Convergence: An Experimental Study*

Wolf Ze'ev Ehrblatt
New York University

Kyle Hyndman
New York University

Erkut Y. Özbay
New York University

Andrew Schotter
New York University

9th December 2005

Abstract

One way to define a Nash equilibrium is by positing a set of beliefs (or conjectures) for each player over (about) the actions of their opponents that has the property that, given these beliefs, when each player best responds, the actions taken confirm the initial beliefs. This rational expectations definition leaves open the question of how beliefs and actions get into this self-confirming state. For example, do beliefs converge to their equilibrium state first and drag actions into alignment or is the process action driven with them converging before beliefs. What we find is that the process of convergence is one where actions converge before beliefs. However, after reaching equilibrium in actions, the beliefs of subjects converge to the degenerate beliefs that place all the weight on the partner's equilibrium action extremely rapidly (within 2 periods on average). We also identify differences between the early converger and the late converger in a group — often it is the case that the early converger plays his part of the Nash action profile long enough to convince his opponent to adhere. Finally, we investigate the process of belief formation and argue that, unlike all of the most common learning models, the belief formation process is one that takes into account not only the payoff of the learner but also those of his opponents.

Keywords: Game Theory, Belief Formation, Learning, Convergence.
JEL Classification: C70, C91, D83, D84

1 Introduction

One way to define a Nash equilibrium is by positing a set of beliefs (or conjectures) for each player over (about) the actions of their opponents that has the property that, given these beliefs, when

*This paper has benefited from comments from participants at the C.E.S.S. Experimental Design Workshop, the NYU-CREED Graduate Student Workshop in Amsterdam and from the 2005 International Meeting of the Economic Sciences Association (Montréal, PQ). We are also grateful to Guillaume Fréchette and Colin Camerer for valuable suggestions. This research was supported by the Center for Experimental Social Science at NYU and by the National Science Foundation under grant SES-0425118 (Schotter). *Corresponding Author:* Andrew Schotter, Department of Economics, New York University, 269 Mercer St. 7th Floor, New York, NY 10003. E-mail: andrew.schotter@nyu.edu

each player best responds the actions taken confirm the initial beliefs. This rational expectations definition leaves open the question of how beliefs and actions get into this self-confirming state. For example, do beliefs converge to their equilibrium state first and drag actions into alignment or is the process action driven with them converging before beliefs? One thing we have learned in our work is that people do not arrive at the equilibrium play of a game by a process of deductive reasoning but rather by induction, observation and learning as the game is iterated. They converge to it as time passes rather than leaping to it spontaneously after a logical reasoning process. Along the path of this convergence we also know that in some learning models people are forward looking and make conjectures (form beliefs) about what their opponent is likely to do.[1]

In this paper we investigate the answers to the following questions:

1) When people playing a finite strategy matrix game converge to the Nash equilibrium of the game, assuming that the game has a unique pure-strategy equilibrium, do their actions converge first and then drag their beliefs into convergence or do their beliefs converge first and pull in their actions?

2) When the actions of players in these circumstances fail to converge to the Nash equilibrium, is it true that their beliefs converge but the players fail to best respond to them or is it simply that their beliefs fail to enter the best response region where the Nash equilibrium beliefs exist?

3) In the final play of a repeated game that converges to a unique pure-strategy Nash equilibrium do the players typically place all the probability mass on the equilibrium strategy of the other player or are they still unsure of what he or she is going to do? If the probability distribution is degenerate, how long does it take to get there after actions have exhibited the Nash equilibrium pattern?

4) Do the answers to questions 1-3 above change when the game played can be solved by the iterative elimination of dominated strategies? Do games that are dominance solvable converge faster than those that are not and do beliefs or actions get to equilibrium first?

5) Is there a consensus across people on the belief formation process, i.e., if we showed the time series of play of a simple two-person game to a set of people and asked them to state their beliefs about how the players are expected to play, would the belief formation process look similar across the observing agents? In other words, do people tend to form their beliefs in a similar way holding the play of the game they are watching constant?

6) Are existing learning models that do not include the payoff received by *both* players descriptive of how people form beliefs?

The answers to these questions are not readily available in the existing literature. The reason is that we do not have a good description of how people converge to the equilibrium of games even when those equilibria are unambiguously defined, i.e. unique, and easily approachable logically

[1]Of course some learning models, such as reinforcement learning, are totally backward looking and do not involve the use of beliefs. Still most of the commonly used learning models can be reinterpreted as belief learning models as Camerer and Ho (1999) indicate in their description of the EWA model which nest belief learning as a special case.

(e.g., can be arrived at by a process of iterative deletion of dominated strategies).[2]

This paper tries to answer these questions by having laboratory subjects play a set of two two-person 3×3 games for twenty periods while beliefs are elicited from them as they play. Both 3×3 games played have a unique pure strategy equilibrium. In one game the equilibrium is dominance solvable while in the other it is not. We chose these games because we expected play to converge to the pure Nash equilibrium (as opposed to games where the equilibrium is only in mixed strategies) and hence the combination of observing actions and beliefs could be informative about the convergence process. After having 32 pairs play one of these games we selected one pair and their associated time series from the dominance-solvable (DSS) and non-dominance solvable (nDSS) game experiments that either converged or did not (DSS.c, DSS.nc, nDSS.c and nDSS.nc) and brought in new subjects to play a "prediction game" in which they were paid, period by period, to predict the next action in these time series.

What we find is that in answer to Question 1, actions converge faster than beliefs. While we will formally define what we mean by this in the next section of the paper, suffice it to mean what it says, on the way to convergence people do not hold "equilibrium beliefs" before they play equilibrium actions. Rather they seem to need an exhibition of equilibrium play in order to believe that the Nash equilibrium is likely to occur. This exhibition is typically given by one of the players, whom we will call the "early converger." Specifically, the early converger plays his/her part of the Nash equilibrium for several periods even when the Nash action is not a best response to the beliefs held at the time. For this player actions clearly converge before beliefs. The other player, the "late converger," seeing the Nash actions played, alters his or her behavior after a while and conforms to Nash as well but the beliefs of the late converger almost simultaneously converge along with his or her actions. After both players play the Nash, in answer to Question 3, beliefs quickly become degenerate. These patterns do not differ much when we compare dominance solvable and non-dominance solvable games.

One striking finding of our experiment is related to Question 2 and to what occurs when play does not converge. Here it appears that for those games that do not converge beliefs rarely enter into the set of beliefs for which the Nash strategy is a best response. In other words, non-convergence of actions is equivalent to the failure of players to ever hold equilibrium beliefs even if, on occasion, they exhibit equilibrium actions. Non convergence appears to be the result of a failure in beliefs and not in the ability of players to best respond.

In answering Question 5 we take advantage of a unique feature of our experimental design that shows individuals the same time series of actions and elicits their beliefs period by period. This allows us to investigate whether people in general tend to update their beliefs in similar ways. We know of no other experiment in which this is done. What we find is that there is a good deal of consensus in the way people update their beliefs. One feature of this updating is that when

[2]Nyarko and Schotter (2002a) elicit beliefs for the repeated play of games with mixed strategy equilibria but due to the nature of the equilibrium play does not converge to the repeated play of one strategy there.

updating observers appear to take account of the payoffs of *both* the player whose actions she is predicting and the payoffs of his/her opponent. This feature is absent in just about all learning models such as reinforcement, fictitious play, and EWA learning models (see, for example, Erev and Roth (1998) and Camerer and Ho (1999)). Hence it goes a long way to explaining the stylized features of elicited belief time series seen in this paper as well as Nyarko and Schotter (2002a,b) where the elicited beliefs of subjects were very volatile, as opposed to the smooth looking belief vectors expected from previous learning models.[3] Two notable exceptions are the work of Stahl and Wilson (1995) and Camerer *et al* (2002). The finding that there is a consensus in the way people update in these games (or at least how observers update) provides hope that we may one day create a convincing theory of belief formation which is behaviorally rich since without a consensus such a theory would have to be based on psychological characteristics in an individual by individual manner.

In this paper we will proceed as follows. In Section 2 we define more precisely what we mean when we say that either beliefs or actions converge. In Section 3 we describe our experiment design and procedures, while in Section 4 we present the results of our experiments. In Section 5, we argue that there is a consensus in the belief formation process. We also identify many regularities that we believe any reasonable model of belief formation must incorporate. Section 6 concludes the paper. Instructions can be found in Appendices A and B, while figures are collected in Appendix C.

2 Definition of Convergence

Consider a finite strategy two person game $\Gamma = \langle N, A_i, \pi_i \rangle$ where N is the set of players indexed i = 1,2, A_i is the set of actions for each player, and π_i is player i's payoff function which maps $A_1 \times A_2$ into a set of real valued payoffs. Assume that this game is to be played T times and that Γ has a unique pure strategy Nash equilibrium $a^* = (a_1^*, a_2^*), a_1^* \in A_1$, $a_2^* \in A_2$ whose payoffs are not Pareto dominated by any other payoffs in the game. In such a case we would expect that a^* would remain the unique pure strategy equilibrium even for the T times repeated play of Γ and that such an equilibrium would be sub-game perfect in the repeated game. We will call a_i^* player i's Nash action. Let $h_t = [(a_1^1, a_2^1) \ldots (a_1^t, a_2^t)]$ be a history of actions in the play of Γ over the first t periods of play.

Let $T_i^a := \{t' \leq T - 2 \,|\, a_i^t = a_i^* \quad \forall t' \leq t \leq T\}$. Now we can define the notion of convergence in actions to Nash equilibrium at period t.

Definition 1. Player i *converges in actions to Nash equilibrium in period* $\bar{t}$ if $T_i^a \neq \emptyset$ and $\bar{t} = \min_{t' \in T_i^a} t'$. Such $\bar{t}$ will be called *the convergence period of actions.* If $T_i^a = \emptyset$ then we will say that player i does not converge in actions to Nash equilibrium.[4]

In other words, player i converges to Nash equilibrium in period $\bar{t}$ if this is the earliest period in

[3] See Nyarko and Schotter (2002b) for an elaboration of this point.
[4] We will sometimes omit the words "to Nash equilibrium."

which player i has played her Nash action and continued to play her Nash action from that period until the end of the game. If there is no such period, it means that the player did not play her Nash action consistently up to period T and thus did not converge. Notice that we put an upper bound on the convergence period, $T-2$, since we do not want to consider a player as converged in period T if that player chose her actions randomly and the action chosen in period T just happened to be her Nash action. While using the upper bound of $T-1$ would have been sufficient, we have decided to be more conservative in our definition of convergence in actions and defined the upper bound as $T-2$.

We will say that a game has *converged in actions to Nash equilibrium* if there are $\bar{t_1}, \bar{t_2}$ such that player 1 has converged in actions in period $\bar{t_1}$ and player 2 has converged in actions at period $\bar{t_2}$.

If player j has K actions then define the K dimensional simplex $\Delta_i = [0,1]^K$ for player i as her beliefs simplex where $b_i^t \in \Delta_i$ defines a K dimensional belief vector at time t for player i; that is, $\sum_{k=1}^{K} b_{i,k}^t = 1$. $B_i \subseteq \Delta_i$ is that subset of beliefs for player i for which her best response to any vector of beliefs in B_i is to play her part in the Nash equilibrium, i.e., $a_i^* = \underset{a_i \in A_i}{\text{argmax}} \sum_k b_{i,k} \pi_i(a_i, a_j^k) \quad \forall b_i \in B_i$. Let $\vec{b_t} = [(b_1^1, b_2^1)....(b_1^t, b_2^t)]$ be a belief history defining the beliefs held by the two players at each point in the history of the game up to period t.

We will define convergence in beliefs similar to the way we defined convergence in actions. First define $T_i^b := \{t' \leq T-1 \,|\, b_i^t \in B_i \quad \forall t' \leq t \leq T\}$. Next define convergence in beliefs.

Definition 2. Player i *converges in beliefs to Nash equilibrium in period* $\tilde{t}$ if $T_i^b \neq \emptyset$ and $\tilde{t} = \min_{t' \in T_i^b} t'$. Such $\tilde{t}$ will be called *the convergence period of beliefs.* If $T_i^b = \emptyset$ then we will say that player i does not converge in beliefs to Nash equilibrium.

A game will be said to have converged in beliefs if there are $\tilde{t_1}, \tilde{t_2}$ such that player 1 has converged in beliefs in period $\tilde{t_1}$ and player 2 has converged in beliefs at period $\tilde{t_2}$.

3 Experimental Design and Procedures

In order to answer the questions posed in the beginning of this paper, we conducted two experiments — the first of which we call the AB experiment (actions and beliefs) and the second, the B experiment (beliefs only). Both experiments were conducted in the laboratory of the Center for Experimental Social Science at New York University. All the participants were NYU undergraduate students recruited via e-mail. Participants received a \$7 show-up fee in addition to the gains from the experiment. Each experiment session took about 1.5 hours to complete. The number of participants in the AB experiment and the B experiment were 64 and 53, respectively, and the average payoff in each experiment was \$19.7 and \$20.9. All the sessions were run using the experimental program z-Tree by Urs Fischbacher (1999).

3.1 The AB Experiment

In the AB experiment subjects played the two 3×3 games shown in Figure 1. Note that each game has a unique pure-strategy equilibrium but that the game in Figure (1.a) is dominance solvable while the game in Figure (1.b) is not dominance solvable. Each game was played for 20 periods with a fixed partner but partners were randomly switched when the game changed. Also, the role of each player was randomly determined. There were four sessions in the AB treatment with the order of the games presented to subjects changing in each one.

The games chosen had the following features:

- A unique Nash equilibrium in pure strategies in the stage game.
- Nash payoffs are on the Pareto frontier.
- Payoffs in the Nash equilibrium were not symmetric.

These properties insured that there was a unique sub-game perfect equilibrium to the 20-period repeated game the subjects played in the lab in which they play the stage-game Nash equilibrium in each period.

At each period, the subjects were asked to make 2 decisions. The first was to choose the action for that period. The second was to state their beliefs regarding their partner's action in that period.[5] The action decision was rewarded according to the relevant game matrix, while the beliefs predictions were rewarded using the Quadratic Scoring Rule (QSR).[6] All payoffs from the action choices and the belief predictions were then summed up to give subjects their final payoff.

3.2 The B Experiment

The AB experiment produced a set of action choices, some converging to the Nash equilibria while others not. In the B experiment new subjects were recruited and brought into the lab. In front of the room there was a screen upon which the period-by-period play of one pair of subjects who played the game in the AB treatment was projected. In other words, we took the time series of actions of a pair in the AB experiment and played it out period by period. In the instructions, the subjects were informed that the games they were about to see were played in the past by NYU undergraduates so that ambiguity regarding the population will be eliminated. The subjects were

[5]While the game subjects face is a repeated one, the beliefs elicited here are only for one period.

[6]Under the assumption the subjects are risk neutral, the use of the QSR should make the subjects state their true beliefs regarding their opponent's action. Sonnemans and Offerman (2001) find that the QSR is incentive compatible and that subjects tend to report their true beliefs when the QSR is used. Nyarko and Schotter (2002a) also use a quadratic scoring rule and offer substantial evidence that subjects best respond to the beliefs they state. Moreover, Wilcox and Feltovich (2000) report that belief elicitation does not always affect subjects' behaviour. However, Rutström and Wilcox (2004) argue that the act of solicitation may focus the attention of the subject on his or her beliefs in a way that may be unnatural.

shown the time series of 2 games (DSS and nDSS). Their task was to predict the actions of one of the players in this game for 20 periods as the actions in the time series were revealed to them period by period. Predictions were rewarded with the same QSR that was used in the AB experiment.

Note that in this experiment subjects do not play a game but are spectators who were asked to make predictions, period by period, about the actions of one of the players whose behavior they were observing. The nice feature of the experiment was that since all subjects observed the same time series we are able to hold the actions observed constant and study the belief formation process of subjects observing the same set of actions. In all other belief elicitation experiments that we know of, subject beliefs are elicited pair by pair so that the observed actions are not controlled. In our design, the actions observed by all subjects are held constant so we can study the belief formation process in isolation and the consensus (if any) of observing subjects about beliefs.

The only other experiments we know of where spectators were used were those of Offerman *et al* (1996) and Huck and Weizsäcker (2002). In the former experiment, spectators were matched to an actual player in a public goods experiment and asked to state beliefs on the contributions of the *other* group members. In the latter experiment, subjects were asked to predict a second group's choice frequencies in a set of lottery-choice tasks. Our design differs from both of these settings by having many spectators view the same choice path and predicting the same player behavior rather than having one spectator be attached to one player.

4 Results

In this section and the next we report the results of our experiments by answering the six research questions stated in the introduction.

Question 1: When people playing a finite strategy matrix game converge to the Nash equilibrium of the game, assuming that the game has a unique pure-strategy equilibrium, do their actions converge first and then drag their beliefs into convergence or do their beliefs converge first and pull in their actions?

Despite the fact that each game our subjects played had a unique pure strategy equilibrium, there was a significant failure on the part of subjects to reach it. Using the definition of convergence in Section 2, we categorized each pair in each game as either converging or non-converging. In the dominance-solvable game only 17 of 32 pairs converged, while in the non-dominance solvable game 16 of 32 pairs converged.

Regarding the relative speed of convergence of beliefs and actions, Table 1 provides clear evidence that actions converged before beliefs in both the dominance solvable and the non-dominance solvable games. For example while, on average, in the dominance solvable game it took 5 periods for players to reach the Nash equilibrium action, their beliefs did not converge until 7.7 periods.

