Available online at www.sciencedirect.com
SCIENCE DIRECT
Games and Economic Behavior 44 (2003) 134–144
GAMES and Economic Behavior
www.elsevier.com/locate/geb

‘Learning’ with no feedback in a competitive guessing game

Roberto A. Weber

Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Received 19 May 2000

Abstract

An assumption underlying current models of learning in games is that learning takes place only through repeated experience of outcomes. This paper experimentally tests this assumption using Nagel’s (1995, Amer. Econ. Rev. 85, 1313–1326) competitive guessing game. The experiment consists of several periods of repeated play under alternative feedback conditions, including no-feedback conditions in which players receive no information between periods. If learning takes place only through reinforcement resulting from experienced outcomes, choices in the no-feedback conditions should not converge towards the Nash equilibrium. While less than under full information, there *is* convergence towards the equilibrium prediction in the no-feedback conditions. Varying priming given to subjects between periods does not affect the results.

JEL classification: C72; C92; D83

Keywords: Learning; Games; Feedback

1. Introduction

The study of learning in games has recently received considerable attention. Since behavior in experimental studies of games changes with experience, the goal of much of this research has been to model this learning process. Several models have been proposed that map feedback on outcomes and payoffs into subsequent choices. Learning in these models takes place only when players receive feedback on prior performance

E-mail address: rweber@andrew.cmu.edu.

0899-8256/03/$ – see front matter
doi:10.1016/S0899-8256(03)00002-2

before choosing again. However, much economic activity takes place with delayed or poor feedback concerning performance.[1]

This paper explores the possibility that a different form of learning may take place with repeated experience. This type of learning occurs when individuals receive no feedback on prior performance before performing a task again, yet still improve their performance. None of the current models of learning in games predict this change in behavior when players receive no new information.

There are many different models of learning in games. Examples include: learning direction theory (Selten and Buchta, 1994), choice reinforcement learning (Erev and Roth, 1998), weighted fictitious play (Cheung and Friedman, 1997), experience-weighted attraction (Camerer and Ho, 1999) and rule learning (Stahl, 2000). In all of these models players' choices are revised only through their observation of previous outcomes. So, for example, in choice reinforcement learning models such as Erev and Roth's a player's propensity to play a particular strategy is updated by the payoff experienced when that strategy is played. In weighted fictitious play and experience-weighted attraction, on the other hand, the propensities are also affected by the foregone payoffs of strategies that were not selected. In Stahl's rule learning model, a rule is a response to observed outcomes and rules are updated using their expected payoff given the actual distribution of other players' choices.[2]

One method for testing the competing models is to vary the information received by subjects concerning outcomes. For instance, choice reinforcement models assume that players only use information about the payoffs to their chosen strategies in making subsequent decisions. Therefore, according to these models, taking away information concerning payoffs to strategies not selected should have no effect on behavior in repeated play. In other models, such as weighted fictitious play and experience-weighted attraction, players use the information on payoffs not received and therefore these models predict that eliminating it will have an effect. This test was performed by Mookherjee and Sopher (1994) and by Van Huyck et al. (2001) who found that reducing such payoff information did have an effect on observed learning paths, which contradicts an important assumption of choice reinforcement learning. A similar study was conducted by Duffy and Feltovich (1999), who found that giving players information on the actions and payoffs of other pairs of players as well on as their own outcomes affected repeated play in the ultimatum game. This phenomenon of "observational learning" also provides evidence against simple choice reinforcement models.

While the above models all show that learning can take place when subjects receive feedback on payoffs and the actions of other players, they ignore the possibility that learning may take place in the absence of this information. Therefore, these models do not account for the possibility that people may learn simply through repeated experience with an environment. For instance, repetition and experience with a set of procedures might

[1] Any activity in which a task has to be performed repeatedly prior to obtaining feedback satisfies this description. An example is preparing several proposals (or papers, projects, etc.), one after another, which each take some time for review. Another example is whenever accurate performance feedback can only be obtained from a supervisor's evaluation, which may occur infrequently.

[2] For detailed discussion of these models and others, see Camerer (2003, Chapter 6).

lead people to obtain insights concerning how best to perform a task, even when they do not receive any feedback. This paper tests whether or not this type of learning from repetition without feedback can take place in a strategic environment.

