Learning Transference Between Dissimilar Symmetric Normal-Form Games

by

Ernan Haruvy
University of Texas at Dallas

and

Dale O. Stahl
University of Texas at Austin

March 2009

ABSTRACT. Learning and adaptive models in the economics literature are built on the premise that actions that have yielded higher payoffs in the past are more likely to be selected in the present. Such action-based learning models are not directly applicable to learning that occurs between situations where the set of actions is not directly comparable from one situation to the next. One way to account for the transfer of learning between such situations or games is to re-label actions based on some common properties. In this work, we examine one framework, known as Level-n, for such a purpose, and combine it with two learning dynamics—Experience Weighted Attraction and Rule Learning—to arrive at predictions for a sequence of ten thrice-played dissimilar games. Using experimental data, we find that when Experience Weighted Attraction is augmented with simple action re-labeling, it performs well in capturing the between-game transference as it affects initial play in each new game. However, only Rule-Learning, without the need for re-labeling, captures the ability of players to learn to reason across games.

Keywords: Learning; Experiments; Transference

1. Introduction.

The question how agents learn across games is, in our view, one of the most important open questions in the theory of learning in games. We examine data from an experiment designed to shed light on how human players behave when faced with a sequence of dissimilar games. By "dissimilar" we mean that there is no re-labeling of the actions that makes the games monotonic transformations of each other (as in Rankin, Van Huyck, and Battalio, 2000[1]). The purpose of this work is to propose approaches to capturing learning in dissimilar games. We pay particular attention to what knowledge is learned from one game to the next and how to best account for such learning. While there have been works that show that humans transfer knowledge from one game to another game (Camerer, Ho snd Weigelt, 1998; Stahl, 2000, 2001, 2003; Chong, Camerer and Ho, 2006; Rankin, Van Huyck, and Battalio 2000[2]; Rapoport, Seale and Winter, 2000; Cooper and Kagel, 2003, 2008; Weber, 2003a, 2003b, 2004), a general learning model is needed to apply in transfer situations.

The field of learning in games has evolved considerably, with many models able to make robust predictions in a variety of interesting games. The focus is typically on a single repeated game where players receive feedback each period about payoffs and the history of play.[3] Different models, although different in details, derive their dynamic predictive power from the fact that the same game is played repeatedly. In contrast, if the players encounter different games, the performance of action labeled "A" in one game has no relevance to the performance

---

[1] In that paper, games which are monotonic transformations of each other are called "similar games".
[2] In that paper, games which are monotonic transformations of each other are called "similar games"
[3] Under action reinforcement learning (Roth and Erev, 1995; Erev and Roth, 1998), each action is reinforced according to the payoff received relative to a dynamic aspiration level. Under belief learning (Fudenberg and Levine, 1998; Cheung and Friedman, 1997, 1998), each player updates his/her belief about the action the other player will choose, and then chooses a (possibly noisy) best-reply to that updated belief. Camerer and Ho (1999) have studied hybrid models that combine reinforcement and belief learning features. See also Anderson, et.al. (2001), Feltovich (2000), Friedman, et.al. (1995), Mookherjee and Sopher (1994, 1997), Nagel (1995), Rapoport,

of the action labeled "A" in a different game. Indeed, if the players encounter a sequence of "sufficiently dissimilar" games so there is no obvious linkage between the actions in subsequent games, then the above-mentioned learning models would predict no learning. On the other hand, it is not unreasonable to expect human subjects to learn something from their experience in a sequence of dissimilar games. For example, they might learn to identify and eliminate dominated strategies, and perhaps even to iterate such eliminations[4], or that might learn to imitate others (Huck et al., 1999).

It might be possible to bridge between different games by re-labeling their actions according to their properties. For example, one could capture learning to avoid dominated actions by labeling dominated actions as "D" in all games and looking at reinforcement over "D" in the sequence of games. While such re-labeling, as we show in this work, can improve the fit of action-based learning models, it cannot account for players increasing their sophistication over time in playing dissimilar games. That is, after playing several different games, players may learn to reason about the games. Such increased sophistication cannot be captured by action-based learning models.

In contrast to action-based learning, the Rule Learning framework of Stahl (1996, 1999, 2000, 2002, 2003, henceforth "Rule Learning"), with an explicit account of *transference* across games, would predict learning in a sequence of dissimilar games. Analogous to iterative elimination of dominated strategies and rationalizability, the Level-n rules have an iterative structure, so a player can learn to switch to higher level rules based on past performance even though the game changes.

et.al. (1997), Sarin and Vahid (1999), Selten (1990, 1991), Stahl (2003), Tang (1998), and Van Huyck, et.al. (1991, 1994, 1996).

In Stahl (1996), Rule Learning was introduced and successfully applied to data from Nagel (1995). Subjects played one guessing game for four periods and there was no opportunity for transfer to a different game.[5] In Stahl (1999), a set of 15 3x3 matrix games was played once simultaneously (in any order, without feedback in between games, and with the possibility of revisiting any game), and then a second time following feedback on all 15 games. The rule-learning model was successfully applied, although there was only one opportunity for transferrence (i.e. between the first and second run of the 15 games). Lastly, Stahl (2000) examined data consisting of two 15-period runs with 5×5 games, and Stahl (2003) examined two 12-period runs with 3×3 games. In these data, a transfer parameter was incorporated and found significant. However, in such designs, transfer of knowledge between games took place only once in the entire experiment.

In contrast to Stahl (2000) and Stahl (2003), in the present experiment, transfer between games takes place 10 times-- one-third of all the periods to be predicted. This gives us a real empirical test of the ability of alternative learning models to capture transfer. Moreover, the current design has transfer taking place between 10 games (as opposed to two in Stahl, 2000), so the kind of transfer that can take place is much richer. Lastly, the current design has the pairing structure that allows us to corroborate the rule-learning results with pair-wise tests. In short, while Rule Learning has the ability to account for transfer, this ability has not been given a serious test prior to the present study.

Section 2 describes the experimental design and the data. Section 3 presents the Rule Learning model adapted to the environment of dissimilar games. Section 4 confronts

---

[4] Weber (2004) shows that such "reflective learning" can occur even without feedback. That learning has to do with the ability to reason about a game and to learn strategies from reflection over time. In this paper, we focus on cognitive learning with feedback.

Experienced Weighted Attraction and the Rule Learning model with the experimental data. Section 5 concludes.

2. The Experiment.

We chose a sequence of ten 4×4 symmetric normal-form games, as displayed in Figure 1. The game payoffs were in tokens, with an exchange rate of $1 per 100 tokens. Four games (3, 5, 8 and10) are dominance solvable, and all distinguish between Level-1, Level-2 and Nash actions. Note that game $i \in \{1, 2, 3, 4, 5\}$ is the same as game $i + 5$, but with the rows (and columns) permuted so the identity is not obvious. Each game was played for three periods before proceeding to the next game. We will call the first 15 periods the "first run", and the second 15 periods with the permuted games the "second run".

A "mean-matching" protocol was used. In each period, a participant's token payoff was determined by her choice and the percentage distribution of the choices of all *other* participants, as follows: the row of the payoff matrix corresponding to the participant's choice was multiplied by the choice distribution of the *other* participants. Payment was made in cash immediately following the session.

