Behavioral Spillovers with Interdependent Institutions: An Experimental Study*

Jenna Bednar[†] Yan Chen[‡] Tracy Xiao Liu[§] Scott Page[¶]

March 15, 2008
(Preliminary Version)

Abstract

We describe laboratory experiments that produce behavioral spillovers across strategic contexts. In these experiments subjects play two distinct games simultaneously with different opponents. We find that behavior is highly context dependent: when games are paired in ensembles, play differs from the isolated controls. Behavior is also influenced by which other game composes the ensemble, and in predicted ways. These results suggest that people do not treat each strategic situation in isolation but instead construct heuristics that they apply across games. The results reject the hypothesis that subjects play games independently; instead the findings imply that the effect of a particular institution on behavior depends upon the full institutional context. The results have implications for any attempt to transport institutions across contexts, including developmental programs and constitutional designs.

*This research was sponsored by grants from the National Science Foundation, the James S McDonnell Foundation, and the United States Air Force Office of Scientific Research MURI. We would like to thank Andrea Jones-Rooy for excellent assistance in the data assembly and analysis. We are also indebted to those who have supplied thoughtful reactions to our project, including Tom Palfrey, Rajiv Sethi, and seminar participants at the annual meeting of the Midwest Political Science Association (Chicago, March 2007), ICER (Turin, June 2007), ESA Asia Pacific Meetings (July 2007), and the Santa Fe Institute (July 2007).

†Department of Political Science, University of Michigan

‡School of Information University of Michigan

§School of Information University of Michigan

¶Center for the Study of Complex Systems, Departments of Political Science and Economics, University of Michigan

1 Introduction

Policies geared toward the economic improvement of developing nations or the establishment of democracy often fail. Part of the reason for this failure may be institutional mismatch. Some scholars have focused on how beliefs and trust can bootstrap market based institutions (e.g. North 2006, Greif 2006). Here, we consider an alternative based on behavioral repertoires. Rather than thinking of actors as optimizing given a set of beliefs, we characterize actors as adapting behavioral repertoires that they apply across multiple strategic contexts. Using an experimental design, we explore the degree to which behaviors necessary for the well functioning of one institution are compatible with the patterns of behavior created by another institutional form.

Our behavioral repertoires approach offers an alternative way to interpret recent studies that reveal systematic inter-population variance in subjects' responses to common incentive structures. This empirical evidence on diverse population level reactions to institutions spans multiple methodologies, including cross national surveys (Ingelhart 1990), laboratory experiments (Henrich et al 2001, 2004), and real world choices (Fernandez and Fogli 2007).

We propose that one reason that different societies react differently to common choice situations is that when confronted with a game, people build from or employ existing behavioral rules. This hypothesis implies behavioral spillovers across game. Conventionally, game theorists have analyzed play in isolated games, and experimental economics has focused on subjects who play single games, where their behavior is observed for isolated strategic environments. If behavior is context dependent, then an analysis of isolated strategic situations cannot explain variation in response to identical incentives. To that end, we construct a set of experiments in which subjects play multiple games simultaneously and examine whether attributes of one game influence behavior in other games.[1] Rather than depend upon diverse population

[1] Our design therefore differs from experiments in which subjects play distinct games sequentially

characteristics (such as wealth, education, risk, etc) to explain behavioral variation, we offer an alternative explanation: the effect of the institutional context.

Bednar and Page 2007 offer a theoretical approach to the study of institutional context on behavior, developing a model of multiple game analysis. In this study we develop laboratory experiments to study the effect of multiple-game play on subject's behavior. Subjects are presented with two games on their screen at the same time and are asked to indicate how they would like to play in each game. Other than having the two games presented on the screen simultaneously, there is no indication to the subject that the games are linked, and the subject is free to develop distinct strategies to each.

We are interested in the possibility that the implicit standard hypothesis from game theory of game independence does not reflect the way that human subjects respond to contextual strategic environments. Instead, we are interested in the possibility of externalities between games. We study two effects: behavioral spillovers between games and cognitive load. We posit specific hypotheses for each type of effect. The results strongly support our hypotheses: We find that behavior varies significantly from control treatments when games are played simultaneously. Furthermore, behavior in one game depends upon what other game is included in the two-game ensemble. This trend suggests that variance is not (exclusively) attributable to cognitive overload, but instead indicates the presence of behavioral externalities. Finally, behavior suggestive of cognitive overload is most present where predicted.

Why does this matter? Even if we can produce these behavioral spillovers in the laboratory, why should social scientists take notice? If we find evidence of the importance of institutional context for behavior, we demonstrate two possibilities with wide ranging implications. First, we provide an incentive-based explanation for some portion of the variation in institutional performance seen empirically. Our results would suggest that ensemble effects—externalities generated by the complete set of incen-

to create framing and learning effects. See for example Frolich and Oppenheimer (1996).

tive structures that an individual confronts—contribute to these differences. Second, to the extent that experiments like ours can show which combinations of institutions can coexist successfully, then they can help to explain why markets and democracies take off in some societies but do not in others. And, in some cases, experiments such as these might even inform the choice over institutions in designing political and economic transitions.

We have organized the paper as follows. In the next section, we propose a model and derive hypotheses testing the null of game independence against our two posited effects, behavioral spillovers and cognitive load. We also include a description of the specific games included in this study. Section 3 describes our experimental design. Sections 4 and 5 report our findings. Section 4 displays the results from our control treatment of games played in isolation. Section 5 presents the ensemble effects, first reporting the changes in behavior between the control and ensemble play, and next examining significant differences between ensembles that share a game in common. In Section 6 we discuss what these findings might mean and comment on potential future directions. In an appendix, we include tables of all behaviors observed in the laboratory and the levels of significant difference between the control and ensemble play. We generally report two behavioral proportions: one that is an average across all rounds, and a second that focuses on later rounds, to get a sense of behavioral adaptation during play. Using these comparisons we are able to offer a test of a learning model (which would hypothesize that play would become increasingly Pareto optimal) against a spillover model, where subjects would increasingly apply heuristics.

2 Model and Hypotheses

We focus on four two-by-two games: the Prisoner's Dilemma (PD), Strong Alternation (SA), Weak Alternation (WA), and a Self Interest game (SI).[2] The individual games

[2]These games are related to the six constructed in Bednar and Page (2007), in a theoretical article that derives behavioral spillovers and cognitive load effects. Their analysis relied on both mathematical theorems—they proved conditions for the existence and efficiency of behavioral externalities—and

belong to a class of two-person two-action games that contain a selfish action (S) and an alternative; in three of the games, this alternative is cooperative. In three of the games, cooperation lowers a player's own payoff and raises the payoff of the other and being selfish does the opposite, so in the one shot game, the unique dominant strategy equilibrium involves both players choosing selfish. In the fourth game, Self-Interest (SI), selfish behavior is both the stage game dominant strategy and Pareto optimal.