The experiment in this paper is similar to previous studies in that it examines an extreme case of payoff information manipulation. This manipulation tests the assumption in all of the above models that learning in games takes place only through reinforcement by the observation of outcomes and payoffs. Using Nagel's (1995) competitive guessing game (see also, Ho et al., 1998), an experiment is conducted in which players receive no information on the actions of other players or on their own payoff. In these no-feedback treatments, all of the above models predict that no systematic change in behavior should take place with repeated play. Any convergence towards the Nash equilibrium would indicate that learning does not only take place through experienced payoffs and observed outcomes, but is also affected by experience with an environment and procedures and by repeatedly thinking about a game.

In addition to the main no-feedback manipulation and to explore how learning in the absence of feedback is determined by the extent to which subjects repeatedly think about strategic aspects of the game—even when they do not find out what others actually did—the no-feedback treatments varied in the extent to which subjects received cues for performing this type of strategic thinking. Specifically, subjects in different no-feedback conditions received different degrees of priming about strategic aspects of the game. If learning without feedback results from insights obtained by repeatedly thinking about and experiencing play of the game, then this priming should lead to higher levels of convergence towards the equilibrium.

The next two sections discuss the game used in the experiment and the experimental design. Section 4 presents the results. The paper ends with conclusions and a discussion of possibilities for future research.

2. The game

The game used in this experiment is one first studied experimentally by Nagel (1995) to determine the number of steps of iterated reasoning being satisfied by subjects. In the game, each of N subjects simultaneously chooses a number in the interval $[0, 100]$. The average of all the players' choices is then computed and multiplied by a parameter p to determine a target number. The player whose number choice is closest to this target number wins a fixed amount X and all other players receive nothing. If more than one player chooses the number closest to the target number, the amount X is divided equally among the winners. While Nagel uses several values of p (including values greater than one), the experiment in this paper uses $p = \frac{2}{3}$, which is one of the values studied by Nagel.

It is easy to see that the unique equilibrium in the game where $p = \frac{2}{3}$ is for everyone to choose 0. For any average of all players' choices, μ, the best response for all players is to choose $p\mu$, resulting in a new average, and this process has a unique fixed point at zero.

Nagel conducted four sessions with $p = \frac{2}{3}$ in each of which the game was repeated four times.[3] At the end of each period, subjects were informed of the choices of all other participants, the average, the target number, and the winning choices. There were 15 to 18 subjects in each session. No subjects chose the equilibrium strategy of 0 in the first period. The average in the first play of the game was 36.73 and several subjects chose numbers greater than $66\frac{2}{3}$, violating weak dominance.

With repeated play, the choices converged towards the Nash equilibrium. The average choice was 24.17, 16.14, and 9.52 in periods 2, 3, and 4, respectively. Thus, subjects in Nagel's experiment appeared to learn since their choices reflected a best response to the average in the previous period. The question this study asks is whether similar convergence toward the equilibrium prediction is observed when subjects receive *no* feedback on payoffs or outcomes. If so, this points to a kind of learning not captured by the current models.

3. Experimental design

In this study, Nagel's game with $p = \frac{2}{3}$ and $X = \$6$ was used in an experiment with four information conditions. In each session, the game was repeated ten times with 8 to 10 players.

In the *control* condition (C), the experimenter wrote the average, target number, and participant number(s) of the winner(s) on a board at the front of the room at the end of each period. This treatment serves to replicate, with feedback, Nagel's results.

In the *no-feedback no-priming* condition (NP), the game was played ten times, but subjects received no feedback at the end of each period. Instead, after the experimenter recorded each subject's choice at the end of period t, he simply stated: "This concludes period t. We will now move on to period $t + 1$. Please make your choice by writing it on your record sheet." At the conclusion of the tenth period, subjects were informed of the average, target number, and participant number of the winner or winners for all ten periods.

In the *no-feedback low-priming* condition (LP), subjects again received no feedback at the end of each period. However, in this treatment, at the end of each period the experimenter told subjects that he had calculated the average and target number and determined who the winner or winners were. The participants were informed that the experimenter had done this, *but were not told the results*, and were then asked to make a choice for the next period.