Participants were seated at private computer terminals separated so that no participant could observe the choices of other participants. The relevant game, or decision matrix, was presented on the computer screen. Each participant could make a choice by clicking the mouse button on any row of the matrix, which then became highlighted. In addition, each participant could make hypotheses about the choices of the other players. An on-screen calculator would then calculate and display the hypothetical payoffs to each available action given each

[5] Camerer, Ho and Weigelt (1998) provided an opportunity for transfer to a new guessing game and found evidence of such a transfer, but no model to account for such transfer was tested.

hypothesis. Participants were allowed to make as many hypothetical calculations and choice revisions as time permitted. Following each one-minute period, each participant was shown the payoff matrix, her choice, the percentage distribution of the choices of all other participants, and her payoff, and was given 30 seconds to contemplate that information before proceeding to the next period.

The experiment consisted of three sessions of 24, 25, and 25 participants playing this sequence of thirty games. The subjects were upper division undergraduate students and non-economics graduate students from the University of Texas at Austin.

3. Models

In many works on learning, initial play has been all but ignored. In some works (e.g., Camerer and Ho, 1999), initial propensities are estimated as free parameters; in others (Erev and Roth, 1998), initial play is fixed. Either approach may be deemed reasonable when the focus is on the dynamics from the initial starting point. However, in predicting play in a sequence of games, dismissing initial play is not an option as this would result in eliminating much of the history of play (one third in our case). Hence, a model of initial play in each game is now central to the modeling of learning. In this work, we pursue the Level-n approach to modeling initial period play (Stahl and Wilson, 1995). By this approach, human decision makers in games fall into various boundedly rational hierarchies of reasoning about what other players will do. A Level-0 player is equally likely to select any of the available strategies. A Level-1 player anticipates that other players will select uniformly among all available strategies. A Level-2 player believes that other players will mostly (noisily) best respond to the Level-1 belief. While there may be higher levels of reasoning, they are typically not statistically significant, with the exception of a Nash equilibrium type. A Nash equilibrium mode of behavior is often significantly present and is typically included in any Level-n framework.

In an environment with new games being introduced sequentially, a complete learning model must accomplish two goals: (1) predict initial play in each new game, and (2) predict dynamics within each new game. One advantage of studying a learning model is that we can look at how the entire past history—not just the most recent period—affects initial game play in a new game. The second advantage is that we can look at how the past affects within-game learning and reasoning.

The first learning model we examine is Experience Weighted Attraction (hereafter EWA, Camerer and Ho, 1999). This model is widely used in the literature and encompasses other common representations. We adapt the model to learning in a sequence of games by re-labeling the actions according to the taxonomy employed by Stahl and Wilson (1995, hereafter SW95). Specifically, we assign each action one of the following auxiliary labels: {Nash, Level-1, Level-2, Other}. Then, at the data-analysis stage, before applying the EWA model, we reorder the rows (and columns) so the first action is the Nash equilibrium strategy, the second is the Level-1 strategy of SW95, the third is the Level-2 strategy of SW95, and the fourth is the remaining strategy. Thus, EWA reinforcement of the first action of the reordered game is equivalent to reinforcement of the Nash equilibrium strategy, reinforcement of the second action of the reordered game is equivalent to reinforcement of the Level-1 strategy, etc. The crucial effect of applying EWA to the reordered game is to permit transference across dissimilar games: i.e. while the performance of action B in the previous game is irrelevant to the performance of action B in a new dissimilar game, the performance of the Nash (Level-1, etc.) action in the previous game could be quite informative about the performance of the Nash (Level-1, etc.) action in the new game.

The second model we examine is Rule Learning (Stahl, 1996, 1999, 2000). Rule Learning model is intended to represent how subjects learn to reason about games. This feature cannot be captured in a re-labeled-strategy EWA model. Moreover, Rule Learning does not require reordering the rows and columns as in the relabeled-strategy EWA model.

3.1. Experience Weighted Attraction

Camerer and Ho (1999) proposed a propensity-based learning model. The idea behind the EWA model is that decision makers evaluate the performance of each possible action in previous periods and update their propensities to use each action accordingly. Actions are reinforced according to past performance, but actions actually selected receive greater reinforcement.

EWA is an individual learning model, which we aggregate to derive a population version. We also restrict attention to the logit form. The underlying "attraction" state variable, $A_j(t)$, can be thought of as the decision maker's propensity to choose action j at time t. We prefer the term propensity, which we will use hereafter. The propensity to play action j among an available set of J actions is updated according to the following dynamic:

$$A_j = [\theta N(t\text{-}1) A_j(t\text{-}1) + \lambda g_j(t\text{-}1)]/N(t), \tag{1}$$

where θ is a constant discount parameter on past propensities, $A_j(t\text{-}1)$ is the propensity to choose action j in the previous period, and λ is scale parameter which converts the monetary payoffs (in whatever units they are given) into propensities. N(t) is called the "experience weight" and denotes the accumulated strength of past propensities over time, where $N(0)=N_0$ and $N(t) = \gamma N(t-1) + 1$. If $N_0 < (1-\gamma)$ and $\gamma \in (0,1)$, then N(t)/N(t+1) declines, putting less weight on new evidence as time passes (the power law of practice), and vice versa. On the other hand, if $\gamma = 0$ or $\gamma = (N_0-1)/N_0$, then N(t) is constant for all t. Since N_0 and λ are then not separately identifiable in the population model, we eliminate N(t) entirely.

The function g(.) is the evaluated payoff function represented by:

$$g_j(t) \equiv \alpha e_j' Up(t) + (1-\alpha)p_j(t)[e_j' Up(t)n/(n-1) - U_{jj}/(n-1)] . \qquad (2)$$

The parameter α (which Camerer and Ho call the "imagination parameter") is the probability that an individual evaluates the past performance of all actions. The vector e_j is a J×1 vector with 1 in the j^{th} element and 0 elsewhere. U is the JxJ payoff matrix for the applicable game, and p(t) is a vector of observed proportion of choices by others in period t.

For initial conditions, Camerer and Ho let $\{A_j(l)\}$ be free parameters. Rather than increase the number of free parameters, we take the approach of insufficient reason, specifying $A_j(0) = 0$ for all j and $p(0) = p^0$.

To adopt EWA to dissimilar games, we re-order the actions so the first action is the Nash equilibrium strategy, the second is the precise Level-1 strategy (best response to a uniform belief over others' choices), the third is the precise best-reply to the Level-1 strategy, and the fourth is the remaining strategy. To allow for transference, for games g = 2,…, 10, we specify the initial attractions as

$$A_{jg}(1) = \tau A_{j,g-1}(4) + (1-\tau)A_{j,g-1}(1). \qquad (3)$$

Note that if $\tau = 1$, EWA learning proceeds across games without interruption, while if $\tau = 0$, no transference at all occurs across games. We restrict attention to the logistic-form of the probabilistic choice function with trembles:

$$\varphi_{jg}(t) = (1-\varepsilon)exp[A_{jg}(t)]/\sum_k exp[A_{kg}(t)] + \varepsilon/J, \text{ for } t \geq 1, \tag{4}$$

where $\varphi_{jg}(t)$ is the probability of choosing action j in game g in period t, and ε is the probability of a tremble to a random action.

3. 2. Rule Learning

Rule Learning was introduced by Stahl (1996) and developed further in Stahl (1999, 2000). Rule Learning entails a specification of behavioral rules, a process for selecting rules, and a process for updating the likelihood of using these rules.