For non-dominance solvable games the comparable numbers are 7 and 8.7.[7] The results, in the

Table 1: Summary Statistics: Convergent Pairs

	DSS	nDSS
Number of pairs	17 (of 32)	16 (of 32)
Mean period of convergence in actions	5.0	7.0
Mean period of convergence in beliefs	7.7	8.7

column labeled "All" in Table 2, of the paired t-test clearly show that beliefs converge after actions for both game types.[8] However, note that looking at the data at this level of aggregation masks an important distinction between so-called *early convergers* and *late convergers* that we discuss below.

Table 2: Results of the Paired t-test (H_0: Beliefs and Actions Converge Simultaneously)

	DSS			nDSS		
	All	Early	Late	All	Early	Late
Mean: Conv Period of Beliefs minus Conv Period of Actions	2.71	5.70	0.69	1.72	3.43	-0.14
t-statistic	4.03	4.54	1.06	3.75	5.16	-0.46
p-value	<0.001	<0.001	0.16	<0.001	< 0.001	0.67
Number of Observations	34	13	13	32	14	14

"Early" uses only those subjects whose actions were first to converge (in a group), while "Late" uses only those subjects whose actions converged second (in a group). In the case of simultaneous convergence, the pair was excluded.

It appears that the process of reaching equilibrium is an action-led process. Pairs do not seem to reach equilibrium by first having their beliefs reach the Nash best response set and then replying appropriately. Rather subjects seem to need proof that their opponent knows what an equilibrium is before they will choose the equilibrium action themselves. Only after this do their beliefs converge. This creates a problem for convergence, however, since if each player is waiting to observe their opponent play the equilibrium action before they do, it would seem as if it would be difficult, if not impossible, to converge. This standoff can only be resolved if one player leads the way and acts first. We call such players "early convergers" (teachers).

To illustrate the difference between early and late convergers consider Figure 2. For both early and late convergers, this figure plots the histogram of the difference between the action convergence period and the beliefs convergence period. A very clear pattern emerges: early convergers' actions almost always converge before beliefs (in fact there are no negative differences), while for late convergers two things are noticeable. First, the difference between when the actions and beliefs converge is much smaller for these subjects. For example, in the DSS game, while the mean

[7] In Table 1 we allowed for final period deviations so long as a clear pattern of convergence emerged beforehand. For example, we labeled a subject as converging if from period 10 to 19 the subject played his/her Nash strategy but defected in period 20.

[8] Results are the same for the Wilcoxon Signed-Rank test and are not reported here.

difference between when actions and beliefs converged for the early convergers was 5.7 periods it was only 0.69 periods for late convergers; a similarly stark difference arises is the nDSS game as well. This is not a surprise since, often times, the early convergers are leading the way and waiting for their opponent to converge. While they are waiting their beliefs are outside of the Nash best response set. Second, while for early convergers in either of our games beliefs always converged after actions, this was not true for 7 of 16 followers in the nDSS game and 2 of 17 in the DSS game. The results of a paired t-test in the columns "Early" and "Late" in Table 2 lend further support to our claim that early convergers' beliefs converge after their actions, while late convergers' actions and beliefs converge almost simultaneously.

This suggests that teaching and learning happen in the games that converged. One player understands what the equilibrium is in the game and plays it for the rest of the session. The early converger sticks with her action even though it is not a best response to her beliefs in order to influence her partner's beliefs, assuming that he will eventually choose the Nash action. Convergence in actions takes some time for the early converger's opponent and thus her beliefs converge after the actions. Meanwhile, the late converger sees that his opponent has chosen the same action for consecutive periods and the realization of the Nash equilibrium occurs at the same time as the convergence of the beliefs. Thus beliefs and actions converge in, roughly, the same period.

Remark 1. What we have just discussed may be loosely thought of as *successful teaching*. That is, one of the players saw the Nash equilibrium, began to play it, despite it not being a best response to his/her beliefs, until eventually his/her opponent also played the Nash action and the game converged. However, there are also many instances of unsuccessful *teaching episodes*. For example, approximately 30% of the players not part of a pair that converged to Nash equilibrium actually played their Nash equilibrium strategy for three or more consecutive periods. Moreover, it was usually the case that their beliefs lied outside of the Nash Best-Response region — perhaps indicating a desire to *teach* the other player the Nash equilibrium. Of course, what makes such episodes unsuccessful is the fact that her opponent did not choose his Nash action. Eventually, this *teacher* simply gave up on playing her Nash strategy and the game did not converge to Nash equilibrium.

4.1 Early Convergers: A Digression

From our discussion above, there appear to be substantial differences between early and late convergers. We would like to know why some players converge on their Nash action relatively early and others converge on their Nash action relatively late. One may conjecture that the strategic role you play (i.e. Row or Column) is a causal factor in determining whether you are likely to be an early converger. Beyond this conjecture, one might ask, for dominance solvable games, does the early converger tend to be the player with fewer steps of iterated elimination of strategies? (The idea here being that he or she may be more able to see where the eventual equilibrium is.) Finally, how many periods are their between the convergence of the early and late convergers?

The answers to these questions are not very informative. First, we find that half of the early convergers are column players and the other half are row players. In other words, your strategic role in the game has little to do with whether you converge first or second. This is true for both the DSS and nDSS games. This also answers our second question since if both row and column player are equally likely to be converge first, then the number of iterated steps of elimination they face can not be a factor. Finally, it appears as if late convergers recognise the Nash equilibrium quite quickly after his opponent has converged on her Nash action: on average, late convergers play the Nash equilibrium 2.8 periods later.[9]

This discussion leaves us with a very unclear idea what differentiates between those that converge early and those that converge late. However, we offer the following conjecture which is that it is the subject whose payoff is below his or her expectations that ultimately converges first to his/her Nash action. More precisely, say that you and I are playing a game and each period, given your expectations about me, you are pleasantly surprised by my actions in that your payoff often exceeds your expected payoff while just the opposite it true for me. In such a case we might expect that the person who is getting the short end of the expectational stick will look more closely at the game and try to lead her opponent to the Nash equilibrium where, in both of the games we used, each player does rather well. In this sense, we may view the early converger as a *teacher*.

To investigate this conjecture we calculate the ratio of the actual payoffs (AP) that players received to their expected payoffs (EP), given their elicited beliefs, in the periods before convergence (or the periods before the early converger played Nash); we denote this ratio by $\frac{\mathrm{AP}}{\mathrm{EP}}$. The motivation for this exercise is as follows. If a player's actual payoff is lower than her expected payoff, then she may devote more attention to *learn* about the game. Note, however, that an important difference between dominance solvable and non-dominance solvable games may arise. In the former, a player may learn that one or more strategies is dominated, and this may help her find her Nash strategy. However, in a non-dominance solvable game, she may simply learn how to best respond to her beliefs (or to her opponent's actions), which will not necessarily lead her to her Nash strategy. In this way, we may expect that the $\frac{AP}{EP}$ ratio will be significantly lower for early convergers only in the dominance solvable game. This ratio was calculated for each subject for each period up to the period in which the subject had converged in actions. We have 18 sets of useable observations, since in 9 cases convergence was in the initial period (for the early converger). In Table 3 we report the average of the means of the $\frac{\mathrm{AP}}{\mathrm{EP}}$ ratio for early and late convergers.[10] As can be seen, for dominance solvable games, the ratio is substantially higher for late than for early convergers, as is our conjecture. Indeed, in Table 4, which reports the results of the Wilcoxon signed rank test, one can see that this difference is statistically significant at the 1% level. Consistent with our intuition,

[9]Of course, there is a selection issue here since we only count those groups that actually converged. In particular, there were instances of unsuccessful teaching episodes in which a *potential* late converger *never* played his Nash action (see Remark 1).

[10]The calculation of the average of means of the $\frac{\mathrm{AP}}{\mathrm{EP}}$ ratio was done as follows: First, the ratio was calculated for each subject for each period prior to convergence. Second, the mean ratio was computed. Finally, the average over all the players' means is reported in the table.

the same result does not hold when we examine non-dominance solvable games. We see from Table 4 that there is no significant difference between early and later convergers.[11] Thus, for dominance solvable games, players whose actual payoffs did not match their expectations, appeared motivated to learn about the game, leading them to converge first.

Table 3: The $\frac{\text{AP}}{\text{EP}}$ Ratio for Early & Late Convergers

	DSS	nDSS	All
Mean $\frac{\text{AP}}{\text{EP}}$ Ratio for Early Convergers	0.66 (9)	0.93 (9)	0.80 (18)
Mean $\frac{\text{AP}}{\text{EP}}$ Ratio for Late Convergers	1.14 (13)	0.89 (14)	0.95 (27)

The number in brackets denotes the number of observations over which the mean was taken.

Table 4: Wilcoxon Signed-Rank Test (H_0: $\frac{\text{AP}}{\text{EP}}_{early} = \frac{\text{AP}}{\text{EP}}_{late}$)

	Up to own convergence period			Up to early convergence period		
	DSS	nDSS	All	DSS	nDSS	All
Mean $\frac{\text{AP}}{\text{EP}}$ Ratio for Early Convergers	0.66	0.93	0.80	0.66	0.93	0.80
Mean $\frac{\text{AP}}{\text{EP}}$ Ratio for Late Convergers	0.98	0.85	0.91	0.88	0.88	0.88
z-statistic	-2.67	1.01	-1.68	-2.07	0.53	-1.29
p-value	<0.01	0.31	0.09	0.04	0.59	0.20

A positive z-statistic implies that the ratio for early convergers is higher than for late convergers, while a negative statistic implies the converse.

Lastly, an interesting pattern arises when comparing the AP/EP ratio of non-converging games to converging ones. The mean $\frac{\text{AP}}{\text{EP}}$ ratio of converging games is 0.916 and the corresponding ratio for non-converging games is 0.983. A Wilcoxon-Mann-Whitney test shows that the null hypothesis of equality in these ratios can be rejected in favor of the alternative hypothesis that the ratio for the non-converging games is higher than for the converging games (z-statistic=2.613, p-value < 0.01).[12] Thus, given the relatively high $\frac{\text{AP}}{\text{EP}}$ ratios, it may explain why these groups did not reach Nash equilibrium — they had little incentive to learn about the game and seek out the Nash equilibrium.

4.2 Comparison of Converging and Non-Converging Games

As we stated above, not all pairs of subjects converged in their play of our experimental games. This was a little surprising since our expectation was that with a 20 period horizon and only one pure strategy equilibrium, each pair would ultimately be able to find their way to it. Luckily they did not since this affords us the opportunity to investigate the non-convergence phenomenon. This motivates our second question:

[11] For late convergers, the entries in Tables 3 and 4 do not match since in the latter we exclude those pairs for which the early converger converged in the first period.

[12] Here there is no significant difference between dominance and non-dominance solvable games.

Question 2: When the actions of players fail to converge to the Nash equilibrium, is it true that their beliefs converge but the players fail to best respond to them or is it simply that their beliefs fail to enter the best response region where the Nash equilibrium beliefs exist?

Before we answer question 2, we investigate whether convergence and non-convergence is heavily path dependent by trying to discover whether the way the game starts (*i.e.,* the subjects period 1 beliefs and actions) determines how they will end up. Table 5 reports Kolmogorov-Smirnov tests on the beliefs and action distributions in the first period between converging and non-converging games. The results in the table show that the difference between converging games and non-converging games does not stem from the initial choice of actions or the initial beliefs since the distribution of those are not significantly different across games and roles when comparing converging and non-converging pairs.

Table 5: Generalized Kolmogorov-Smirnov Test:

	Difference Between Converging and Non-Converging Players			
	DSS		nDSS	
	Row Players	Column Players	Row Players	Column Players
Test Statistic: First Period Beliefs	0.163 (0.983)	0.349 (0.277)	0.375 (0.198)	0.234 (0.763)
Test Statistic: First Period Actions	0.222 (0.778)	0.413 (0.100)	0.250 (0.633)	0.188 (0.912)

p-values are in parentheses.

Since the initial period beliefs or actions does not determine if a game will converge or not, we need to look for other explanations. It is our claim that failure to converge is a result of a failure of the subjects to ever hold beliefs in the Nash Best-Response set. This is a striking result. To illustrate what we mean, consider Figures 3 and 4 where we present the simplex of beliefs and the time paths of beliefs of four subjects whose play failed to converge in either the DSS or nDSS game. In each figure, the point (0,0) represents the case in which a player holds degenerate beliefs that her opponent will play his Nash strategy. The beliefs on the two non-Nash actions are then given by a point in the (x, y) plane and the area enclosed by the dashed line represents the Nash Best-Response set. That is, if beliefs lie inside this set, it is a best response for the player to choose *her* Nash action. Finally, the points labeled S and F depict where the subject's beliefs started and finished in the simplex; when it is not clear, an arrow points to the direction in which the subject updated his/her beliefs. What is obvious is that for these players, who are very representative of all players who failed to converge (as we will see later), throughout the entire game their beliefs almost never entered into the Nash Best-Response set. In other words, if subjects are capable of best responding and best respond to the beliefs we elicited, then this is clear proof that failure to converge is a result of a failure in beliefs not in actions.

To contrast this to representative convergent pairs, consider Figures 5 and 6 where we show a typical sequence of beliefs for players whose actions converged. Here, after some initial periods outside the Nash best-response set, the beliefs of the subjects enter the set and very shortly become degenerate. To give a more aggregate picture of the beliefs of non-convergent subjects (beyond the four presented above) consider Figures 7 — 10. In these graphs the probability simplex is represented and the best response region (BR region) is drawn. The graphs plotted all the beliefs of all the players for each game, role and classification (convergent or non-convergent). It is obvious that in the non-converging games, it was rare that the beliefs were inside the BR region (see also Table 6). For example, in the nDSS game where play did not converge there were only 80 out of 640 instances where beliefs fell in the Nash Best Response region, while in the nDSS game where convergence occurred, the numbers were 435 out of 640. For the DSS game the numbers were 112 out of 600 for the non-convergent games and 462 out of 680 for the convergent games.

Table 6: Average Number of Periods With Beliefs Inside the Best-Response Region

	DSS	nDSS
Mean # of Periods in BR Region Players that converged to Nash	13.588	13.594
Mean # of Periods in BR Region Players that did not converge to Nash	$3.73\bar{3}$	2.500

Finally, further support for our explanation of non-convergence can be found by calculating the probability of staying in the BR region conditional on the beliefs being in the BR region in previous period - $P(b_i^t \in BR_i | b_i^{t-1} \in BR_i)$. For the non-converging games this probability is 37.7% while for the converging games this probability is 76.7%.[13] These probabilities are significantly different from each other.[14] The difference in probability implies that in order to converge to equilibrium, the players had to "believe" in equilibrium; *i.e.,* convergence had to supported by stable beliefs that one's opponent will choose her Nash action in coming periods.

The fact that the beliefs of convergent and non-convergent pairs are very different is not the entire story. One might argue that subjects in convergent pairs were simply better at best responding; however, claim that this is not the case. To demonstrate that there was no differential ability of convergent pairs to best respond we could simply compare the percentage of time convergence pairs best respond to that same percentage for non-convergent pairs. However, this would bias the comparison in favor of the convergent pairs since, by definition, once they converge they best respond while non-convergent pairs never converge. In order to control for the effect of the convergence, we calculate the percentage of best response for the converging games excluding periods in which the games converged in actions or beliefs (which ever came first) and compare this to the

[13]For the converging games we report a more conservative estimate - $P(b_i^t \in BR_i | b_i^{t-1} \in BR_i \quad and \quad b_i^{t-2} \notin BR_i)$, i.e. the probability conditional on the beliefs in previous period being in the BR region *and* that two periods prior the beliefs were *outside* the BR region. We do so in order to eliminate the effect that convergence in beliefs has on the probability. If we calculate the probability in the same way we calculated the probability for the non-converging games, it would be 96.1%.

[14]The t-statistic for the two sample proportion test is -6.410 ($p - value < 0.001$).

percentage of best responses for non-convergent pairs. Doing this we see that subjects best respond 51.6% and 51.9% of the time in the DSS and the nDSS converging games and 52.2% and 60% of the time in the non-convergent games. We cannot reject the hypothesis that these proportions are identical at the 5% level.

These results are summarized in Table 7 which reports the empirical frequencies with which the subjects have chosen the best-response action to their stated beliefs. For convergent games, the first number uses the entire sample, while the bracketed number uses only those periods up to convergence. Table 7 clearly shows that even when subjects did not converge to the Nash equilibrium, they played the best response to their belief at least half of the time. When not playing the best response action, players were more likely to choose the second best option. This leads us to the conclusion that games did not converge because players' beliefs fail to enter the BR region and not because non-convergent subjects were relatively incapable of best responding.

Table 7: Frequencies of Best-Response Behavior (%)

	DSS		nDSS	
	Convergent	Non-convergent	Convergent	Non-convergent
Action was a best-response	76.2 (53.8)	52.2	75.9 (51.9)	60.0
Action was a 2^{nd} best-response	18.1 (26.9)	29.0	17.0 (25.8)	25.8
Action was a 3^{rd} best-response	5.7 (19.2)	18.8	7.1 (12.6)	14.2

The number in parentheses is the empirical frequency using data only up to the period of convergence.

Question 3: In the final play of a repeated game that converges to a unique pure-strategy Nash equilibrium do the players typically place all the probability mass on the equilibrium strategy of the other player or are they typically still a little unsure of what he or she is going to do? If the probability distribution is degenerate, how long does it take to get there after actions have exhibited the Nash equilibrium pattern?

