Uncommon Knowledge: Behavior and Learning in the "Dirty Faces" Game

Roberto A. Weber*

Department of Social and Decision Sciences

Carnegie Mellon University

Pittsburgh, PA 15211

rweber@andrew.cmu.edu

January 24, 2000

Abstract

This paper presents and examines, as a Bayesian game, the Dirty Faces problem discussed by Littlewood (1953). The equilibrium prediction of behavior in this game

*The author has benefited greatly from comments and suggestions from Colin Camerer, Richard McKelvey, Charles Plott, Rosemarie Nagel, Yuval Rottenstreich, Tom Palfrey, participants at the Economic Science Association meetings in Tucson and Russell Sage Behavioral Economics summer institute, and two anonymous referees. The financial support of the National Science Foundation (Grant #SBR-9223701 to Caltech; P.I.'s: Thomas Palfrey and Richard McKelvey) and the Russell Sage Foundation is also gratefully acknowledged. Comments are welcome.

makes extreme assumptions on the rationality of the players and on their beliefs concerning the rationality of others. Common knowledge of rationality among the players is required for the solution to arise in the general form of the game. The exact number of steps of iterated rationality necessary for equilibrium to arise, however, depends on the number of players of a particular type.

An experiment is used to test the actual behavior of subjects. While behavior at the group level is inconsistent with the game-theoretic prediction, individual level behavior shows a greater degree of consistency with theory. Three models that incorporate error into players' actions all do better than the standard theory in predicting behavior. Finally, a second set of experiments find support for the hypothesis that learning occurs with experience.

1 Introduction

It is often the case that standard game theory makes unrealistic predictions concerning the behavior of players in a game. Support for this observation comes from the extensive experimental work on games. Instead of behaving according to theoretical predictions in these situations, it is frequently the case that subjects rely on other aspects of games or modify their selection principles according to unprescribed rules. For example, subjects' behavior may be affected by the way in which a game is framed or presented,[1] by other principles such

[1]See, for example, Schotter, Weigelt and Wilson (1994) and Bacharach and Bernasconi (1997). Furthermore, Cooper, DeJong, Forsythe and Ross (1993); Güth, Huck, and Rapoport (1995a and 1995b); Rapoport and Fuller (1995); and Camerer, Knez and Weber (1996), all show that differences in timing of a game may

as fairness or altruism,[2] or by a lack of common knowledge of complete rationality among the players. It is with the latter of these cases with which this paper is primarily concerned.

Common knowledge of perfect rationality among the players in a game is important in much of equilibrium analysis. For example, iterated strict dominance, rationalizability, and backward induction rely on the assumption that it is commonly known that players are rational and will not play dominated strategies.[3]

This paper studies the Dirty Faces Game, a variant of an example originally developed by Littlewood (1953), in which players' knowledge and common knowledge of an event result in different equilibrium outcomes, similarly to Rubinstein's (1989) electronic mail game. Moreover, this equilibrium also relies on the presence of common knowledge of rationality among the players. The purpose of this study is to examine the behavior of actual participants in this game and compare it to the behavior predicted by the theory.

The paper is organized as follows. The next section discusses experiments similarly intended to address the question of whether or not common knowledge of rationality is satisfied in the laboratory. Immediately following, the Dirty Faces Game will be discussed in detail

---

affect play even without the usually assumed differences in information. These results are striking in that the timing of moves alone is generally considered inconsequential.

[2]McKelvey and Palfrey (1992) find support for this hypothesis in experimental tests of the centipede game. Furthermore, there is much evidence for this idea in the results of ultimatum and dictator games. The results of studies on these games are discussed extensively in Camerer and Thaler (1995).

[3]See Bernheim (1984), Pearce (1984), and Aumann (1995).

and the experimental design used to test the game will be presented. The final sections will present the results of two sets of experiments and provide a discussion of these as well as possibilities for future research.

2 Previous experiments on common knowledge of rationality

It is surprising that, in spite of the importance of common knowledge of rationality in game theory and the close relation between this field and experimental economics, more laboratory work has not been conducted to test this assumption. This section reviews the existing experimental work.