A *behavioral rule* maps information about the game and the history of play to a probability distribution over available actions, interpreted as a decision maker's probabilistic choice. The basic idea of Rule Learning is that rules which perform well are more likely to be used in the future than worse performers.

The second element of Rule Learning is a probability measure over the rules: $\varphi(\rho,t)$ denotes the probability of using rule ρ in period t. Because of the non-negativity restriction on probability measures, it is more convenient to specify the learning dynamics in terms of a transformation of φ that is unrestricted in sign. To this end, we define $w(\rho,t)$ as the *log-propensity* to use rule ρ in period t, such that

$$\varphi(\rho,t) \equiv exp(w(\rho,t)/[\int_R exp(w(z,t)dz] . \tag{5}$$

Given a space of behavioral rules R (described in detail below) and probabilities φ, the induced probability distribution over actions for period t is

$$p(t) \equiv \int_R \rho d\varphi(\rho,t) \,. \tag{6}$$

The third element of the model is the equation of motion. The Law of Effect states that rules which perform well are more likely to be used in the future. This law is captured by the following dynamic which gives the updated log propensity after period t:

$$w(\rho,t^{+}) = \beta_0 w(\rho,t) + \beta_1 \rho' U^t p^t, \; for \; t \geq 1 \,, \tag{7}$$

where β_0 is the inertia parameter, and β_1 scales the expected utility that rule ρ would have yielded in period t.

Conceptually, there are infinitely many rules players could use (just about any behavioral model ever published could be thought of as a rule). So the approach here is not to try to guess all possible rules players could use, but rather to choose a limited number of salient rules that encompass a larger space of rules. Recall the Level-n types introduced earlier in this section to account for initial play. Our dynamic behavioral rules are based on the same concept, which makes it easy to combine the dynamic model and initial play model into a single framework. Specifically, a Level-1 player believes other players will behave as they have in the past, whereas a Level-2 player believes other players are Level-1 players. The Nash player believes other players are Nash players.

Unlike the model of initial play, where players' types are exogenous, here players are free to change their behavior based on the observed past success of the various rules. That is, a player takes the past performance of each rule to create "evidence" for and against the available rules. The player then tends to choose the rule with the highest net favorable evidence (the exact model and mechanism will be presented shortly). Each rule has different evidence which serves as a score for that rule, where higher score increases the probability of choosing that rule in the next period. Evidence is not action-specific (there is no evidence for action A, or B, or C), but rather rule-specific. Evidence is not accumulated over periods, but there is inertia in the propensities that result from evidence, so evidence from a given period influences periods in the future through that inertia.

We now define the evidence-based rules, and assign parameters and specify functional forms. In the initial period of each game, the Level-1 player holds the belief, p^0, that all available actions by other players are equally likely to be chosen. The Level-1 evidence in later periods is based on the belief that other players will behave as they did in the past, using simple distributed-lag forecasting with lag parameter θ:

$$q^t(\theta) \equiv (1-\theta)q^{t-1}(\theta) + \theta\, p^{t-1} . \tag{8}$$

where p^0 is the uniform distribution and p^{t-1} is the empirical distribution of play in period t-1. The "Level-1" evidence in favor of action j is the expected utility payoff to action j given belief $q^t(\theta)$.

The second kind of evidence is based on the Level-2 player who believes all other players are Level-1 players. The "Level-2" evidence in favor of action j is the expected utility payoff given this belief.

The third mode of behavior is Nash equilibrium. The “Nash equilibrium evidence” in favor of action j is the expected utility payoff assuming others play the (mixed) Nash equilibrium strategy.[6]

[6] Since the experiment design entails only games with a unique pure-strategy Nash equilibrium, the issue of multiple Nash equilibria can be ignored. For an approach to equilibrium selection in a similar setting see Haruvy and Stahl (2004).

So far we have defined three kinds of evidence: $\{y_1, y_2, y_3\}$. Each type of evidence is weighted to arrive at the overall evidence. Each player assesses the weighted evidence with some error, and chooses the action that from his/her perspective has the greatest net favorable evidence. For the mapping between evidence and action, we opt for the multinomial logit specification because of its computational advantages when it comes to empirical estimation. Letting $\nu_1 \geq 0$ denote the scalar weight associated Level-1 evidence, $\nu_2 \geq 0$ for Level-2 evidence, and $\nu_3 \geq 0$ for Nash evidence, the quaduple $(\nu_1, \nu_2, \nu_3, \theta)$ completely defines a (probabilistic) rule. Thus, we have a four-dimensional space of evidence-based rules.

Next, we represent behavior that is random in the first period and "follows the herd" in subsequent periods. Following the herd does not mean exactly replicating the most recent past, but rather following the past with inertia as represented by $q^t(\theta)$. Hence, eq(8) represents herd behavior as well as the beliefs of Level-1 types.

Finally, we allow for uniform trembles by introducing the uniformly random rule. Thus, the base model consists of a four-dimensional space of evidence-based rules, a herd rule, and uniform trembles; together, this is the space *R* of rules in eq(6).

Since this theory is about rules that use game information as input, we should be able to predict behavior in a temporal sequence that involves a variety of games. For instance, suppose an experiment consists of one run with one game for *T* periods, followed by a second run with another game for *T* periods. How is learning about the rules during the first run transferred to the second run with the new game? A natural assumption would be that the log-propensities at the end of the first game are simply carried forward to the new game. Another extreme assumption would be that the new game is perceived as a totally different situation so the log-propensities revert to their initial state. We opt for a convex combination,

$$w(\rho,T+1) = (1-\tau)w(\rho,1) + \tau w(\rho,T^+) , \qquad (9)$$

where τ is the *transference* parameter, and $w(\rho, T^{+})$ is the log-propensity that eq(6) would give to rule ρ if it were to be encountered in period *T+1*. If $\tau = 0$, there is no transference, so period *T+1* has the same initial log-propensity as period 1; and if $\tau = 1$, there is complete transference, so the first period of the second run has the log-propensity that would prevail if it were period *T+1* of the same game. This specification extends the model to any number of runs with different games without requiring additional parameters.

Note that in contrast to the EWA model, this specification of Rule Learning does not use re-labeling of the actions. The definitions of the evidence-based rules themselves implicitly identify the Nash, and Level-1 and 2 actions for the first period of play of a game. On the other hand, due to the distributed-lag function $q^{t}(\theta)$, the action the Level-1 (and Level-2) rule selects in periods 2 and 3 may differ from the action the Level-1 (and Level-2) rule selects in period 1. Further, $q^{1}(\theta) = p^{0}$ (the uniform distribution) always, so the herd has zero transference (in contrast to EWA).[7]

The entire Rule Learning framework involves 10 parameters: $\beta \equiv (\bar{\nu}_1, \bar{\nu}_2, \bar{\nu}_3, \bar{\theta}, \sigma, \delta_h, \varepsilon, \beta_0, \beta_1, \tau)$. The first four parameters $(\bar{\nu}_1, \bar{\nu}_2, \bar{\nu}_3, \bar{\theta})$ represent the mean of the participant's initial propensity over the evidence-based rules, and σ is the standard deviation of that propensity; the next two parameters (σ, δ_h) are the initial propensities of the herd and tremble rules respectively; β_0 is an inertia parameter; β_1 is a scaling parameter; and τ is the transference parameter for the initial propensity of the subsequent runs.

