Games" and in Bowles and Gintis (1986), who discuss the importance of the full spectrum of games---e.g. games played within the family, the state, and the economy---to understanding individual behavior.

[3] 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).

spite of optimizing motivations, but because of how those motivations are affected by incentives, cognitive constraints, and institutional precedents.[4] Thus, agents in different environments may play the same game differently.

We first examine the empirical evidence of behavior commonly attributed to culture in order to define more precise characteristics of behavior and review the existing formal literature's ability to explain these characteristics. In Section 2, we define a framework for modeling behavior in a multiple game environment. To provide intuition for our main results, in Section 3 we include an example. In the example, the agents in two communities of agents play different ensembles of games. While both communities play the Prisoner's Dilemma, behavior in that game depends upon the other game that the communities play. Agents in the first community play a different strategy than the agents in the second community. In Section 4, we state results from both mathematical and computational analyses of agents playing games ensembles drawn from a class of cooperation games. Our findings echo the characteristics of culture behavior that we describe in Section 1. We conclude with a discussion of how our framework may be able to explain some of the puzzling empirical findings.

1 EVIDENCE AND DESIDERATA

That culture exists and matters is indisputable. Across disciplines, scholars rely on culture as the basis for regional exceptionalism, and there are numerous empirical studies demonstrating culture's impact on the choices made by individuals and communities. Cultural differences correlate with diversity in important activities, including values and goals (Inglehart 1977, 1990, 1997) and political participation (Almond & Verba 1963; Inglehart 1997).

In studies of primitive and advanced societies, experimental social scientists attribute behavioral variety to cultural differences. 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. Experimental results from fifteen small-scale societies provide further evidence (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. A related literature on corporate culture demonstrates behavioral similarities of workers within firms (Ouchi and Wilkins 1985).

Evidence from psychology and cognitive science also points to the relevance of culture to

[4] 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 a benchmark, rather than descriptively accurate, there is support for looking for richer alternatives or amendments. The three most promising approaches all build upon 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 Leher 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.

behavior. Social psychologist Richard Nisbett and colleagues (2001) have found empirical evidence that thought processes differ between people living in Asian and western communities. In experiments, as well as reviews of the historical and social records, Nisbett and colleagues find strong support that Asians tend to have more holistic interpretations of the world, while westerners tend to be more atomistic in their thinking. They hypothesize that the differences partly arise due to different social systems. Perhaps most compelling is the recent neurological research suggesting that human cultures influence the formation of neural architecture. Our genes may encode scaffolding for the neuronal connections that form but the particulars will be partially determined by our experiences. In other words, our cognitive and social environments play a role in how our brains develop (Quartz and Sejnowski 2002).

These studies show that within a culture individuals behave and think similarly but differ across cultures. A second body of research demonstrates their profound effect on institutional performance (Putnam 1993, Chong 2000, Elster et al 1998, Huntington 1981, Putnam 1973, Schaffer 1998, Greif 1994, Litwack 1991, North 1990, Hay & Shleifer 1998). These effects cannot be dismissed as temporary given that cultural behavior proves remarkably persistent (Jennings & Niemi 1981; Jennings & Stoker unpub; Jennings, Stoker & Bowers unpub) and tends to change slowly, through generational replacement (Inglehart 1997 and Nisbett et al 2001). Behavioral patterns that emerged from the "mists of the Dark Ages" (Putnam 1993:180) doom or buoy democracy today. Relatedly, post-communist transitions to markets and democracy have struggled because those societies lacked appropriate norms and could not quickly "grow" them (Litwack 1991, Ordeshook & Shvetsova 1996, Pitt & Khandker 1998).

The empirical evidence suggests a variety of specific behavioral patterns linked to culture. First, if cultural behavior exists, then it should generate an identifiable set of behavioral traits within a single individual. Secondly, the data generated through survey research, as well as the results from experiments in behavioral economics, are evidence that individuals like others within their community, but may behave differently from those in other communities. Finally, the studies in the institutional performance literature suggest the existence of suboptimal behavior, and a lag in response to incentive adjustment. These elements can be summarized in a five-item "wish list."