McKelvey and Page (1990) experimentally tested their previous result (McKelvey and Page, 1986) concerning the use of an aggregate statistic to convey information concerning individuals' posterior probabilities of an event and the convergence of these probabilities.[4] Their results show that, while participants in the experiment used the publicly announced

[4] In the earlier paper, McKelvey and Page (1986) showed that results by Aumann (1976) and Geanakoplos and Polemarchakis (1982) concerning the equivalence of posterior probabilities under common knowledge can be maintained when only an aggregate statistic of the posteriors becomes common knowledge. This statistic need only satisfy stochastic regularity (which, for example, is satisfied by a statistic given by the mean of the posterior probabilities). Nielsen, Brandenburger, Geanakoplos, McKelvey and Page (1990) further extend this result to show that it holds not only for the conditional probabilities of an event, but equally for the conditional expectations of a random variable. This result is also proven to hold when an aggregate statistic of the conditional expectations become common knowledge.

mean to update their beliefs, they did not do so in a manner consistent with the perfect Bayesian updating that the model suggests.[5]

Stahl and Wilson (1994 and 1995) used a series of games with varying properties to estimate the level of rationality for each of the participants, using their actions in the experiment. According to the models on which the experiments are based, players' types are determined by their action selection process and by their beliefs of the types of others. In one of these models, Level-0 type players simply randomize among their strategies, Level-1 type players assume that all other players are Level-0 types and respond accordingly, and Level-2 types believe that all other players are Level-0 and Level-1 types. A final type of players is described as behaving according to Nash equilibrium theory. Their results indicated that the number of Level-0 types in the experiment was negligible and that participants generally behaved as though they belonged to the higher types. They found support for the existence of a large percentage of sophisticated Nash types in the population.

In ongoing research, Costa-Gomes, Crawford, and Broseta (1998) use Mouselab technology[6] to record which payoff information subjects access when playing several two-player normal

[5]Hanson (1996) points out flaws in McKelvey and Page's design, which invalidate the proof that the claimed Bayes-Nash equilibrium to the game is, in fact, such an equilibrium. Moreover, Hanson indicates that the fact that the Bayes-Nash equilibria to the game are not known invalidates the comparison between actual and predicted behavior.

[6]Mouselab is a computer interface in which payoff information is concealed from subjects until they use the mouse to reveal it. Thus, the experimenter can control the amount of information revealed to a subject at one time and, more importantly, can record the patterns in which subjects look up information. For more information on Mouselab, see (Camerer, et al. (1993).

form games, and then use this information and subjects' decisions to determine their level of sophistication. While the analysis of the data generated by viewing patterns has not been completed, the strategy choices of subjects indicate that no more than two steps of iterated dominance are being satisfied.

In perhaps the most clever design addressed at measuring the level of rationality of players, as well as their beliefs concerning the rationality of others, Nagel (1995) studied a game previously discussed by Moulin (1986).[7] In these experiments, subjects were asked to select a number between 0 and 100. The average of these numbers was then computed, as well as a target number. The target number was the mean of the participants' choices multiplied by a constant, $0 < p < 1$, and the participant whose number was closest to this target number would then win the game and a predetermined prize. The unique Nash solution to this game, under the assumption that there is common knowledge between the participants of perfect rationality among the players, is for everyone to pick 0. The game is such, however, that if this necessary common knowledge assumption is not satisfied, it may no longer be optimal to behave according to the Nash prediction.[8] Nagel finds that it is never the case, for different values of p, that all subjects pick the equilibrium in the first play of the game. Instead, choices tend to be significantly higher, indicating that subjects are either not perfectly rational themselves, or are not certain that the infinite hierarchy of knowledge of

[7]A variant of this game has also been studied by Ho, Weigelt and Camerer (1998).

[8]Note that it may be the case that all of the participants are aware of the unique equilibrium. However, if they are not sure that everyone else is aware of it, or that everyone else is aware that everyone is aware of it, then the failure of the Nash prediction may still occur.

rationality among the players is satisfied.

A further interesting aspect of Nagel's experiments, however, is that they make it possible to measure to what extent the hierarchy of subjects' beliefs over the rationality of other players is satisfied. For example, a choice greater than $100p$ indicates that a particular subject is not behaving rationally, since the target number can never be above this value. Furthermore, only a subject who does not believe that everyone else is rational would choose a number in the range $(100p^2, 100p)$, since this implies the belief that at least one of the other subjects will choose a value greater than $100p$. Hence, every choice greater than zero violates some form of the iterative process necessary to arrive at common knowledge of rationality. Nagel's results indicate that, while the majority of subjects do not exhibit violations of rationality in the first period, they only proceed through the above iterative process to two or three steps. Nonetheless, repeated play leads to adjustment towards the equilibrium in all cases.