---

[7] We will address later how endowing the herd with transference possibilities can improve the performance of rule learning.

4. Results.

4.1. Pairwise Comparisons of Games

Recall that the experiment design repeated each of the first five games in permuted form, with four games (12 periods) in-between. This allows us to compare behavior for a game in the first run with behavior in the (permuted) game in the second run. By “pair i" we mean games i and i+5.

Camerer, Ho and Weigelt suggest that learning transfer (between p-beauty contest games with different p’s) is most prominently manifested in faster speeds of convergence for players with experience with a previous different game. We observe a similar pattern here. Figure 2 shows the change in Nash equilibrium play between periods 1 and 2 of each game as a percentage of the change between periods 1 and 3. From figure 2, we see that in each pair, relatively more change takes place from period 1 to period 2 in the second run. This suggests faster learning and possibly increased sophistication. The hypothesis that the proportion playing Nash equilibrium strategy in the last period of the second run of a pair is greater than the last period of the first run has a p-value of 0.0001 (14 d.f.). The hypothesis that the proportion playing the Level-1 strategy in the last period of the second run of a pair is smaller than the last period of the first run has a p-value of 0.0014 (14 d.f.).

Figure 3 shows the differences from the first to the second runs in the proportions playing Level-1, Level-2, and Nash for each pair and each period. The summary of the results in the fourth panel clearly shows the rise of Nash equilibrium play and the decline of Level-1 play between the two runs. Note that the largest changes in proportions of Level-1, Level-2 and Nash occur in period 2, as compared to period 1 and period 3. These differences are not significant between period 2 and period 3, but are statistically significant at the 6% and 5% level (2-sided t-test) between period 1 and period 2 for Level-1 and Nash, respectively (but not Level-2). The difference between period 1 and combined period 2 and 3 is significant at under 4% for both Level-1 and Nash.

4.2. Experience Weighted Attraction

Our maximum likelihood estimation of the population EWA model yielded an estimate for γ of exactly 0, which renders N_0 and β not separately identifiable. Therefore, we dropped γ and N_0 from the population EWA model. The following table gives the ML estimates of the population EWA model.

Table 1 Population EWA model with re-labeling and transference

	MLE	5%	95%
θ	0.600	0.426	0.756
α	0.848	0.695	1.000
ε	0.0381	0.000	0.069
λ	0.0652	0.055	0.077
τ	0.0402	0.028	0.057
LL	-1951.48		

To test the hypothesis of no-transference, we set $\tau = 0$. The maximized LL decreases to -1985.18. Twice this difference is distributed Chi-square with 1 degree of freedom, and has a p-value $< 10^{-17}$; thus, we reject the null hypothesis of no transference in the EWA framework. For completeness we also tested the hypothesis of 100% transference by setting $\tau = 1$ and re-estimating the model. Not surprisingly, the maximized LL dramatically decreases to -2098.59, thereby strongly rejecting the hypothesis of 100% transference in the EWA framework. The LL for EWA with no relabeling is -1985.18. This is the same LL that we report for our test of no-transference. In other words, the non relabeled EWA comes out identical to no-transference EWA, which is not very surprising (the relabeling is the only meaningful transference that could take place here). However, it does demonstrate that relabeling is crucial for the success of EWA here.

4.3. Rule Learning

We estimated the 10 parameters of the Rule Learning model using all the data from the three sessions. We pool over games in order to try to identify regular features of learning dynamics that are general and not game-specific, so we can be more confident that these features will be important in predicting out-of-sample behavior. From previous studies (Stahl, 2000, 2001, 2003; Stahl and Haruvy, 2002), we have a strong prior that $\bar{\nu}_3 = 0$ and $\beta_0 = 1$. Indeed, these parameter values are within the bootstrapped confidence intervals for the 10-parameter model. Imposing our prior decreases the maximized log-likelihood by only 3.05. Since our prior is on the boundary of the parameter space, the regularity conditions for the asymptotic Chi-square test are not satisfied. Therefore, we bootstrapped the likelihood ratio statistic, and we cannot reject our prior at the 5% level. Moreover, the likelihood difference of 3.05 is not significant according to the Bayesian Information Criterion[8]. We henceforth focus on the restricted 8-parameter model with $\bar{\nu}_3 = 0$ and $\beta_0 = 1$.

[8] The Bayesian Information Criterion score is -2 LL + (# of free parameters) ln(# of obs). Given any two estimated models, the model with the lower value of BIC is the one to be preferred. The maximized LLs for 8 and 10 parameter models are -1855.12 and -1852.07. Hence, the BIC scores are 3771.9 and 3781.2. Note that the 8-parameter model would have a higher BIC score with just 21 observations, whereas our data is composed of 74 subjects, so even considering each subject's 30 choices as one observation, we would prefer the 8-parameter model.

The ML parameter estimates of the 8-parameter model are given in Table 2. The maximized LL is -1855.12, compared to the entropy[9] of the data of -1651.02, gives a psuedo-R^2 of 0.89. The Pearson Chi-square statistic for the entire data set is 428.17. The Root-Mean Squared Error of the predicted versus actual choice frequencies averaged over all games is 0.095.

Table 2. Parameter Estimates for Rule Learning

	MLE	**5%**	**95%**
σ	1.915	1.783	2.106
δ_h	0.484	0.386	0.584
$\bar{\nu}_1$	1.082	1.081	1.082
$\bar{\nu}_2$	0.000	0.000	0.000
θ	1.000	1.000	1.000
β_1	0.00648	0.00527	0.0213
τ	1.000	0.310	1.000
ε	0.000	0.000	0.000
LL	-1855.12		

Note that six of these ten ML estimates are on the boundary of their respective theoretical domain. Thus, *only four* of the parameters are determining the fit of the data: (i) the standard deviation of the initial distribution of propensities (σ), (ii) the initial precision of the Level-1 rule ($\bar{\nu}_1$), (iii) the initial probability of the herd rule (δ_h), and (iv) the reinforcement scaling parameter (β_1). It is also noteworthy that the ML estimate of the transference parameter (τ) is exactly 1. In other words, there is 100% transference of rule propensities across the dissimilar games.

[9] This is the likelihood ceiling, computed by plugging in the observed empirical frequencies.

The central hypothesis to consider is whether there is any transference and/or learning of rules whatsoever. To test the hypothesis of no-transference, we set $\tau = 0$ and re-estimate the model. The maximized LL decreases to -1889.83. Twice this difference is distributed Chi-square with 1 degree of freedom, and has a p-value $< 10^{-16}$; thus, we can strongly reject the null hypothesis of no transference. To test the hypothesis of no-rule-learning, we set $\beta_0 = 1$ and $\beta_1 = 0$, and re-estimate the model. The maximized LL decreases to -1889.83.[10] Twice this difference is distributed Chi-square with 2 degrees of freedom, and has a p-value $< 10^{-15}$; thus, we can strongly reject the null hypothesis of no rule learning.

One way to assess what is learned is to compute the implied beliefs over types:

$$q_k(t) \equiv \overline{v}_k(t)/[\overline{v}_1(t) + \overline{v}_2(t) + \overline{v}_3(t)]. \tag{10}$$

For $k \in \{1, 2, 3\}$, $q_k(t)$ can be interpreted as a representative participant's belief (probability) that the other participants use the Level-0, Level-1, and Nash rules respectively. At the beginning of period 1, q(1) = (0.442, 0.101, 0.101), and it changes smoothly to q(30) = (0.421, 0.104, 0.175). There is a slight decline in the belief that others are Level-0 types, and a corresponding increase in the belief that others are Nash types.