One of the most striking results regarding the convergence of beliefs is the speed with which beliefs become degenerate on the Nash action; *i.e.,* the belief that the opponent will play his Nash action with 100% probability, from the moment the opponent converges to his Nash action (see Figure 11). The average number of periods is 2.6 while the median is 2 periods. This fact demonstrates, perhaps, that subjects know what the equilibrium of the game is before they play it and only need assurance from their opponent that he or she understands as well. Once that assurance is received (*i.e.,* once the Nash play is observed) beliefs quickly converge.

Question 4: Do the answers to questions 1 — 3 above change when the game played can be solved by the iterative elimination of dominated strategies? Do games that are dominance solvable converge faster than those that are not and do beliefs or actions get to equilibrium first?

In the previous section, we combined the data from all the experiments and reported the differences between convergent and non convergent games. However, the two games were different in their structure. Dominance solvable games have a "natural" way to reach the equilibrium, through elimination of dominated strategies, while in non-dominance solvable games do not. In spite of this conceptual difference, we find few empirical differences between DSS and nDSS games — either in the way they are played or in the beliefs formed by the subjects playing them. For example, as we have seen before, when subjects fail to converge, their beliefs fail to enter the Nash best-response set and this is true no matter what type of game they are playing. Once convergence in actions occurs for both players, then in both types of games beliefs quickly become degenerate. One difference that does exist is that DSS games seem to converge faster than nDSS games. This can be seen in Table 8 where we see the mean period of convergence for early convergers is 3.78 and 6.83 for DSS and nDSS games, respectively, and for late convergers it is 5.75 and 8.31 for DSS and nDSS games. Despite these differences in means, the results of the Kolmogorov-Smirnov test comparing the distributions of convergence periods between the game types, also reported in Table 8, shows no significant difference in the distribution of convergence period when considering teacher and followers separately.

Table 8: Kolmogorov-Smirnov Test: Period of Convergence

	Mean Period of Convergence		K-S Test Diff. b/w DSS & nDSS
	DSS	nDSS	
Early Convergers	3.78	5.75	0.2708 (0.495)
Late Convergers	6.83	8.31	0.3542 (0.188)

The first number in the third column reports the test statistic while the second number in parenthesis reports the p-value.

5 Consensus in Beliefs

Question 5: Is there a consensus across people on the belief formation process; *i.e.,* if we showed the time series of play of a simple two-person game to a set of people and asked them to state their beliefs about how the players are expected to play, would the belief formation process look similar across the observing agents? In other words, do people tend to form their beliefs in a similar way holding the play of the game they are watching constant?

Economists have paid considerable attention to the process of belief formation (see, *e.g.*, Camerer 2003, Ch. 6). A successful model of belief formation, however, would have to be fairly universal and employed by a wide variety of people since if there were excessive heterogeneity across people about the way they form their beliefs, attempting to build one general model of this process would be a vacuous exercise.

To answer this question we use the data generated by our Belief-Only experiment where we show subjects the time series of four particular games played by four pairs in our Belief-and-

Actions experiment. The beliefs we measure are then those of the observers who watch these time series as they unfold and offer their best guess as to what the player being observed will do in the next period. As we have said before, this design has the advantage of holding the time series of actions constant across observers and seeing whether they respond in similar ways to what they are observing. In other words, we are looking for a consensus in the belief formation process.

This question is very important. If it turns out that the degree of consensus is very minimal, it suggests that there is little hope in writing one general belief-formation model. Therefore, we view a positive answer to the first question as a necessary hurdle to pass before addressing the second question.

5.1 Consensus Indices

There are many ways in which one may try to uncover *consensus* in beliefs. In this section we propose two different measures whose purpose is to capture two different aspects of the consensus problem. We call these two consensus measures the "dynamic" or "directional" measure and the "static" measure, and we will attempt to describe them below. To explain what we mean by these consensus measures let us make an analogy between a flock of birds migrating and a set of beliefs. Each bird at any point in time is characterized by a point in some three dimensional space just as the beliefs of our observers are a simplex. As the birds fly they are subject to various shocks (bolts of lightning, claps of thunder, *etc.*) just as our observers see various outcomes of the game they are watching. One question that we can ask is do the birds respond to these shocks in a similar manner, i.e., do they move away or towards them similarly. This is analogous to what we will call our dynamic belief consensus measure since it asks how birds (subjects) alter their position (change their beliefs) in response to a shock (observing an outcome). So this is consensus about how to change one's position.

Alternatively, one can view consensus as a measure of how closely packed the birds (beliefs) are. The less space between the birds (the more closely packed beliefs are) the more consensus exists. This is our static consensus measure which we will expand on below. Note that each measure measures a different aspect of the consensus problem both of which are important for descriptive purposes. Note that it is possible for a set of beliefs to show considerable consensus on one of these measures but little on the other. They measure different things.

5.1.1 Dynamic or Directional Consensus Measures

As stated above, one way to measure whether people form beliefs in a similar manner is to ask and measure whether they respond to information in a similar way. That is, do observers' beliefs move in a similar direction and/or magnitude between periods $t-1$ and t after seeing the outcome of play in round $t-1$? Since we are interested in how subjects change their beliefs here and not in the levels, this is an exercise in directional learning and the consensus across people about it (see, *e.g.*, Selten and Stoecker 1986, Selten and Buchta 1999). Subjects may have completely different priors about what action is likely to have been chosen, but conditional on the observed history, it is reasonable to imagine that most subjects may update their beliefs in the same direction. To get an idea of whether or not this is true, we posit a consensus index at each time period t.

Suppose we have the beliefs of N players. Define $B_j^i(t)$ denote the belief of observer i for strategy j at time t and let $\Delta B_j^i(t) = B_j^i(t) - B_j^i(t-1)$ for $t = 2, \ldots, 20$. Finally define the variable:

$$D_j^i(t) = \begin{cases} 1, & \Delta B_j^i(t) > 0 \\ 0, & \Delta B_j^i(t) = 0 \\ -1, & \Delta B_j^i(t) < 0 \end{cases} \tag{1}$$

That is $D_j^i(t)$ takes value 1 if observer i increased the weight on strategy j from time $t-1$ to time t, -1 if the weight was decreased and zero if the weight was unchanged.

We will define the following notion of consensus for all players i and $t = 2, \ldots, 20$:

$$C^i(t) = \frac{1}{12(N-1)} \sum_{k=1}^{N} \sum_{j=1}^{3} (D_j^i(t) - D_j^k(t))^2 \tag{2}$$

Therefore, for each observer i we take the squared difference between the directional change in i's beliefs and the directional change in player k's beliefs and sum over all players k and all strategies. The normalization $\frac{1}{12(N-1)}$ ensures that $C^i(t) \in [0, 1]$. Finally then, our consensus index is simply the average over the indices for all players:

$$C(t) = \frac{1}{N} \sum_{i=i}^{N} C^i(t) \tag{3}$$

The indices are plotted in Figure 12 for each of the four games in our beliefs only experiment. The solid line hovering around 0.5 can be thought of as a benchmark for the amount of consensus. Specifically, it measures the degree of consensus that would arise if all subjects formed beliefs in an i.i.d. fashion in each period.[15] The dotted lines denote the 95% confidence interval. In all games, the consensus indices are well below the lower confidence band. This indicates to us, at the very least, that beliefs are not all random and there is some degree of consensus in how beliefs are updated.

Look now at the indices for the specific games. A couple of points can be made. First, until around period 8, the indices for the four different games look very similar. It is only after period 8 when there appears to be a growing consensus for the game DSS, early convergence; however, this is not unexpected, since the actual players in this game converge to the Nash equilibrium in period 6. Indeed, the beliefs of almost all observers quickly converged to degenerate beliefs on

[15]Specifically, we computed the consensus index for each of 500 replications in which 20 subjects formed beliefs in an i.i.d. fashion for 20 periods and then averaged over all replications.

the Nash strategy after which there was no change in their belief vectors. Second, for the other games, two of which did not converge to Nash equilibrium and one of which converged in period 16, there is nothing distinguishable in the indices. Obviously, since there was no, or at least late, convergence there was less consensus shown than in the DSS early convergence game. Still, the amount of dynamic consensus shown was significantly less than what we would have seen if beliefs were updated randomly.

5.1.2 Static Consensus

Another way to measure consensus and whether people respond to outcomes in a similar fashion is to measure how closely packed their beliefs are and how the diversity of their beliefs changes as they observe the history of the game they are either playing or watching. Obviously, we might expect more consensus amongst people when they share similar beliefs so that the less variance in their beliefs at a point in time the more consensus exists amongst them.

To measure this type of static consensus we used a simple Euclidean metric. At any round t we simply calculated the mean belief of the set of observers watching the game and calculated the mean Euclidean distance of the observers from the mean belief. The smaller this mean distance the greater the consensus. Also, if in round $t-1$ an action was observed that caused this mean distance to decrease between period $t-1$ and t, then the outcome observed was consensus increasing (or consensus building). This is a static measure because it simply looks at the level of beliefs and measures their closeness and is defined statically on beliefs in each round rather than dynamically using data from the change in beliefs across two adjacent rounds.

The results of our measure of static consensus are presented in Figure 13(a - d). The solid line is the mean (Euclidean) distance of observer beliefs from the group mean, while the dashed line is distance between the belief of the person who actually played these games and the group mean. As we can see, except in the early convergence game (panel (b)) observing the time series of actions over the 20 periods of the game was not a consensus building experience. In fact, in all cases except for the game that converged early, consensus was typically higher in round 1 than in almost all later rounds indicating that subjects tended to be confused by the actions they saw played during the game.[16] This does not contradict our previous finding that a fair amount of dynamic consensus existed since that measure simply says that people reacted to outcomes in the game by moving generally in the same direction while our static measure says that when they did so it did not imply that their beliefs got closer together.

One interesting feature of Figure 13 is how well the mean belief of the group of observers tracks the belief of the player who actually played the game. More precisely, the distance between that player's belief and the mean of the group is always less than the mean distance of the members of the group to the group mean. This means that using the mean belief of the observing group is a

[16] The horizontal line has an intercept at the consensus measure in the first period. Therefore, above (below) this line, consensus is worse (better) than in the first period.

good statistic for the belief of the player playing the game.

This result is interesting for several reasons. First it gives us some confidence that the beliefs we elicited are meaningful since, despite the variance in them, their mean seems to track the beliefs of the person playing the game rather well. Second, it yields some hope for the use of focus groups or other survey methods whose aim is to measure whether consumers are interested in a particular product or simply their confidence about the economy, *etc.* This follows since the mean of solicited opinions tracks the mean of the population being studied; therefore such surveyed beliefs may be a useful (and cheap) input into macro policy making.

5.2 Belief Updating: Learning Models

Question 6: Are existing learning models that do not include the payoff received by *both* players descriptive of how people form beliefs?

Figure 12 tells us, at a minimum, that changes in beliefs are *not* random. In particular, subjects, observing a common history, tend to increase their beliefs in the same direction. However, it does not inform us about the direction in which beliefs move, nor what determines any changes in beliefs. One way to make such predictions, however, is to subscribe to one of the often used learning models that economists have become fond of. For example, if one were to subscribe to a noisy fictitious play model employed by Fudenberg and Levine (1998), then at each point in time one would assume that the probability that a player chooses an action at period t is equal to the fraction of times she chooses that strategy up until period $t-1$. After forming beliefs in this manner, the player would be assumed to noisily best respond. If one subscribed to a reinforcement-type model (Erev and Roth (1998), then the choice of a strategy today would be proportional to how heavily one was reinforced or rewarded for that choice in the past. If one were an EWA-type learner (see Camerer and Ho (1999), then how attractive a strategy is for a player, (which determines how likely that person is to choose that strategy) will depend on how she was rewarded for choosing it as well as the counter-factual payoff she would have received by choosing an alternative strategy. Finally, if one were a follower of Cheung and Friedman (1997) then, instead of giving equal weights to all past choices of your opponent, you would use geometrically declining γ-weights, and weight recent history more heavily.

The point of interest, however, is that all of these belief formation processes give no weight to the payoff history of one's opponent. For example, reinforcement learning is totally egocentric in that it cares only about how heavily I am reinforced for taking certain actions without giving any consideration to how my opponent has done. EWA, while allowing counter-factuals, also ignore how my opponent does as does all of the other models mentioned. The only models we know of that takes one's opponent's payoffs into account is the strategic teaching model of Camerer *et al* (2002) and Stahl and Wilson (1995).

While we do not offer a full fledged model of belief formation here, we do attempt the demonstrate that we think these existing models are myopic because of this omission. For example, say that you and I are repeatedly playing a 3×3 non-zero sum game as our subjects do in this experiment. At round t I choose a row and you choose a column and I get a great payoff but you do miserably; *i.e.,* you get the minimum payoff in your column. While I am happy (my payoff was higher than my expectations and I have no regret given your choice) I will probably recognize that you are heavily disappointed and regretful. This recognition should lead me to think that you will not repeat your action again especially if you think that I will since I did so well. However, as this iterative logic works itself out, it leads to a belief formation process that is not nested in any of the above mentioned models. For example, in a reinforcement learning model, if I choose a row and receive the highest payoff available in it, an observer watching this game and using a reinforcement learning model will increase the probability that I will use this strategy again regardless of how my opponent did. Obviously I may not if I assume that my opponent will change his or her behavior forcing me to adapt to that expectation.

Given these remarks we proceed as follows: First we will consider two well studied models of belief formation, the Cheung-Friedman (1997) and the Camerer-Ho (1999) models and consider how well they described the data in our experiments. While we find that they do have a considerable predictive ability, they ultimately fail to explain a key feature of the belief data – the feature that changes in subject beliefs depend on the payoffs of both subjects and not just one. This correlation is not predicted by these models but is certainly a feature of our data.

First consider Cheung and Friedman's (1997) model of γ-weighted beliefs. Under this weighted fictitious play model, if a player observes strategy s_t^i being played in period t, then her belief for strategy i will increase in period $t+1$ and her beliefs on all other strategies will decrease. Therefore, if the observers in our sample believe that the actual players of the game behave in this fashion, the beliefs on the chosen action should (weakly) increase, while the beliefs on *all* unchosen actions should (weakly) decrease. In Table 9, we present the frequency of observations which were consistent with this view in each of our four games.

Table 9: Testing γ-weighted Beliefs; Fraction of Obs. Consistent With Theory

	(1)	(2)	(3)	(4)
Game	**↑ Chosen**	**↓ Both Un Chosen**	**(1) & (2)**	**(2) Conditional On (1)**
nDSS (N.C.)	72.4%	50.0%	50.0%	69.1%
DSS (N.C.)	66.8%	44.7%	44.7%	66.9%
nDSS (CON)	82.0% **(81.8%)**	61.4% **(60.4%)**	61.4% **(60.4%)**	74.9% **(73.8%)**
DSS (CON)	88.5% **(76.4%)**	73.7% **(43.0%)**	73.7% **(43.0%)**	83.3% **(56.3%)**

Column (1): Frequency with which observers increase their belief on last period's chosen action.

Column (2): Frequency with which observers decrease their beliefs on both of last period's unchosen actions.

Column (1) gives the empirical frequency with which observers (weakly) increased their beliefs on the action which was chosen in the previous period, while the column (2) gives the empirical frequency with which observers (weakly) decrease their beliefs on both of the actions which were *not* chosen in previous period. For the two convergent games, we provide two numbers in each cell; the first is the empirical frequency over all periods, while the second is the empirical frequency over those periods *before* the actual player converged to his/her Nash strategy. As the reader can see, for the two games which did not converge to Nash equilibrium, approximately 70% of the time, observers increase their belief on the chosen action. Moreover, between 76% and 89% of the time (depending on the sample), in the convergent games, was it the case that observers increased their belief on the action which was chosen last period. This suggests that there is often inertia in the belief updating process, something which seems rather myopic given that last period's chosen action may not have been the *correct* action to take.[17] This is just one component of the predictions from a γ-weighted belief updating model. Looking at column (2) gives us the complete picture. If we restrict attention to the period before convergence,[18] we see that *at most* 60% of the observations are fully consistent with the predictions of γ-weighted beliefs.

Now consider Camerer and Ho's (1999) model of experience weighted attraction. Specifically, recall that attractions for each strategy, j, and player i are updated according to:

$$\begin{aligned} A_i^j(t) &= \frac{\phi N(t-1)A_i^j(t-1)+[\delta+(1-\delta)\mathbb{I}(s_i^j,s_i(t))]\pi(s_i^j,s_{-i}(t))}{N(t)} \\ N(t) &= \rho N(t-1)+1 \end{aligned} \qquad (4)$$

Furthermore, recall that Camerer and Ho (1999) interpret $\delta \in [0,1]$ as an *imagination* parameter. The payoffs of all actions, whether chosen or not, contribute to the updating of attractions. If the action was not chosen, the contribution is $\delta\pi(s_i^j, s_{-i}(t))$, while if it was chosen, the *undiscounted* payoff enters. Consider the case in which $\delta = 1$ so that players have full imagination and therefore equally weight all possible payoffs, whether realized or not. In this case, the attraction will increase the most on the action that *would have* received the highest payoff, given the strategy of the opponent, while the attraction will increase the least (or even decrease) on the action that *would have* received the lowest payoff, given the strategy profile of the opponent. Therefore, if observers think actions are chosen according to EWA (and that $\delta = 1$), they should *increase* their belief on the action that would have received the highest payoff and decrease their belief on the action that would have received the lowest payoff. The empirical frequencies are given in Table 10. Here the first column gives the empirical frequency with which the belief on the action that would have given the highest payoff weakly increases and the second column gives the empirical frequency with which the belief on the action that would have received the lowest payoff weakly decreases.