(1) Intra-individual consistency: As an individual moves from task to task it responds similarly.
(2) Inter-agent consistency: individuals within the same community, encountering the same problems, will act like one another.
(3) Contextual effects: Individuals from different communities may react differently to the same problem or phenomenon.
(4) Suboptimal behavior: The strategy employed by individuals within a community may be suboptimal, where individuals could benefit more by acting in a different way. Formally, the behaviors are not equilibrium strategies in the repeated game or if they are equilibrium strategies, the resulting equilibrium does not belong to the set of Pareto efficient equilibria.
(5) Behavioral stickiness: individuals may not immediately alter their behavior despite changes to their incentives.

Our goal is to explain the existence, and better yet, emergence, of these characteristics. To date, the formal literature has pursued a different objective when incorporating culture into its models, and therefore has not been able to explain how these five behavioral characteristics of culture might arise. Most of the formal literature that includes culture does so in order to type agents, to demonstrate the effect of different utility functions or belief systems. Relatedly, culture can be appealed to as an external factor that refines equilibria by making some more focal than others. Two excellent examples of this work are Greif 1994 and Fearon & Laitin 1996. The approach helps us to understand the *effects* of culture, but it cannot explain its *emergence*. In the context of the above wish list, these models explain what happens if traits #2-5 are true, rather than how these cultural aspects would appear.

Another approach takes primordial behaviors as given (Axelrod 1997). In these models, cultures are vectors of agent attributes. Shared culture---common attributes---emerges because agents become more like those they interact with, and they are more likely to interact with those sharing similar attributes. As an example, Axelrod's model generates local homogeneity but---interestingly---global polarization of attributes: despite a tendency to become more alike, distinct "cultures" emerge and persist, and where inter-cultural differences exist, they are by-products of the intra-community dynamic. Axelrod's does not model strategic behavior, but does explain the evolution of similarity between individuals, approximating #2 on the list above, as well as differences between communities, or #3.

None of the aforementioned models capture behavioral stickiness (#5). In the long run, incentives may well rule the day. Individuals, and perforce, entire societies may grope their way toward efficient behaviors. In this case, the role of culture would be limited to influencing the choice from among the efficient equilibria. In the short run, and in the medium run, behavior may be conditioned on history as much as it is by incentives (Pierson 2000, North 1990).

Finally, there is a large literature on the evolution of reciprocity, altruism, and reciprocity (see Henrich forthcoming for a partial survey). These models rely almost exclusively on a repeated game framework with both individual and group level selection. The modelers often distinguish between cultural transmission and genetic transmission. The former implies greater forces for within group conformity which undoubtedly contribute to the incredible diversity across human cultures, and addresses #2 on our list. Along similar lines, Bendor and Swistak (1997, 2001) define norms as strategies that are collectively enforced and also assume conforming pressures in their models of the evolution of norms.

As impressive as these studies are, they 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? Under what conditions does culture lead someone to act suboptimally?

In addition to suggesting specific behavioral characteristics to explain, the reality described by empirical research points toward the consideration of a broader behavioral context, one that includes multiple strategic contexts. This multiple-game approach is a departure from canonical game theory, which tends to study behavior within a single game. True, repeated game theory models consider multiple games, but these games are all the same and are played sequentially, not simultaneously. The empirical and historical literature on cultural behavior suggests a need for horizontal interdependence of games. Such an approach allows us to investigate how the social, political, and economic dimensions of people's lives affect the decisions people make in specific situations.

2 A GAMES THEORY FRAMEWORK

Our goal is to derive the specific behavioral characteristics listed above without compromising the logical rigor and emphasis on incentives assumed by game theory. We maintain the rationality assumption: our agents are self-interested, and although they have cognitive constraints, they possess sufficient capacity to play each game optimally. To capture the contextual factors suggested by the empirical literature, we design a framework of multiple games that agents play simultaneously, adopting a distinct strategy for each.5

2.1 The Games

The agents in our framework play games that belong 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 a player's own payoff and raises the payoff of the other and being selfish does the opposite, so for both players being selfish is the unique pure strategy equilibrium. In the final two games, these conditions hold only if both players cooperate. In each of these games, there exists a unique pure strategy equilibrium in which one player cooperates and one is selfish.