The first game is a standard Prisoner's Dilemma, where the stage game has a dominant strategy equilibrium, (S, S), which is Pareto dominated by (C, C). Note that (C, C) also maximizes the joint payoff of the two players.

Prisoner's Dilemma: (PD)

	C	S
C	7, 7	2,10
S	10,2	4,4

In the second and third games, Strong Alternation (SA) and Weak Alternation (WA), while (S, S) remains the dominant strategy equilibrium for the stage game, agents do best (i.e., maximize joint payoff) in repeated play by alternating between the off-diagonals, (C, S) and (S, C). In Strong Alternation, the incentives to alternate are much stronger than in Weak Alternation. The alternation games are a distant cousin to the conventional Battle of the Sexes, a game where agents are rewarded for coordinating their behavior, but they prefer opposite behaviors. In our alternation games, four behaviors are rewarded with positive payoffs: CC, SS, and the alternating strategies of CS then SC and SC then CS. Coordinating on CC or SS is much less taxing than working out an alternating behavior, and the positive payoffs for each reduce the focality of an alternating equilibrium.

Strong Alternation: (SA)

	C	S
C	7, 7	4,14
S	14,4	5,5

on computational agent based models (Miller and Page 2007). Their agent based models showed that simple learning rules could locate the proposed equilibria. That paper provides a theoretical foundation for the current paper: Here we test whether the phenomena derived within models and generated by artificial agents can be produced in a laboratory with real people.

Weak Alternation: (WA)		C	S
	C	7, 7	4,11
	S	11,4	5,5

In the final game, Self Interest (SI), the dominant strategy equilibrium, (S, S), also Pareto dominates all other outcomes. Furthermore, in the stage game, S uniformly dominates C.

Self Interest: (SI)		C	S
	C	7, 7	2,9
	S	9,2	10,10

The implicit null hypothesis in our investigations is of game independence: play in one game should not be affected by the existence of another game to play. If the independence hypothesis is correct then we should see no difference between behaviors in the control studies (games played in isolation) and when games are presented to subjects as part of ensembles. Our experimental design tests the existence of two types of ensemble effects: *behavioral spillovers* and *cognitive load*. If games presented within ensembles create *behavioral spillovers*, subjects will respond as if they are developing heuristics that they apply across games. In particular, dominant behavior in one game will influence choice in another. With cognitive effects, subjects' propensity to play Pareto dominant behavior will decline as suboptimal play increases.

These effects may be manifested in two ways: (1) control/ensemble, where we compare behavior in a game played in isolation against the behavior in that game when it is paired with other games, and (2) ensemble/ensemble, comparing behavior in a game when it is matched with different games. Our investigations center on two questions: first, does behavior differ between the control—isolated game play—and ensemble, where the game is paired with others? Second, does the ensemble play depend upon *which* other game is in the ensemble? A positive answer to either of these questions would fail to support the null hypothesis of independent play, and support instead the hypothesized contextual dependence of game play.

H1: Pareto Dominant Behavior *In the control treatments, the subject's choice will approach the Pareto dominant behavior in each game.*

In the individual game treatments, we anticipated that three distinct behaviors would emerge: in the PD game, cooperation induced by fear of punishment; in the SI game, selfishness, given its dominance; and in SA and WA an alternating form of cooperation, where subjects alternate between the cooperative and selfish actions. Weak Alternation has weaker incentives, so Pareto optimal behavior may not be as likely as with Strong Alternation. Each of these behaviors is Pareto optimal in the stage game (in the Alternation games, alternation is Pareto optimal over two rounds of play).

H2–5: Behavioral Spillovers *Subject's choice of action in a particular game will be influenced by the other game in the ensemble, particularly biasing choice toward the other game's Pareto optimal behavior.*

Specifically, we expect:

- **[H2]:** Games paired with Self-Interest will exhibit more selfishness.
- **[H3]:** Games paired with the Prisoner's Dilemma will exhibit more cooperation.
- **[H4]:** Games paired with Strong Alternation or Weak Alternation will exhibit more alternation, with more significant effects in Strong Alternation.
- **[H5]:** Behavioral spillovers will increase over the series.

The final behavior spillover hypothesis is not a specific prediction, but rather an interest to test two competing hypotheses: growing inclination to apply heuristics versus a standard learning model where subjects grope their way toward payoff-maximizing play. The latter alternative is related to the independence hypothesis. **H5** is supported if spillover effects increase through the series; we may interpret this as evidence of a growing prediliction to apply a successful strategy elsewhere. This effect may be particularly strong if the strategy requires orchestrated coordination, such as with WA or SA. However, this tendency may be overcome as the rounds progress, and the players gain experience: initial cross-application of one game's successful strategy may diminish as the subject is able to turn attention to solving the other game in the ensemble. We therefore offer the fourth as a "test" hypothesis; not a prediction, but an inquiry.

Support for these hypotheses is evidence that behavior from one game in the ensemble bleeds over into other games. Note that to guard against coordination focality as an explanation for this phenomenon, players are matched with different opponents for each of the two games, a design feature that we discuss in more detail when describing the experiments.

Secondly, we developed a partial ordering of the four games based upon the strength of the incentives for the subjects to reach Pareto optimal play. We focus on the Pareto optimal outcome that gives both players identical payoffs. In the Self Interest game the dominant strategy equilibrium is also Pareto optimal, so this is the simplest game. In the other three games, repeated game strategies are necessary to sustain the Pareto optimal outcome. The coordinated equilibrium of Alternation requires that the two players synchronize their behaviors; it is more difficult to coordinate on this behavior than the relatively easier mutual cooperation that sustains the Pareto optimal outcome in PD. Given this logic, and given that the incentives to alternate are only slightly greater than cooperation, Weak Alternation is the most difficult game to play efficiently. We do not rank SA and PD; while PD has the coordination advantage, the incentives in SA are strong. In sum, we posit that Self Interest is the easiest game to play and Weak Alternation the most difficult, and this relative difficulty will be reflected in subjects' behavior.

H6–7: Cognitive Load *Subject's choice of action will be affected by the difficulty of the game.*

- **[H6]:** Game play should correlate with ease of play, with behavior in SI the most efficient and WA the least.

