Carnegie Mellon University
Research Showcase @ CMU

Department of Social and Decision Sciences | Dietrich College of Humanities and Social Sciences

4-1-2003

Learning and Transfer of Learning with No Feedback: An Experimental Test Across Games

Roberto A. Weber
Carnegie Mellon University, rweber@andrew.cmu.edu

Follow this and additional works at: http://repository.cmu.edu/sds

This Article is brought to you for free and open access by the Dietrich College of Humanities and Social Sciences at Research Showcase @ CMU. It has been accepted for inclusion in Department of Social and Decision Sciences by an authorized administrator of Research Showcase @ CMU. For more information, please contact research-showcase@andrew.cmu.edu.

Learning and transfer of learning with no feedback: An experimental test across games

Roberto A. Weber*
Social and Decision Sciences
Carnegie Mellon University
Pittsburgh, PA 15213

April 1, 2003

*This draft is preliminary and incomplete; please do not quote, circulate or criticize without the author's permission. Thanks to participants at the 2002 Economic Science Association meetings in Boston, the 2003 Public Choice / Economic Science Association meetings in Nashville, and to seminar participants at Carnegie Mellon and the University of Pittsburgh for helpful comments and suggestions. Thanks also to Jason Dana for valuable research assistance and for writing the software for the experiments. Alternate title "Learning with no Feedback 2: The Electric Boogaloo."

Abstract

This paper explores the extent to which people learn in repeated games without feedback, and the extent to which this learning transfers to new games. Current theories of learning model learning as adjustment in behavior in response to feedback about outcomes and payoffs and largely ignore the possibility that learning may take place in the absence of such feedback. An earlier paper (Weber, in press) demonstrates that convergence towards equilibrium can occur even in the absence of any feedback between plays of the game. However, this previous work demonstrates this "no-feedback" learning using a special game with properties that might facilitate this kind of learning and does not clearly identify mechanisms through which this learning might take place. The current paper extends this previous work. Using an experiment in which subjects play four games without feedback, the experiment allows a stronger test of "no-feedback learning" as well as a better understanding of the mechanism by which this learning occurs. The results show that, in every game, players behavior adjusts over time in the direction of Nash equilibrium. Subjects also appear to acquire principles useful in playing games – specifically, iterated dominance – and then apply these principles when playing subsequent games, demonstrating transfer of learning.

1 Introduction

The study of learning is an important area of research in economics and other social sciences. A considerable amount of this literature explores how individuals learn to make choices in strategic environments. This body of research – which usually explores both empirically and theoretically how subjects in laboratory experiments learn to play games – is important because of its application to strategic learning outside the laboratory. For instance, results on experiments in which subjects make quantity decisions in a Cournot oligopoly environment shed light on how strategic learning might lead to outcomes in real markets (e.g., Davis, 1999, Hück, Norman, and Oechssler, 1999).

A large number of experimental studies demonstrate that in most games subjects do not initially play equilibrium strategies, but that with repetition their behavior converges towards equilibrium (see Camerer, 2003). Several models attempt to provide a theoretical basis for this regularity (e.g., Camerer and Ho, 1999; Erev and Roth, 1998). While these models vary in how they assume learning takes place, they all share the feature that learning occurs through repeated observation of outcomes and payoffs. Players observe how well they did – or how well they could have done by behaving differently – and then adjust their behavior in subsequent plays of the game. These models ignore the possibility that a different kind of learning may simultaneously occur, absent of the effect of any feedback. Such "no-feedback" learning might occur because of increased familiarity with the strategic environment, because of the serendipitous discovery of important applicable principles, or because of introspection that occurs from repeatedly making choices – all in the absence of any feedback. The existence of this kind of learning might indicate that the current learning models exaggerate the influence of feedback-based learning on strategic behavior.[1]

This paper explores the possibility that learning may occur in the absence of feedback, using an experiment in which subjects play simple normal form games repeatedly but without any feedback between repetitions In all four games, behavior adjusts in the direction of Nash equilibrium – though in one of the games this is not surprising because of multiple equilibria – and violations of iterated dominance decrease in the three games that are dominance solvable. Equally importantly, there is clear evidence of transfer of learning across games: subjects are less likely to violate either one or two steps of iterated dominance in games played later than in those played earlier.

There are at least three reasons why the possibility of feedback-free learning is important for economics and game theory. First, current models of learning and dynamic behavior in repeated games all assume that change in behavior only occurs through adjustment in response to feedback. If the goal of these models is to accurately describe how behavior changes with repetition in strategic environments or to predict how people will learn in new games, then the occurrence of learning not predicted by these models poses a challenge for their descriptive accuracy and for the validity of the underlying mechanisms on which they

[1] It is worth noting that there exist models that allow subjects to behave more rationally as a result of introspection (e.g., Capra, 2003; Goeree and Holt, 2002; MacLeod, 2002). However, in these models it is assumed that all the introspection – or feedback-less learning – takes place before the first play of the game.

are based. Therefore, it is worthwhile to explore if – and how – this kind of no-feedback learning might occur.

Second, there are numerous examples of environments in the real world where people might make several decisions without receiving feedback about success or failure. For instance, employees in a firm might only receive performance feedback intermittently and after performing activities for considerable time. Someone working at a new task – such recording important infomration that might be used by others in the future – might try several approaches to the task (without receiving feedback about which one works best) before settling on one that he or she thinks is likely to be the best. Alternatively, in strategic environments, it might be the case that a strategic player has to make several decisions before finding out what an opponent has done. For instance, in foreign policy, a country's decision makers might make decisions over months or years before finding out what other countries' counterparts were doing. Therefore, understanding the kind of learning that is possible in these feedback-less environments – or whether learning is possible at all – is important.