These six games were chosen with the restriction that one strategy be cooperative and one be selfish and so as to allow maximal behavioral diversity. Each of the four boxes in the two by two normal form representation is the efficient repeated game equilibrium for one of the games, as is alternation between boxes. Further, one of the games, Knife Edge, has two efficient repeated game equilibria, one in a single box and one that requires alternation.

insert Table 1 here

The first game is a standard Prisoner's Dilemma. In the second game, Alternation (AL), agents do best in repeated play by alternating between the off-diagonals. In the third game, which we call Knife Edge (KE), the players receives the same average payoff when both alternate between being selfish and cooperative as they do if each 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. For each of the final two games, Top Right (TR) and Bottom Left (BL), the upper right and lower left cells are both equilibrium outcomes. Each of these last two games is biased in favor of one of the two players. Asymmetric coordination games are useful 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

[5] Our games theory/game ensemble framework should not be confused with Tsebelis's (1990) work on "nested games" in which agents play a single strategy in multiple games with variable payoffs. Our methodology is related to Samuelson (2000) but he considers agents who play only two games who partition their attention. We allow sharing of cognitive subroutines.

We encode the strategies that the agents use to play these games using finite state automata.[6] Finite state automata consist of three parts: mental states *M*, rules for how to transition between these mental states *T*, and initial states *I*. The mental states are numbered and each one prescribes an action, such as "be selfish" or "be cooperative." The number of states in the agent's automata serves as a proxy for their cognitive capacity. 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."

Combinations of states and transitions can be thought of as *cognitive subroutines.* This construction allows us to describe the components of the cultural phenomena with respect to the mental states and cognitive subroutines of the agents. In effect, we can open up the agents' heads, view their cognitive architectures, and construct comparisons based on formal measures.

Two State Automata: A two state automaton has just two mental states, denoted by 0 and 1. These 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 a designation of the initial state followed by a description of each state and its transitions. A state can be described by a three-tuple, such as (C, 0, 1). This three tuple tells the player to "choose action 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)}. The agent begins in state 0 in which he cooperates. Thereafter, the agent stays in state zero unless his opponent acts selfishly. In which case, the agent moves to state 1, where the agent plays S.

In addition to TFT many familiar strategies can be encoded as two state automata. In Table 2, we describe those strategies that are either needed in the proof of claims or that emerged from our computational model.

Insert Table 2 here

Many strategies have a "suspicious" analog where the agent chooses S rather than C in the first period. We 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.[7]

---

[6] 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.

[7] 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

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.[8]

2.3 From Automaton to Agent

Agents play ensembles of games with other members of their community. The automata game playing literature has focused almost exclusively on agents that play a single game so there is no operational difference between an automata and an agent's strategy. But our multiple game analysis requires the automata differentiate between the games as well as define possibly distinct strategies for each. We implement this feature by adding a separate initial state for each game in an agent's ensemble. 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, and a single two state automaton can represent both Grim Trigger and All Selfish, if it assigns different initial states to each game.

Suppose that an ensemble includes two games: the Prisoner's Dilemma and the Self-Interest game. An automata representing an agent would include an initial state for each game as well as mental states and transition mappings. It is possible that an automata with only two states of capacity could play GRIM in the PD and ALL S in the SI game. This can be accomplished using the same mental states but different initial states for the two games. This particular two game automata can be written: {PD =0, SI = 1, (C,0,1), (S,1,1)}. State 1 is used as the starting point in SI. The strategy for the PD game begins in state 0 and plays C until the other agent plays S, at which point the strategy moves to state 1 and the agent plays S.

In our analysis, we differentiate between row agents and column agents and number each type from 1 to *N*. We assume that row agents play only with column agents and vice versa. Each row (column) agent plays five neighboring column (row) agents. Row agent 21 will play column agents 19, 20, 21, 22, and 23. To maintain symmetry, we place the agents on a circle, i.e. if there

---

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.

[8] 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, 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.

are 100 agents, then row agent 50 will play 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. People interact with friends and associations. They generally do not interact with the entire population with partners chosen randomly. 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 AN EXAMPLE OF EMERGENT CULTURE IN GAME ENSEMBLES

We present the first results of our model in the form of an example to allow us to provide a richer description of the phenomena that occur. In this example, we assume two communities of agents, each playing a two game ensemble. All agents within the same community share the same ensemble of games, and interact with their neighbors playing only those games.

In this example, both communities are assumed to lie adjacent to large bodies of water that supply the communities with two types of fish, coho and skimmers. Coho are large and require cooperation to land and skimmers are small, easily landed by an individual. In each community, people fish in two person boats containing one man and one woman. Each person has a choice of two actions: fish for coho, which we denote by C, or fish for skimmers, which we denote by S. If the two agents choose action C, they get each get a positive payoff. However, each has an incentive to deviate and fish for skimmers and then jump in and help if by chance the other person hooks a coho. But, if both fish for skimmers, they are worse off because no coho are caught. This, then, is a classic Prisoner's Dilemma.