The major dynamic change is a dramatic decrease in the propensity of the herd rule – from 48.4% to 7.3%, and the corresponding increase in the evidence-based rules (Level-1, Level-2 and Nash). Thus, while the distribution over these evidence-based rules does not change substantially, the aggregate propensity to use them increases dramatically.

[10] In the previous test of $\tau = 0$, the ML estimate for β_1 was 0, so the two maximized LL values are the same. Curiously, without the possibility of transference, rule learning does not help explain the data: all the behavior change appears to be due to herding and belief learning.

As noted, this Rule Learning model does not utilize re-labeling, in contrast to the EWA model we estimated. Consequently, herd behavior in the first period is specified to be the uniform distribution. It is easy to modify this assumption, by using re-labeling and allow transference across games analogous to eq(7):

$$q(T+1, \theta) = (1-\lambda)q(1, \theta) + \lambda q(T^{+}, \theta), \tag{11}$$

where λ is transference parameter for the herd. If $\lambda = 0$, there is no transference and re-labeling is superfluous; if $\lambda = 1$, there is 100% transference. Estimating this enhanced Rule Learning model yields a maximized LL of -1848.20, and an estimate of 0.023 for λ. While the improvement in LL is statistically significant (p-value = 0.0002), the small rate of herd transference indicates that the behavioral effect is almost negligible.

4.4. Comparison of Rule-leaning to EWA

Though Rule Learning has a much better maximized LL fit than EWA (-1855.12 vs. 1947.61), the purpose of this paper is not to horse-race rule-learning against EWA[11], but rather to show two useful approaches to capturing the transfer of learning across dissimilar games.

The advantage of EWA is that it is easy to implement and has been heavily publicized so working programs are readily available. The modification proposed here is relatively minor and involves no necessary changes to the program (one could simply re-label the choices in the data to reflect the new labeling). In addition, the same re-labeling approach we used for EWA could be easily applied to other simple learning models in the exact same way described for EWA. The main disadvantage is that the rules being transferred are not discovered in the program but

[11] Moreover, the original EWA is a model of individual learning, whereas we have modified it to be a model of population learning.

are rather imposed on it. In contrast, in rule-learning, the building blocks for the rule space are pre-defined but the resulting space of rules is broad and flexible.

Nevertheless, one might wish to compare EWA and Rule-Learning on a game-by-game basis to gauge the relative performance of each. Table 3 shows the maximized LL fit for Rule Learning and EWA for each of the 30 periods of the experiment. Column 1 indicates the game (1-10). The second column indicates the difference in log likelihood between rule-learning and the modified EWA. We see from the table that with very few exceptions, Rule Learning outperforms EWA in most games. However, in game 6, EWA stands out as a better performer. A closer inspection shows that the bulk of the improvement in game 6 comes from period 1 of that game. This suggests the strength of the relabeling approach. Note that Rule Learning outperforms on first periods as well; just not as much.

Table 3. Game-by-game comparisons of Rule Learning and EWA.

Game	RL LL minus EWA LL
1	2.23
2	4.25
3	25.59
4	2.56
5	18.49
6	-10.25
7	6.20
8	24.67
9	5.44
10	17.18
Sum	96.36

5. Conclusions

In recent years, a growing body of literature (Rankin, Van Huyck, and Battalio 2000[12], Rapoport, Seale and Winter, 2000; Cooper and Kagel, 2003, 2008; Weber, 2003a, 2003b, 2004) has introduced settings in which decision makers have experience with a class of non-identical games. When non-identical games are not clear transformations of each other, we call them "dissimilar games." In the experimental literature, it appears that subjects learn to reason about dissimilar games and may even increase sophistication over time (e.g., Cooper and Kagel, 2008).[13] In the present work, we proposed two general frameworks to deal with learning in dissimilar games. The first framework is to treat strategies as actions are treated in commonly used action reinforcement learning models. We chose EWA to illustrate the effectiveness of this approach. In the second proposed framework, decision makers increase their sophistication over time. We chose the Rule Learning model of Stahl (1996, 1999, 2000) to illustrate this approach.

We selected a sequence of ten normal-form games in a sample consisting of 74 participants. By design, unmodified action-reinforcement learning models would predict no learning between these games. In contrast, relabeled action reinforcement models, with the labels representing Level-n strategies have the potential to make useful learning predictions.

An interesting insight comes from pairwise comparisons of permuted games within the sequence of dissimilar games. A majority of subjects exhibit Level-1 behavior in the initial period of a game, but not in later periods. While there is some learning towards the Nash action in both runs of a game, the bulk of the increase in Nash behavior occurs earlier in the second run. Using the relabeled EWA model, we find that relabeling is crucial for predictive success of the model. The transfer between games in EWA, while statistically significant, is nevertheless estimated to be small in magnitude. In contrast, when Rule Learning (even without re-labeling)

[12] In that paper, games which are monotonic transformations of each other are called "similar games"

[13] Increased sophistication with experience has also been shown in learning over similar games. This pattern—first periods seem similar but experienced subjects exhibit faster convergence—has been demonstrated in various settings, including asset markets (Dufwenberg et al. 2005) and prisoner dilemma games (Bereby-Meyer and Roth, 2006).

is used to capture the dynamics, transference of the knowledge learned between games is estimated at 100%. Statistical tests also strongly rejected the null hypothesis of no rule learning.

Whether one prefers EWA with relabeling or rule-learning, both approaches indicate that the participants in our experiments learned abstract aspects of the games which were transferable to subsequent dissimilar games. In both approaches, transference of propensities (attractions) between games takes the form of a convex combination. And both approaches reveal an increase in “depth of reasoning” as a result of learning.

References

Anderson, S., J. Goeree, and C. Holt (2001), "Minimum-Effort Coordination Games: Stochastic Potential and Logit Equilibrium," *Games and Economic Behavior,* **34**, 177-199.

Bereby-Meyer, Yoella, and Alvin E. Roth (2006) "Learning in Noisy Games: Partial Reinforcement and the Sustainability of Cooperation." *American Economic Review,* **96**(4), 1029-1042.

Camerer, C. and T. Ho (1999), "Experience-Weighted Attraction Learning in Normal Form Games," *Econometrica*, **67**, 827-874.

Camerer, C., T. Ho and K. Weigelt (1998), Iterated Dominance and Iterated Best Response in Experimental "p-Beauty Contests", *American Economic Review,* **88**(4), 947-969.

Cheung, Y. and D. Friedman (1997), "Individual Learning in Normal Form Games: Some Laboratory Results," *Games and Economic Behavior,* **19**, 46-76.

Cheung, Y-W, and D. Friedman (1998), "Comparison of Learning and Replicator Dynamics Using Experimental Data," *Journal of Economic Behavior and Organization*, **35**, 263-280.

Chong, J-K., C. Camerer and T. Ho (2006), A Learning-based Model of Repeated Games with Incomplete Information, *Games and Economic Behavior*, **55**(2), 340-371

Cooper, D. J. and J. H. Kagel (2003), "Lessons Learned: Generalizing Learning across Games," *American Economic Review*, **93**(2), 202-207

Cooper, D. J. and J. H. Kagel (2008), Learning and Transfer in Signaling Games, *Economic Theory*, **34**, 415-439.