- **[H7]:** Cognitive load spillover effects will be most prevalent in games played with more difficult games. In particular, subjects will exhibit more sub-optimal behavior in SA and PD when either game is played with WA.

These final hypotheses convey our interest in the experiment's ability to reveal limitations in the cognitive processing of subjects playing multiple games simultaneously. While there are no design features to the experiments that would preclude the subjects from optimizing in each game, we believe that when subjects are asked to solve two games simultaneously they will not be as efficient as they are when playing an isolated game. In particular, with [**H7**] when an ensemble contains WA, we predict more *selfish* behavior in PD or SA—an outcome that is third best. The relative ease of playing SI means that it should generally be free of contextual effects, and play in the control treatments should approximate play in the ensemble treatments [**H6**].

3 Experimental Design

Our experiments consist of four control sessions, each of which consists of a single game, and 14 treatment sessions, each of which consists of a pair of games. This experimental design enables us to determine the effects of ensemble on behavior by comparing the ensemble with the corresponding control sessions and to compare behavior across ensembles.

The control sessions follow the standard protocol of infinitely repeated games in the laboratory. We have one 12-player session for each of the single games. Participants are randomly matched into pairs at the beginning of each session, and play the same match for the entire experiment. In each session, participants first play the game for 200 rounds. After round 200, whether the game will continue to the next round depends on the "throw of the die" that is determined by the computer's random number generator. At the end of each round after round 200, with 90% chance, the game will continue to the next round. With 10% chance, the game stops. In other words, we implement an infinitely repeated game, with a discount factor of 1 for the first 200 rounds, and 0.9 thereafter. With the chosen discount factors, (C, C) can be sustained as a repeated game equilibrium in PD, SA and WA. With 12 players in each control session, we have 6 independent observations for each single game.

In the ensemble treatment, we again use twelve players in each session. Within each session, at the beginning, each player is randomly matched with two other participants, both of whom will be her matches for the entire experiment. She plays two distinct games with each of these people. This design allows us to analyze whether or not behavior in one game is influenced by the nature of the other game. As in the control sessions, we implement an infinitely repeated game, with a discount factor of 1 for the first 200 rounds, and 0.9 thereafter. Within each session, the 12 players are partitioned into independent groups of 4 each,[3] yielding 3 independent observations. As the two games are displayed side by side, we conduct two independent sessions for each game ensemble, changing the order of the display to avoid the order effect within each round. For example, for the game ensemble of SA and WA, we display SA as the left game in one session, and WA as the left game in another session. This way, if a player always makes decisions from left to right, we have a balanced number of observations for each order.

We used z-Tree to program our experiments. As z-Tree does not record the mouse movements within each stage, we ran two additional sessions with ensembles, (SI, WA) and (WA, SI), where we use the software Morae to record the mouse movement. These two sessions will allow us to determine the order of decisions within each round. The (SI, WA) session has 12 subjects, while the (WA, SI) has only eight subjects.[4]

[Table 1 about here.]

Table 1 reports the features of experimental sessions, including the name of the game, the number of players in each session, the number of independent pairs for each

[3]The matching protocol is the following: $\underbrace{4-2-1-3}$, $\underbrace{6-5-7-8}$, $\underbrace{10-9-11-12}$ form three independent groups, each with four participants positioned on a circle, and each participant plays her left and right match.

[4]We recruited for twelve subjects, however, only eight showed up.

control session, the ensemble of games, the number of players in each session, as well as the number of independent groups in each ensemble session.

Overall, 18 independent computerized sessions were conducted in the RCGD lab at the University of Michigan from March to October 2007, yielding a total of 212 subjects. Our subjects were students from the University of Michigan, recruited by email from a subject pool for economic experiments.[5] Participants were allowed to participate in only one session. Each ensemble treatment session lasted approximately 90 minutes, whereas each control session lasted about 45 minutes. The exchange rate was set to 100 tokens for $1. In addition, each participant was paid a $5 show-up fee. Average earnings per participant were $37.49 for those in the treatment sessions and $22.77 for those in the control sessions. Data are available from the authors upon request.

4 Results

We first outline the basic results from the control sessions on the individual games. We then report results from the ensembles, comparing behavior in the ensemble with that in the control, as well as behavior across ensembles.

4.1 Control Sessions

In this subsection, we report the results from the control sessions. In addition to supporting the derivation of our ranking of the games for the cognitive load hypotheses, these results provide a benchmark from which we can interpret the ensemble results.

As the games are implemented as infinitely repeated games, there can be many repeated game strategies. In the analysis, we restrict ourselves to three simple repeated games strategies, SS, CC and ALT. (Elaborate ...) When we represent repeated game strategies with automata, these strategies are the simplest, i.e., whose automaton representation has the least number of states. There have been empirical evidence that

[5] Graduate students from the Economics Department are excluded from the list.

simple repeated game strategies are more likely to be chosen (Baron and Kalai 1993). In addition to the simplicity argument, data in the control sessions also support our focus on simple strategies.

[Figure 2 about here.]

We first present the time series data for each pair in each of the control sessions. Figure 2 presents behavior in the Self Interest game. In this game, all six pairs converged to the Pareto dominant equilibrium quickly and stayed there. We speculate that this is because of the uniform dominance property of the dominant strategy equilibrium. Additionally, subjects required significantly less time to play Self Interest than any other game ($p < 0.01$ for every pairwise comparison, one-sided permutation tests.) Based upon the uniform dominance property of the unique Pareto efficient payoffs and the length of time it took for subjects to complete the game, we posit that SI would be the easiest to play efficiently. It also provides a strong and clean context.

[Figure 1 about here.]

Figure 1 presents behavior in the Prisoner's Dilemma game. In this game, over half of the pairs established CC, the efficient outcome, which is consistent with findings from previous experiments (Andreoni and Miller 1993). Curiously, one pair also alternated for a fair number of rounds. As a "context" this game does not establish as strong a behavioral norm as the Self Interest or Strong Alteration games. Based upon this finding, we anticipate that PD will have a weaker behavioral pull than either SI or SA. The difficulty of learning to cooperate in the PD game may limit its spillover effects on play in Strong Alternation for the simple reason that play in SA may already evolve to some stable outcome before play in the PD game has developed a common behavior. In short, behavior that doesn't yet exist cannot spill into another context.

[Figure 3 about here.]

Figure 3 presents behavior in the Strong Alternation game, where 5/6 of the pairs successfully established the alternation outcomes. Pair 2 also tempted alternation on and off during the experiment. This game also provides a strong benchmark (or context?).