In addition to fishing, members of the first community can also paddle their boats to a nearby island to harvest bananas and mangos. Gathering bananas (C) is hard work while harvesting mangos (S) is easy. Bananas, though, are more nutritious. If both agents gather bananas, they get moderate payoffs: the hard work of harvesting bananas is balanced off by the increase in nutrition. They get identical payoffs if both pick mangos: the easier work is balanced by the loss of nutritional value. However, if one of them picks bananas and the other picks mangos, the first gets a negative payoff and the second gets a high payoff, because the second shares in the bananas that the first picked without expending the effort. Specific numerical values for the payoffs for these two games are given in the table below.

Table 3: Games Ensemble For The First Community

Fishing Game (PD)

		Women C	**Women** S
Men	C	(4, 4)	(-2, 6)
	S	(6, -2)	(2, 2)

Harvesting Game (ALT)

		Women C	**Women** S
Men	C	(2, 2)	(-2, 10)
	S	(10, -2)	(2, 2)

The agents can distinguish between these two games. They know when they are fishing and when they are harvesting fruit. They can thus play distinct strategies in each. However, they have limited cognitive capacity. They must use two game, two state automata. This restricts the strategies that they can employ in the two games and could prevent them from receiving the optimal unconstrained payoffs.

In our computational model of this two game ensemble, we find that in this community the agents evolve a "do unto others" behavioral culture. In the fishing game, both men and women start out by playing C. Thereafter, the strategy "do unto others" becomes Tit for Tat. Using this strategy the agents support (C, C) in the Fishing Game. In the Harvesting Game, we see the following behavior. On the first day, the women harvest mangos (S) and the men gather bananas (C). On subsequent days, each invokes the "do unto others" rule. The result is an alternation between (C, S) and (S, C) in the Harvesting Game. The logic in the second game works like this: my partner was nice (or selfish) last period, so I'll be nice (or selfish) this period. This evolved behavior can also be shown to be optimal for two game, two state automata.

Members of the second community also play the fishing game and can go to a nearby island to gather food. The difference is that their food gathering game has a different payoff structure. They can pick cherries or berries (C or S). The berry patch is small; two people get in one another's way, with the clumsy men tending to knock over the women's buckets. Furthermore, the men can't climb the cherry trees alone, but are able to pick cherries if the women are around to help. The Picking Game is equivalent to the game we call BL in our class of games. The table below shows the payoff matrices for the Picking Game as well as the Fishing Game.

Table 4: Games Ensemble For The Second Community

Fishing Game (PD)

		Women C	**Women** S
Men	C	(4, 4)	(-2, 6)
	S	(6, -2)	(2, 2)

Picking Game (BL)

		Women C	**Women** S
Men	C	(4, 2)	(0, 4)
	S	(6, 2)	(2, 0)

Given their inability to pick cherries, the men have a dominant strategy in the Picking Game, namely S. As a result, the women in this community have an iteratively dominant strategy: C. Loosely speaking, we can interpret the behavior of the men in this game as selfish and the behavior of the women as deferential.

In the Fishing Game, we might expect the same results that we saw in the first community: that the men and women would learn to play Tit for Tat in order to sustain cooperation, but this does not happen. In fact, it *cannot* happen, as we show formally later in this paper. The strategies "deferring to the men" and "doing unto others" are quite distinct behaviors, and the agents do not possess sufficient cognitive capacity to do both. What we find when the Fishing Game is paired with the Picking Game is that the men and women both end up choosing S in the Fishing Game. This choice of S in the Fishing Game results from a cognitive shadow cast by the Picking Game.

This example highlights several phenomena emerge that correspond to the cultural behavior in Section 1's wish list. First, each of the two communities exhibits trait #1, *intra-individual consistency*. In the first community, agents use the same cognitive subroutine, "do unto others," in the two games to play different strategies even though they could have used other strategies and still located the efficient equilibrium. In the second, the men play selfishly in both games despite the fact that they could do better by cooperating in the Fishing Game.