Dufwenberg, M., T. Lindqvist and E. Moore (2005). "Bubbles & Experience: An Experiment," *American Economic Review,* **95** (2005), 1731-37.

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**, 848-881.

Feltovich, N. (2000), "Reinforcement-Based vs. Beliefs-Based Learning in Experimental Asymmetric-Information Games," *Econometrica*, **60**, 605-641.

Friedman, D., D. Massaro and M. Cohen (1995), "A Comparison of Learning Models," *J. of Mathematical Psychology*, **39**, 164-178.

Fudenberg, D. and D. K. Levine (1998), *The Theory of Learning in Games*, MIT Press.

Haruvy, E., and D. Stahl (2004), "Deductive versus Inductive Equilibrium Selection: Experimental Results," *J. of Econ. Beh. and Org.*, **53**, 2004, 319-332.

Huck, Steffen, Hans-Theo Normann and Jorg Oechssler (1999), Learning in Cournot Oligopoly – an Experiment, *The Economic Journal*, **109**, 80 - 95

Mookherjee, D. and B. Sopher (1994), "Learning Behavior in an Experimental Matching Pennies Game," *Games and Economic Behavior*, **7**, 62-91.

________ (1997), "Learning and Decision Costs in Experimental Constant Sum Games," *Games and Economic Behavior*, **19**, 97-132.

Nagel, R. (1995). "Unraveling in Guessing Games: An experimental study," *American Economic Review,* **85**, 1313-1326.

Rankin, Frederick W., John B. Van Huyck and Raymond Battalio (2000) "Strategic Similarity and Emergent Conventions: Evidence from Similar Stag Hunt Games" *Games and Economic Behavior,* **32**, 315-337.

Rapoport, A., I. Eren, E. Abraham, and D. Olson (1997), "Randomization and Adaptive Learning in a Simplified Poker Game," *Organizational Behavior and Human Decision Processes*, **69**, 31-49.

Rapoport, A., D. A. Seale, E. Winter (2000), "An experimental study of coordination and learning in iterated two-market entry game," *Economic Theory*, **16**(3), 1432-1479s

Roth, A. and I. Erev (1995), "Learning in Extensive Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term," *Games and Economic Behavior,* **8**, 164-212.

Sarin, R. and F. Vahid (1999) "Payoff Assessments without Probabilities: A Simple Dynamic Model of Choice," *Games and Economic Behavior*, **28**, 294-309.

Selten, R. (1990), "Anticipatory Learning in Games," in *Game Equilibrium Models. Vol. I: Evolution and Game Dynamics* (R. Selten, ed), Springer-Verlag.

Selten, R. (1991), "Evolution, Learning, and Economic Behavior," *Games and Economic Behavior*, **3**, 3-24.

Stahl, D. (1996), "Boundedly Rational Rule Learning in a Guessing Game," *Games and Economic Behavior,* **16**, 303-330.

Stahl, D. (1999), "Evidence Based Rules and Learning in Symmetric Normal Form Games," *International Journal of Game Theory* **28**, 111-130.

Stahl, D. (2000), "Rule Learning in Symmetric Normal-Form Games: Theory and Evidence," *Games and Economic Behavior,* **32**, 105-138.

Stahl, D. (2001), "Population Rule Learning in Symmetric Normal-Form Games: Theory and Evidence," *J. of Econ. Org and Behavior,* **45**, 19-35.

Stahl, D. (2003), "Action-Reinforcement Learning versus Rule Learning," *Greek Review of Economics,* **22**, 27-56, also in *Transfer in Knowledge in Economic Decision Making*, D. Zizzo (ed), Houndmills and New York: Palgrave Mcmillan, 2005.

Stahl, D. and E. Haruvy (2002), "Aspiration-based and Reciprocity-based Rules in Learning Dynamics for Symmetric Normal-Form Games," *J. of Math. Psych,* **46**, 531-553.

Stahl, D. and P. Wilson (1995), "On Players Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, **10**, 218-254.

Tang, F-F. (1998), *Anticipatory Learning in Two-Person Games: An Experimental Study*, Lecture Series in Mathematical and Economic Systems, Springer-Verlag (in press).

Van Huyck, J., R. Battalio and R. Beil (1991), "Strategic Uncertainty, Equilibrium Selection Principles, and Coordination Failures in Average Opinion Games," *Quart. J. of Economics*, **106**, 885-911.

Van Huyck, J., J. Cook, and R. Battalio (1994), "Selection Dynamics, Asymptotic Stability, and Adaptive Behavior," *J. of Political Economy*, **102**, 975-1005.

Van Huyck, J., R. Battalio and F. Rankin (1996), "Evidence on Learning in Coordination Games," mimeo.

Weber, R. A. (2003a). "Learning with no feedback in a competitive guessing game." *Games and Economic Behavior,* **44**(1): 134-144.

Weber, R. (2003b), "Learning and transfer of learning with no feedback: An experimental test across games," working paper

Weber, Roberto (2004), "Reflective Learning and Transfer of Learning in Games Played Repeatedly without Feedback," working paper.

Figure 1. The Games.[14]