Third, a common methodological practice in experimental economics is for subjects to play several games once each without any feedback, and to then treat each of the choices made by a subject as an independent one-shot play of the game (e.g., Guyer and Rapoport, 1972; Stahl and Wilson, 1994 & 1995; Costa-Gomes, Crawford and Broseta, 2002). However, if learning without feedback does occur – and if this kind of learning transfers across games – then a basic assumption underlying this method is not true and analysis of these data need to allow for the possibility that subjects learn while playing the games.

At least three previous experimental papers provide support for the hypothesis that learning can occur without feedback. Weber (in press) demonstrates that learning can occur in the absence of feedback. In experiments in which subjects play Nagel's (1995) competitive guessing game 10 times without any feedback between rounds, actions converge towards the unique Nash equilibrium prediction. While convergence towards the equilibrium is greater in sessions in which subjects receive feedback, the results nonetheless clearly indicate that adjustment in behavior consistent with learning occurs. In a market entry coordination games, Rapoport, Seale and Parco (2002) also find that equilibrium play increases without feedback. The magnitude of this result, however, is small. Finally, Grether (1980) finds that subjects' posterior probabilities are less random and more consistent with Bayesian updating if they have prior experience with a decision, even if they did not receive feedback.

While previous research points to the possibility of feedback-free learning, it also leaves important questions unanswered. First, the one study that directly tests this kind of learning – and provides the strongest evidence of its occurence – does so using an experiment that was specifically designed to give no-feedback learning the best chance of occurring. In Weber (in press), Nagel's competitive guessing game was used because it seemed like the most likely one to produce the phenomenon. Moreover, the no-feedback treatments were conducted under several "priming" conditions under which subjects were forced to consider strategic aspects of the game (such as others' choices), though this priming appeared not to affect behavior. Finally, the experiments were conducted using students at the California Institute of Technology, among which quantitative ability – which might help in figuring out the

equilibrium – is particularly high. Therefore, in order to truly demonstrate feedback-free learning as a robust phenomenon, it is necessary to demonstrate the result holds for other games and populations.

A second, and related, question is under what conditions no-feedback learning occurs. The competitive guessing game has several properties that might facilitate no feedback learning – it is symmetric, dominance-solvable, and has a unique equilibrium on the boundary of the strategy space. Without conducting further study, it is impossible to determine which, if any, of these factors facilitates no-feedback learning. Demonstrating that this kind of learning occurs in games with some of these properties but not in games with others would allow us to better understand how feedback-free learning occurs.

2 Theories of no-feedback learning

This section briefly describes three simple theories of how no-feedback learning might occur that all account for the behavior of subjects in a previous study of no-feedback learning (Weber, in press). The theories differ in their predictions of the conditions under which feedback-free learning should occur and are therefore empirically testable.

"Consensus" reasoning. People think that others are likely to take the same actions as themselves, resulting in a "false consensus" (Ross, Greene & House, 1977) – or "non-false consensus" depending on one's view of the irrationality of the behavior (see Dawes, 1990). Therefore, one possible explanation of no-feedback learning is that subjects simply take an action, treat their action as representative of the choices of other subjects, and then best respond to this belief. This kind of reasoning would predict convergence towards equilibrium in the competitive guessing game. Note, however, that it would not apply to games that are not symmetric and would therefore predict the absence of feedback-free learning in these environments.

"A-ha" learning. Playing rationally in a game often requires the understanding of certain principles that select or eliminate certain strategies. For instance, in the competitive guessing game, a subset of the strategy space is eliminated if a subject obeys dominance, and a further subset is eliminated if a subject obeys the second step of iterated dominance (i.e., believes others do not play dominated strategies). Therefore, in games that are dominance solvable – such as the competitive guessing game – making choices repeatedly, and thinking about the possible consequences of these choices, might lead to the acquisition of these principles and convergence towards Nash equilibrium. This kind of learning is possible for any of several possible principles that subjects might use to play games. Some concepts, such as dominance are more intuitive and easier to understand and, therefore, seem more likely to be acquired by this process (than, say, equilibrium refinements such as divinity). Therefore, since iterated dominance is perhaps the most relevant principle for deducing the unique Nash equilibrium in the competitive guessing game, tests of this theory will focus on whether subjects learn to apply iterated dominance. Note also that this theory of no-feedback learning implies transfer of learning across games: once a subject has acquired a principle it is likely to be applied to making choices in subsequent games where it is useful.

"Virtual feedback" learning. A final theory is that subjects behave as if they observed the actions of others (perhaps because they become better at guessing what others are doing), indicating that learning should be in the direction of what it would be under feedback. Weber and Camerer (in press) find that when subjects play a coordination game sequentially, but without observability of prior moves, the results indicate a small degree of "virtual observability," meaning that subjects behave as if they observed the actions of previous choosers. Perhaps no-feedback learning results from a similar process whereby subjects infer something about others' (previous) choices based on the fact that these choices were made previously, even though they were unobservable. Note that this kind of no-feedback learning would predict convergence towards equilibrium in coordination games, while neither of the previous theories would do so.

The three theories above simply represent three plausible explanations of how the no-feedback learning observed in the competitive guessing game might occur. They are not exclusive, since it might be the case that a combination of the three produces the change in behavior. At the same time, there are likely other explanations as well, and perhaps the correct theory lies outside those mentioned above. However, before proceeding with new experiments that further explore no-feedback learning, it is helpful to develop simple theoretical framework that the experiments can be used to test. Therefore, the experiment in the next section was designed with at least partly the goal of testing between the above three theories.