Secondly, all of the agents in the first community played Tit for Tat in the Fishing Game and all agents played alternation in the Harvesting Game, exhibiting #2, *inter-agent consistency*. The combined effect of #1 & #2, intra- and inter-agent cultural behavior, creates a community of agents who act consistently and similarly. This intra-community similarity lies in the intersection of many definitions of culture, but it not complete. The consistency must be explained by the particulars of the environment.

A third cultural phenomenon is the persistence of behavioral differences between cultures. If this did not exist, suboptimality and consistency might be explained by genetics. In our example, the agents in the first community cooperated by fishing for coho in the Fishing Game, but in the second community, the women and men both fished for skimmers. The diversity of behavior in the game is an exhibition of #3, *contextual effects*. The ensemble of the Fishing Game and the Harvesting Game creates a different cognitive context than the ensemble consisting of the Fishing Game and the Picking Game.

The fact that different communities can and often do adopt different behaviors underlies much of the aforementioned game theoretic literature on culture as an equilibrium selection device. In those models, if there are multiple efficient equilibria---to greet people, we could bow, shake hands, or kiss---culture selects among them. Typically, game-theoretic models do not allow for culture to create *suboptimal behavior*, the fourth behavioral phenomena that we link to culture.

In the second community, the women and men fish for skimmers in the Fishing Game. Although to do so is an equilibrium behavior in the one shot game, it is not optimal in the repeated game, demonstrating a potential derivation of suboptimal behavior. Obviously, suboptimal behavior can occur for many other reasons as well. These include inertia, cognitive biases, and improper beliefs. Our emphasis here on culture effects induced by the binding cognitive constraint is meant as an additional source of such biases, not as a replacement.

4 RESULTS

We organize our results around the framework generated by the wish list of Section 1. We present our results in two sections. First, we mathematically derive conditions that support #3, *contextual effects*, and #4, *suboptimal behavior*. While these claims are instructive, they can only derive the best possible outcomes for our agents; they are not capable of demonstrating what is most likely to emerge from a dynamic model of interacting agents. With the agent-based model we can counter the obvious criticism of equilibrium models that they only prove the existence of equilibria, not that agents ever actually reach these equilibria. In section 4.2 we present the results of our agent based model demonstrating the emergence of the aspects of cultural behavior defined in the wish list.

4.1 Conditions that Support *Contextual Effects* and *Suboptimal Behavior*

We first establish 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 have the cognitive capacity to play any one game optimally. Within the context of standard game-theoretic analysis, which isolates attention to single games, the outcomes selected by these agents would not be compromised by the cognitive bounds.

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

Proof: All proofs are in the appendix.

Notice that the proof states that the equilibrium exists in the class of automata. In models in which automata represent agents, the equilibria are defined with respect to automata. Each agent chooses an automata which is optimal within the class of automata given the choice of automata by the other agents. When the agents play multiple games simultaneously, they no longer always play each game optimally. In the next three claims, we establish the conditions that lead to #4, *suboptimal behavior*.

Claim 2: Two State, Two Game Ensembles and Suboptimality *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/Bl, and TR/BL, but not for the other six two game ensembles.*

The combination of the Top, Right game or Bottom, Left with the Prisoner's Dilemma, Alternation, or Knife-Edge games creates too complex an environment for a two state automaton to play optimally. To maintain (S, C) as an equilibrium in Bottom, Left, 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 Prisoner's Dilemma, they would need two other states for a total of three. The next claim states a similar result for the three game ensembles.

Claim 3: Two State, Three Game Ensembles and Suboptimality *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 claim shows that even with three state automata, the cognitive constraint binds and affects behavior in certain contexts.

Claim 4: Three State, Three Game Ensembles and Suboptimality *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}.*

Claim 4 describes ensembles that force our agents into suboptimal play despite their increased cognitive capacity. Again, the cognitive complexity demanded from the optimal play of the Top, Right and Bottom, Left games drains the capacity of the cognitively-limited agent; here, a three-state agent cannot play many ensembles of three games optimally.

These last three claims also demonstrate the existence of #3, *contextual effects*: agent behavior in one game is affected by play in another. Play in any one game cannot be fully understood without considering its strategic context, the other games that the agent plays at the same time.