In general, for the non-convergent games, the predictions of EWA beliefs explain more of the data than does γ-weighted beliefs; however, for the convergent games, EWA beliefs do substantially

[17]Inertia is not the ideal word, since the beliefs on last period's chosen action may actually be quite low.

[18]Of course, for the two non-convergent games, this implies that we make use of the entire sample.

Table 10: Testing EWA Beliefs; Fraction of Obs. Consistent With Theory

Game	(1) ↑ Best Payoff	(2) ↓ Worst Payoff	(3) (1) & (2)	(4) (2) Conditional on (1)
nDSS (N.C.)	76.1%	75.8%	63.7%	83.7%
DSS (N.C.)	73.9%	68.2%	59.7%	80.8%
nDSS (CON)	70.8% **(68.6%)**	73.7% **(70.8%)**	57.6% **(53.8%)**	81.3% **(78.5%)**
DSS (CON)	82.5% **(53.3%)**	84.5% **(67.9%)**	77.4% **(44.2%)**	93.8% **(83.0%)**

Column (1): Frequency with which observers increase their belief the action that would have given the highest payoff.

Column (2): Frequency with which observers decrease their belief on the action that would have given the lowest payoff.

worse. However, we believe that there is an explanation for this result. Consider the convergent DSS game. Observers were asked to state beliefs on the action choices of the column player. This game converged in period 6 and the action choices of the row and column player for the first six periods were $\{(1,2),(1,1),(1,1),(1,1),(3,2),(3,1)\}$, where $(3,1)$ is the Nash equilibrium strategy pair. In particular, in periods 2, 3 and 4 the column player was playing her Nash action, despite the fact that action 2 is a best response to the row's choice of action 1. Therefore, if the observers believed that the column player was trying to *teach* the row player, in order to get convergence to the Nash equilibrium, they would not necessarily increase their beliefs on action 2, giving more weight instead to action 1, despite it not being a best response to the row player's choice.

The γ-weighted belief model predicts that the beliefs on *both* of the unchosen actions should decrease. However, as Table 9 shows, at most 60% of the observations are consistent with this model. We may then wish to uncover more about how beliefs on unchosen actions are updated. This is the content of Table 11. The first column reports the empirical frequency with which subjects (weakly) increase their belief on the unchosen action with the higher payoff, while the second column reports the empirical frequency with which subjects (weakly) decrease their belief on the unchosen action with the lower payoff and the third is the intersection of the first two columns. The most robust feature apparent in Table 11 is that the beliefs on the unchosen action with the lower payoff decrease a substantial proportion of the time, as EWA would suggest. As a final remark, note that we conducted a similar analysis using the data from the beliefs and actions experiment; the results are largely the same and are not presented here.

5.3 Belief Updating: Logit Regressions

Given the descriptive statistics reported above, we would like to try to disentangle all of the varying effects more formally; for example, it may well be that the action chosen last period, was also a

Table 11: Updating Beliefs on Unchosen Actions

	(1)	(2)	(3)
Game	**↑ Unchosen Higher Payoff**	**↓ Unchosen Lower Payoff**	**(1) & (2)**
nDSS (N.C.)	63.2%	74.5%	53.7%
DSS (N.C.)	67.4%	72.4%	53.9%
nDSS (CON)	58.1% **(58.5%)**	81.7% **(80.3%)**	51.2% **(51.3%)**
DSS (CON)	63.2% **(47.9%)**	84.5% **(67.9%)**	58.5% **(40.6%)**

Column (1): Frequency with which observers increase their belief on last period's unchosen action with the *higher* payoff.

Column (2): Frequency with which observers decrease their beliefs on last period's unchosen actions with the *lower* payoff.

best response to the opponent's choice. It is important to separately identify these effects. Also, the descriptive statistics above speak only of changes in beliefs and not levels. While we do not immediately go far from this, we can say more. For example, whether or not an observer increases her belief on last period's chosen action may be related to how *surprising* the realized action choice was. Consider an observer with a very low initial belief on action j; she may be much more likely to increase her belief on action j (assuming it was chosen last period) than an observer with a high initial belief. Let y_{it} be the event that observer i increases her belief at time t on the action chosen in the previous period. We posit the following random effects model:

$$\begin{aligned} Y_{it} &= \alpha + \beta_1 D[MAX_t] + \beta_2 L_{it} + u_i + \epsilon_{it} \\ y_{it} &= 1 \iff Y_{it} \geq 0 \end{aligned} \tag{5}$$

where Y_{it} is an unobserved variable and $D[MAX]$ is a dummy variable which takes the value 1 if the action chosen in the previous period gave the maximal payoff, given the action choice of the other player and L_{it} represents the level of the belief on the chosen action.[19] Given the results of the descriptive analysis above and our intuition, we expect that $\alpha > 0$, $\beta_1 > 0$ and $\beta_2 < 0$. We assume that ϵ_{it} is a logistic disturbance term. The results of this exercise are given in Table 12. In brackets, below the estimated coefficients, are the z-statistics.[20]

Notice that the variable $D[MAX]$ is only significantly positive for the non-convergent games and the convergent DSS game (using the full sample). In the latter case, when we restrict attention to the pre-convergence periods, it becomes negative and is no longer significant. On the other hand,

[19]To be clear, suppose that we are predicting the *row* player's actions. The variable $D[MAX]$ takes value 1 if the *row* player received her maximal payoff last period and 0 otherwise.

[20]Hausman tests for fixed versus random effects were conducted in all cases, with a fixed effects specification being preferred in nDSS (NC), DSS (C) and nDSS (C), pre-conv. However, since we are mostly interested in the signs and significance of the variables, non of which are changed when using the random effects specification, we do not report results for the fixed effects estimation.

Table 12: Logit Regressions

	DSS (NC)	nDSS (NC)	DSS (C)	DSS (C) Pre-conv	nDSS (C)	nDSS (C) Pre-conv
constant	2.12 *(6.88)*	1.98 *(6.23)*	1.62 *(4.42)*	2.67 *(5.02)*	2.56 *(7.99)*	2.71 *(7.57)*
D[MAX]	0.94 *(3.33)*	1.34 *(4.62)*	1.71 *(4.82)*	-0.15 *(-0.28)*	-0.12 *(-0.53)*	-0.27 *(-1.04)*
Level	-3.94 *(-6.84)*	-3.19 *(-5.62)*	-0.67 *(-1.31)*	-3.40 *(-3.61)*	-1.61 *(-4.27)*	-1.81 *(-4.31)*
σ_u	0.57	0.60	1.04	0.93	0.90	0.99
ρ	0.089 *(p = 0.007)*	0.100 *(p = 0.005)*	0.246 *(p < 0.001)*	0.209 *(p = 0.026)*	0.197 *(p < 0.001)*	0.228 *(p < 0.001)*
L.L.	-206.09	-193.79	-201.68	-80.83	-279.05	-233.95

The dependent variable is a dummy which takes value one if the belief on last period's chosen action is increased and zero otherwise.

for the nDSS game, the coefficient is negative but not significant. For the levels of beliefs on last period's chosen action, we find a significantly negative effect in all cases, except over the full sample of the DSS game. That we lose significance in the convergent DSS game over the full sample is not surprising, since the game converged to Nash equilibrium in period 6 and beliefs for almost all players quickly became degenerate on the Nash action choice. Therefore, since we only insist on weak increases in beliefs, we have that $y_{it} = 1$ and $L_{it} = 1$ for much of the sample. Finally, notice that in all cases, a likelihood ratio test rejects the null hypothesis that $\rho = 0$, in most cases, with p-values below 1%. This means that a significant portion of the overall variation comes from the observer-specific disturbance u_i.[21]

In our empirical work, we also experimented with a number of other variables that we thought may influence beliefs, such as a dummy for whether the chosen action gave the *second* highest payoff, the payoff difference between the maximal payoff given the other player's action choice and the actual payoff experienced as well as the cumulative average payoff from a given strategy up to time t, given the choices of the other player in all previous periods. Rarely did these extra variables add much in the way of explanatory power and so we do not report the results here.

5.4 Belief Updating: Do Opponent Payoffs Matter?

In the logit regressions reported above, we saw that subjects generally increased their belief on last period's chosen action if that payoff was a best response to their choice. This view is consistent with reinforcement learning or EWA in that a player's experienced payoff affects the likelihood that he/she will choose a particular action. However, it may be that one's opponent's payoff affects the

[21]To see this note that $\rho = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_\epsilon^2}$.

likelihood that he/she will choose a particular action. Consider the following example. Suppose that the row player is forming beliefs over the action choices of the column player. Moreover, suppose that the particular action combination implied that the row player received her *worst* possible payoff, while the column player received his *middle* payoff. If the column player thinks that the row player is very likely to change actions (since the chosen action turned out to be very bad for the row player), this will affect column's action decision. Therefore, if the row player is aware of this, her beliefs may be different in this situation than if, say, she had received her *middle* or *best* payoff.

Consider Table 13, which tabulates how the beliefs of row players about the column players' actions move given the 9 possible payoff realization combinations. This table is constructed by pooling all of the row player data from the actions and beliefs experiment. In the left-hand pane, the cell corresponding to (C. Max, R. Med) indicates that 57% of the time (66 in total) the row player (weakly) increased her belief on the action chosen by the column last period, given that the column player received his maximal payoff while the row player received her middle payoff.[22]

Table 13: The Actual Players: Row Players' Beliefs About Column

Frequency With Which Belief on Last Period's
Action is Increased Given Outcome

	R. Max	R. Med	R. Min
C. Max	84%	57%	57%
C. Med	61%	79%	70%
C. Min	75%	72%	69%

Table 14: The Actual Players: Column Players' Beliefs About Row

Frequency With Which Belief on Last Period's
Action is Increased Given Outcome

	C. Max	C. Med	C. Min
R. Max	72%	63%	75%
R. Med	78%	83%	88%
R. Min	50%	69%	72%

We now wish to see if the belief formation process depends on the joint realizations of payoffs. We conduct the Fisher Exact Test to test for independence. For example, suppose that the row player is predicting the actions of the column player. We ask, "Given that the column player obtained her Max (Med or Min) payoff, does the fraction of times in which beliefs increase depend on the payoff rank, Max, Med or Min, of the row player?" The p-values of the Fisher Exact Test are given in Table 16.

[22]Of course, 43% of the time, the row player (strictly) decreased her belief on column's last period action.

Table 15: The Observers: Beliefs About Column

Frequency With Which Belief on Last Period's
Action is Increased Given Outcome

	C. Max	C. Med	C. Min
R. Max	91%	75%	71%
R. Med	79%	78%	80%
R. Min	-	88%	62%

Table 16: Fisher Exact Test: A Test of Independence

	Max	Med	Min
Row on Col	< 0.001	0.002	0.849
Col on Row	0.027	0.287	0.065
Obs on Col	< 0.001	0.256	0.053

Indeed, we see that in 4 of 9 cases, we can reject the null hypothesis of independent random samples at the 5% level of significance and in two other cases, we can reject the null hypothesis at the 10% level. Therefore, it appears that, for example, when the row player is forming beliefs about the column player, she takes into consideration the payoff that *she herself* receives since that is likely to affect how the column players reasons. To see this more clearly, consider the first row of Table 13. When both players are obtaining their maximal payoff (and so, playing Nash), the row player increases her belief that the column player will play his Nash strategy again 84% of the time. On the other hand, when the row player is obtaining her middle or lowest payoff, she increases her belief on column taking the same action only 57% and 58% of the time, respectively.

Finally, this evidence is strongest when the player receives the maximal or minimal payoff and less clear when the player receives her middle payoff. This makes sense; there may be strong incentives to change actions after one receives the minimal payoff and very weak incentives after one receives the maximal payoff. In either case, these clear incentives (strong or weak) may translate to beliefs in a clean way. However, when one receives her middle payoff, the incentives are not clear and so how this translates into beliefs may also not be clear; in this case, subjects may simply default to a particular rule for which independence holds.

6 Conclusions

This paper has attempted to investigate the process through which people playing games converge to an equilibrium — a state where their beliefs about the actions of their opponents are confirmed. By conducting a set of experiments we have learned that the convergence process is an action-led process in that the actions of players reach the equilibrium before their beliefs. However, one key ingredient for convergence is the presence of a "teacher" who chooses the Nash action in an effort to teach his opponent what to expect. Failure to reach equilibrium appears to be a failure in beliefs and not in best response behavior in that for such games the beliefs of players almost never enter that subset of the belief simplex where Nash actions are best responses. Also, players in games that fail to converge to Nash equilibria tend to best respond equally as often as do their convergent cohorts. We have investigated the process of belief formation and found that while there is some consensus in the dynamics of how people update their beliefs, the actual belief vectors vary wildly and do not exhibit a tendency to converge. Finally, we present evidence that suggests that a truly successful belief formation model must be one that includes a role for the payoff that one's opponent receives rather than focusing exclusively on a player's own past payoffs as is true of all reinforcement models, EWA, and weighted historical belief models.

References

[1] Camerer, C. and Ho, T-H. "Experience-weighted Attraction Learning in Normal Form Games" vol. 67, issue 4, *Econometrica*, 1999, pp 827-874.

[2] Camerer, C., T. Ho, and J. Chong (2002), "Sophisticated EWA Learning and Strategic Teaching in Repeated Games," *Journal of Economic Theory*, 104, 137-188.

[3] Cheung, Y.-W., and D. Friedman (1997), "Individual Learning in Normal Form Games," *Games and Economic Behavior*, 19, 46-76.

[4] Costa-Gomes, Miguel A. and Weizsäcker, Georg, "Stated Beliefs and Play in Normal-Form Games" (May 22, 2004)

[5] Costa-Gomes, Miguel, Vincent Crawford, and Bruno Broseta (2001), "Cognition and Behavior in Normal-Form Games: An Experimental Study," *Econometrica*, 69, 1193-1235.

[6] Crawford, V. "Adaptive Dynamics in Coordination Games," Econometrica, 63 (1995), 103-143.

[7] Dufwenberg, Martin, and Uri Gneezy (2000), "Measuring Beliefs in an Experimental Lost-Wallet Game," *Games and Economic Behavior*, 30, 163-182.

[8] Erev, I. and A. Roth, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed-strategy Equilibria," *The American Economic Review*, 88 (1998), 848-881.

[9] Fudenberg, D. and D. K. Levine (1998), "The Theory of Learning in Games," Cambridge, MA: MIT Press.

[10] Haruvy, Ernan (2002) "Identification and Testing of Modes in Beliefs," *Journal of Mathematical Psychology*, 46, 88-109.

[11] Hopkins, Ed (2002), "Two Competing Models of How People Learn in Games." *Econometrica* 70 (6), 2141-2166.

[12] Huck, Steffen, and Georg Weizsäcker (2002), "Do Players Correctly Estimate What Others Do? Evidence of Conservatism in Beliefs," *Journal of Economic Behavior and Organization*, 47, 71-85.

[13] Feltovich, Nick (2000) "Reinforcement-based vs. Belief-based Learning Models in Experimental Asymmetric-information Games," *Econometrica* 68 (3), 605-641.

[14] Fischbacher, Urs (1999) "z-Tree - Zurich Toolbox for Readymade Economic Experiments — Experimenter's Manual," Working Paper Nr. 21, Institute for Empirical Research in Economics, University of Zurich.

[15] Mason, Charles, and Owen Phillips (2001), "Dynamic Learning in a Two-Person Experimental Game," *Journal of Economic Dynamics and Control*, 25, 1305 1344.

[16] McKelvey, Richard and Thomas Palfrey (1995), "Quantal Response Equilibrium for Normal Form Games," *Games and Economic Behavior*, 10, 6-38.

[17] Nyarko, Yaw and Andrew Schotter (2002a), "An Experimental Study of Belief Learning Using Real Beliefs," *Econometrica*, 70, 971-1005.

[18] Nyarko, Yaw and Andrew Schotter (2002b), "Comparing Learning Models Using Ideal Micro-data Sets," mimeo.

[19] Offerman, Theo and Joep Sonnemans (2001), "Is the Quadratic Scoring Rule Behaviorally Incentive Compatible?" Working paper.

[20] Offerman, Theo, Joep Sonnemans and Arthur Schram (1996), "Value Orientations, Expectations and Voluntary Contributions in Public Goods," *Economic Journal* 106, 817-45

[21] Rutström, E. Elisabet and Nathaniel T. Wilcox (2004), "Learning and Belief Elicitation: Observer Effects," Working paper.

[22] Salmon, Timothy C. (2001), "An Evaluation of Econometric Models of Adaptive Learning," *Econometrica*, 69 (6), 1597-1628.

[23] Selten, R., and J. Buchta (1999), "Experimental Sealed-Bid First Price Auctions with Directly Observed Bid Functions." In D. Budescu, I. Erev, and R. Zwick (eds.): *Games and Human Behavior: Essays in the Honor of Amnon Rapoport*. Lawrenz Associates Mahwah NJ.

[24] Selten, R., and R. Stoecker (1986), "End Behavior in Sequences of Finite Prisoners Dilemma Supergames: A Learning Theory Approach." *Journal of Economic Behavior and Organization*, 7, 47-70.