One concern with interpreting this adjustment, however, is that a decrease in subjects' choices across periods may not represent an attainment of a higher level of iterative rationality. It may instead be the case that subjects are using a simple heuristic to guide their adjustment. For example, using a simple rule such as "choose p times the previous period's target number" will lead to convergence towards the Nash equilibrium in a manner similar to that observed in Nagel's experiments. However, it is hard to argue that such convergence indicates that more steps of the iterative process are being performed or that common knowledge of rationality is being approximated.

As the above experimental results indicate, certain aspects of the analysis underlying standard game theory may be unrealistic in that common knowledge of rationality among players may be not be present, at least initially, in an actual setting where games are played.

3 The Dirty Faces Game

The game which this paper concentrates on is one which is frequently present in the discussion of common knowledge and iterative reasoning. Littlewood (1953) presents the problem as follows:[9]

> Three ladies, A, B, C in a railway carriage all have dirty faces and are all laughing. It suddenly flashes on A: why doesn't B realize C is laughing at her?–Heavens! *I* must be laughable. (Formally: if I, A, am not laughable, B will be arguing: if I, B, am not laughable, C has nothing to laugh at. Since B does not so argue, I, A, must be laughable.)

[9] While this problem is present in much of the literature in several forms and under different names (see, for instance, Binmore and Brandenburger (1988) and Geanakoplos (1993)), this paper models it as a game in a manner similar to the way it is presented in Fudenberg and Tirole (1993).

> ... But further, what has not got into the books so far as I know, there is an extension, in principle, to n ladies, all dirty and all laughing. There is an induction: in the (n+1)-situation A argues: if I am not laughable, B, C, ... constitute an n-situation and B would stop laughing, but does not.

The game can include any finite number of players, n. Each player's face is either "dirty" or "clean", as determined by nature. Players are aware of the states of the faces of other players, but not of their own face. In the situation, then, where all of the players have an incentive to choose a certain action when and only when they are certain that their face is dirty, an equilibrium can arise where, if all of the players' faces are dirty, they all act after exactly n periods, and no one acts before that time. As will be explained in more detail subsequently, however, this result relies significantly on the presence of common knowledge about the fact that at least one face is dirty and on common knowledge of rationality.

Formally, let $N = \{1, \cdots, n\}$. Assume that, for each player, $i \in N$, $x_i \in \{X, O\}$ represents that player's type as determined by chance according to a commonly known probability p. That is, with probability p, nature draws a player's type to be X, and with probability $1-p$ the player is determined to be of type O. Each player's type is thus determined identically and independently of the types of the other players.

Assume next that the game consists of up to $T \geq n$ periods, in which each player chooses one of two actions, $\{U, D\}$, and that the game is over after any period in which any player chooses D. Further assume that players are faced with the following payoff table in each period of the game:

		Type	
		X	O
Action	U	0	0
	D	α	$-\beta$

Table 1: Generic payoff table for dirty faces game

Hence, for α and β such that $p\alpha < (1-p)\beta$, a player will act when and only when she is certain that her type is X.[10]

In the game presented above, if at the end of each period all players observe the actions of all others, no player can learn anything about her own type in any period and should, therefore, never choose D, even when all the n players are of type X. This is because of the independence of player types and is true as long as the information each player receives at the beginning of the game is only the types of the other $n-1$ players. Thus, the equilibrium to the above game is for all players to select U in all periods.

Assume now, however, that a public announcement is made to the entire group of n players at the beginning of the game. In this announcement, all players are informed of whether there is at least one player whose type is X. Looking again at the case where all of the players are of type X, this announcement does not provide any of the players with new

[10]Assuming, not unreasonably, that she is either risk-neutral or risk-averse. This will be discussed in more detail later.

information, since they could already observe that everyone else is of type X and, therefore, that there is at least one individual whose type is X. What this announcement does accomplish, however, is to make this previously known fact common knowledge to all of the participants.[11].