The final treatment was the same as the LP condition, with one exception. In this *no-feedback high-priming* condition (HP), subjects were also given no feedback until the end of the session, but after the experimenter calculated the average in each period, participants were instructed to write down their guess of the value of the average. While this guess did not provide subjects with any new information and their earnings were not affected by the

[3] Nagel's results have been replicated by Ho et al. (1998) using $p = 0.7$. For a detailed survey of existing research using this game, with varying parameters and payoffs, see Nagel (1998).

accuracy of their guess, it was introduced to aid participants in thinking about the fact that they wanted to best respond to their expectation of the average.[4]

The experiment was conducted using graduate and undergraduate students at the California Institute of Technology with little or no formal training in game theory. At the end of each session, subjects were privately paid their earnings in all ten periods plus a $7 participation bonus. Three sessions were conducted for each treatment ($n = 30$ in NP, $n = 28$ in both HP and LP, and $n = 26$ in C). Each session lasted 30 to 45 minutes.

Under the null hypothesis that learning does not take place without feedback, there should be no change in subjects' behavior across periods in the no-feedback treatments, but we should see convergence towards the equilibrium prediction in treatment C. On the other hand, if learning does take place by subjects simply gaining experience with the environment and having to think repeatedly about the game, we should see convergence towards 0 in all three no-feedback treatments. Finally, if priming subjects to think about strategic aspects of the game leads them to perform better the iterative reasoning required for equilibrium behavior to arise, even in the absence of feedback, then convergence towards the equilibrium in the no-feedback treatments should be greatest in the HP treatment and lowest in the NP treatment.

4. Results

Table 1 presents the mean and median choice by period for each treatment; the medians are also represented in Fig. 1.[5] In period 1, subjects' choices are very similar in all three no-feedback treatments (NP mean = 33.4, LP mean = 31.0, HP mean = 31.6) and there is no significant difference in behavior between the three conditions (NP–LP: $t_{56} = 0.49$, NP–HP: $t_{56} = 0.34$, LP–HP: $t_{54} = 0.12$).[6] The period 1 choices in the control condition are lower than in the other three treatments (mean = 24.6), but this difference is not significant for both the LP and HP treatments ($t_{52} = 1.35$ for both C–LP and C–HP comparisons) and only marginally significant for the NP treatment ($t_{54} = 1.70$, $p < 0.10$).

Figure 2 displays the cumulative frequency of first period choices in all four conditions. There is no significant difference between the frequencies using a two-tailed Kolmogorov-Smirnov test.[7] Thus, while initial choices differ slightly between treatments, these differences are not significant.

In all four treatments, the mean and median choices decreased between periods 1 and 10. As expected, the greatest decrease was in treatment C (decrease in mean = 18.1 (74%);

[4] The author originally conducted sessions with only treatments C, LP, and HP. Based on these results, which indicated more "learning" in condition HP than in condition LP, an anonymous reviewer suggested conducting the NP condition. As the results will indicate, this additional treatment does not alter the main result associated with the presence of learning without feedback, but does affect the interpretation of the role of priming.

[5] The full data set of subject choices and instructions are available from the author's web site: http://www.andrew.cmu.edu/user/rweber/.

[6] Note that these results are similar to, but slightly lower than, Nagel's first period results (mean = 36.7).

[7] The relevant test statistics are: C–LP: $D_{26,28} = 0.28$; C–HP: $D_{26,28} = 0.28$; C–NP: $D_{26,30} = 0.19$; LP–HP: $D_{28,28} = 0.18$; LP–NP: $D_{28,30} = 0.15$; HP–NP: $D_{28,30} = 0.20$.

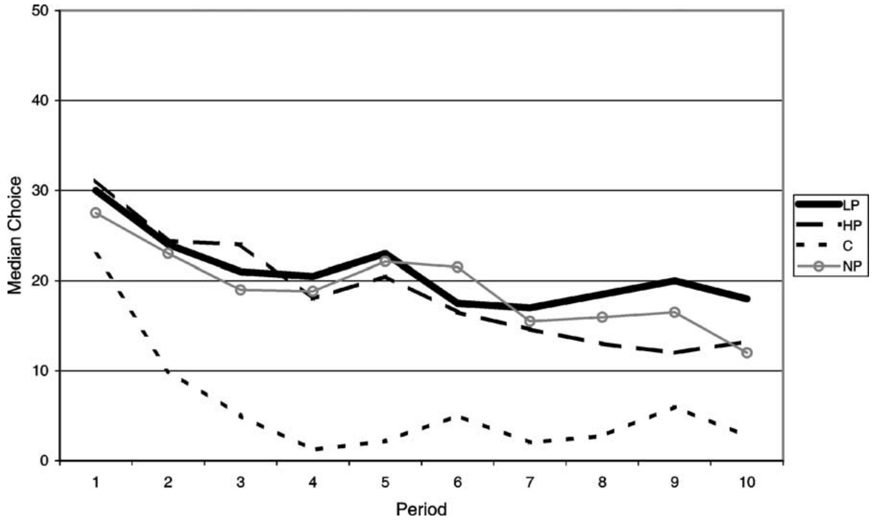

Fig. 1. Medians across periods.