[25] Sonnemans, Joep and Theo Offerman (2001), "Is the Quadratic Scoring Rule Behaviorally Incentive Compatible?" Working Paper.

[26] Stahl, Dale, and Paul Wilson (1995), "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10, 218-254.

[27] Von Neumann, J. and O. Morgenstern (1944). Theory of Games and Economic Behavior. Princeton, N.J., Princeton University Press.

[28] Weizsäcker, Georg (2003), "Ignoring the Rationality of Others: Evidence from Experimental Normal Form Games," *Games and Economic Behavior*, 44, 145-171.

[29] Wilcox, Nathaniel, and Nick Feltovich (2000), "Thinking like a Game Theorist: Comment," mimeo, University of Houston.

A Actions and Beliefs Experiment Instructions

The following instructions were used for the beliefs and actions experiment:

General Instructions

Welcome and thank you for coming today to participate in this experiment. The purpose of this experiment is to learn how people behave in certain very simple settings.

After this experiment, another experiment will take place. The precise details of that experiment will be explained to you at the appropriate time. At the beginning of this experiment, you will be randomly paired with another participant and you will remain paired with this person for the duration of this **first** experiment. However, at no point in time will it ever be revealed to you with whom you are matched. Depending on your choices and those of your partner you will earn money, which will be paid at the end of the experiment. The exact method of calculating your final payment will be described below.

We ask that you remain silent throughout the experiment. If, at any time, you have a question, please ask the session coordinator. Failure to comply with these instructions means that you will be asked to leave the experiment and all earnings will be forfeited.

In the experiment it is more convenient to work with points rather than dollars. At the end of the experiment, the total number of points earned will be converted to dollars. The exact conversion factor is the following:

$$200 \text{ points} = \$1.00$$

Decision Problem

In this experiment, you and your partner will play a game for a total of 20 periods. You will be asked two things. In each period, you will be asked to choose an action. Your action choice and the action choice of your partner will determine the bulk of your payoff for each period. However, before making your action choice, you will be asked to *predict* the action choice that your partner will make. We now explain, in detail, both of the decisions you must make.

The Game

On your computer screen, you will be presented with the following representation of the game that you and your partner will play. The game is exactly the same for all 20 periods.

	$A1$	$A2$	$A3$
$A1$	$12, 83$	$39, 56$	$42, 45$
$A2$	$24, 12$	$12, 42$	$58, 76$
$A3$	$89, 47$	$33, 94$	$44, 59$

In each period, both you and your partner will simultaneously choose one of three actions, labeled $A1$, $A2$ and $A3$. The actions that you and your partner take in each period, as well as your position as either the **row** or

column player, determine the payoffs for that period. Each of the nine boxes above represent the nine possible action combinations. In each box, the **first** entry represents the payoff for the **row** player, while the **second** entry represents the payoff for the **column** player.

For example, suppose that you are the *row* player and chose action $A3$ in the current period; suppose also that your partner, the *column* player, chose action $A1$. Then you will earn 89 experimental points for the current period and your partner will earn 47 experimental points. As another example, suppose that you are the column player and took action $A2$, while your partner, the row player, took action $A1$. In this case, you, the column player, will earn 56 points and your partner, the row player, will earn 39 points.

On the computer screen it will be clearly marked whether you are the ***row*** *player or the* ***column*** *player. Moreover, your position will not change in* ***any*** *of the 20 periods which comprise this first experiment. That is, if you are the row player in period 1, you will continue to be the row player in each of periods 2 through 20.*

Predicting Other People's Choices

Prior to choosing an action in each period, you will be given the opportunity to earn additional money by predicting the choices of your partner in the game. On your computer screen, each of you will be asked the following three questions:

- On a scale from 0 to 100, how likely do you think it is that your partner will take action $A1$?

- On a scale from 0 to 100, how likely do you think it is that your partner will take action $A2$?

- On a scale from 0 to 100, how likely do you think it is that your partner will take action $A3$?

Your response to each question must be a number between 0 and 100. Moreover, the sum of the three numbers that you provide ***must*** *be* ***exactly*** *100.*

For example, suppose that you think there is a 30% chance that your partner will take action $A1$, a 25% chance that your partner will take action $A2$ and a 45% chance that your partner will take action $A3$. In this case, you will enter 30 in the first box on the screen in the bottom left corner and 25 and 45 in the second and third, respectively. The exact computer screen you will see is given below.

You will earn experimental points for your predictions according to a specific payoff function, which we now explain. Suppose your predictions are as in the above example. Furthermore, suppose that in the current period your partner actually chose $A2$. In that case your payoff for predicting your partner's action will be:

$$\text{Payoff} = 5[2 - (\frac{30}{100})^2 - (1 - \frac{25}{100})^2 - (\frac{45}{100})^2]$$

In other words, we will give you a fixed amount of 10 points from which we will subtract an amount which depends on how inaccurate your prediction was. To do this, we find out what choice your pair member has made. We then take the number you assigned to that choice, in this case 25% on $A2$, subtract it from 100%, square it and multiply by 5. Next, we take the number you assigned to the choices not made by your pair member, in this case the 30% you assigned to $A1$ and the 45% you assigned to $A3$, square them and multiply by 5. These three squared numbers will then be subtracted from the 10 points we initially gave you to determine your final point payoff. Your point payoff will then be converted into dollars at the same conversion factor as given above.

Note that since your prediction is made before you know what your partner has actually chosen, the best thing you can do to maximize the expected size of your prediction payoff is to simply state your true beliefs about what you think you partner will do. Any other prediction will decrease the

amount you can expect to earn as a prediction payoff.

Note also that you cannot lose points from making predictions, you can only earn more points. The worst thing that could happen is that you predict that your partner will choose one particular action (*e.g.,* $A2$) with 100% certainty *but* it turns out that your partner actually chose a different action (*e.g.,* $A3$). In this case, you will earn 0 points. In all other situations, you will earn a strictly positive number of points.

The Computer Screen

In each period, you will see the following computer screen. At the very top of the screen, you will see which period you are in, how many periods in total there are and the time remaining to make your decision. In the experiment, you will have 1 minute to state your predictions regarding what you think the other player will do and choose an action.

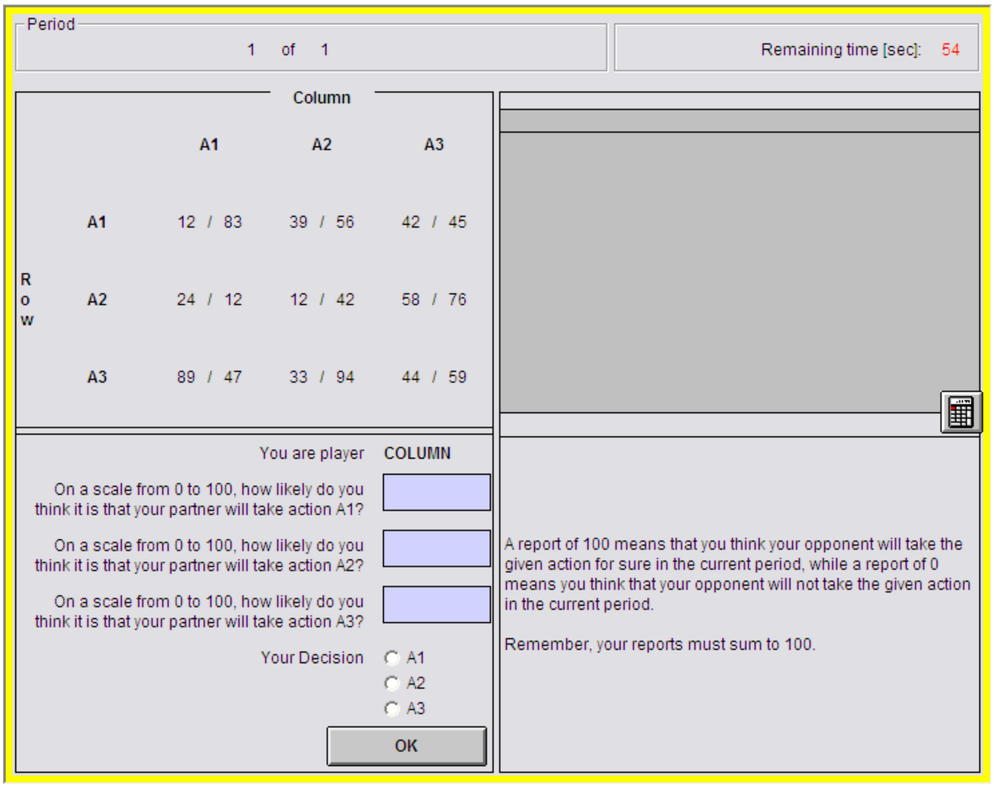
In the top left portion of the screen, you will see the game that you are playing and the payoffs to both you and your partner for each combination of actions taken. Recall that the first number in every box gives the payoff to the **row** player and the second number gives the payoff to the **column** player for each of the nine possible action combinations.

The top right portion of the screen shows the outcomes from all previous periods. In particular, you are able to see what action you took in all previous periods and what action your partner took in all previous periods.

The bottom left corner of the screen is where you will make your decisions. The first thing that you will see is whether you are the row player or the column player. Below that are the three questions asking you to predict the action choice of your partner. **Your responses to these three questions must each be numbers between 0 and 100 and the three numbers must sum to 100. Your response may contain at most 1 number after the decimal point.**

The bottom right portion of the screen has a couple of reminders that you may wish to refer to during the experiment. You will also see a calculator button. By pressing this button, the computer's calculator appears, which can be used as a check that your predictions add up to 100.

After all of your decisions have been made, click on the **OK** button. Once both you and your partner have pressed OK, you will be taken to a new screen where you may review the action that you took, learn the action taken by your partner and find out your payoff from your action choices for that period. In the bottom right corner of this screen, you may press **continue**. Once both you and your partner have done so, you will be returned to the main screen, where a new period, exactly the same as the previous, will begin. There are 20 periods in total.



Final Payment

Your final payment for the experiment will be determined as follows. We will sum up the number of experimental points earned in each period for your action choices as well as for your predictions regarding your partner's behavior. The total number of points will then be converted back into dollars at the rate of $1 = 200 experimental points. This will be combined with your $7 participation fee to come up with your final payment. Payments will be made privately at the conclusion of the two experiments.

B Beliefs Only Experiment Instructions

The following instructions were used for the beliefs only experiment:

General Instructions

Welcome and thank you for coming today to participate in this experiment. The purpose of this experiment is to learn how people make decisions in certain very simple settings.

After this experiment, another experiment will take place. The precise details of that experiment will be explained to you at the appropriate time. Depending on your choices you will earn money, which will be paid at the end of the experiment. The exact method of calculating your final payment will be described below.

We ask that you remain silent throughout the experiment. If, at any time, you have a question, please ask the session coordinator. Failure to comply with these instructions means that you will be asked to leave the experiment and all earnings will be forfeited.

In the experiment it is more convenient to work with points rather than dollars. At the end of the experiment, the total number of points earned will be converted to dollars. The exact conversion factor is the following:

$$20 \text{ points} = \$1.00$$

A Previous Experiment

In a previous experiment, we had two subjects play the following game for 20 periods.

	A1	*A2*	*A3*
A1	51, 30	35, 43	93, 21
A2	35, 21	25, 16	32, 94
A3	68, 72	45, 69	13, 62

One of the subjects had the role of the **row** player, while the other had the role of the **column** player. In each of 20 periods, the two subjects **simultaneously** chose an action — either A1, A2 or A3. The actions taken by the row and column players in each period determine the payoffs for that period. Each of the nine boxes above represent the nine possible action combinations. In each box, the **first** entry represents the payoff for the **row** player, while the **second** entry represents the payoff for the **column** player.

To understand how to calculate the payoffs for this game, suppose that the row player chose **A2** and the column player chose **A3**. In this case, the row player would have earned 32 points and the column player would have earned 94 points.

The subjects who have played this game before were recruited just as you were today by the CESS lab recruiting program. Hence they are NYU undergraduates just as you are. They played the game for 20 periods and we have recorded their choices in each of the 20 periods of their interaction. That means that in each of the 20 periods the row player has made one of his or her three possible choices A1, A2, or A3 as has the column chooser. Your task in this experiment is to predict the actions of the **COLUMN** player in each of the 20 periods of his or her interaction with the row player he or she was matched with. We stress that these two subjects were paired with each other for the entire 20 periods. We will now explain this task to you in more detail as well as how you will be paid for your decisions.

Predicting Other People's Choices

In each period, but before learning what actually happened, you will be asked the following three questions which will appear on the computer screen in front of you:

- On a scale from 0 to 100, how likely do you think it is that the COLUMN player will take action $A1$?
- On a scale from 0 to 100, how likely do you think it is that the COLUMN player will take action $A2$?
- On a scale from 0 to 100, how likely do you think it is that the COLUMN player will take action $A3$?

Your response to each question must be a number between 0 and 100. Moreover, the sum of the three numbers that you provide ***must*** *be* ***exactly*** *100.*

For example, suppose that you think there is a 30% chance that the COLUMN player will take action $A1$, a 25% chance that the COLUMN player will take action $A2$ and a 45% chance that the COLUMN player will take action $A3$. In this case, you will enter 30 in the first box on the left-hand side of the screen, 25 in the second box and 45 in the and third box. The exact computer screen you will see is given below.

After you have submitted your predictions, you will be taken to a waiting screen on which you will see the actions actually chosen by both the ROW and the COLUMN players. Based on your predictions and the action actually chosen by the COLUMN player, you will earn experimental points according to a specific payoff function, which we now explain. Suppose your predictions are as in the above example. Furthermore, suppose that in the current period the COLUMN player actually chose $A2$. In that case your payoff for predicting the COLUMN player's action will be:

$$\text{Payoff} = 5[2 - (\frac{30}{100})^2 - (1 - \frac{25}{100})^2 - (\frac{45}{100})^2]$$

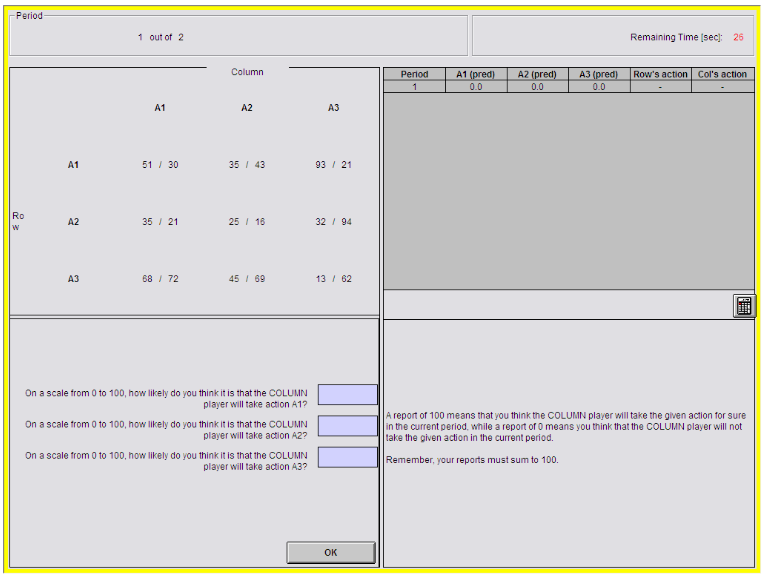
In other words, we will give you a fixed amount of 10 points from which we will subtract an amount which depends on how inaccurate your prediction was. To do this, we find out what choice the COLUMN player made. We then take the number you assigned to that choice – in this case 25% on $A2$ – subtract it from 100%, square it and multiply by 5. Next, we take the number you assigned to the choices not made by the COLUMN player – in this case the 30% you assigned to $A1$ and the 45% you assigned to $A3$ – square them and multiply by 5. These three squared numbers will then be subtracted from the 25 points we initially gave you to determine your final point payoff. Your point payoff will then be converted into dollars at the conversion factor as given above.

Note that since your prediction is made before you know the choices of both the row and column players, the best thing you can do to maximize the expected size of your prediction payoff is to simply state your true prediction about what you think the ROW player will do. Any other prediction will decrease the amount you can expect to earn as a payoff.

Note also that you cannot lose points from making predictions. The worst thing that could happen is you predict that the COLUMN player will choose one particular action (*e.g.,* $A2$) with 100% certainty *but* it turns out that the COLUMN player actually chose a different action (*e.g.,* $A3$). In this case, you will earn 0 points. In all other situations, you will earn a strictly positive number of points.