Following the announcement, players now should be able to determine their own true type by the actions of the other players. Specifically, let k be the number of players of type X and let $K \subseteq N$ be the set of all players of type X. Then, for any n and k such that $0 < k \leq n$, in the unique Nash equilibrium to the game $s_i = U$ in all periods for which $t < k$ and, in period $t = k$, $s_i = U$ for all $i \notin K$ and $s_j = D$ for all $j \in K$.[12]

To see why this is true, remember that if $k \geq 1$, the announcement makes this common knowledge at the beginning of period 1. It is also true that, $\forall t$, if at the beginning of the period it is common knowledge that $k \geq t$, then if $k = t$, every player $i \in K$ will know that her type is X (since they see the $k - 1$ other players of type X) and will select D while all the players not in K will choose U. If, on the other hand, $k > t$, then all the players will observe at least t other players of type X, everyone will choose U, and this will make it commonly known that $k > t$ at the end of the period. Thus, everyone will choose U in all

[11]In Littlewood's example, the announcement is replaced by the fact that all three women are laughing. Hence, since the laughter is observable by all of the women, who can in turn observe that they can all observe the laughter, it is commonly known that there is at least one person whose face is dirty when at least one person laughs out loud. In this example, the action D corresponds to ceasing to laugh, which any of the ladies would do immediately when she realized that her face was dirty.

[12]In the case where $k = 0$, the announcement will make it common knowledge that there are no players of type X and all players will select U in every period.

periods t, such that $t < k$, and in period $t = k$, $s_i = D \; \forall i \in K$ and $s_j = U \; \forall j \notin K$.

An important consideration, however, is that common knowledge of rationality among the players is overwhelmingly important in the above analysis. Otherwise, for example, the failure of any player to choose D in Period 1 (when $k > 1$) might be attributed to a lack of rationality rather than to the fact that no player observes two players of type O. In this case, the next step in which it is commonly known that there are at least two players of type X is not reached.

As an illustration of this point, consider the case where $N = \{1, 2\}$ and let $\omega = XX$. Further, define the events

$$\begin{aligned} \mathbf{R} &\equiv \text{Everyone is rational;} \\ \mathcal{K}^1(\mathbf{R}) &\equiv \text{Everyone knows that everyone is rational;} \\ \mathcal{K}^2(\mathbf{R}) &\equiv \text{Everyone knows that everyone knows,} \\ &\quad\;\, \text{that everyone is rational;} \end{aligned}$$

and so on, so that $\mathcal{K}^l(\mathbf{R})$ corresponds to l iterations of the knowledge process. Now, if $\mathbf{R}$ holds, then it will be the case that each of the two players will choose D in Period 1 if and only if they observe the type of the other player to be O. However, this is not sufficient for the above equilibrium to hold. In order for it to then become common knowledge that both players are of type X, it must be the case that the event $\mathbf{R}\&\mathcal{K}^1(\mathbf{R})$ is true. Otherwise, the fact that the other player did not choose D might be attributed to a lack of rationality. However, if $\mathbf{R}\&\mathcal{K}^1(\mathbf{R})$ is satisfied, then this is sufficient for both players to become aware of the true state and choose D in Period 2.

The three player case proceeds similarly. Assume, as before, that the true state is $\omega = XXX$. Then, if $\mathbf{R}$ is true, players will choose D if and only if they observe two players of type O. As long as $\mathbf{R}\&\mathcal{K}^1(\mathbf{R})$ is also true, given that everyone chose U in Period 1, it will be known by everyone that there are at least two players whose type is X. Thus, since $\mathbf{R}$ holds, each player will choose D if and only if they observe one player of type X and one player of type O. Note, however, that in order for everyone to know that everyone knows that there are at least two players of type X, it is necessary that $\mathbf{R}\&\mathcal{K}^1(\mathbf{R})\&\mathcal{K}^2(\mathbf{R})$ be true. Therefore, this must also be the case if, after observing that everyone chose U, it is to be known by all players that the true state is XXX. It must be the case that $\mathbf{R}\&\mathcal{K}^1(\mathbf{R})\&\mathcal{K}^2(\mathbf{R})$ is satisfied in order for the predicted equilibrium to arise when the true state is XXX.