4.2 Agent Based Model Results

We now present an overview of the results that emerged from our agent-based model. Since many of these findings have been established mathematically, we refrain from a detailed statistically analysis of the agent based model. As a baseline, we first present the results of the single game ensembles in Table 5. This table presents the findings from one hundred random trials of the model. As is evident from the table, consistent with Claim 1, the agents prove capable of locating the utilitarian social welfare maximizing equilibria. Thus, our agent-based model demonstrates that a simple learning dynamic rule suffices for the communities of agents to evolve strategies that are optimal within their constraints.

insert Table 5 here

Since it is equally important how the agents achieve these payoffs, we include the strategies in the table. In the Prisoner's Dilemma, our agents evolve TFT and GRIM, and unsurprisingly, in the Self-Interest game (SI), where Selfish is the dominant strategy, agents evolve automata that play ALL S.

We now turn to the multiple-game simulations. In the Appendix is Table 6, which summarizes the results from computational experiments on each of the fifteen two game ensembles. We have divided these ensembles into three classes. In the first class are those ensembles in which behavior is largely unchanged from the single game context. The second class contains those ensembles in which behavior and in some cases the equilibrium chosen changes but the outcomes are still efficient. In the third class are those ensembles within which behavior is suboptimal for at least one game.

As we did in our example from the previous section, we emphasize those results that correspond to the five cultural behavioral phenomena we defined in section 1. In addition, we comment briefly on a sixth phenomenon that has been identified in the empirical, experimental, and theoretical literatures on culture, *frequency dependence*: how often a game is played influences how likely the agents evolve optimal strategies.

(1) Intra-individual consistency

We highlight three findings relevant to intra-individual consistency. First, when the Self-Interest game (SI) is paired with the Prisoner's Dilemma (PD) game, the agents cooperate in the PD using GRIM, while in the Self-Interest game they play ALL S. These strategies, GRIM and ALL S, are behaviorally consistent in their unwillingness to forgive. This combination evolves because the SI game has a Pareto efficient dominant strategy. 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. Moreover, it is easy to show that *no two state automaton agent of the type we create is capable of playing GRIM and TFT simultaneously.*

We also see intra-agent cultural behavior in the PD/AL ensemble, one type of agent (row or column) plays TFT in both games and the other type plays TFT in the Prisoner's Dilemma and STFT in the Alternation game. This was the ensemble we considered in the first community in our example from section 3. The agents evolve a "do unto others" cognitive subroutine. In each of these two ensembles, intra-agent consistency emerges because mental states and cognitive subroutines (states and transitions) are shared within an agent across games, invoking identical cognitive processes in distinct contexts.

Agents do not always evolve intra-agent consistency. Agents confronting an ensemble of TR and BL assign a unique state to each. No states are shared across the games and hence the agents share no cognitive subroutines. In this ensemble, the cognitive constraint does not impinge on behavior, and the optimal actions in the games differ for both of the agents so they must assign different states to each. This same logic applies to a lesser extent to the TR/SI and the BL/SI ensembles. In these ensembles, one type of agent exhibits no intra-agent consistency for the same reasons as in the TR/BL ensemble, but the other type can use the same mental state, namely S, for both games.

(2) Inter-Agent Consistency

Within the context of our model, when agents in a community use common mental states and cognitive subroutines across games, it is evidence of inter-agent consistency. Inter-agent cultural behavior is not evident in single game ensembles. In the Self Interest, Top Right and Bottom Left games, all agents tend to choose the same strategy because the strategies that generate the utilitarian social welfare maximizing are unique. In the other three games where it could occur, we see little inter-agent consistency. In the Prisoner's Dilemma both TFT and GRIM emerge in most runs. In the Knife Edge (KE) and Alternation Game (AL) agents evolve pairs of strategies that alternate between the upper right and lower left such as TFT for the row agents and Suspicious Switch After Cooperate (SSAC) for the column agents. Agents need not use the same subroutine as long as they alternate, so the particular strategies can be arbitrary.

Inter-agent cultural behavior becomes more likely with larger ensembles. Recall from our discussion of intra-agent cultural behavior that when the Prisoner's Dilemma is paired with the Self Interest game almost all of the agents play GRIM in the PD and All S in SI. Therefore, in this ensemble we also see inter-agent cultural behavior. In the three game ensembles, inter-agent cultural behavior was even more pronounced. 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 can play (C, C) in PD, they can alternate between (S, C) and (C, S) in AL, and they can play (S, S) in SI.