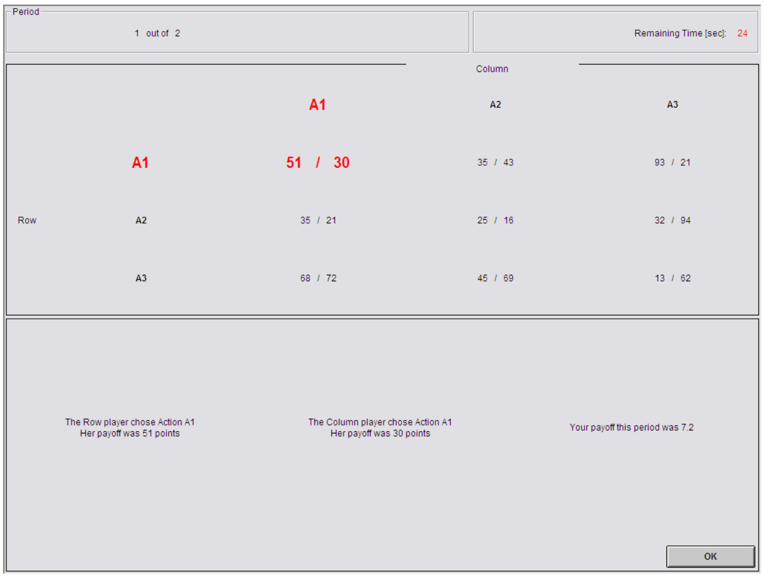
The Computer Screen

On your computer screen, in each period you will see the following screen:



You make your predictions by entering a response to each question on the bottom left-hand side of the computer screen. To submit your predictions simply press [OK]; you will then be taken to a waiting screen, which will be shown below. **Your responses to these three questions must each be numbers between 0 and 100 and the three numbers must sum to 100. Your response may contain at most 1 number after the decimal point.** On the bottom right-hand side, you will see a reminder message as well as all of your previous predictions and a calculator button, while on the upper right-hand side of the computer screen you will see the actions chosen by the ROW and COLUMN players in each of the **previous** periods as well as your past prediction.

After you have made your predictions, you will be taken to a waiting screen. On this screen, you will see the actions that the ROW and COLUMN players **actually** made for that period as well as the number of experimental points they earned for that period. You will also see the number of points that you earned for making your predictions.



In this example, the row player chose action A1 and the column player also chose A1. For this period, the ROW player earned 51 points while the COLUMN player earned 30 points. At the beginning of the next round, at the right-hand side of the screen, it will be marked that each player chose A1 in period 1.

This concludes one round. In every round, except the 20th, a new round will proceed in exactly the same manner.

Final Payment

Your final payment for the experiment will be determined as follows. We will sum the number of points you earned in each of the 20 rounds that you played. This number will then be converted back into dollars at the rate of $1 = 20 points. This will be combined with your $7 participation fee to come up with your final payment. Payments will be made privately at the conclusion of the two experiments.

C Figures

Figure 1: Games Used in the Experiments

	$A1$	$A2$	$A3$
$A1$	51, 30	35, 43	93, 21
$A2$	35, 21	25, 16	32, 94
$A3$	**68, 72**	45, 69	13, 62

(1.a) DSS

	$A1$	$A2$	$A3$
$A1$	12, 83	39, 56	42, 45
$A2$	24, 12	12, 42	**58, 76**
$A3$	89, 47	33, 94	44, 59

(1.b) nDSS

Figure 2: Difference Between Convergence Periods Beliefs and Actions - Empirical Distributions

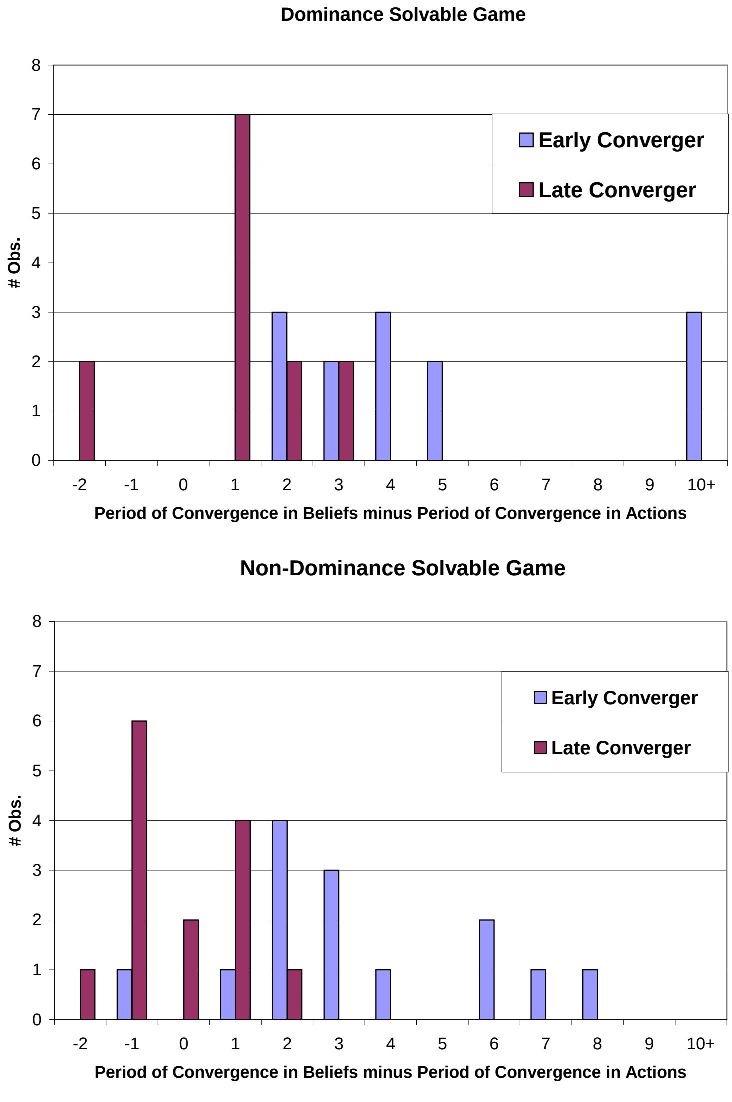
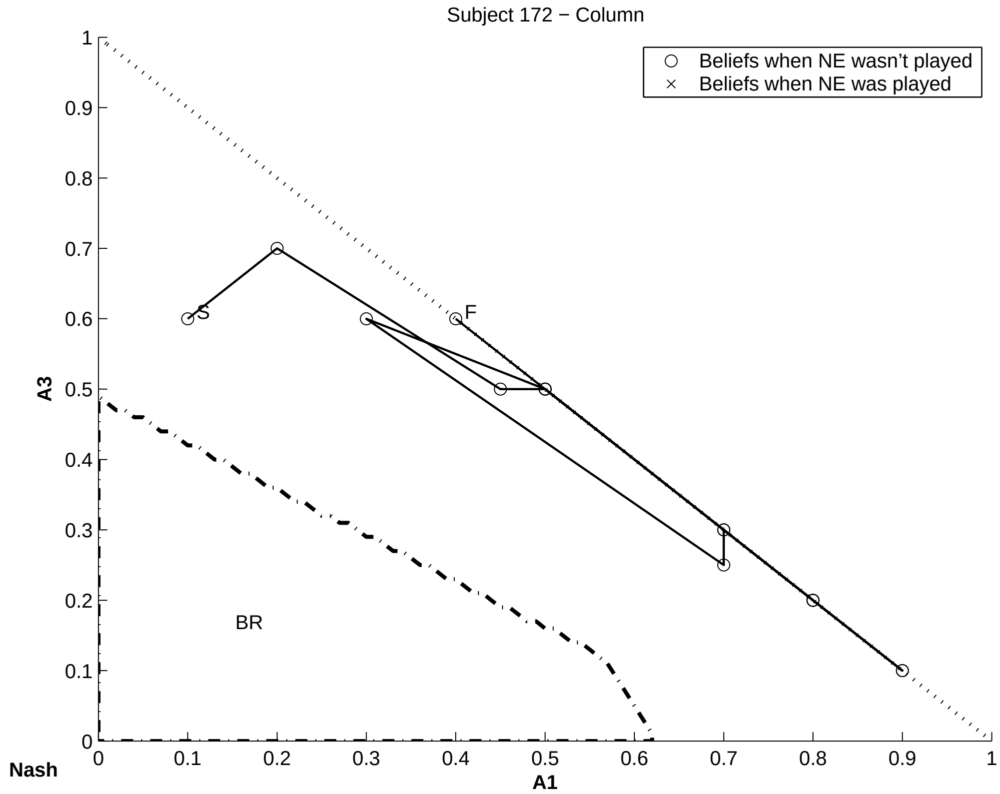


Figure 3: Belief Data - nDSS(non-converging)



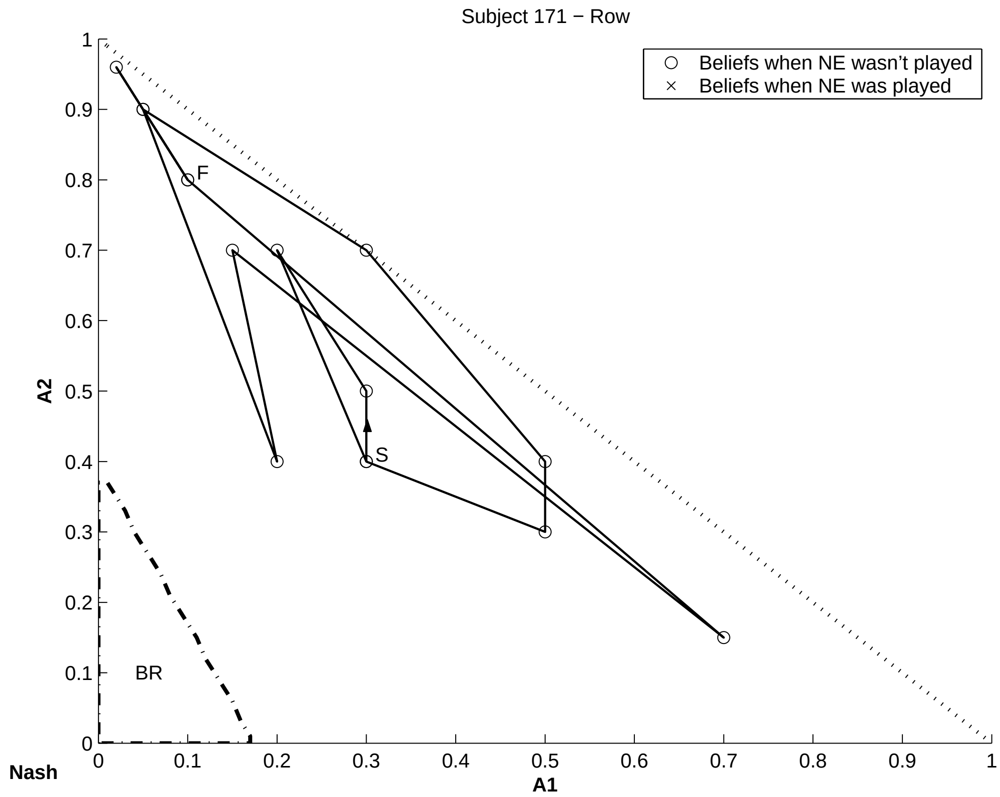
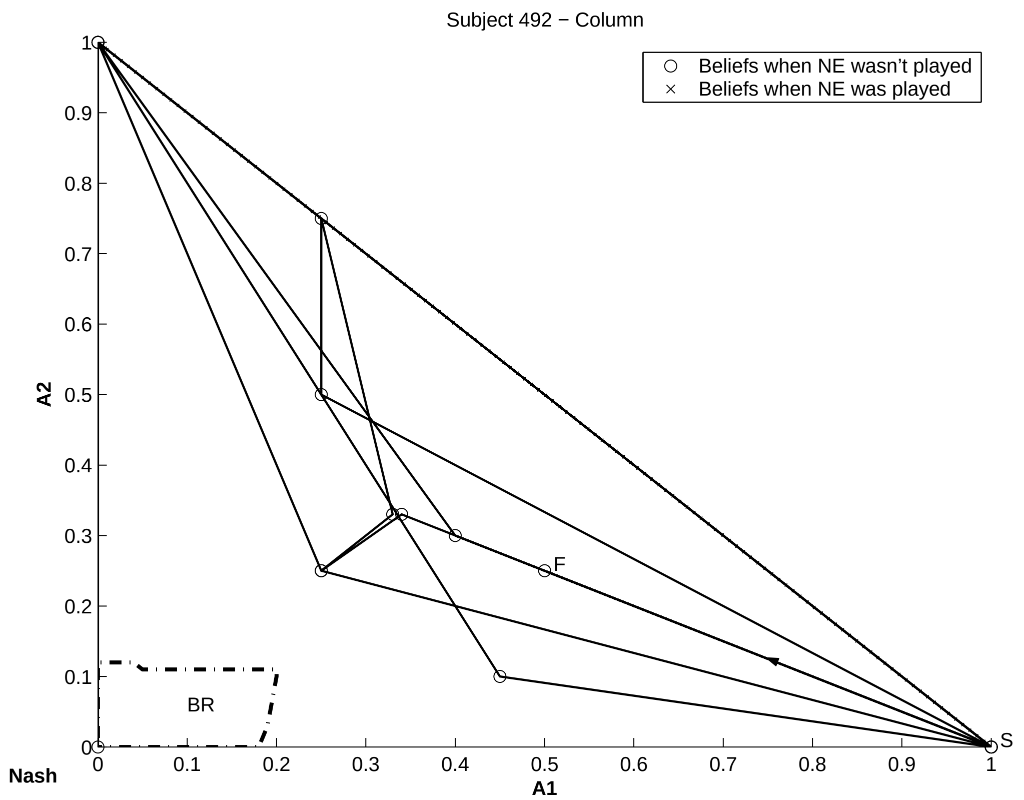


Figure 4: Belief Data - DSS(non-converging)



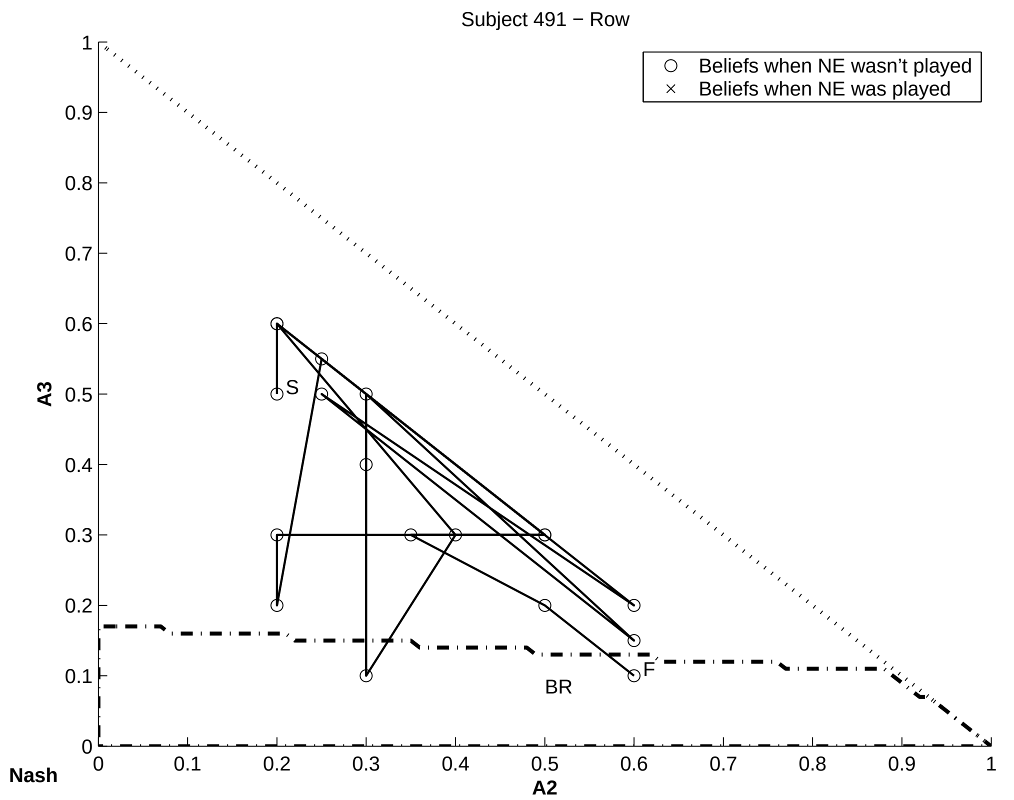
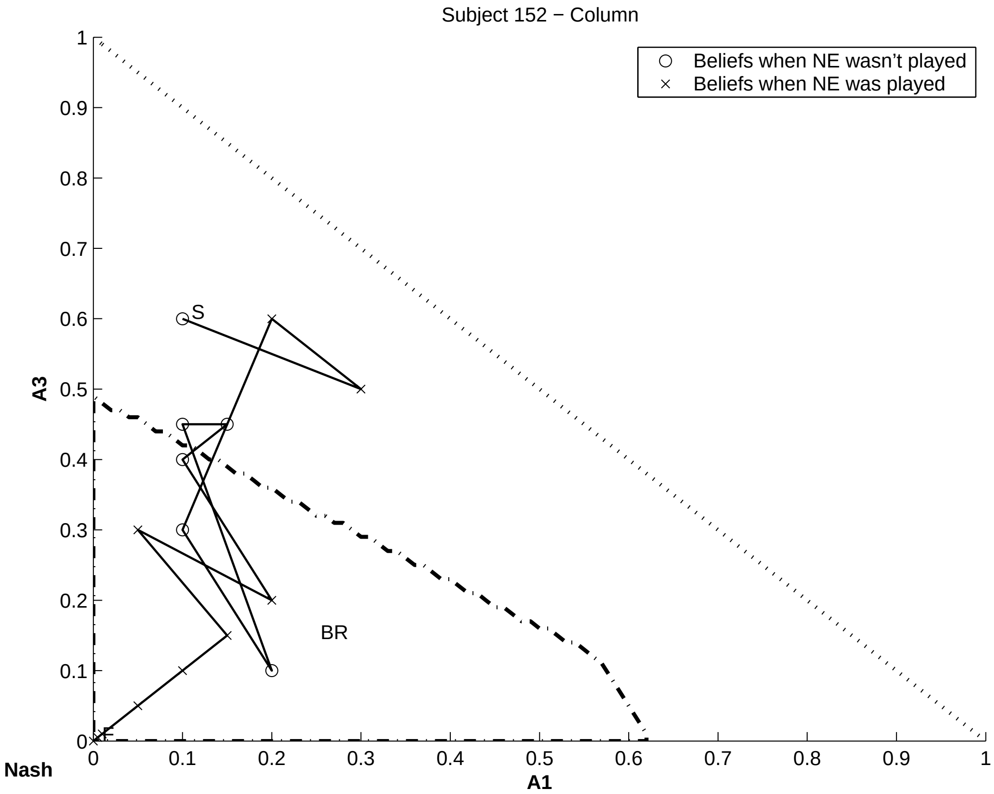