This result can be generalized, by induction, to the case where there are n players and all of them are of type X. In order for the correct equilibrium to arise in this case, it is sufficient that $\mathbf{R}\&\mathcal{K}^1(\mathbf{R})\&\cdots\&\mathcal{K}^{n-1}(\mathbf{R})$ is true.

Note, however, that the above statement need not be true in order for the equilibrium outcome to arise in other cases. For example, in the $n = 3$ case, all that is necessary when $\omega \in \{XOO, OXO, OOX\}$ is that $\mathbf{R}$ hold. Thus, if everyone is rational, the one player who observes two players of type O will know her own type and choose D in the first period and the other two players will select U, regardless of whether they think the other players are rational or not. Furthermore, if the true state is in the set $\{XXO, XOX, OXX\}$, then all that is necessary is that $\mathbf{R}\&\mathcal{K}^1(\mathbf{R})$ be true.

More generally, regardless of n, in the case where there are exactly k players whose type is X, what is sufficient for the correct equilibrium to arise is that $\mathbf{R}\&\mathcal{K}^1(\mathbf{R})\&\cdots\&\mathcal{K}^{k-1}(\mathbf{R})$.[13] Nonetheless, the complexity of the problem is not the same in all situations where k is equal. To see this, compare, using $n = 2$ and $n = 3$ as examples, the case where $k = 2$ and the true states are XX and XXO. In the two player game, the following are necessary: 1) 1 is rational, 2) 2 is rational, 3) 1 knows that 2 is rational, and 4) 2 knows that 1 is rational. However, in the case where there are three players, it must be true that: 1) 1 is rational, 2) 2 is rational, 3) 3 is rational, 4) 1 knows that 2 is rational, 5) 1 knows that 3 is rational, 6) 2 knows that 1 is rational, 7) 2 knows that 3 is rational, 8) 3 knows that 1 is rational, and 9) 3 knows that 2 is rational. In both cases, however, it is only the condition $\mathbf{R}\&\mathcal{K}^1(\mathbf{R})$ which is

[13] In the trivial case where $n = 1$, this is still true since all that is necessary is $\mathbf{R}$ so that when the announcement is made, the single player will immediately know her type and will act accordingly.

being satisfied. Therefore, the number of conditions which must be satisfied in the problem grows exponentially in the number of players.[14]

This paper tests the behavior of actual subjects playing the dirty faces game. A couple of hypotheses are examined. First is the hypothesis that subjects will conform to the theoretical prediction above. If this is not supported, two additional questions arise: 1) do subjects exhibit learning in repeated play of the game? and 2) what kinds of theories best describe the behavior?. Finally, the previous paragraph discusses how the number of conditions that need to be satisfied is affected by varying the group size – holding k (the number of players of type X) constant. Conducting experiments with different group sizes allows a test of whether subjects' behavior is across group sizes when k is the same.

4 Experiment 1: Behavior in the dirty faces game

4.1 Experimental Design

In order to test the situations where the necessary iterated levels of knowledge are the fewest, experiments were conducted using the $n = 2$ and $n = 3$ cases. The choice of parameters α, β, and p proved to be a more difficult decision. This was because the condition $p\alpha < (1-p)\beta$

[14]However, if you consider the fact that players can ignore completely the actions (and rationality and knowledge) of players which they observe to be of type O, then the problems are identical in complexity. In the XXO case, for example, Players 1 and 2 can ignore Player 3 since it is only the assumptions about the other player of type X which are significant. If Player 3 is not rational, however, it could still be possible for her to choose D, leading to a non-equilibrium outcome.

had to be satisfied while it was also a goal to maximize p in order to produce the most instances of the case where all players are of type X and to minimize the occurrence of the trivial situation where no players are of type X.[15] Furthermore, α had to be made large enough so that subjects would stand to earn a significant, or at least reasonable, sum by choosing D once they knew that their type was X. However, increasing α also meant that β had to be increased so that the above inequality would hold. Since it was not desirable to have the possibility of negative earnings for the experiment for any subject[16], the value of β could not be increased unboundedly.

These considerations resulted in the choice of parameters being $p = 0.8$, $\alpha =$ \$1.00, and $\beta =$ \$5.00, and the resulting payoff table presented in Table 2. It was decided that the possibility of negative earnings in an experiment would be compensated for by a participation bonus.