[Figure 4 about here.]

Lastly, Figure 4 presents the dynamics from the Weak Alternation game. In this game, only two out of six pairs develops an alternating behavior, two pairs cooperate, one (pair 4) converge to SS, and the last pair (pair 6) does not seem to have converged to a stable outcome. In sum, none of the simple strategies emerges as the dominant choice of the subjects. Therefore, we speculate that, while subject behavior in WA is more likely to be influenced by the other game in an ensemble, when paired with other games, it might increase the subjects' cognitive load.

[Table 2 about here.]

We now summarize the results in the control sessions. Table 2 reports the proportion of simple strategies in each game over the entire series. If we restrict ourselves to the 101-200-round blocks, or the last 100-round block, the results largely hold.

Result 1. *In the control sessions, the proportion of simple Pareto optimal strategy play (SS in SI, CC in PD, ALT in SA and WA) is significantly higher than any other simple strategies in SI, PD and SA, while for WA, there is no significant difference in the proportions of simple strategies played.*

Support. *Table 2 reports the proportion of simple strategy play in each of the four games, and the corresponding p-value for the one-sided permutation tests for each pairwise comparison. For SI, PD and SA, all pairwise comparisons are significant at the five-percent or one-percent levels, while for WA, none of the pairwise comparisons is significant at the five-percent level.*

We next compare the efficiency of each game. Efficiency is defined as the actual total payoffs of the two players divided by the maximum joint payoffs.

[Table 4 about here.]

Table 4 presents the average efficiency achieved in each control session, as well as the p-values for pairwise comparisons using the permutation tests. In the analysis, the average efficiency achieved by each pair in a session is an independent observation.

Result 2. *In the control sessions, the SI game generates significantly higher efficiency than any other game, while pairwise efficiency comparison is not significant between the PD, SA and WA games.*

Support. *In Table 4, each pairwise comparison between SI and the other three games is significant at the 1% or 5% level, while none of the other pairwise comparisons is significant at the 10% level.*

From the control sessions, in SI, the Pareto efficient stage game Nash equilibrium is played 99.86 percent of the times overall, and 100 percent in the last 100 rounds.

We can characterize the game in terms of simplicity, i.e., the number of automata needed in playing the simple strategies. The both SI and PD are simpler than SA and WA. If we use decision time as a proxy for mental activities needed to play the game, SI take less time per round than any other game, followed by PD, which is in turn followed by SA and WA.

4.2 Ensemble Effects

We present our results comparing both control and ensemble play and behavioral differences in particular games compared across ensembles. To briefly summarize, we find, as expected, evidence of behavioral spillovers across games. In some cases, those spillovers are stronger than we anticipated and in others weaker. For example, when Weak Alternation is paired with Self Interest, players prove far more likely to take the same action period after period in Weak Alternation. In contrast, in the pairing of Weak Alternation and the Prisoners' Dilemma, we do not get as much play of CC in the Weak Alternation game as we expected. Both of these results are likely to represent the confounding effect of cognitive load on behavioral spillovers, which we anticipated to be most significant with WA.

4.2.1 Comparing Control and Ensemble

Our anticipation was that subjects would play particular games differently between the control treatments, where they played a single game, and when that game appeared as part of an ensemble. This prediction emerges from the two core hypotheses: both cognitive taxes and behavioral spillovers will affect play in ensembles. We do not rule out the possibility that subjects use supergame strategies (to be elaborated). Here we highlight some results from the data that support these expectations.

[Table 5 about here.]

Result 3 (Effect of SI in Ensembles)**.** *In (SI, **PD**) and (SI, **SA**), the proportion of SS in PD and SA is significantly higher than the corresponding PD and SA control groups.*

Support. *Table 5 reports the proportion of SS in controls and ensembles (note still have 91-190 rounds). Table 10 reports p-value for the one-sided Montel Carlo Permutation Tests. The null hypothesis is that the proportions of different strategies are the same between different PD(SA) ensembles and their control groups. We reject the null in favor of the alternative hypothesis that the proportion of SS is (weakly) significantly higher in the ensemble (*$p = 0.064$ *(all rounds) and* 0.028 *(last 100 rounds) for SA;* $p = 0.066$ *(all rounds) and* 0.091 *(last 100 rounds) for PD).*

The Self Interest game is the easiest game to play; whether in control or ensemble treatments, subjects quickly converged on selfish play, with over 99% of selfishness across the rounds. This behavior confirmed our predictions: we did not anticipate behavioral spillovers to affect play in the Self Interest game. However, we did posit strong behavioral effects of SI on other games [**H2**] precisely because of its ease of play (not a very good argument, not from theory ...). See Table 5; the proportion of selfish behavior exhibited in both PD and SA more than double when these games are paired with SI. The effect on the Prisoner's Dilemma is particularly salient given the game's ubiquity in the social sciences: it appears highly vulnerable to selfish contextual influence, and subjects do not recover in later rounds but instead maintain the selfish behavior. The change in behavior in WA is not significant, but as WA did not exhibit strong behavioral tendencies when played in isolation, the lack of significance does not surprise us.

Result 4. *Compared to the corresponding control sessions, in (**SA**, PD) and (**SA**, SI), the proportion of ALT in SA is significantly less than that in SA played in isolation.*

[Table 7 about here.]

Support. *Table 7 presents the proportion of simple strategies in the control sessions of PD and SA and in the ensembles, (**SA**, PD), (**SA**, SI). The last three column reports p-value for the one-sided permutation tests. The null hypothesis that the proportion of simple strategies is the same between the ensemble and the corresponding control is rejected in favor of the alternative that the proportion of ALT in the two ensembles is significantly less than that in SA played in isolation.*

In Strong Alternation, subjects alternate less ($p = 0.083$) and cooperate more ($p = 0.049$) when SA is paired with PD [**H3**]. The initial cooperation increase with PD loses significance in the later rounds, suggesting that subjects initially applied a heuristic from PD, but then learned the more difficult alternating form of cooperation as the rounds progressed (failing to support [**H5**]). Other significant results may be noted by examining the behavioral tables attached at the end of the paper.

Finally, we note the effect of cognitive load. We hypothesized that when either PD or SA was played with WA, subjects would play PD and SA suboptimally, including increased selfish behavior [**H7**], i.e., resorting to the dominant strategy in the stage game. Findings confirm these expectations, as demonstrated in Table 6. When PD is paired with WA, subjects played selfish nearly as often as when PD was paired with SI; for SA, the effect of pairing SA and WA was even greater than with SA and SI. Again, we want to draw particular attention to the vulnerability of the Prisoner's Dilemma; in this case cognitive load appears to make cooperative behavior less likely to emerge.