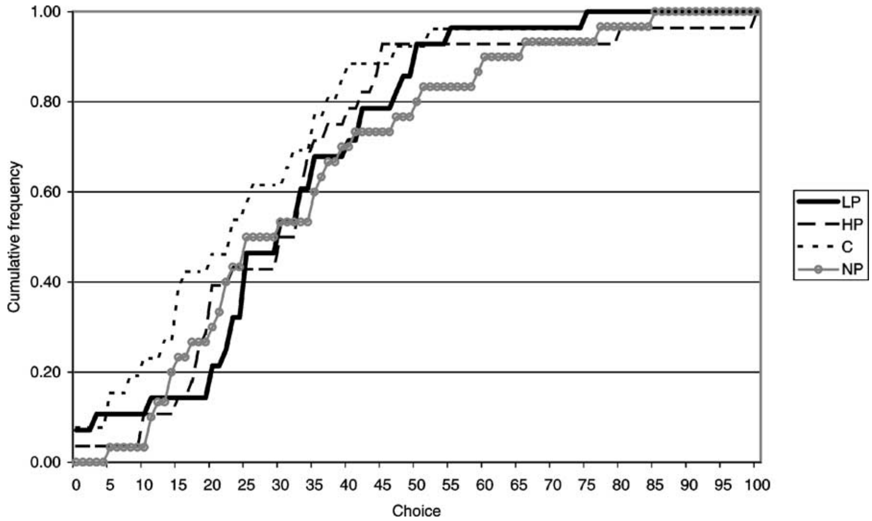

Fig. 2. Cumulative choice frequency—All treatments (period 1).

decrease in median = 20.2 (88%)).[8] Choices also decreased in the three no-feedback treatments, though not as much as in treatment C. In fact, the magnitude of the decrease

[8] The failure of the mean and median to decrease monotonically in treatment C was due to strategic behavior on the part of at least one subject in each session. Following a rapid initial convergence towards zero, these subjects made high choices (usually 100) in an attempt to cause other subjects to respond with choices that were too high in the following period. That this was the reasoning behind this behavior was determined in informal debriefing at the end of the session. However, the success of this strategy is questionable since no subject that chose 100 had the winning choice in the following period.

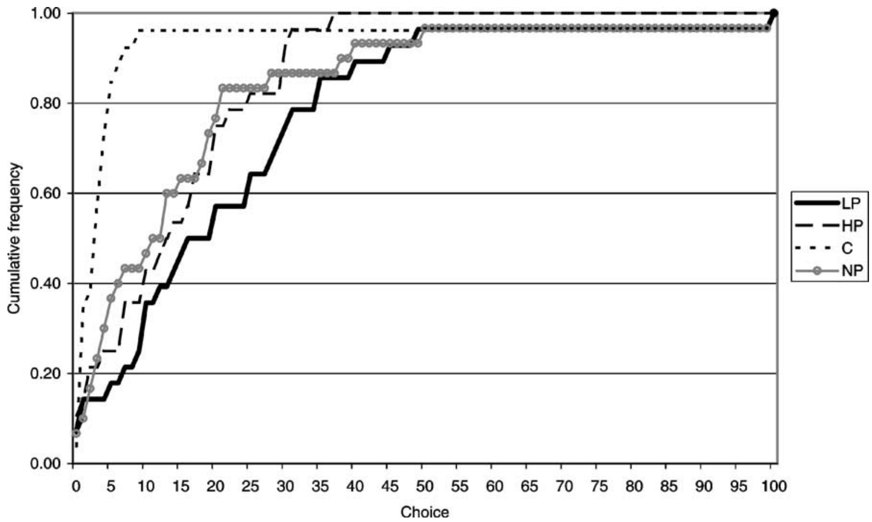

Fig. 3. Cumulative choice frequency—All treatments (period 10).