3 Experimental Design

The four games used in the experiment are presented in Table 1. The payoffs represent points, each of which equaled \$0.50 in earnings. The games have varying properties. Two of the games (A and B) are $2x2$, while the remaining two (C and D) are $3x3$. Only two games are symmetric (A and C).

Game A is a stag-hunt coordination game with three equilibria. The two pure strategy equilibria are a Pareto-optimal equilibrium (Top, $Left$) and a risk-dominant equilibrium ($Bottom$, $Right$) (Harsanyi and Selten, 1988). There is also an equilibrium in mixed strategies where Top and $Left$ are chosen with probability 0.555.

Game A

	Left	Right
Top	90, 90	0, 50
Bottom	50, 0	50, 50

Game B

	Left	Right
Top	75, 35	40, 40
Bottom	65, 100	35, 10

Game C

	Left	Middle	Right
Top	30, 30	35, 40	100, 35
Middle	40, 35	45, 45	65, 40
Bottom	35, 100	40, 65	60, 60

Game D

	Left	Middle	Right
Top	50, 50	40, 75	75, 55
Middle	20, 25	50, 65	65, 45
Bottom	90, 55	25, 30	60, 80

Table 1. Games used in the experiment

Games B, C, and D each have a unique equilibrium, which arises through iterated deletion of dominated strategies. To find the equilibrium for Game B, first recognize that *Bottom* should never be played (step 1), from which it follows that the column player should never play *Left* (step 2), giving rise to the equilibrium prediction of *Top*/*Right*. For Game C, which is symmetric, both *Bottom* and *Right* are dominated and once these strategies are eliminated (step 1), *Top* and *Left* should no longer be played (step 2), resulting in the equilibrium where both players choose *Middle*. In Game D, the column player's choice of *Left* is dominated. Once this strategy is removed (step 1), *Bottom* should never be played (step 2) and, as a consequence of this, the column player should never play *Right* (step 3). Once the column player has eliminated both *Left* and *Right*, the row player should no longer play *Top* (step 4). Thus, the equilibrium is also where both players choose *Middle*.

The different properties of the four games allow a test of the three theories discussed in the previous section. "Consensus" learning applies only to symmetric games, predicting that we should observe convergence to equilibrium in the two symmetric games (A and C). On the other hand, if subjects are learning exhibiting "a-ha learning" and acquiring principles such as dominance, then we should observe similar decreases in violations of dominance across the three games where it applies. Similarly, subjects might acquire principles such as

Pareto-optimality, risk dominance, or mixed-strategy equilibrium indicating how behavior might change in Game A, though these seem less intuitive. Finally, if subjects are learning through virtual feedback, then subjects' choices should exhibit convergence towards actions that are best responses to the choices of others in early rounds.

There were four sessions. In each session a group of 18 to 10 subjects played the above four games for 20 rounds each. The sequence of games is presented in Table 2.

	Game 1	Game 2	Game 3	Game 4	N
Session 1	D	A	C	B	18
Session 2	C	B	A	D	18
Session 3	A	D	B	C	18
Session 4	B	C	D	A	20

Table 2. Sequence of games by session

Subjects' roles (Row or Column) were fixed for the entire experiment. Each subject was anonymously paired with someone of the opposite role for the duration of the experiment. Matching and role assignment were done by random assignment of participant numbers.

Each subject sat a computer monitor. At the beginning of each session, subjects received extensive instruction in how to interpret game matrices with generic payoffs.[2] They were quizzed about this knowledge and the experiment proceeded only after every subject correctly answered the questions and the experimenter publicly read the correct answers.

At the beginning of each play of 20 rounds, the computer monitor displayed the game matrix for those rounds. The experimenter read all of the payoffs in the matrix aloud. Subjects then proceeded to play 20 rounds of the game by clicking on a choice. After each choice, the computer screen froze, displaying the subject's choice and the matrix for 20 seconds.

Subjects were graduate and undergraduate students at the University of Pittsburgh recruited from an e-mail list of participants interested in participating in decision experiments. The experiment lasted approximately 1.5 hours. Subjects were informed at the beginning of the experiment that their earnings would be determined exclusively by the points they accumulated during the 80 rounds of play. At the end of the experiment, subjects were privately paid one at a time. Earnings ranged from \$11.25 to \$28.95, with average earnings of \$21.36.

4 Results

We begin by exploring the results for each game, pooling the results across sessions, to test whether learning occurs during the 20 rounds that a game is played. We will then turn our attention to whether there is any transfer of learning across games.

[2]Instructions are in the appendix.

4.1 Learning within games

The aggregate choice frequencies – by 5-period blocks – are presented in Table 3.[3]

While behavior appears to change across rounds, we are interested in systematic changes consistent with learning. Therefore, we can explore whether choices converge towards equilibrium with no-feedback repetition. In Game A, average choices lie between the Pareto-optimal and mixed strategy equilibria for all four blocks of periods. The frequency of $Top/Left$ choices decreases from the first block to the last one (primarily in the first 10 periods), indicating that subjects are moving away from the Pareto-optimal equilibrium and towards the mixed strategy and risk-dominant equilibria. Using a paired t-test to compare the mean of subjects' choices in the first block with the last reveals a significant difference ($t_{73} = 3.22$, $p < 0.005$ two-tailed). While behavior converges towards equilibrium with repetition, this is not surprising since any change in behavior would do so.