Figure 5: Belief Data - nDSS(converging)



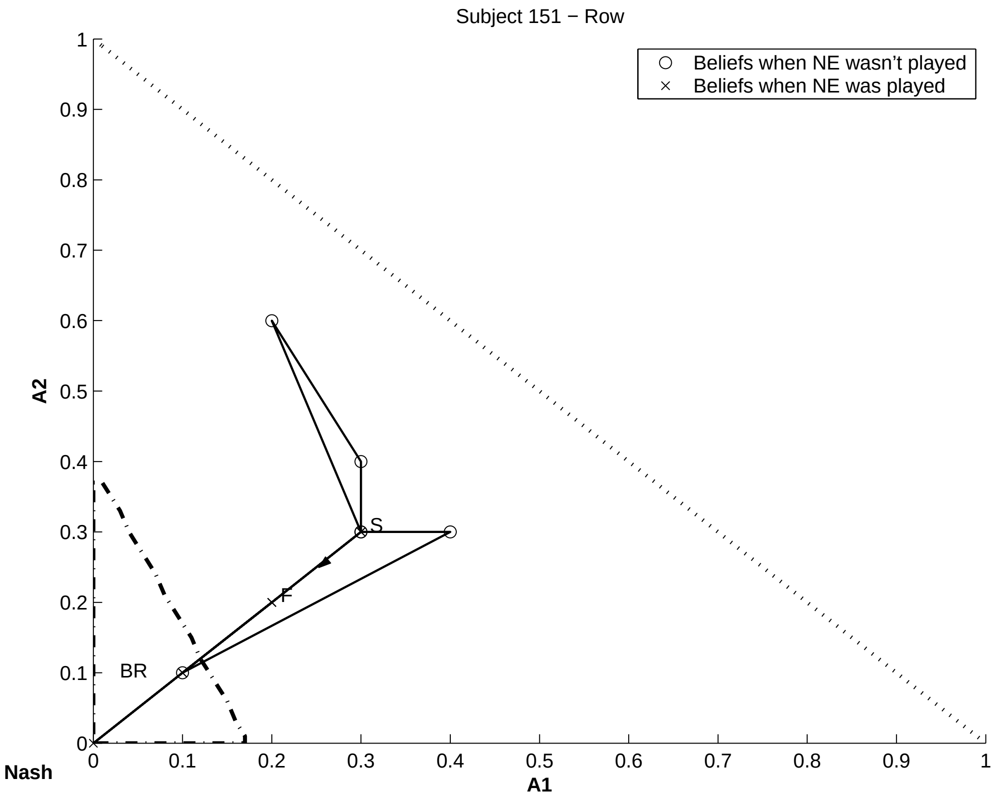


Figure 6: Belief Data - DSS(converging)

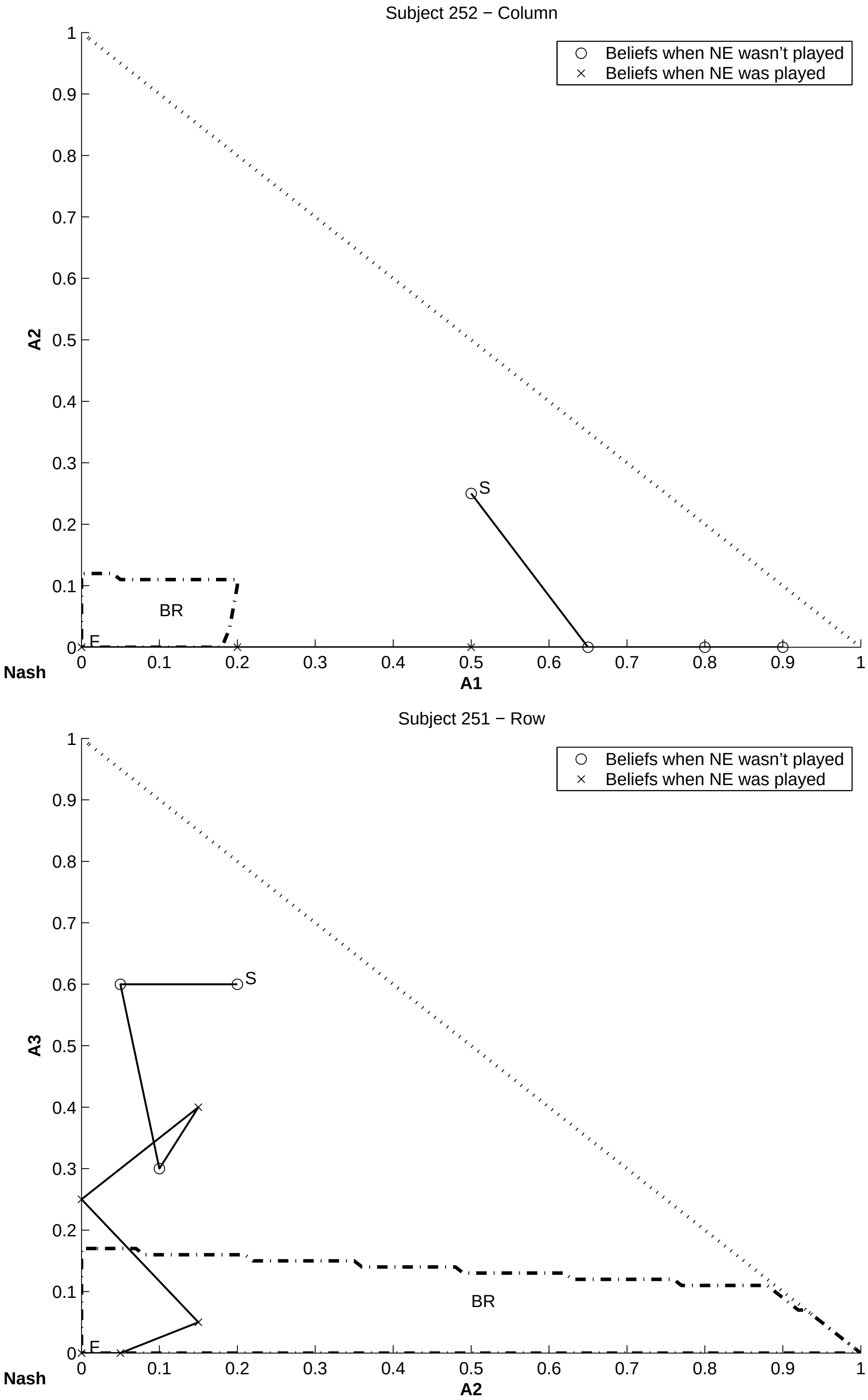
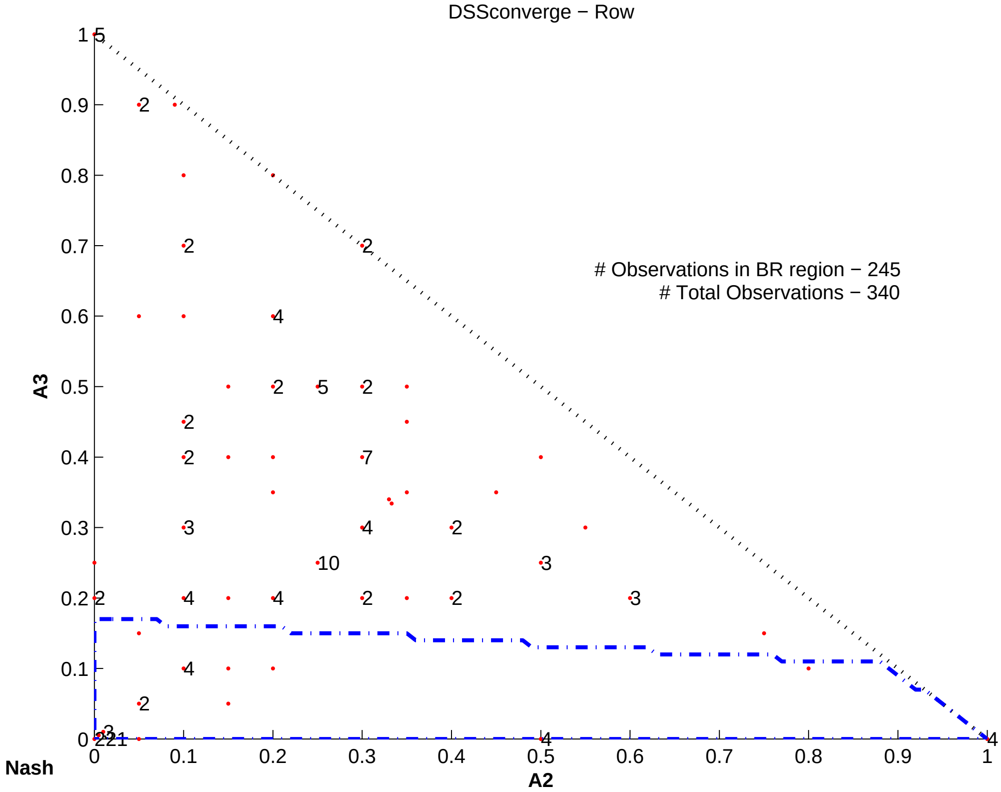


Figure 7: Belief data - DSS(converging)



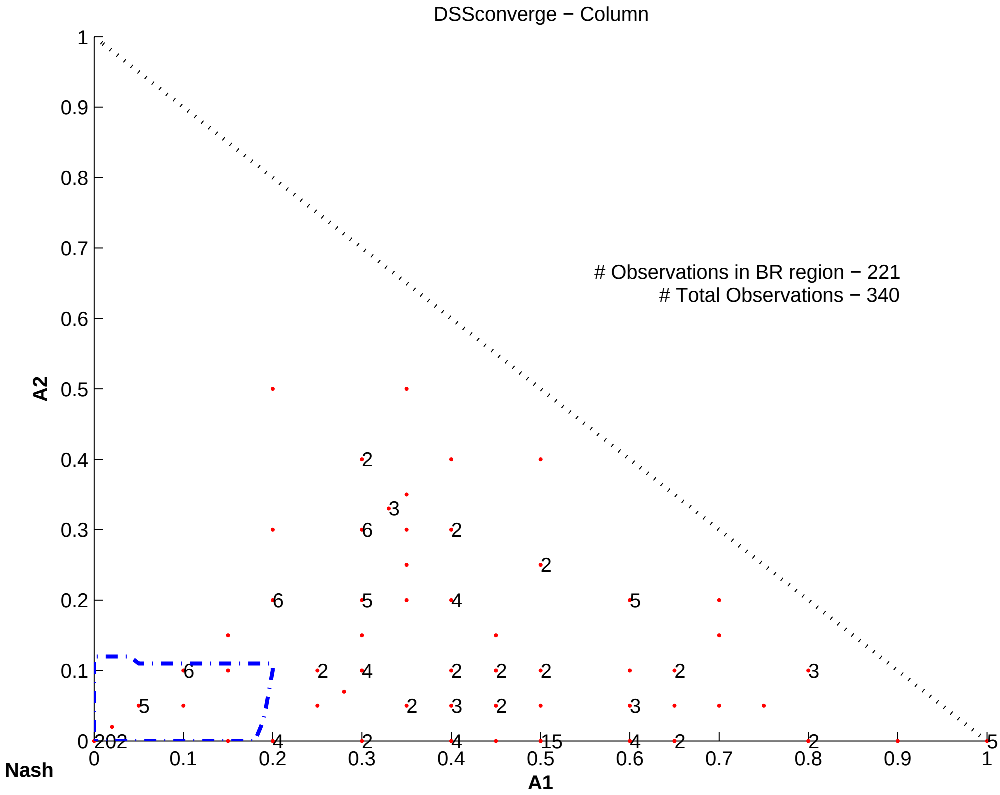


Figure 8: Belief data - DSS(non-converging)

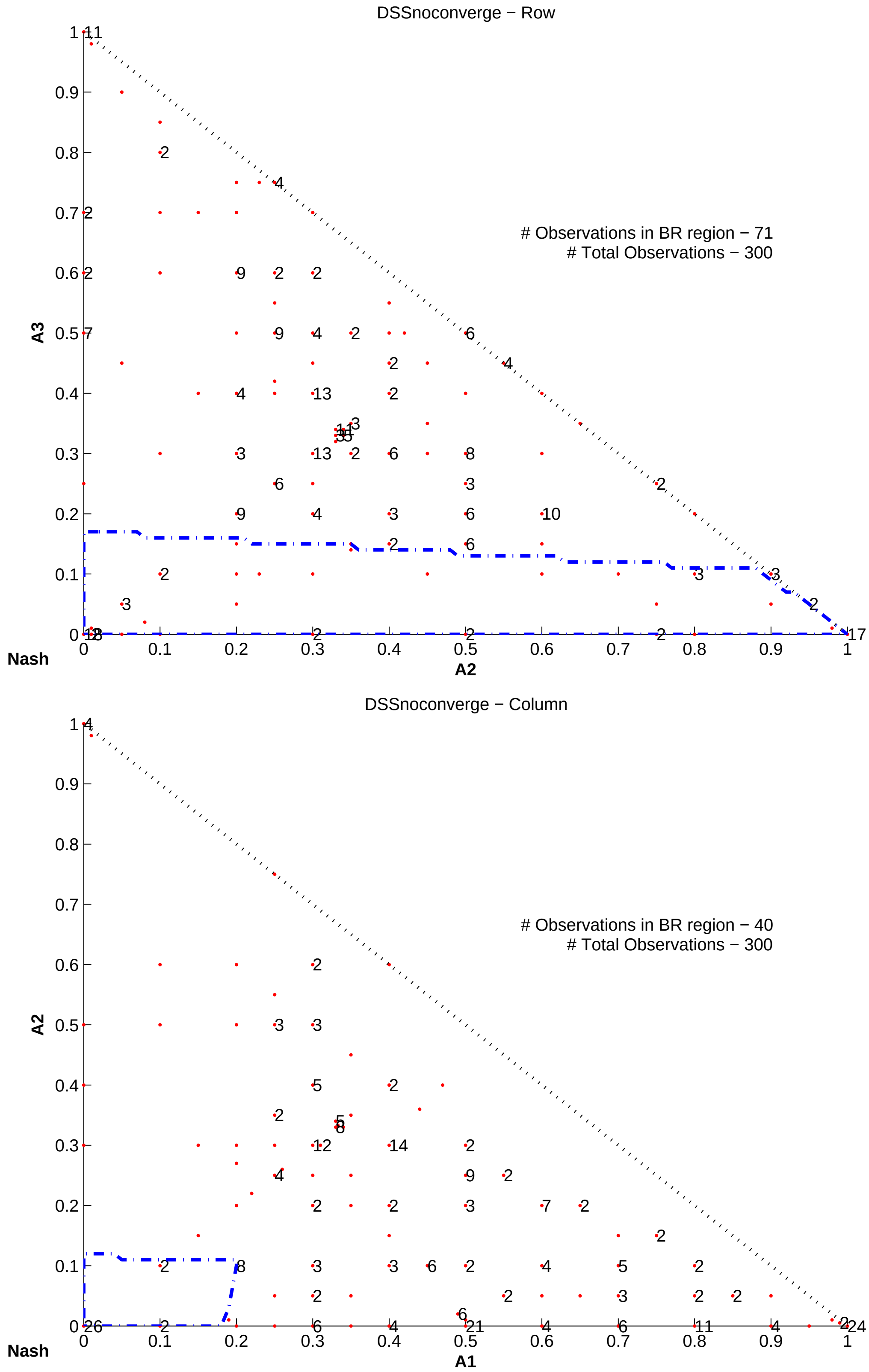
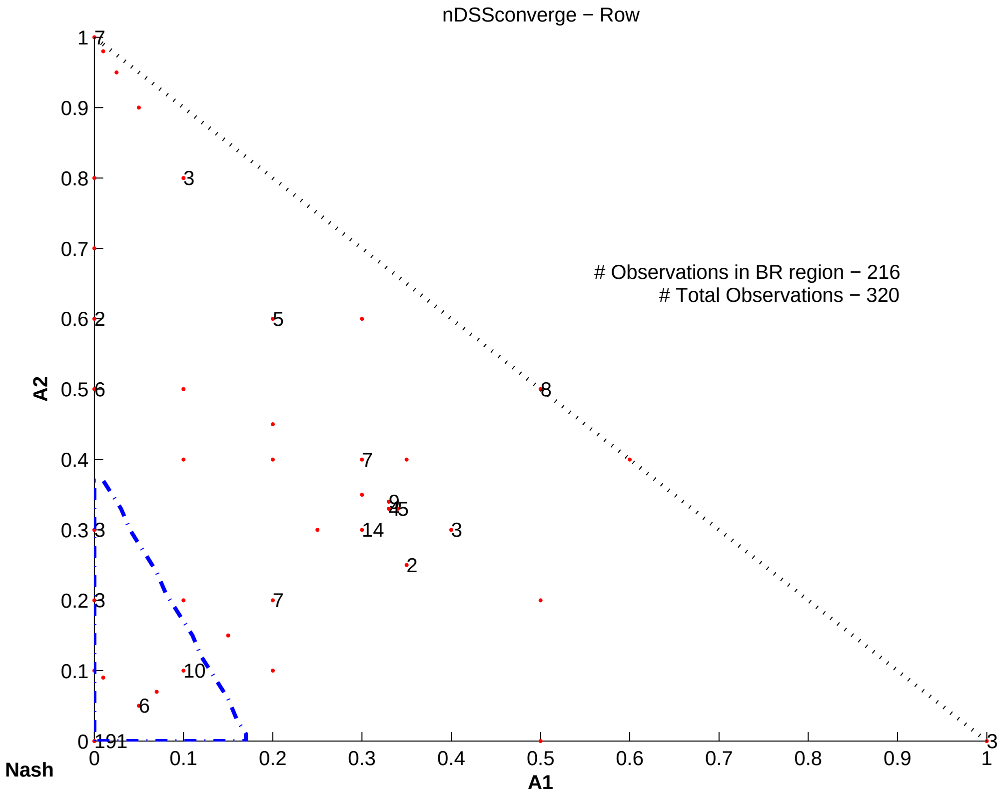