Result 5. *In (**SA**, WA), the proportion of SS is significantly higher and the proportion of ALT is significantly lower that that of the respective strategies in SA alone. Furthermore, (**PD**, WA), the proportion of SS is significantly higher than that in PD alone.*

[Table 8 about here.]

Support. *Table 8 presents the proportion of simple strategies in the control sessions of SA and PD as well as in the ensemble, (**SA**, WA), (**PD**, WA). The last three column reports p-value for the one-sided permutation tests. The null hypothesis that the proportion of simple strategies is the same between the ensemble and the corresponding control is rejected in favor of H_1 that the proportion of SS is significantly higher in the ensemble. Furthermore, the proportion of ALT is significantly lower in (SA, WA) compared to SA alone.*

[Table 9 about here.]

Lastly, we compare the efficiency in the control and ensemble sessions. Table 9 presents the average efficiency in the control and ensemble sessions. For each ensemble, we present the efficiency of each game in the ensemble, as well as the overall ensemble efficiency.

Result 6. *Efficiency in SI, PD and SA control sessions is higher than that of the corresponding games in ensemble sessions. The following comparisons are significant at the 5% level: SI > (**SI**, SA), SI > (**SI**, WA), and SA > (**SA**, WA).*

[Table 10 about here.]

Support. *Table 10 presents the p-values of one-sided permutation tests comparing efficiency between control and ensemble sessions (top panel) and between ensemble sessions (bottom panel).*

Result 6 follows directly from the comparison of behavior between the control and ensembles.

4.2.2 Comparing Behavior Between Ensembles

A second method for investigating the presence of behavioral spillovers and cognitive overload is to compare behavior between ensembles. The effect of behavioral spillovers on play in the PD is most significantly seen between ensembles. We also highlight key results in WA and SA play.

Result 7. *Comparing (**PD**, SA) and (**PD**, SI), subjects alternated more in PD when paired with SA (28% versus 5%) and pairs played selfishly more in SI (45% versus 24%).*

Support. *One-sided permutation tests comparing the proportion of ALT in (**PD**, SA) and (**PD**, SI) yield* $p = 0.10$*. Similarly, one-sided permutation tests comparing the proportion of SS in PD between the two ensembles yield* $p = 0.10$*.*

Result 8. *Comparing (**PD**, SA) and (**PD**, WA), subjects alternated in PD more with SA (22% versus 2%) and played the PD selfishly significantly more with WA (38% versus 24%).*

Support. *One-sided permutation tests comparing the proportion of ALT in (**PD**, SA) and (**PD**, WA) yield* $p = 0.09$*. Similarly, one-sided permutation tests comparing the proportion of SS in PD between the two ensembles yield* $p = 0.04$*.*

Result 9. *Comparing (**WA**, PD) and (**WA**, SA), subjects alternated more in WA when also playing SA (37% versus 18%) and cooperated significantly more when WA was paired with PD (29% versus 11%).*

Support. *One-sided permutation tests comparing the proportion of ALT in (**WA**, PD) and (**WA**, SA) yield* $p = 0.10$. *Similarly, one-sided permutation tests comparing the proportion of CC in WA between the two ensembles yield* $p = 0.022$.

Lastly, comparing (**SA**, PD) and (**SA**, SI), subjects played CC more often in SA when paired with PD (14% versus 7%). In later rounds, the effect grows even more pronounced: while they continue to play CC in SA when it is paired with PD, where SA is paired with SI subjects shift from CC to ALT, so that the final CC percentages are 14% in (**SA**, PD) versus 1% in (**SA**, SI), a difference significant at the .01 level [**H3**].

In general, the experimental results agree with our hypotheses: game independence is not supported, but instead subjects are influenced by contextual effects of behavioral spillovers and cognitive load. The exception is [**H5**]: We did not find unambiguous support for increasing cross-application of heuristics; arguably, behavioral spillovers decrease over time, in support of a learning hypothesis (To be completed).

5 Discussion

In this paper, we describe an experimental study to test for ensemble effects in game playing behavior. Our study reveals strong evidence of behavioral spillovers that depend in predictable ways on features of the games in the ensemble. In particular, if subjects play one game in an ensemble that encourages selfishness or cooperation, then they are more likely to exhibit that behavior in the other game in their ensemble. We also see evidence of cognitive overload when ensembles include the Weak Alternation game.

Our findings that show how a person's behavior in a given game depends on the ensemble of strategic situations that the person faces calls into question both the theoretical and empirical analysis of isolated games as well as the standard mechanism design assumption that incentives can be considered independent of the broader behavioral context. They also provides a possible path toward greater understanding of cross organizational and cross national differences in the performance of specific institutions or game forms.

In future work, we hope to consider an experimental design that adds games to the ensemble sequentially. To experimentalists, who worry about contaminated subject pools, a finding of sequential effects would not be a huge surprise. However, if we can learn how people play one in one game depends upon they other games they face, then if we know the current ensemble of strategic choices that a population confronts, we can gain insights into how they might behave when confronted with a new institutionalized game form - such as a market or democratic mechanism. Moreover, evidence of sequential behavioral spillovers would imply that potential for a theory of institutional path dependence based on behavioral spillovers (Page 2006).

As with any laboratory experiment, our results may not translate to the larger world. People often rely on contextual clues to behave differently in distinct situations. Thus, people can act altruistically to their children but competitively at work. We do not deny the human capacity to bracket contexts and act accordingly. However, we believe that such contextual bracketing requires cognitive effort and that, in general, people will seek out consistent behaviors that apply across multiple settings. Our experiments support that hypothesis.

A second potential criticism pertains to the simplicity of the games we consider. Would these effects continue to hold for more complex games embedded in a richer institutional and cultural context? We cannot answer that question in a laboratory. But the fact that the game ensembles can influence behaviors in individual games in a relatively sterile laboratory would seem to suggest that such effects might also exist in the real world as well.