across 10 periods in the no-feedback treatments was roughly the same as the decrease in one period with feedback. Among the no-feedback treatments, the greatest decrease was observed in treatments HP (decrease in mean = 17.2 (54%); decrease in median = 17.9 (56%)) and NP (decrease in mean = 17.1 (51%); decrease in median = 15.5 (56%)). However, Fig. 1 reveals that the difference between these two treatments and treatment LP (decrease in mean = 8.8 (28%); decrease in median = 12.0 (40%)) is due mainly to differences in the final few periods.[9]

Figure 3 presents the cumulative frequency of choices for period 10 in all conditions. Comparing Figs. 2 and 3, the distribution of lower choices is greater in period 10 than in period 1 for all four conditions, indicating that subjects are revising their choices towards the equilibrium prediction even when they do not receive any feedback. However, the frequency of lower choices is higher in treatment C than in the other two treatments and these differences are significant using a one-tailed Kolmogorov–Smirnov test (C–NP: $X^2(2) = 15.54$, $p < 0.001$, C–LP: $X^2(2) = 27.30$, $p < 0.001$; C–HP: $X^2(2) = 21.72$, $p < 0.001$). Comparisons between the three no-feedback treatments indicate that none of the differences in tenth period choices are significant.

Additional evidence that subjects' choices are decreasing between periods 1 and 10 can be seen in Table 2. Table 2 presents, for each condition, the number of subjects who changed their choices between periods 1 and 10 in each direction. For instance, between

[9] Analysis of the standard deviation of choices is complicated by a few subjects' choices of 100. When these choices are excluded, the standard deviations generally decrease across periods in all four treatments. The decrease in the three no-feedback conditions is from about 18–20 in periods 1 and 2 to 12–14 in periods 9 and 10. When choices of 100 are excluded for treatment C, the decrease in standard deviation is from 18 in period 1 to about 3 in the final two periods. Thus, while the dispersion of choices also decreases in all four treatments, the convergence is not surprisingly greater with feedback.

Table 1
Summary of outcomes by treatment

Period	C		NP		LP		HP	
	Average	Median	Average	Median	Average	Median	Average	Median
1	24.6	22.9	33.4	27.5	31.0	30.0	31.6	31.1
2	16.4	10.0	32.6	23.0	27.2	24.0	28.8	24.4
3	6.7	5.0	23.7	19.0	24.4	21.0	28.3	24.0
4	6.2	1.2	22.6	18.8	19.6	20.5	22.4	18.0
5	12.1	2.2	21.6	22.1	22.5	23.0	22.1	20.5
6	5.4	5.0	24.1	21.5	18.4	17.5	17.4	16.5
7	9.6	2.0	20.7	15.5	18.2	17.0	17.1	14.6
8	11.2	2.8	18.0	16.0	18.0	18.5	17.3	13.0
9	8.4	6.0	17.3	16.5	19.0	20.0	15.7	12.0
10	6.5	2.7	16.3	12.0	22.2	18.0	14.4	13.3
diff. (P1 − P5)	12.5	20.7	11.8	5.4	8.4	7.0	9.5	10.6
diff. (P1 − P10)	18.1	20.2	17.1	15.5	8.8	12.0	17.2	17.9

Table 2
Direction of changes in subjects' choices between periods 1 and 10

	C		NP		LP		HP	
Choice increased	3	11.5%	2	6.6%	8	28.6%	0	0.0%
Choice unchanged	1	3.8%	3	10.0%	1	3.6%	4	14.3%
Choice decreased	22	84.6%	25	83.3%	19	67.9%	24	85.7%

periods 1 and 10, eight subjects in the LP condition increased their choices, 19 decreased their choices, and one subject chose the same number. Using a sign test, the null hypothesis that the underlying distribution of choices in condition LP is unchanged can be rejected ($p < 0.03$). The changes observed in the other three no-feedback conditions are even more extreme. No subjects in the HP treatment increased their choices between periods 1 and 10 while 24 subjects decreased their choices. In the NP treatment, only 2 subjects increased their choices and 25 lowered their choices (both significant at $p < 0.001$). Finally, as expected more subjects in treatment C also lowered (22) than increased (3) their choices ($p < 0.001$).[10]

The above results indicate that convergence towards equilibrium occurs in all four conditions.[11] This convergence is greater in the feedback control treatment than in the three no-feedback conditions, indicating that while behavior resembling learning does take place without feedback, the process is stronger when outcomes are revealed.