Figure 9: Belief data - nDSS(converging)



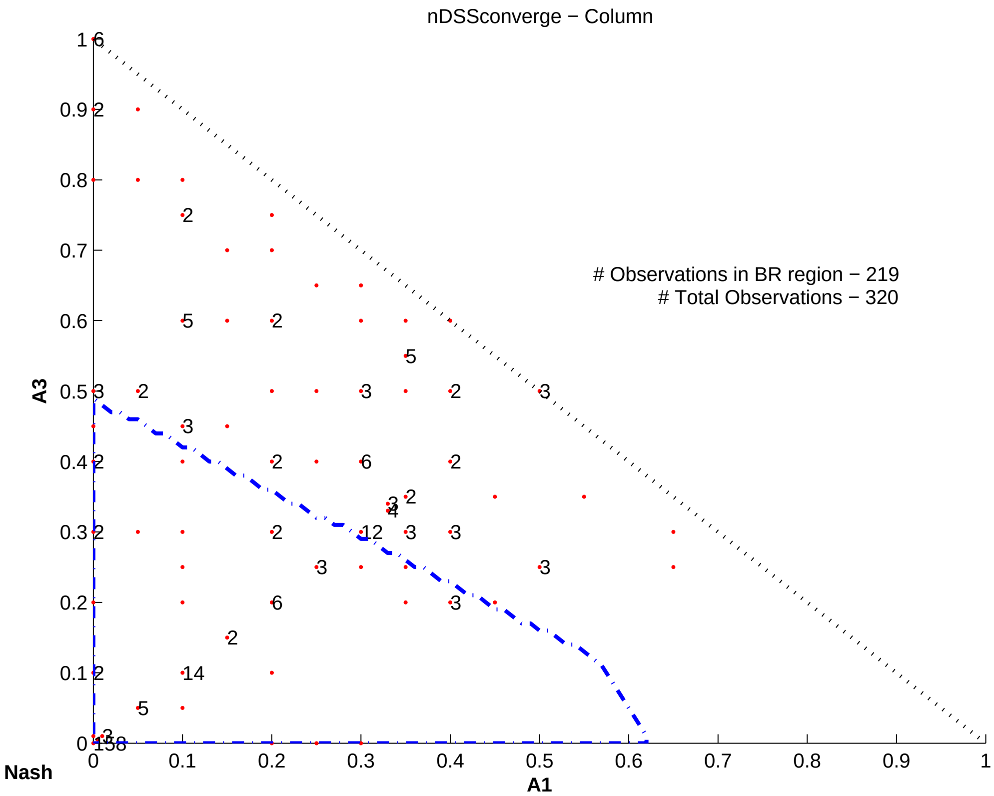
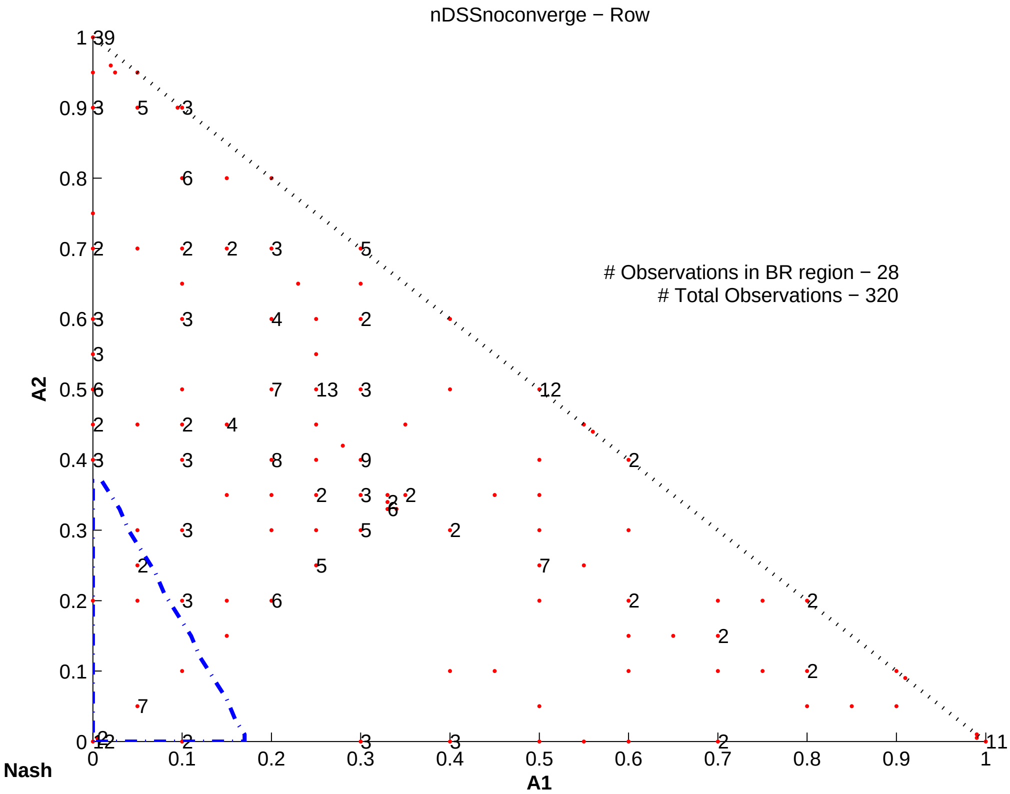


Figure 10: Belief data - nDSS(non-converging)



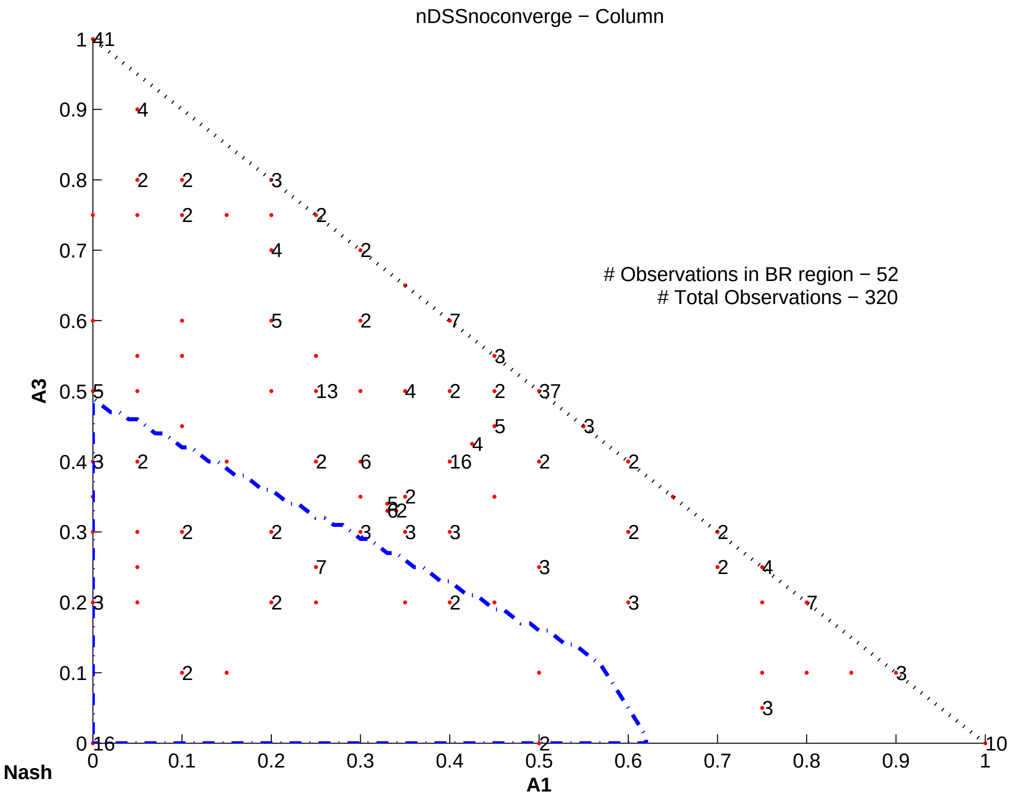


Figure 11: Periods until beliefs degenerate

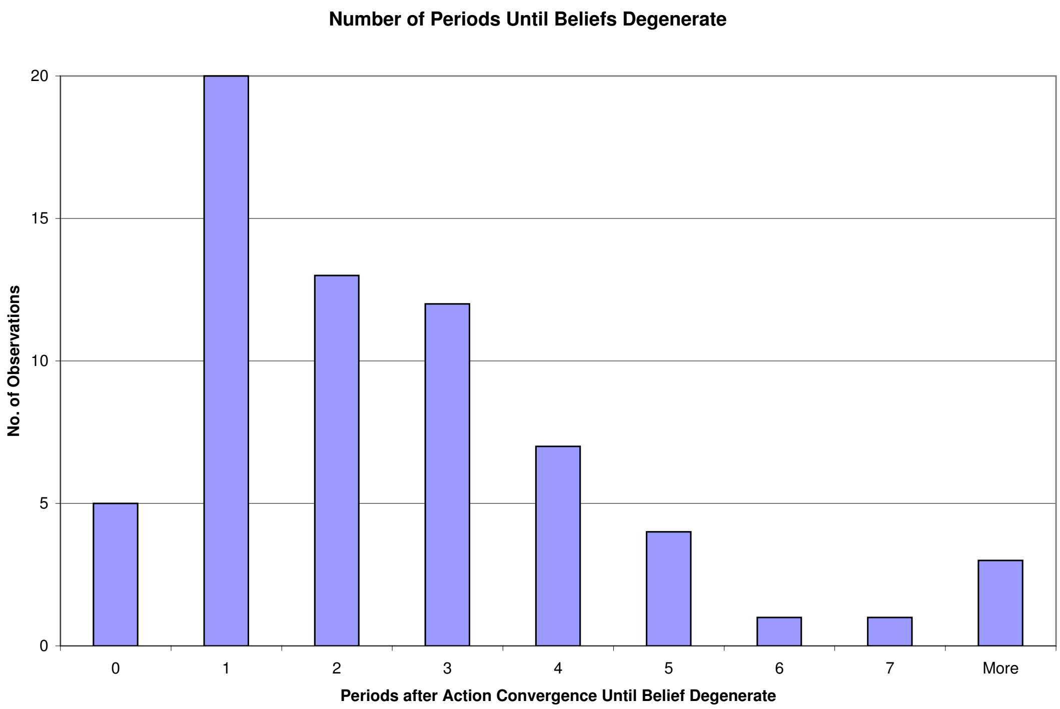


Figure 12: Consensus Index

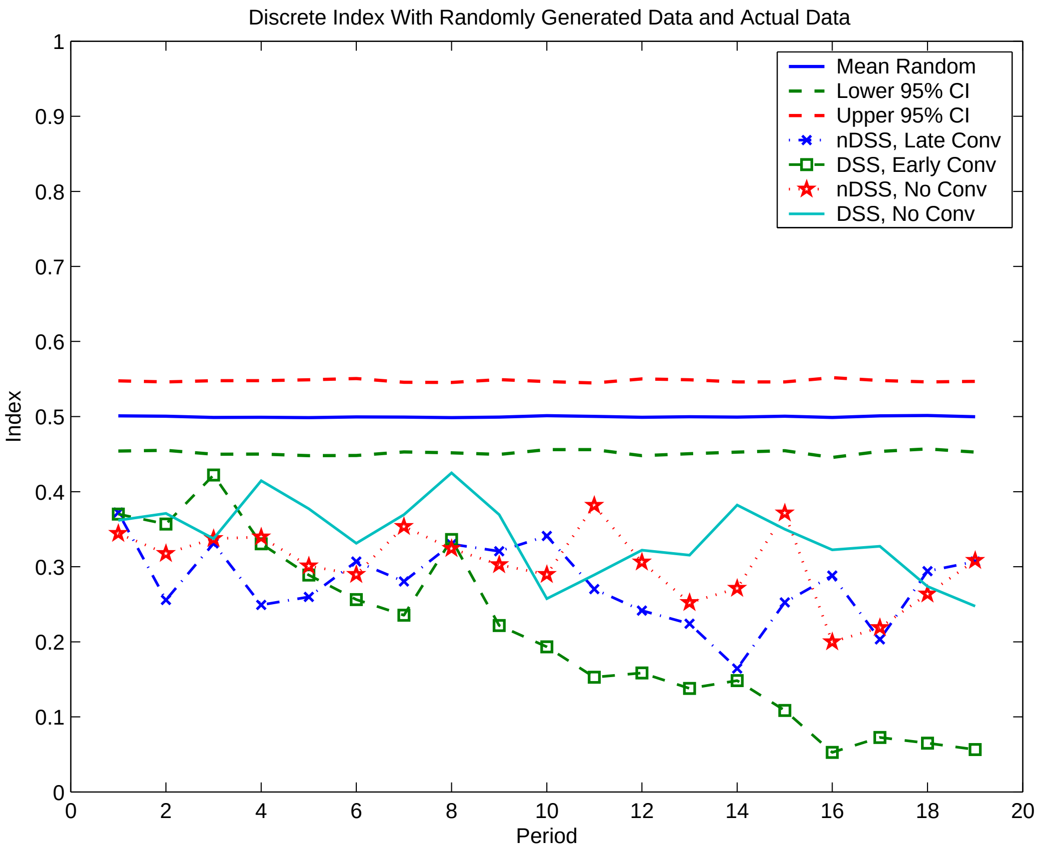
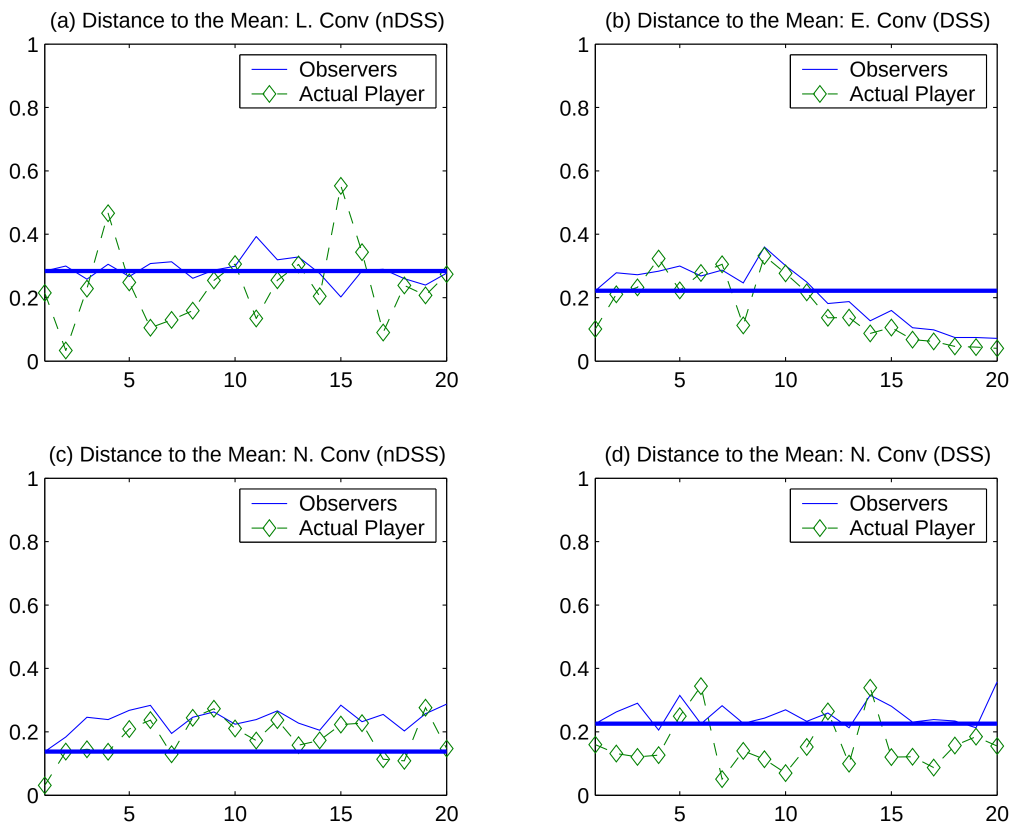


Figure 13: Static Consensus



L. Conv = Late Convergence, E. Conv = Early Convergence & N. Conv = non-convergent