To summarize, these experiments demonstrate that significant ensemble effects

Table 1: Features of Experimental Sessions

Control		Ensemble Treatment		
Game: n	Pairs	(Left, Right):	n	Groups
		(PD, WA):	12	3
		(WA, PD):	12	3
		(PD, SI):	12	3
PD: 12	6	(SI, PD):	12	3
		(SA, WA):	12	3
		(WA, SA):	12	3
		(SA, PD):	12	3
SA: 12	6	(PD, SA):	12	3
		(SI, WA):	$12 + 12$	6
		(WA, SI):	$12 + 8$	5
		(SI, PD):	12	3
SI: 12	6	(PD, SI):	12	3
WA: 12	6			
Total: 48	24		164	

can be produced in the lab. And, more importantly, these ensemble effects can be predicted based upon the the attributes of the games. Subjects with incentives to behave cooperatively (resp. selfishly) in one game, tend to behave similarly in another game <u>even if that behavior is neither efficient nor an equilibrium.</u> This finding suggests that when we consider the performance of an institution, we should look not only at the outcome that it produces but also at the behaviors that produce that outcome as those behaviors may well influence play in other settings. The creation of a cooperative culture may well depend on creating multiple institutions that create strong incentives for cooperation so that cooperative behavior can spill over into other contexts as well.

Table 2: Average Proportion of Simple Strategies in Control Sessions

Game	% Simple Strategies			P-value of Pairwise Comparisons		
	CC	SS	ALT	CC vs. SS	CC vs. ALT	SS vs. ALT
SI	0.00	**99.86**	0.00	0.0000	1(Two-sided)	0.0000
PD	**55.68**	17.82	15.44	0.0386	0.0313	0.3885
SA	5.02	14.81	**71.12**	0.0402	0.0000	0.0008
WA	33.18	22.51	**36.14**	0.3082	0.4304	0.3174

Table 3: Average Proportion of Pareto Optimal Behaviors in Control Experiments

	SS in SI	ALT in SA	CC in PD	ALT in WA
whole series	99.86	71.11	55.68	36.14
rounds 91-190	100.00	84.00	58.17	35.00

Table 4: Average Efficiency in Control Sessions

Game	Average Efficiency	Pairwise Comparison		
		Game1	Game2	P-value
SI	99.94	SI	PD	0.0018
		SI	SA	0.0000
		SI	WA	0.0138
PD	88.58	PD	SA	0.2639
SA	92.30	PD	WA	0.3831
WA	90.29	SA	WA	0.3506

Table 5: Proportion of Selfish Behavior in Controls and with Self Interest (* p<0.1; **p<0.05)

	Whole Series		Rounds 91-190	
	Control	w/SI	Control	w/SI
SI	99.86		100.00	
PD	17.82	44.55*	14.50	44.50*
SA	14.81	32.39 (p:0.1197)	7.33	32.50**
WA	22.51	28.62	23.50	24.34

Table 6: Proportion of Selfish Behavior in Controls and with WA or SI (* p<0.1, ** p<0.05)

	Whole Series			Rounds 91-190		
	Control	w/WA	w/SI	Control	w/WA	w/SI
PD	17.82	37.65*	44.55*	14.50	39.40*	44.50*
SA	14.81	39.39**	32.39 (p:0.1197)	7.33	38.00**	32.50**

Table 7: Control vs. Ensemble: (SA, PD) and (SA, SI)

Control	% Simple Strategies			Ensemble	% Simple Strategies			Control vs. Ensemble		
	CC	SS	ALT		CC	SS	ALT	CC	SS	ALT
PD	55.68	17.82	15.44	(PD,SI)	43.60	44.54	4.95	0.283	0.067	0.246
				(PD,SA)	38.81	24.03	21.93	0.224	0.291	0.357
SA	5.02	14.81	71.12	(SA,SI)	6.75	32.39	47.92	0.312	0.066	0.089
				(SA, PD)	13.96	24.92	47.20	0.049	0.188	0.083

Table 8: Control vs. Ensemble: (SA, WA) and (PD, WA)

Control	% Simple Strategies			Ensemble	% Simple Strategies			Control vs. Ensemble		
	CC	SS	ALT		CC	SS	ALT	CC	SS	ALT
SA	5.02	14.81	71.12	(SA,WA)	9.90	39.39	37.76	0.106	0.045	0.038
PD	55.68	17.82	15.44	(PD,WA)	40.32	39.20	8.69	0.182	0.027	0.401

Table 9: Efficiency in Ensemble

Control	Efficiency in Control	Ensemble (Game1, Game2)	Average Efficiency in Ensemble		
			Game1	Game2	Ensemble
SI	99.94	(SI,PD)	99.69	79.22	91.26
		(SI,SA)	99.66	84.10	92.29
PD	88.58	(SI,WA)	99.55	87.69	94.46
		(PD,SA)	84.39	85.82	85.20
SA	92.30	(PD,WA)	80.27	84.66	82.54
WA	90.29	(SA,WA)	80.29	85.96	82.87

Table 10: P-values of Efficiency Comparison

Control	Ensemble					
	(SI,PD)	(SI,SA)	(SI,WA)	(PD,SA)	(PD,WA)	(SA,WA)
SI	0.112	0.000	0.004			
PD	0.113			0.230	0.059	
SA		0.057		0.096		0.033
WA			0.291		0.140	0.272
Ensemble	(SI,PD)	(SI,SA)	(SI,WA)	(PD,SA)	(PD,WA)	(SA,WA)
(SI,PD)		0.383	0.096	0.082	0.010	
(SI,SA)			0.136	0.042		0.031
(SI,WA)						
(PD,SA)					0.230	0.321
(PD,WA)			0.000			0.467
(SA,WA)			0.003			

Appendix: Behavioral Data

Table 11: Behavior in Self-Interest (All Rounds / Rounds 91-190)

	Control	w/PD	w/SA	w/WA
Selfish	99.86/100.00	99.32/100.00	99.24/99.84	98.55/98.75
Cooperative	0.07/0.00	0.00/0.00	0.00/0.00	0.12/0.09
Alternating	0.00/0.00	0.00/0.00	0.00/0.00	0.28/0.42

Table 12: Behavior in Prisoner's Dilemma (All Rounds / Rounds 91-190) (* p<0.1)

	Control	w/SI	w/SA	w/WA
Selfish	17.82/14.5	44.55*/44.50*	24.03/19.25	37.65*/39.40*
Cooperative	55.68/58.17	43.60/46.67	38.81/43.75	42.15/43.77
Alternating	15.44/20.50	4.95*/4.34*	21.93/24.67	1.89/1.09

Table 13: Behavior in Strong Alternation (All Rounds / Rounds 91-190) (* p<0.1, ** p<0.05)

	Control	w/SI	w/PD	w/WA
Selfish	14.81/7.33	32.39/32.50*	24.92/20.34	39.39**/38.00**
Cooperative	5.02/4.17	6.75/1.58	13.96*/14.08	9.90/6.00
Alternating	71.11/84.00	47.92/55.42*	47.20*/54.75*	37.65**/44.25**