[10] While it is surprising that any subjects increased their choices in treatment C, two of the three increases were by subjects who initially chose zero. The other increase was by a subject who chose 100 in the final period, possibly out of frustration at not having won in previous periods.

[11] Figure 1 clearly indicates that choices in none of the conditions appear to be likely to reach exactly zero with more repetition. Therefore, the phrase "convergence towards zero" does not imply that choices are likely to reach the equilibrium with more repetition, but rather that play across ten periods indicates that choices move in the direction of zero.

The second question this experiment addresses is whether or not priming subjects to think about the strategic aspect of the game increases the convergence towards the equilibrium. While there is some evidence that “learning” occurs to a greater extent in the high-priming condition than in the low-priming condition, the effect of priming is reversed when comparing low-priming and no-priming. Moreover, there is no consistent significant difference in convergence between the three no-feedback conditions and the differences between treatment LP and the other two no-feedback treatments appear to result only from behavior in the final few periods. As the comparisons from Fig. 3 indicate, choices in the final period do not differ significantly between the three conditions. In addition, when the direction of change in Table 2 is limited to decrease vs. no decrease, there are no significant differences between any of the four conditions in a Fisher Exact test (also providing further support for the main hypothesis of learning without feedback).

5. Conclusion

The experiment reported in this paper addresses whether convergence towards equilibrium behavior can occur in repeated play of games without *any* feedback between periods. The results from all three no-feedback treatments provide strong support for this hypothesis, pointing to a form of learning not captured by the models in the literature. In all three conditions, choices decreased with repeated play. While the convergence was greater in treatment C, in which subjects received feedback, the fact that it took place at all in the other treatments indicates that learning can take place in the absence of information about outcomes and payoffs.[12]

A second question was whether priming (such as announcing that the mean had been calculated or requiring subjects to write down a guess of the value of the mean) would increase convergence towards the equilibrium. The results provide little support for this hypothesis.

These results show that something resembling learning can take place when subjects play games repeatedly, even when they receive no feedback on payoffs or the choices of other players. These results are important because they reveal the possibility that the majority of the current learning models are misspecified in that they only take into account learning through adaptation. Since these models ignore the kind of learning that takes place in the absence of feedback, parameter estimates for “feedback learning” might be biased.

An interesting question meriting further study is to what games this result can be extended. The game used in this experiment has a unique solution on the boundary of the strategy space which is the only strategy to survive iterated deletion of dominated strategies. It is not clear whether similar convergence toward equilibrium behavior would take place in games without this property. In particular, this result might be more difficult to extend to games with multiple equilibria requiring coordinated behavior—such as Battle-

[12] One possible concern is that the no-feedback phenomenon might be unique to Caltech students. However, two sessions ($n = 14$) were conducted at Stanford using the LP treatment. Mean (median) choices in these sessions also decreased by 18.4 (22.5), indicating that the phenomenon is not unique to the population in this study.

of-the-sexes or pure-coordination games.[13] On the other hand, subjects may learn to play the equilibrium without feedback in a large number games that require some insight or careful thought to determine the solution, such as the Dirty Faces game (see Weber, 2001) or variants of the Monty Hall problem (see Friedman, 1998).

Another important question has to do with how to model the "learning" taking place. As stated above, the results of the experiment raise the possibility that models that account for learning only through response to feedback concerning outcomes—which is the case with current models of learning—might be misspecified in the extent to which they attribute learning to adaptation to previous outcomes when part of this learning takes place even when outcomes are not revealed.[14] One possibility is that players may be using their own actions as an estimate of what other players may be doing, and may then best respond to this estimate. Such a model of "false consensus" learning, rooted in social psychology (Ross et al., 1977; Dawes, 1990), would predict the type of learning observed in this experiment. However, this approach would be proven incorrect if no-feedback learning occurs when no such reinforcement is possible. Instead, it might be the case that subjects are simply finding it easier to solve for equilibrium behavior with repeated experience.[15] For instance, repeated exposure to the game (and the environment in which the game is presented) may lower the cognitive costs of figuring out the equilibrium. While modeling this type of adjustment appears difficult, the above results indicate that it is necessary to do so.