In Game B, the unique equilibrium is $Top/Right$. The frequency of both of these choices increases from the first block to the last. When we compare the average of subjects' choice frequencies in the first block with the last block, we see that subjects play equilibrium strategies significantly more in the last 5 periods than in the first 5 ($t_{73} = 1.56$, $p < 0.07$, one-tailed).

The unique equilibrium in Game C is for both players to choose $Middle$. As Table 3 indicates, the frequency with which subjects played the equilibrium strategy increased across all five-period blocks. A comparison of subjects' choices in the first five periods with the last five reveals a significant difference ($t_{73} = 2.11$, $p < 0.02$, one-tailed).

Finally, for Game D the equilibrium is also $Middle/Middle$. The frequency with which subjects played these strategies increased from the first block to the last, though this increase is not monotonic. The difference in the frequency of equilibrium choices between the first five periods and last five periods is significant ($t_{73} = 1.79$, $p < 0.04$, one-tailed).

[3]The full data set of choices is available on the author's website: *http: www.andrew.cmu.eduřweber.*

Periods	1-5	6-10	11-15	16-20
Game A				
Top/Left	0.765	0.668	0.665	0.659
Bottom/Right	0.235	0.332	0.335	0.341
Game B				
Top	0.741	0.768	0.827	0.811
Bottom	0.259	0.232	0.173	0.189
Left	0.654	0.654	0.611	0.627
Right	0.346	0.346	0.389	0.373
Equilibrium	0.543	0.557	0.608	0.592
Game C				
Top/Left	0.200	0.186	0.224	0.189
Middle (Equil.)	0.603	0.624	0.630	0.678
Bottom/Right	0.197	0.189	0.146	0.132
Game D				
Top	0.530	0.508	0.584	0.589
Middle	0.092	0.151	0.103	0.119
Bottom	0.378	0.341	0.314	0.292
Left	0.076	0.049	0.049	0.049
Middle	0.243	0.308	0.270	0.314
Right	0.681	0.643	0.681	0.638
Equilibrium	0.168	0.230	0.186	0.216

Table 3. Choice frequencies by 5-period blocks

Subjects' behavior changed significantly in *all* four games. The total frequency of choices consistent with equilibrium increased in *all* three games for which it could do so nontrivially. In the remaining game subjects appeared to move away from the Pareto-optimal equilibrium towards the risk-dominant and mixed strategy equilibria. While it is clear that subjects are learning – their behavior shows systematic change in the direction of the predictions of rationality – how or what they are learning is not as clear at first glance.

To better understand the no-feedback learning taking place, it is worthwhile to consider the three theories of no feedback learning discussed earlier in the paper. The "virtual feedback" learning theory is not supported by the data. Subjects' choices change in Game A, where this theory predicts that they should. However, based on the distribution of choices in the first period (where 82 percent of subjects chose $Top/Left$), virtual feedback predicts that the frequency of choices should *increase*, since this is what would happen if actions were observable. In fact, the opposite occurs. Therefore, the trend in Game A clearly goes against that predicted by virtual feedback.

Similarly, "consensus" learning predicts that the frequency of $Top/Left$ choices should increase in Game A (subjects who choose $Top/Left$ initially – or who do so with probability greater than 0.55 – should increase the frequency with which they make this choice subsequently). Moreover, consensus learning predicts that learning should not occur in the non-symmetric games, which is not supported by the data.

This leaves the question of whether "a-ha" learning can adequately explain the data. The change in behavior in Game A can be trivially explained by this theory, since any change in behavior can be explained by subjects acquiring a principle (such as Pareto-dominance or risk-dominance). A better test of this theory might be to explore how subjects learn in the three dominance-solvable games. The principle that best applies to these games is iterated dominance. "A-ha" learning applies if subjects learn how to apply up to four steps of iterated dominance. Therefore, subjects should learn to apply these steps similarly across the three games.

If subjects acquire the first step of iterated dominance, the prediction is clear: the frequency of *Bottom* choices in Game B, *Bottom/Right* choices in Game C, and $Left$ choices in Game D should all decrease since these strategies are dominated. A comparison of the first five periods with the last five in Table 3 indicates that this is true for all three games. These changes are significant for Game B ($t_{36} = 2.49$, $p < 0.01$, one-tailed), Game C ($t_{73} = 2.34$, $p < 0.02$, one-tailed), but not for Game D ($t_{36} = 1.15$, $p < 0.13$, one-tailed), though the frequency of choices in the last game starts off very close to the boundary (0.076), not leaving very much room for adjustment. A joint test pooling across the three games is highly significant ($t_{147} = 3.16$, $p < 0.001$, one-tailed).

Acquiring the second step of iterated dominance means subjects should be less likely to play $Left$ in Game B, $Top/Left$ in Game C, and *Bottom* in Game D. Looking at Table 3, we see that all three of these predictions are also supported in a comparison of the first five rounds versus the last five. This difference is not significant for Game B ($t_{36} = 0.48$ or for Game C ($t_{73} = 0.42$), but is significant for Game D ($t_{36} = 2.21$, $p < 0.04$, one-tailed). A pooled comparison across games is significant ($t_{147} = 1.58$, $p < 0.06$, one-tailed).

The third and fourth steps of iterated dominace only apply to Game D. The prediction for the third step is that column players should be less likely to play *Right*. The difference between the frequency of *Right* choices in the first block and the last is in the correct direction, but is not significant ($t_{36} = 0.84$). Finally, the prediction for the fourth step – that Row players should play Top less often – does not hold, since the change is in the opposite direction.