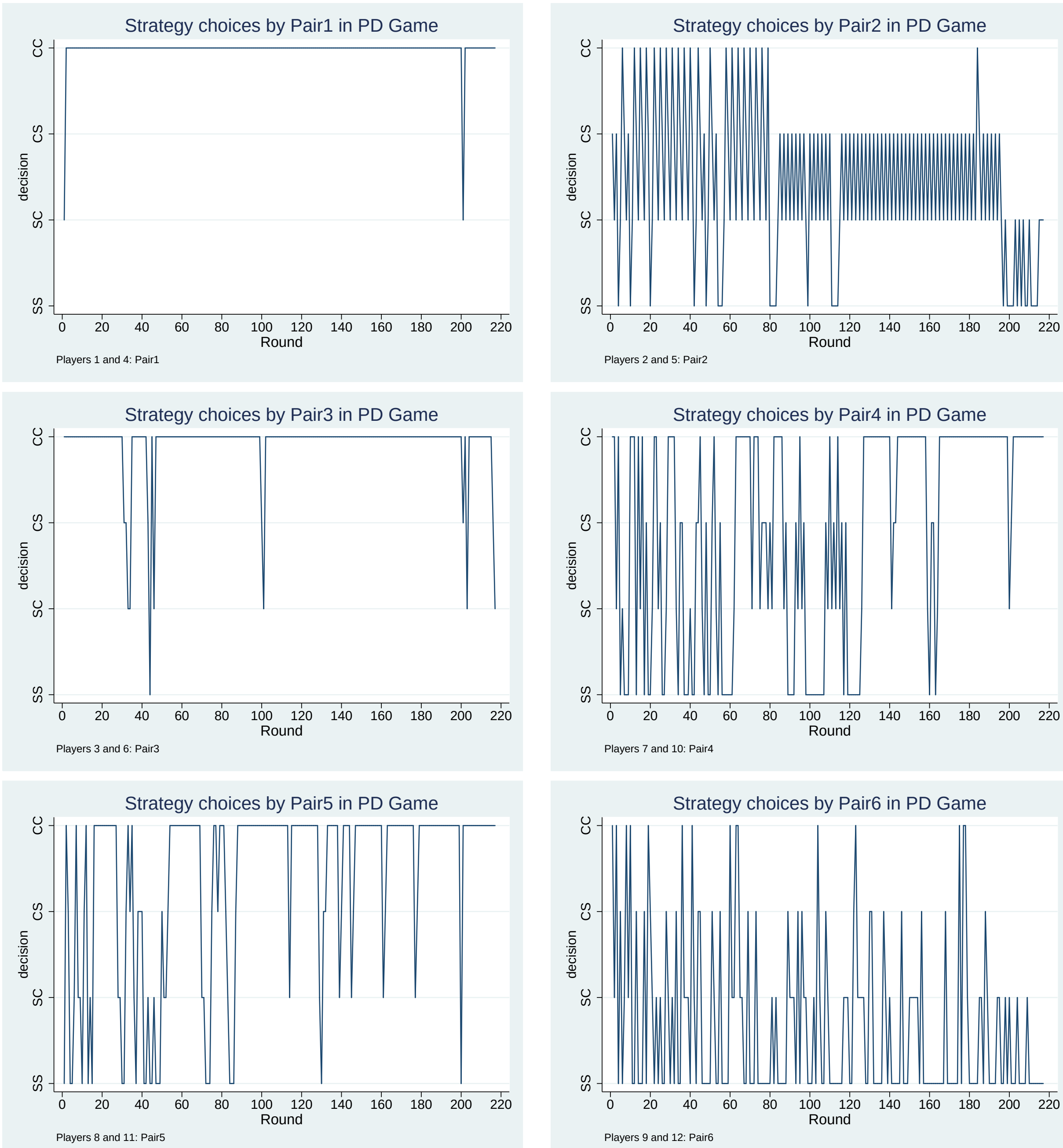

Figure 1: Control: Behavior in Prisoner's Dilemma

Figure 2: Control: Behavior in Self Interest

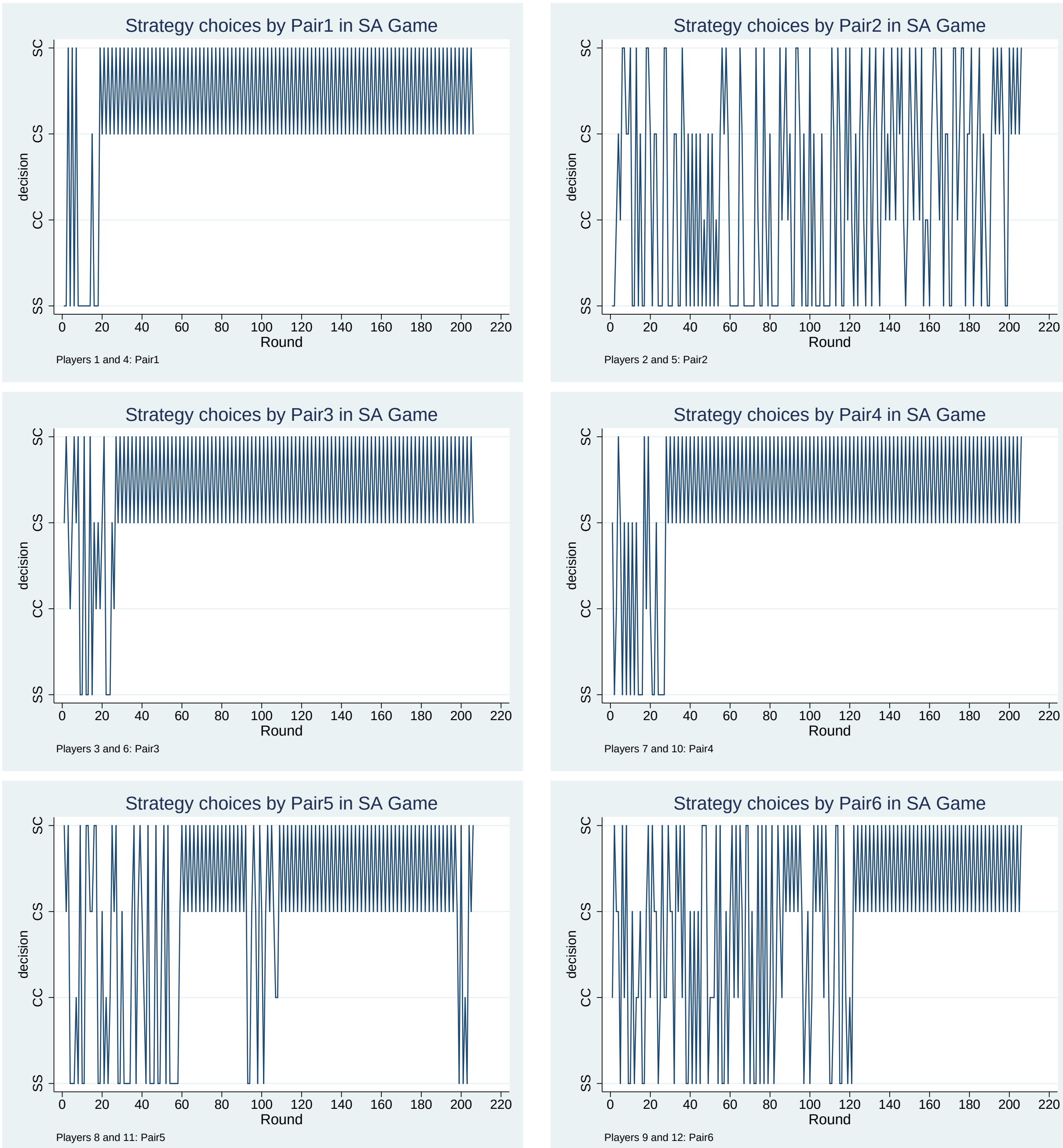

Figure 3: Control: Behavior in Strong Alternation

Figure 4: Control: Behavior in Weak Alternation

Table 14: Behavior in Weak Alternation (All Rounds / Rounds 91-190) (* p<0.1, ** p<0.05)

	Control	w/SI	w/PD	w/SA
Selfish	22.51/23.50	28.62/24.34	38.23/38.73	39.83/37.83
Cooperative	33.18/36.00	33.10/31.84	28.91/31.79	11.44*/8.75**
Alternating	36.14/35.00	30.98/41.58	17.93/17.56	36.67/44.84

Table 15: Proportion of Selfish Behavior (* p<0.1, ** p<0.05)

	Whole Series					Rounds 91-190				
	Control	w/SI	w/PD	w/SA	w/WA	Control	w/SI	w/PD	w/SA	w/WA
SI	99.86		99.32	99.24	98.55	100.00		100.00	99.84	98.75
PD	17.82	44.55*		24.03	37.65*	14.50	44.50*		19.25	39.40*
SA	14.81	32.39 p:0.1197	24.92		39.39**	7.33	32.50 p:0.0515	20.34		38.00**
WA	22.51	28.62	38.23	39.83		23.50	24.34	38.73	37.83	

Table 16: Proportion of Cooperative Behavior (* p<0.1)

	Whole Series					Rounds 91-190				
	Control	w/SI	w/PD	w/SA	w/WA	Control	w/SI	w/PD	w/SA	w/WA
SI	0.07		0.00	0.00	0.12	0.00		0.00	0.00	0.09
PD	55.68	43.60		38.81	42.15	58.17	46.67		43.75	43.77
SA	5.02	6.75	13.96*		9.9	4.17	1.58	14.08		6.00
WA	33.18	33.10	28.91	11.44 p:0.0664		36.00	31.84	31.79	8.75	

Table 17: Proportion of Alternating Behavior (* p<0.1, ** p<0.05)

	Whole Series					Rounds 91-190				
	Control	w/SI	w/PD	w/SA	w/WA	Control	w/SI	w/PD	w/SA	w/WA
SI	0.00		0.00	0.00	0.28	0.00		0.00	0.00	0.42
PD	15.44	4.95 p:0.1039		21.93	1.89	20.50	4.34 p:0.1061		24.67	1.09
SA	71.11	47.92 p:0.1081	47.20*		37.65**	84.00	55.42 p:0.1001	54.75*		44.25**
WA	36.14	30.98	17.93	36.67		35.00	41.58	17.56	44.84	

References

[1] Andreoni, James and John H. Miller (1993). Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence. *The Economic Journal*, Vol. 103, No. 418, pp. 570–585.

[2] Bednar, Jenna and Scott E Page (2007). Can Game(s) Theory Explain Culture?: The Emergence of Cultural Behavior in Multiple Games *Rationality and Society* 19(1):65–97.