Acknowledgments

This research was funded by NSF grant SBR-9631627 to Richard McKelvey and Tom Palfrey. The author thanks Tom Palfrey, Colin Camerer, Linda Babcock, Gary Bolton, Amnon Rapoport, Al Roth, and several anonymous referees for valuable comments and suggestions. Thanks also to participants at the 1998 Economic Science Association meetings in Tucson, the 1998 Society for Judgment and Decision Making meetings in Dallas, the 2000 World Congress of the Game Theory Society in Bilbao, and seminar participants at Carnegie Mellon for helpful comments and discussion.

References

Camerer, C., 2003. Behavioral Game Theory: Experiments in Strategic Interaction. Princeton Univ. Press, Princeton, NJ.

[13] However, Rapoport et al. (2002) report an experiment using a "market entry" coordination game in which there is also evidence of learning without feedback.

[14] It is possible to include a time-trend effect in many current models that captures this effect, as was pointed out by an anonymous referee. However, this trend is difficult to justify in several of the models that are based on the notion of "propensities" for strategies that are revised by knowledge about realized payoffs and actions of others. Building a time trend into these models basically argues that there is an unspecified effect that influences learning but is not part of the main mechanism assumed to underlie the learning in the model.

[15] Plott (1996) provides an informal theory of subject behavior in experiments and discusses how repeated choices and practice (in addition to feedback) lead subjects to behavior more consistent with the predictions of rational choice.

Camerer, C., Ho, T., 1999. Experience-weighted attraction learning in normal form games. Econometrica 67 (4), 827–874.

Cheung, Y., Friedman, D., 1997. Individual learning in normal form games: Some laboratory results. Games Econ. Behav. 19, 46–76.

Dawes, R.M., 1990. The potential nonfalsity of the false consensus effect. In: Hogarth, R.M. (Ed.), Insights in Decision Making. Univ. of Chicago Press, Chicago.

Duffy, J., Feltovich, N., 1999. Does observation of others affect learning in strategic environments? An experimental study. Int. J. Game Theory 28, 131–152.

Erev, I., Roth, A., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. Amer. Econ. Rev. 88 (4), 848–881.

Friedman, D., 1998. Monty Hall's three doors: Construction and deconstruction of a choice anomaly. Amer. Econ. Rev. 88 (4), 933–946.

Ho, T., Camerer, C., Weigelt, K., 1998. Iterated dominance and iterated best response in experimental p-beauty contests. Amer. Econ. Rev. 88 (4), 947–969.

Mookherjee, D., Sopher, B., 1994. Learning behavior in an experimental matching pennies game. Games Econ. Behav. 7, 62–91.

Nagel, R., 1995. Experimental results on interactive competitive guessing. Amer. Econ. Rev. 85, 1313–1326.

Nagel, R., 1998. A survey on experimental 'beauty contest games': Bounded rationality and learning. In: Budescu, D., Erev, I., Zwick, R. (Eds.), Games and Human Behavior: Essays in Honor of Amnon Rapoport. Lawrence Erlbaum Associates, New Jersey.

Plott, C.R., 1996. Rational individual behavior in markets and social choice processes: The discovered preference hypothesis. In: Arrow, K., Colombatto, E., Perlaman, M., Schmidt, C. (Eds.), The Rational Foundations of Economic Behavior. Macmillan, London.

Rapoport, A., Seale, D., Parco, J., 2002. Coordination in the aggregate without common knowledge or outcome information. In: Zwick, R., Rapoport, A. (Eds.), Experimental Business Research. Kluwer, Norwell, MA.

Ross, L., Greene, D., House, P., 1977. The "false consensus effect": An egocentric bias in social perception and attribution process. J. Exper. Soc. Psych. 13, 279–301.

Selten, R., Buchta, J., 1994. Experimental sealed bid first price auctions with directly observed bid functions. Discussion Paper B-270. University of Bonn.

Stahl, D.O., 2000. Rule learning in symmetric normal-form games: Theory and evidence. Games Econ. Behav. 32, 105–138.

Van Huyck, J., Battalio, R., Rankin, F.W., 2001. Selection dynamics and adaptive behavior without much information. Unpublished manuscript.

Weber, R.A., 2001. Learning and behavior in the 'dirty faces' game. Exper. Econ. 4 (3), 229–242.