The results above shed light on what kind of learning is taking place, and how behavior changes in the direction of equilibrium in Games B, C, and D. It appears that subjects acquire up to two steps of iterated dominace with feedback-free repetition. However, if subjects are really learning how to apply these principles, then we would expect that their acquisition to transfer across games. In fact, evidence of transfer of learning would constitue convincing evidence that subjects have indeed learned how to play games better.

4.2 Transfer of learning across games

We now turn to exploring whether learning transfers across games. Games B, C, and D all have in common that iterated deletion of dominated strategies implies that certain choices should not be made. As we find in the previous section, subjects' behavior implies the acquisition of these strategies across the 20 periods in which they play the games. An equally important question, however, is whether this learning spills over into later games in which it can be applied. This is important for two reasons. First, if we are interested in testing whether "a-ha" learning truly holds – and subjects learn to use principles such as iterated dominance wih feedback-free repetition – then a good test of this theory is that the realization of how to use a principle should imply that principle is subsequently possesed by a subject in other instances where it may be useful. Second, demonstrating this kind of spillover would indicate that subjects are really learning, since their play improves (in the sense that they no longer violate dominance or iterated dominance) across games.

To test whether or not learning transfers across games, we compare the frequency with which subjects make choices that violate steps of iterated dominance by games' positions within sessions. Since iterated dominance does not apply to Game A, we consider only Games B, C, and D. Table 4 presents, by game position, the frequency with which subjects violated different steps of iterated dominance.

Table 4 compares behavior in exactly the same three games.[4] Therefore, a decrease in the frequency of violations means that subjects are more often making choices consistent with that principle. The frequency of violations of the first two steps of iterated dominance clearly decreases with game position. For both D1 and D2, roughly 10 percent less choices violated the principle in the last game first than in the first, and a comparison of these frequencies (by subject) is significant (D1: $t_{71} = 1.39$, $p < 0.09$, one-tailed; D2: $t_{71} = 1.47$, $p < 0.08$, one-tailed). However, for both the third and fourth steps of iterated dominance, we see that there is no decrease in the frequency with which violations occur, indicating that subjects are not acquiring these principles, which is consistent with the game-level results. This is also expected because Game D is only encountered once by each subject, eliminating the possiblity of acquiring the principle in one game and then applying it in a subsequent one.

	1st game	2nd game	3rd game	4th game
D1 violations	0.205 (37)	0.150 (38)	0.123 (37)	0.122 (36)
D2 violations	0.393 (37)	0.339 (38)	0.347 (37)	0.286 (36)
D3 violations	0.606 (9)	0.656 (9)	0.695 (10)	0.683 (9)
D4 violations	0.339 (9)	0.661 (9)	0.465 (10)	0.756 (9)

D1: First step of iterated dominance (B in Game B, B/R in Game C, L in Game D)

D2: Second step of iterated dominance (L in Game B, T/L in Game C, B in Game D)

D3: Third step of iterated dominance (R in Game D)

[4]The number of subjects in each cell varies because session 4 had a different number of participants (20) than the other sessions (18).

D4: Fourth step of iterated dominance (T in Game D)

Number of subjects making relevant decision in parentheses

Table 4. Violations of iterated dominance by game position

Another implication of "a-ha" learning is that subjects who acquire a principle should never subsequently violate that principle. To explore this, we considered the number of subjects in the first and last games who violated each of the four steps of iterated dominance. For the first step, 12 of 37 subjects (who had an opportunity to do so) never played a dominated strategy in their first game. When we look at the fourth game, however, this proportion goes up to 21 of 36 ($p < 0.03$, Fisher's exact). Similarly, for the second step, every subject plays a strategy that is dominated once the opponents dominated strategies are removed at least once in their first game (0 subjects with no violations). However, in the last game, 13 of 36 subjects never play such strategies ($p < 0.001$, Fisher's exact).[5] Therefore, we again have evidence that subjects are learning to play consistently with up to two steps of iterated rationality.

The decrease in violations of iterated rationality appears to be at least party driven by a few subjects who acquire the principle and do not subsequently violate it. This raises the possibility of heterogeneity in no-feedback learning. To explore this possibility, we can look at the individual choices made by subjects accross the experiment.

During the experiment, every subject made 80 choices. Of these choices, 40 choices allowed the opportunity to violate the first step of iterated dominance and 40 allowed the opporunity to violate the second step.[6] Therefore we can explore – across a subject's 40 choices – when, if at all, that subject stopped violating the first two steps of iterated dominance.

Figure 1 reports the distribution of points (in the 40 play sequence for each subject) after which a subject no longer violated the first or second step of iterated dominance. Recall that the 40 periods represented on the horizontal axis correspond to the period during which a subject could violate each principle, and therefore do not represent the same 40 periods for each subject or for each principle. Each bar represents the number of subjects whose last violation of the principle was in that period. Note that there is substantial heterogeneity in these cutpoints. There are substantial proportions of subjects who never violate dominance or the second step of iterated dominance as well as several who continue to violate them in the last period. The first step also appears to be easier to acquire than the second step, since many more subjects never violated D1 than D2, but many more subjects continue violating D2 in the last period than D1. In fact, the modal behavior is to never violate D1, but to

[5] When we look at the third and fourth step, we find that only one subject never plays a violating strategy in the first and fourth games (one of nine subjects never plays *Top* in Game D in the first game).

[6] Row players could violate the first step of iterated dominance when choosing in Games B and C and the second step when choosing in Games C and D. Column players could violate the first step when choosing in Games C and D and the second step when choosing in Games B and C.

continue to violate D2 until the last period.