Game 1

	A	B	C	D	
A	60	15	100	90	L1
B	80	70	80	0	NE
C	90	15	35	0	L2
D	65	10	5	70	DOM

Game 2

	A	B	C	D	
A	10	95	20	0	L2
B	90	45	100	55	L1
C	95	5	40	15	L3
D	60	70	15	95	NE

Game 3

	A	B	C	D	
A	25	80	95	15	NE
B	15	80	80	100	L1
C	15	90	75	50	L2
D	5	15	70	0	DOM

Game 4

	A	B	C	D	
A	5	55	95	70	L1
B	30	80	15	80	NE
C	0	10	90	50	DOM
D	100	10	15	55	L2

Game 5

	A	B	C	D	
A	60	20	0	40	DOM
B	100	65	30	25	L1
C	20	50	40	70	NE
D	10	70	30	25	L2

Game 6

	A	B	C	D	
A	35	0	15	90	L2
B	5	70	10	65	DOM
C	80	0	70	80	NE
D	100	90	15	60	L1

Game 7

	A	B	C	D	
A	45	55	100	90	L1
B	70	95	15	60	NE
C	5	15	40	95	L3
D	95	0	20	10	L2

Game 8

	A	B	C	D	
A	0	15	70	5	DOM
B	100	80	80	15	L1
C	50	90	75	15	L2
D	15	80	95	25	NE

Game 9

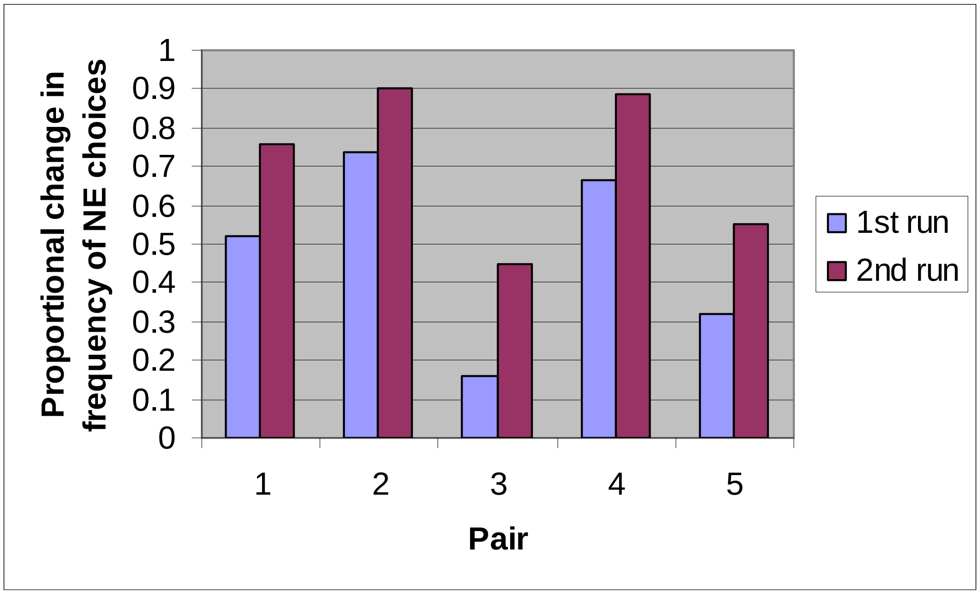
	A	B	C	D	
A	80	80	30	15	NE
B	10	55	100	10	L2
C	55	70	5	95	L1
D	10	50	0	90	DOM

Game 10

	A	B	C	D	
A	25	30	10	70	L2
B	70	40	20	50	NE
C	40	0	60	20	DOM
D	25	30	100	65	L1

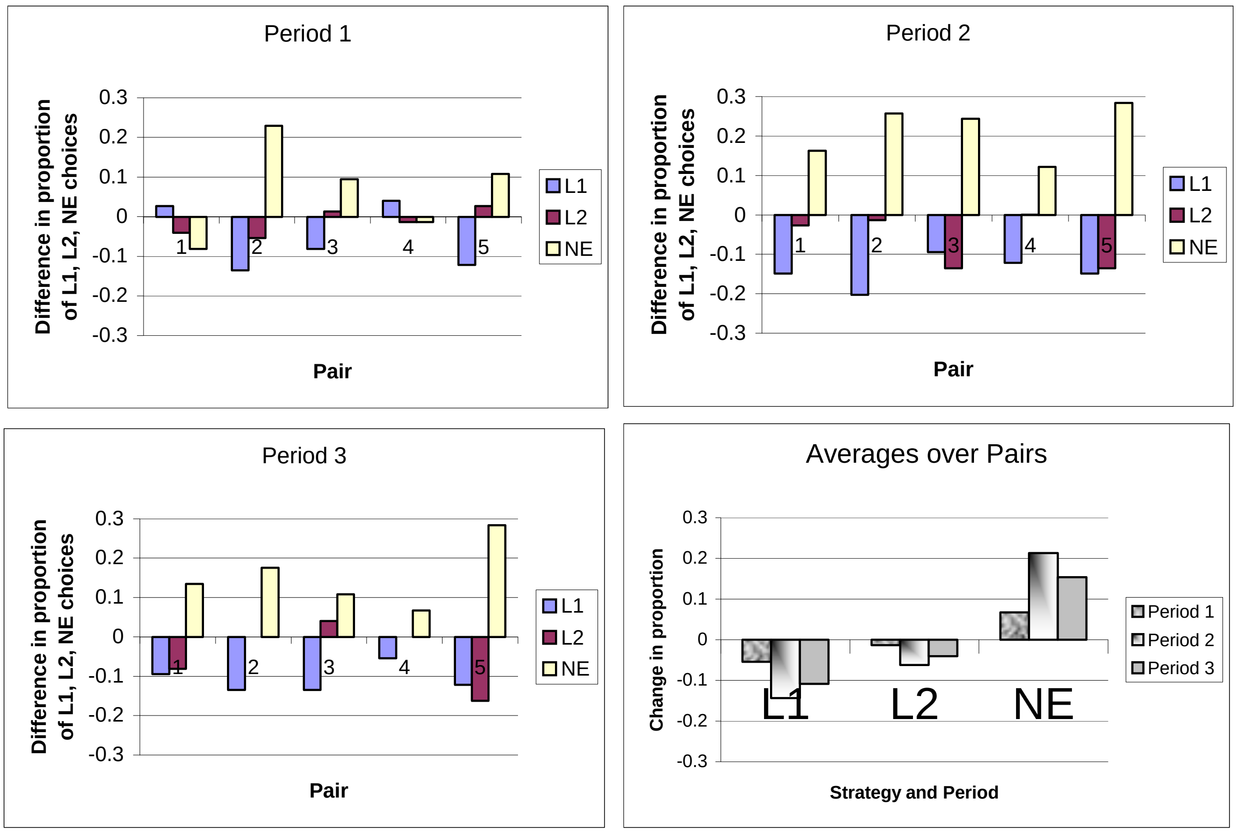
[14] "Ln" denotes Level-n, NE denotes Nash equilibrium and DOM denotes dominated strategy.

Figure 2. Change in proportion of population choosing Nash equilibrium play between periods 1 and 2 as a proportion of the overall change from period 1 to 3.[15]



[15] y = (Proportion choosing Nash equilibrium in period 2 - Proportion choosing Nash equilibrium in period 1) / (Proportion choosing Nash equilibrium in period 3 - Proportion choosing Nash equilibrium in period 1).

Figure 3. Differences from 1st to 2nd run in proportions playing Level-1(L1), Level-2 (L2), and Nash equilibrium (NE) actions by pair and period.



Appendix A Instructions

Welcome. This is an experiment about economic decision making. If you follow the instructions carefully you might earn a considerable amount of money which will be paid at the end of the experiment in private and in cash. It is important that during the experiment you remain silent. If you have questions or need assistance, raise your hand but do not speak. During the experiment, you and all other participants will make 30 decisions, each worth up to $1.00 (exchange rate is 100 tokens = $1). Hence, it is possible to earn up to $30 in this experiment. Each decision you face will be described by a MATRIX, consisting of 16 numbers arranged in 4 rows and 4 columns.

The rows indicate your possible choices; the columns indicate the possible choices of all other participants in this room. The numbers in the MATRIX, along with your choices and the choices of all OTHER participants in this session, determine your TOKEN earnings for each decision. Each participant will face exactly the same matrices and will have the same information. Your decision and those of all the other participants will determine your TOKEN earnings.

How token earnings are computed

Press QUIT and click on the DEMO button. Suppose the matrix you are facing is the following:

	A	B	C	D
A	30	100	50	20
B	40	0	90	40
C	50	75	20	30
D	10	10	10	10

Suppose 20% chose A, 20% chose B, 50% chose C, and 10% chose D. You chose Row A. We first write down the choice labels A, B, C, and D. Underneath them we write down the percentage choices of others. And underneath the percentage choices we write down the numbers of row A (your choice). We then multiply each column. Finally, we add up the results:

	A	**B**	**C**	**D**
% of Others' Choices :	**20%**	**20%**	**50%**	**10%**
Your Row Choice: A	**30**	**100**	**50**	**20**
Product of Each column	**6**	**20**	**25**	**2**

Sum of the bottom row = 6 + 20 + 25+ 2 = 53. Hence, your payoff for choosing A given the other participants' percentages would be 53 tokens.