[3] Bykowsky, Mark M. Robert J. and Cull John O. Ledyard (2000) Mutually Destructive Bidding: The FCC Auction Design Problem *Journal of Regulatory Economics* 17(3) 205–228.

[4] Camerer, C. & Ho, T.H. (1999) Experience Weighted Attraction Learning in Normal Form Games. *Econometrica* 67: 827–874.

[5] Chen, Yan and Fang-Fang Tang (1998). Learning and Incentive Compatible Mechanisms for Public Goods Provision: An Experimental Study. *Journal of Political Economy* 106:3 633–662.

[6] Erev, Ido and A.E. Roth (1998) Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria, *American Economic Review*, 88,4, September, 848-881.

[7] Fernandez, Rachel and Alessandra Fogli (2006) Fertility: The Role of Culture and Family Experience *Journal of the European Economic Association*

[8] rohlich, Norman and Joe A. Oppenheimer, (1996) Experiencing Impartiality to Invoke Fairness in the n-PD. *Public Choice*. 86, 117-135

[9] Grief, Avner (2006) *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade* Cambridge University Press.

[10] Guiso, Luigi, Paola Sapienza and Luigi Zingales (2006) Does Culture Affect Economic Outcomes? *Journal of Economic Perspectives* 20(2) 23–48.

[11] Henrich, Joseph, Robert Boyd, Sam Bowles, Colin Camerer, Herbert Gintis, Richard McElreath, and Ernst Fehr.(2001) In search of Homo Economicus: Experiments in 15 Small-Scale Societies. *American Economic Review*, 91(2): 73–79.

[12] Henrich, Joseph, Robert Boyd, Sam Bowles, Colin Camerer, Herbert Gintis, Richard McElreath, and Ernst Fehr. (2004) *Foundations of Human Sociality*. GET REMAINDER OF CITATION.

[13] Inglehart, Ronald (1990) *Culture Shift in Advanced Industrial Society*, Princeton University Press.

[14] Kahneman, Daniel, and Amos Tversky (1979) Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47 (1979), 263–291

[15] Scott E. Page (2006) Path Dependence, *Quarterly Journal of Political Science*: Vol. 1:No 1, pp 87–115

[16] Roth, A.E., (1990) New Physicians: A Natural Experiment in Market Organization, *Science*, 250, 1524–152