Insert Figure 1

Since the acquistion of the first step of iterated dominance possibly facilitates (or may even be necessary for) the acquisition of the second step, we explore the relationship between the two. A comparison of individual subject behavior is presented in Table 5. The table classifies all 74 subjects into categories depending on their pattern of violations of the two principles over the 40 periods in which the principle applied. For each principle, subjects are classified as never violating, "learning" (not violating the principle in the last five periods in which they could do so), or "never learning" (violating the principle at least once in the last five periods) the principle.[7]

The table reveals that only a small number of subjects (4) never violated either step, but a larger number of subjects (16) continued to violate both steps in the final five periods in which they could do so.

	Never violate D2	"Learn" D2	"Never learn" D2	Total
Never violate D1	4	12	7	23
"Learn" D1	1	13	18	32
"Never learn" D1	2	1	16	19
Total	7	26	41	74

D1: First step of iterated dominance (B in Game B, B/R in Game C, L in Game D)

D2: Second step of iterated dominance (L in Game B, T/L in Game C, B in Game D)

Table 5. Violations of iterated dominance by subject

Interestingly, the proportion of subjects who never violated D1 but violated D2 at least once (19 of 23, 83%) is considerably higher than the proportion who never violated D2 but violated D1 at least once (3 of 7, 43%). This difference – significant at the $p < 0.06$ level in a Fisher's exact test – indicates that it is much more likely that a subject who knew D2 also knew D1 than vice versa. Similarly, of the 19 subjects who are classified as never learning D1 only 3 (16%) learned D2, but of the 41 subjects who never learned D2, 25 (61%) learned D1. This difference is also significant ($p < 0.01$) in a Fisher's exact test. Both of these

[7] Of course, this classification method (choosing the 5th from last period as an arbitrary cutoff) is arbitrary and imperfect. Two other classification methods were used: 1) Using the last period instead of the fifth from last and 2) Classifying as "learners" only those players for whom the last run of no violations can be rejected (at the 0.05 or 0.1 level) as resulting from the distribution that generated that subject's previous choices. The main result discussed with respect to Table 5 are unchanged when using these classifications. Moreover, these classification methods similarly suffer from arbitrariness and potentially mistaken classifications.

comparisons indicate that it is considerably easier to acquire D1 without acquiring D2 than it is to acquire D2 without acquiring D1.

Table 6 presents further evidence that the aqcuisition of D1 is a necessary condition for the acquisition of D2, but not vice versa. The table reports the frequency of choices that violated each principle for each category of subjects. Not surprisingly, the frequency of violations increases for subjects who never violated, learned and never learned that particular principle. However, when we compare across principles, the frequency of D2 violations is lowest for subjects who never violated D1 (0.205), in the middle for subjects who learned D1 (0.366), and highest for subjects who never learned D1 (0.467). This monotonic relationship does not hold for D1 violations (0.139, 0.056, 0.211). This is consistent with the above results indicating that acquisition of the first step of iterated dominance is necessary for the acquisition of the second step.

Classification	Frequency of D1 violations	Frequency of D2 violations	N
Never violated D1	0.000	0.205	23
"Learned" D1	0.122	0.366	32
"Never learned" D1	0.378	0.467	19
Never violated D2	0.139	0.000	7
"Learned" D2	0.056	0.241	26
"Never learned" D2	0.211	0.464	41

D1: First step of iterated dominance (B in Game B, B/R in Game C, L in Game D)

D2: Second step of iterated dominance (L in Game B, T/L in Game C, B in Game D)

Table 6. Frequencies of violations of iterated dominance by classification

Both the first and second steps of iterated reasoning are more likely to be applied by subjects later in the experiment than earlier, which is consistent with subjects' learning these principles during the experiment. Moreover, we also find that the acquisition of the second step is more likely for subjects who have acquired the first step, which is consistent with a theory of no-feedback learning whereby subjects realize how to apply principles with feedback-free repetition.

5 Conclusion

The experiment reveals substantial support for feedback-free learning. In all four games, behavior converges towards a Nash equilibrium (in three games, this is towards the unique equilibrium). Moreover, there is clear evidence (from Games B, C, and D) that subjects acquire the ability to use one and two steps of iterated dominance, and that this acquistion varies considerably across subjects. There is also sustantial change in behavior in Game A, consistent with the acquisition of risk-dominance or mixed strategy equilibrium, but it is

difficult to draw any firm conclusions about what principle subjects are applying based on the results of one game.

Perhaps one of the most striking pieces of evidence in support of no-feedback learning is the clear evidence of transfer of learning across games. Subjects significantly make less choices that are ruled out by one and two steps of iterated dominance in later games than in earlier games. This is especially surprising given that clear evidence of transfer of learning across games has only been found in a handful of experiments *with feedback* (e.g., Cooper and Kagel, 2001; Ho, Camerer and Weigelt, 1998).

There are three main reasons why the above results are important. First, the clear support for the acquistion on the part of subjects of up to two levels of iterated dominance sheds light on how feedback-free learninng is taking place. By repeatedly making choices, some subjects are realizing the implications of a choice never being better than another, and what this implies for responding to the choices of others.

Second, the ability to apply principles to novel situations is truly evidence of learning. Therefore, while the magnitude of the learning reported above is weak compared to the adjustment in games with feedback, the fact that subjects learn not to violate important principles of game theory implies that they have actually learned something about playing games. This is not necessarily the case in experiments with feedback, where adjustment in behavior may be due to increased attraction to certain strategies (as a function of how those strategies have fared in the past) but may not actually involve any real learning to the extent that the subject's ability to think as a game theorist has changed. The experiments above show much clearer evidence of this kind of learning.