The payoffs you would have earned for the other row choices are calculated the same way. We will quickly work out your payoff had you chosen row B. We first write down the choice labels A, B, C, and D. Underneath them we write down the percentage choices of others (same as before). And underneath the percentage choices we write down the numbers of row B (your choice). We then multiply each column. Finally, we add up the results:

	A	**B**	**C**	**D**
% of Others' Choices :	**20%**	**20%**	**50%**	**10%**
Your Row Choice: B	**40**	**0**	**90**	**40**
Product of Each column	**8**	**0**	**45**	**4**

Sum of the bottom row = 8 + 0 + 45+ 4 = 57. Hence, your payoff for choosing B given the other participants' percentages would be 57 tokens.

Entering Hypotheses

During the experiment you will have a computer interface to calculate hypothetical payoffs.

To demonstrate this, enter 20, 20, 50 and 10 in the four white boxes in the bottom labeled A, B, C, D and click on CALCULATOR.

Four numbers now appear to the right of the matrix under the word PAYOFF, corresponding to your payoffs for each row exactly as we computed.

Suppose the actual choice percentages were as in this example but suppose your hypothesis was that all other participants would choose B. Hence, you entered (0, 100, 0, 0) as your hypothesis. Do so and click on CALCULATOR.

Notice that the computed payoffs indicate that choice A would give you the largest payoff (i.e. 100). In reality, entering this hypothesis cannot change anyone else's ACTUAL choices.

Therefore, given the actual choices of everyone else your payoff from choosing A would be ONLY 53, not 100.

The point is that ***the more your hypothesis differs from the actual percentage of other participants, the more the computed hypothetical payoffs will differ from the actual token earnings, row by row.***

What are the hypothesis boxes good for?

1. *By entering different hypotheses and calculating hypothetical payoffs to these hypotheses, you can explore how the actual choices (including your own) will affect your token earnings. In other words, you can answer "what if" questions.*
2. *You can enter your best guess about the percentage of others choosing each row and use the computed token earnings to guide your choice.*
3. *Between periods you can enter the actual choice frequencies of others, which you will be given, and use the calculator to verify your token earnings.*

Making a Choice
We will now demonstrate **how you make a choice**. Move the mouse cursor to the row you wish to choose in the yellow matrix and click the left mouse button. The row you clicked on will change color to an orange/pink color indicating your choice.
Make a choice now by clicking on ANY row of the yellow matrix.
Change your choice now by clicking on ANY OTHER row of the matrix.
Notice that it is not necessary for you to do any hypothetical calculations before making a choice.

Summary

To Enter or Change a Hypothesis: click inside a white box under the Matrix. Use the keyboard to enter a number. All hypotheses are in terms of percentages and hence must sum to 100. Caution: The hypothetical payoffs will NOT match your hypothesis UNLESS you click on CALCULATOR.

To Calculate Hypothetical Payoffs: Once the white boxes contain your hypothesis and total 100, click on CALCULATOR.

To Make a Choice: Click on the desired row of the matrix and that row will turn pink indicating your choice. To change your choice, simply click on a different row.

To Review Instructions: click on INSTRUCTIONS. To return to the main screen, click on “QUIT INSTRUCTIONS” and move the mouse a bit.

<u>Warning</u>: If you fail to make a choice for any period, you will earn $0 in that period and be penalized $5 in Stage II.

At any time during the experiment, you can display this summary page by clicking the INSTRUCTIONS button, and then QUIT INSTRUCTIONS to return.

Practice Session

You will now have a 40 second timed practice session. You will have 40 seconds to make choices and practice making hypothetical calculations on the Demo matrix. The clock at the bottom right of your screen will count down from 40 seconds to 0 seconds. A ***15-Second*** *warning will appear when only 15 seconds remains for you to make your decision. Otherwise the screen will look exactly as during the Demo session. You should practice making hypothetical calculations, making choices, and revising choices.*

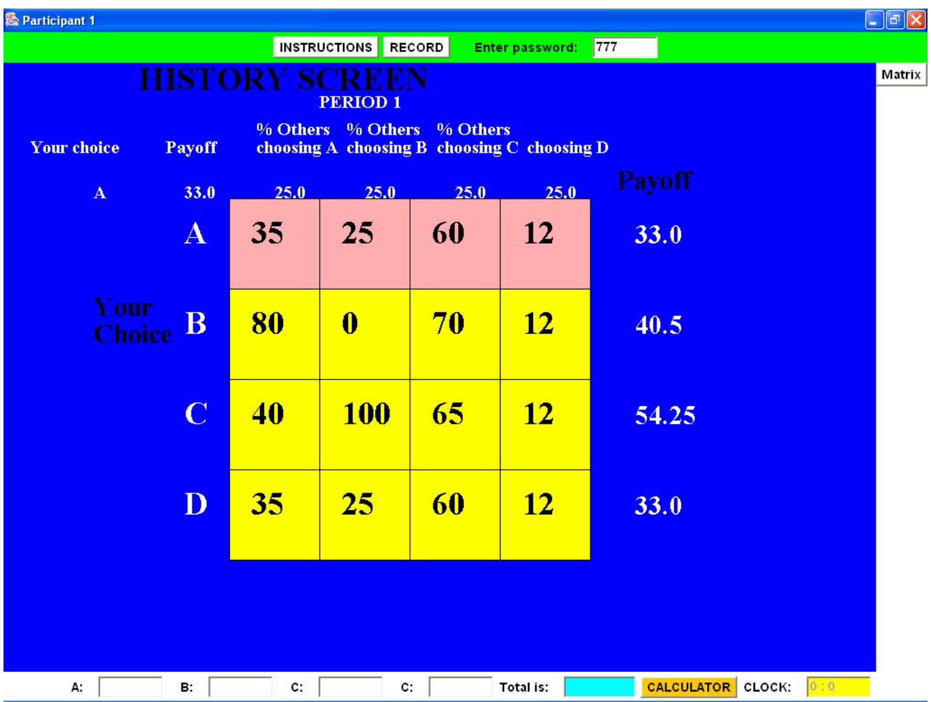
QUIT the Instructions now. Click on the password box on top of your screen, and enter "555". If you do not have 555 entered yet, please raise your hand. The clock starts counting down immediately after you click the DEMO button. Click on the DEMO button now please.

History Screen and the remainder of the experiment

After each period you see a History Screen, like the one displayed below. Your choice will still be highlighted, and the percentage choices of all OTHER participants will be listed above the Matrix. To the right of the Matrix is the computed actual payoff for each row using the ***actual*** *choices of the OTHER participants. Note that your payoff as displayed above the Matrix corresponds to the computed payoff to the right of the highlighted row.*

After each period, you will have about 30 seconds during which you can ponder the results and do more hypothetical calculations on this Matrix before proceeding.

In the experiment, you will face 10 different matrices, each repeated for3 periods. Both the matrix and the period numbers will be displayed on your screen.



Quiz Participant Number ______________

Make sure you put your participant number on the quiz. Your participant number is located on the very top of your screen. Please read the questions carefully, follow the directions exactly. You must use the Demo screen to answer some of these questions. You have 4 minutes in which to complete this quiz.

1. Suppose 50% of the other people in the room chose B and 50% chose D, what would be your payoff if you chose A?

    ____

2. With the same percentages above, which choice would give you the <u>highest</u> possible payoff?

    ____

3. With the same percentages above, which choice would give you the <u>lowest</u> possible payoff?

    ____

4. If the actual percentages of people's choices are as above but you change your hypothesis, can you earn more money?

    ____