(3) Contextual Effects

Contextual effects refer to how a strategy in one game is affected by the other games in the ensemble. If contextual effects are present, we should see agents in different communities using different mental states and cognitive subroutines in the same game. This often occurs. Consider the two game ensembles that include the Knife Edge (KE) game. When played alone, the cooperative outcome (C, C) rarely emerged in KE. However, when paired with the PD, the cooperative outcome in the KE game was supported over 35% of the time and when Knife Edge was paired with the Alternation game, (C, C) never arose in the KE. 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 emerges from the learning dynamic, possibly because once the agents find the efficient equilibrium strategy for PD (or AL), then with a single mutation in the initial state, that same strategy will lead to an efficient equilibrium in the Knife Edge game.

The most compelling evidence for emergent contextual effects on strategies and outcomes occurred in the three-game ensembles. In these ensembles, we ran jus twenty runs. In the PD/KE/AL ensemble, the agents played KE and AL identically in each of 20 runs. Similarly, when KE and AL are paired with SI, in all but one run of twenty runs, 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 that in two game ensembles when PD is paired with SI, agents tend to play GRIM in the Prisoner's Dilemma. The fact that strategies in KE vary in the same way should not be surprising.

(4) Suboptimal Behavior

For the two game ensembles, we focus on the six pairings identified in Claim 2 where the agents cannot evolve utilitarian social welfare maximizing strategies. Cognitive subroutine sharing dominates and we see a bias toward simpler strategies. In fact, no sophisticated strategies evolve. Agents play (S, S) in the Prisoners' Dilemma when it belongs to an ensemble with either Top Right or Bottom Left. Furthermore, the Self-Interest game is unaffected in any of the ensembles. Even when we increase the number of other games, we still find that (S, S) is almost always played in SI.

In the three game ensembles in which the cognitive constraint binds (Claim 3), we found a larger 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 casts a behavioral shadow across other games stifling cooperation. The effect of being able to win big in one game appears to squelch any incentive to abandon that behavior for a norm of cooperation.

(5) Behavioral Stickiness and Frequency Dependence

We conclude our results section with a brief discussion of whether changing the frequency of a game's appearance can lead agents to evolve different strategies and on the implications of behavioral stickiness. To investigate the first hypothesis, we constructed several ten game ensembles. The first ensemble contained five PD games, one SI game, two AL games, one TR and one 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 (fewer 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 the ensemble structure can be modified, behavior will change, perhaps even in games that remain constant within the ensemble. 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 context 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. Our framework provides a theory for their experimental findings. People who play lots of games that demand cooperation, like the Prisoner's Dilemma, 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.

We next turn to behavioral stickiness, a puzzling phenomenon. Even when behavior does shift, it appears to do so slowly but completely, in patterns that permeate the full fabric of people's lives. 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 ensembles 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) analysis of candidate selection for Parliamentary elections in seventeenth-century England, which is often suggested as an example of the effect of culture. 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 more modern form contestation. Ferejohn shows 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. Using the logic of game theory, both changes made the cooperative agreements, necessary for the alternating equilibrium, less likely.

While Ferejohn's argument could not be made without a rich interpretation of history, it rests solely on efficiency. His logic does not support a role for cultural behavior. 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 necessarily include fuzzy math. But evidence indicates that is quite possible that the transition lagged by a generation or more; with 15 contests in 1604, when the changes began to take place, to 91 in the Long Parliament, starting in 1629 (Ferejohn 1991:292). The delayed evolution is consistent with survey evidence of intergenerational change. Our framework 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.

6 CONCLUSION

We develop a framework in which agents with cognitive constraints play multiple games. The ensemble of games influences behavior in individual games in ways that suggest the emergence of culture. We formally define five components of cultural behavior: *intra-individual consistency, inter-agent consistency, contextual effects, suboptimal behavior,* and *behavioral stickiness.* We also use our framework to demonstrate frequency dependence. We derive conditions that support all of these behaviors. Four emerge without any reliance on cognitive constraints.

These results depend on the particular games that we used in our ensemble only in so far that those games enforced the potential for a diversity of behaviors. They would not have emerged had the optimal behavior been the same in all games, but if that were true, we would not expect cultural behavior to emerge. Incentives would dominate. Nor are these results driven entirely by the cognitive constraints: Claim 1 established that the agents, even when limited to two cognitive states, are sufficiently sophisticated to evolve optimal strategies in each game. Instead, it is the fact that these cognitively constrained agents play multiple games simultaneously that generates the patterns of behavior that define cultural behavior.[9]

[9] Our results offer strong support for modeling any behavior with a multiple-games approach. Behavior in a specific context may well depend upon the other games people play, particularly when games have multiple equilibria.