Finally, the experiments raise an important methodological question. The practice of repetition without feedback has long been used by those doing experiments in game theory as a way of recreating the one-shot play of several games (e.g, Stahl and Wilson, 1994 & 1995; Rapoport and Guyer, 19??; Costa-Gomes, Crawford and Broseta, 2002). However, transfer of learning in the above no-feedback experiments implies that this assumption may not be valid.

Of course, there is still considerable work to do. While the experiments reported here answer several questions about how learning can take place without feedback, the exploration of this phenomenon must still be extended to other games. Moreover, while there is clear evidence of the acquisition of iterated dominance in the above results, this only raises the question of what other kinds of principles can be acquired in the absence of feedback.

6 References

References

[1] Camerer, C. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction.* Princeton, NJ: Princeton University Press.

[2] Camerer, C. and T. Ho. 1999. Experience-weighted attraction learning in normal form games. *Econometrica* 67(4): 827-874.

[3] Capra, M. 2003.

[4] Cooper, D. and J. Kagel. 2001. Learning and transfer in signaling games. Case Western Reserve University Department of Economics working paper.

[5] Costa-Gomes, M., V. Crawford and B. Broseta. 2001. Cognition and behavior in normal-form games: An experimental study. *Econometrica*, 69(5): 1193-1235.

[6] Davis, 1999

[7] Dawes, R. M. 1990. The potential nonfalsity of the false consensus effect. In *Insights in Decision Making*, ed. R. M. Hogarth. Chicago: University of Chicago Press.

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

[9] Goeree and Holt. 2002.

[10] Grether, D. M. 1980 Bayes Rule as a descriptive model: The representativeness heuristic. *The Quarterly Journal of Economics*, 96(3): 537-557.

[11] Guyer and Rapoport, 1972.

[12] Harsanyi, J. and R. Selten. 1988. *A General Theory of Equilibrium Selection in Games.* Cambridge, MA: MIT Press.

[13] Ho, T., C. Camerer and K. Weigelt, K. 1998. Iterated dominance and iterated best response in experimental p-beauty contests. *The American Economic Review*, 88: 947-969.

[14] Hück, Norman, and Oechssler, 1999

[15] MacLeod, 2002.

[16] Nagel, R. 1995. Experimental results on interactive competitive guessing. *American Economic Review* 85: 1313-1326.

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

[18] Ross, L., D. Greene and P. House. 1977. The false consensus effect: An egocentric bias in social perception and attributional processes. *Journal of Experimental Social Psychology*, 13: 279-301.

[19] Stahl, D. O. II, and Paul W. Wilson. 1994. Experimental Evidence on Players' Models of Other Players. *Journal of Economic Behavior and Organization* 25: 309-327.

[20] Stahl, D. O. II, and Paul W. Wilson. 1995 On Players' Models of Other Players: A New Theory and Experimental Evidence. *Games and Economic Behavior* 10: 218-254.

[21] Weber, R. A. In press. "Learning" without feedback in a competitive guessing game. *Games and Economic Behavior*.

[22] Weber, R. A. and C. F. Camerer. In press. Timing and virtual observability in ultimatum bargaining and "weak link" coordination games. *Experimental Economics*.

7 Appendix

Instructions

This is an experiment in decision making. Several institutions have provided funds for this research. You will be paid for your participation. The exact amount you receive will be determined during the experiment and will depend on your decisions and the decisions of others. This amount will be paid to you in cash at the conclusion of the experiment. If you have a question during the experiment, raise your hand and an experimenter will assist you. Please do not talk, exclaim, or try to communicate with other participants during the experiment. Participants violating the rules will be asked to leave the experiment and will not be paid.

In this experiment, you will play a series of simple games in which you will accumulate money based on your choices and the choices of one other person in the experiment. You will not know the identity of this other person, either during or after the experiment, just as the other person will not know your identity.

In the experiment, you will play four different games. You will play each game a total of 20 times. Each time you play a game, you may accumulate points based on your choice and the choice of the person you are playing with. At the end of the experiment, you will be paid an amount (in cents) equal to $\frac{1}{2}$ times the number of points you have accumulated. Therefore, each point you accumulate playing the games is worth $\frac{1}{2}$ cents.

In each game, there will be two players. One of these players will be the row player, and the other will be the column player. Below is an example of the kind of game you will play:

		Columns choice	
		Left	Right
Rows	Top	6, 3	1, 8
choice	Bottom	4, 5	7, 2

Each of the games you will play will be presented similarly to the table above, though the numbers above are just for example.

In all the games you will play, there are two players, one labeled "Column" and the other labeled "Row." In the game above, each player has two choices (Row can choose "Top" or "Bottom" and Column can choose "Left" or "Right"). This means that there are four possible combinations of the players choices and four possible outcomes (Top-Left, Top-Right, Bottom-Left, Bottom-Right). Inside the square that corresponds to each of these outcomes are the payoffs for each player. These are the points that each player will receive if that outcome occurs. The number on the lower left of the square represents the payoff for the Row player, while the number on the upper right represents the payoff for the Column player.

Therefore, in the game above:
If Row picks "Top" and Column picks "Left," then Row receives 6 points and Column receives 3 points,

If Row picks "Top" and Column picks "Right," then Row receives 1 point and Column receives 8 points,
If Row picks "Bottom" and Column picks "Left," then Row receives 4 points and Column receives 5 points,
If Row picks "Bottom" and Column picks "Right," then Row receives 7 points and Column receives 2 points.

In the experiment, you will play four games similar to this one. In some of the games, players may have three choices instead of two. If this is the case, then there will be three rows and three columns. Otherwise, the way that the game is presented will be the same. You will play each game 20 times, and will then move on to the next game.

During the experiment, you will be able to make choices by using the mouse to click on a button that corresponds to the choice you wish to make. After clicking, you will be asked to confirm your choice. After confirming this choice, you will not find out what the other person did or how many points you received. You will simply wait 20 seconds and will then make another choice. However, each choice you make counts towards the total number of points you receive and therefore affects the amount of money you will receive at the end of the experiment. You will just not know how much you have received until the end of the experiment.

Are there any questions before we start with the experiment?

Once we are ready to begin, please take an envelope from the experimenter. Inside the envelope, you will find a card with your participant number. This is your participant number for the experiment. This participant number is private and should not be shared with anyone. Your participant number will be the same for the entire experiment.

If your participant number is odd, then you will be a Row player for the entire experiment. If your participant number is even, then you will be a Column player for the entire experiment.

We will now play each of the four games 20 times. After each time that you play a game, please write down your choice on the Record Sheet on the next page. After playing the last of the four games, you will find out how many points you accumulated during the experiment. At the end of the experiment, you will receive your payoffs (in cents) equal to $\frac{1}{2}$ times the number of points you have accumulated.

We are now ready to begin the experiment. There should be no talking from this point on. If you have a question, please raise your hand. When proceeding through the experiment, you should wait at the end of each game for the experimenter to instruct you to proceed to the next game.
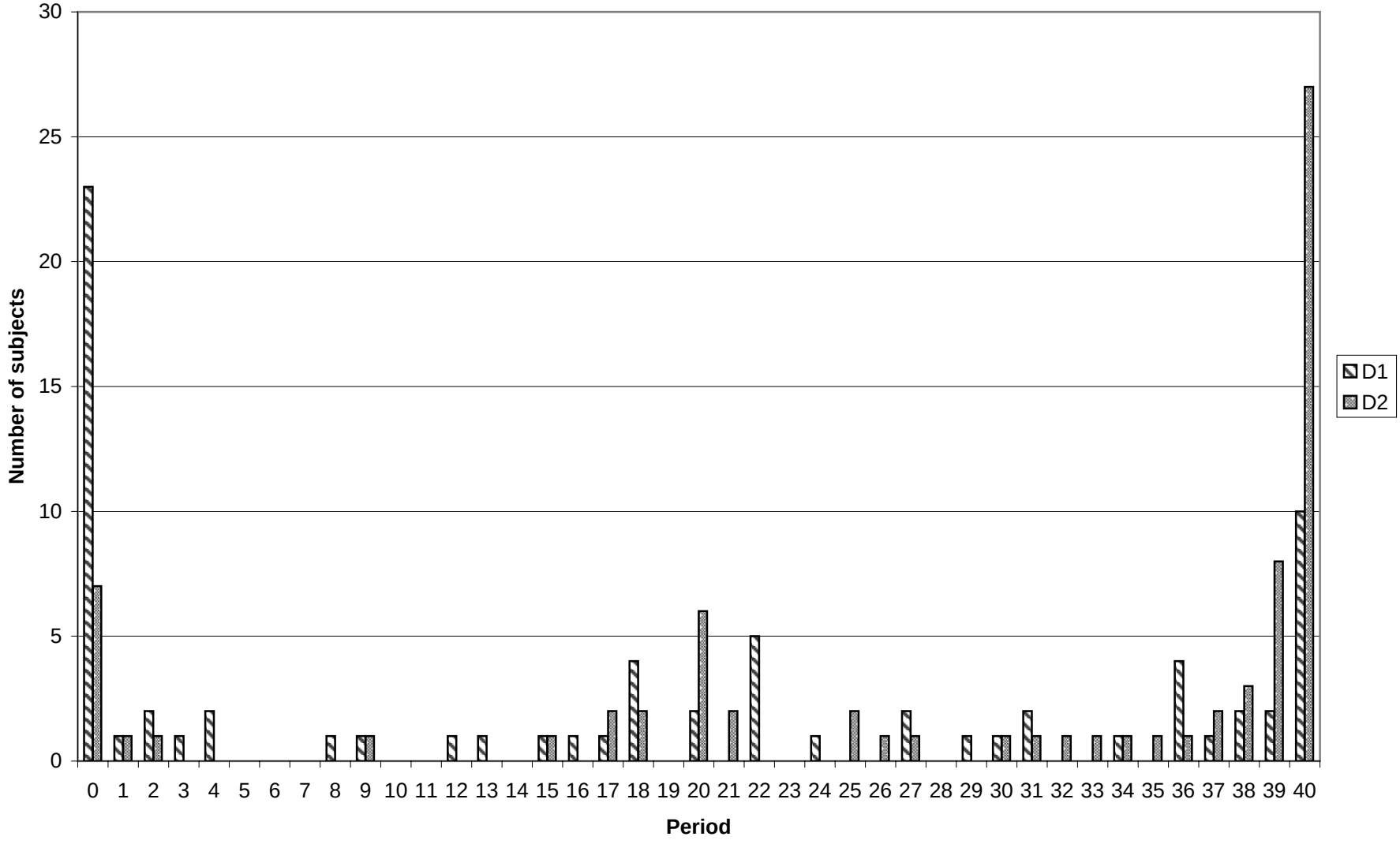
Figure 1. Frequency of subjects' last period of violation of D1 and D2