Based on our results, we would argue that cultural behavior need not be primordial, but adapts to fit the ensemble of problems, a conclusion that agrees with the findings of Nisbett et al., discussed above. Thus our framework also has implications for the historical development of institutions. The set of existing institutions can influence later choices and the success of a society or a firm (Greif 1994, Stinchcombe 1965, Swaminathan 1996, and Boeker, 1989, Cohen and Bacdayan 1994). Institutional tinkering can alter behavior, although it may take some time, and there may be institutional path dependence, a subject we intend to pursue in future work. The introduction of market based or democratic mechanisms need not immediately (or ever!) create a culture that will make those mechanisms successful.

Our framework omits those aspects of culture that relate to symbolism, belief systems, and norms. 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 are be cognitively costly for some societies may be less so for others. 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.

Our framework could be applied in both the laboratory and the field. Laboratory experiments designed with 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 models with more descriptively accurate representations of human behavior such as neural nets or classifier systems, we might identify those institutions that create cognitive complementarities or more rigid behaviors.

The framework can be modified in many directions, including noise and alternative representations of bounded rationality (see Rubinstein 1998). Other definitions of cognitive costs could be employed, such as counting the number of games that are played differently. A model could be developed of interactions between communities. Our framework suggests an alternative explanation for clashes between cultures. 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).

Finally, much of the work on culture is thick and descriptive. The application of our model to historical cases could prove 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 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.

APPENDIX:

(insert Table 6 here)

Proof of Claims

In the mathematical results that we describe, we assume 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, the player 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 column player must play C toget the (C, S) outcome. If she 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. ***TR and BL****:* Row players choose C (resp. S), Column players choose S (resp. C).

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 and KE/SI:*** Row and column players choose GRIM for the PD and ALL S for SI. ALL S can be created from GRIM. Both are equilibrium strategies for the individual games. ***AL/KE:*** Row players choose TFT and 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. ***SI/TR and SI/BL:*** Row and column players each have two one-state automata. The row players use All S for the SI game and All C for the TR (resp BL) game. The column players use an All S for the SI game and an All S (resp C) for the TR game. ***TR/BL:*** Row and column players each have two one state automatas. The row players use an All C for the TR game and an All S for the BL game. The column players do the opposite.

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. Suppose that the transition mappings from that state map back to that state when the column agent chooses S. This strategy can be exploited if the column agent plays All S. Therefore, the row agents must 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 plays C and one state that plays 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. 1997. “The Dissemination of Culture: A Model with Local Convergence and Global Polarization.” *Journal of Conflict Resolution* 41:203-226.

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.

Bendor, Jonathon and Piotr Swistak. 2001. "The Evolution of Norms." *American Journal of Sociology* 106(6): 1493-1545.

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. *Econometric*a 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 (forthcoming) "Cultural Group Selection, Coevolutionary Processes and Large-Scale Cooperation" *Journal of Economic Behavior & Organization*

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.
Long, Herbert 1958 "The Local Community as an Ecology of Games" *American Journal of Sociology* 64: 251-261.
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, 6)
	S	(6, 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)	(0, 4)
	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 5: 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 All S
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, and 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 6: Simulation Results with Two-Game Ensembles

Ensemble	Behaviors and Equilibria Similar to Single Game Ensembles
SI / TR	(S,S) in SI and (C,S) in TR
SI / BL	(S,S) in SI and (S,C) in BL
TR / BL	(C,S) in TR and (S,C) in BL
KE / SI	85% use an alternating strategy in KE, of this 70% SAC in KE
AL / KE	No (C, C) in KE; approximately 70% SAC and 25%TFT
Ensemble	**Efficient Equilibria with Contextual Effects**
PD / AL	100% TFT in PD when efficient equilibrium found
PD / KE	35% play (C,C) in both games
PD / SI	100% GRIM when efficient equilibrium found in PD
AL / SI	65% TFT in AL, 35% SAC
Ensemble	**Suboptimal Behavior**
PD / TR	(S, S) in PD;
PD / BL	(S, S) in PD
AL / TR	(S,S) in AL
AL / BL	(S,S) in AL
KE / BL	(S,S) in KE
KE / TR	(S,S) in KE