The expected monetary value of an uniformed selection of D, therefore, was -\$0.20, while choosing U would always yield \$0.00. While this difference does not appear large, the belief that this gamble will not be taken by many participants finds further support in the extensive results indicating that subjects choose according to a value function which places

[15] It could be argued that the case where only one player is of type X is also trivial, since this player, upon hearing the announcement is immediately aware of her type and should therefore choose D. This requires, however, that the player be rational, which is a testable assumption.

[16] Kahneman and Tversky (1979) show that the behavior of subjects' preferences over losses is substantially different from that of preferences over similar gains. Furthermore, implementing negative participant earnings in an experiment would have created additional design problems.

greater weight on losses than on gains, and which is concave for gains.[17]

Thus, since increasing the difference would have to be accompanied either by a decrease in p or by a higher participation bonus, the difference created by these parameters was judged to be sufficient.

The random determination of players' types according to p was implemented using a ten-sided die. A roll of the die resulting in 1 or 2 meant that a player was of type O, while the remaining faces corresponded to the type X. Thus, in the two player case, the probabilities of obtaining the situation where there are 0, 1, and 2 players of type X were, 0.04, 0.32, and 0.64, respectively, while for the three player case, the probabilities of having 0, 1, 2, and 3 players of type X were 0.008, 0.096, 0.384, and 0.512, respectively. Hence, for both treatments, the desired outcome was the most likely and the expected occurrence of the trivial problem was minimized.

		Type	
		X	O
Action	U	0.00	0.00
	D	1.00	-5.00

Table 2: Payoffs for Experiment 1

In order to convince the subjects that the process by which types were determined was

[17] See Kahneman and Tversky (1979). Moreover, Kahneman and Tversky (1992) construct and estimate parameters for value and weight functions for lotteries Using their parameter estimates, the gamble under consideration in this experiment has a certainty equivalent of -\$1.78.

not controlled by the experimenter and possibly predetermined, a monitor was randomly selected from among the participants. This was done by having each subject roll the same die which would later be used to determine the types of the participants and selecting the participant who rolled the highest number. In this way, subjects were also given the opportunity to observe rolls of the die.

Each session consisted of three rounds, where each round corresponded to a new game and a new draw of player types. Each round consisted of $n+1$ periods. The first round was labelled as a practice round, intended to familiarize the subjects with the procedure of the experiment. For this round, the payoffs were divided by 5 and no announcement was made regarding the players' types.

An interesting design problem arose in selecting the procedure by which to inform subjects of the types of the other players. The original problem as presented by Littlewood considers the situation where a player's face is either dirty or clean and, therefore, this state can be observed by everyone in the room other than the player herself. From an experimental standpoint, however, such a design implies that subjects, in observing each other, might be able to determine their own type by observing non-verbal cues obtained from other subjects, such as a look of astonishment. Therefore, the following design was used.

In order to prevent subjects from observing the identity of those they were playing with, each session consisted of at least two groups and the identity of subjects in a particular group was not revealed. Before the first round, subjects were randomly assigned a participant number which indicated to them their group, denoted by a letter, and their player number within the group. Thus, subjects were only made aware of the participant numbers of the other players in their group, and not of their identity.

At the beginning of each round, the monitor rolled the die to determine the type of each participant while hidden from the other participants behind a screen. For each round the monitor was given a new Type Sheet, which contained all of the participant numbers along with a blank box next to each number. After each roll, the monitor would record either an "X" or an "O" in the box corresponding to that participant. The sheet would then be placed inside of a simple display box. The display box consisted of a cardboard box and a cardboard sheet with flaps. The Type Sheet was placed face up inside the box and the cardboard sheet was placed over it. In this manner, each of the flaps could then be raised to show only the type of any desired subject. In the two experimental rounds, the experimenter, at this point, made one of two possible announcements at the front of the room for each group. In the case where, say for Group A, all of the participants were of type O, the experimenter would announce, "There are no participants of type X in Group A". Otherwise, the announcement would be, "There is at least one participant of type X in Group A". Thus, common knowledge about this important piece of information was established for each group.

Following the announcements, the experimenter then proceeded to each subject and lifted the flaps so that they could observe the types of the other one or two participants in their group. Each subject then recorded this information on their Record Sheet for that round. Following that, the experiment proceeded to the first period.

In each period, every subject would select one of two actions, either "Up" or "Down". The subject recorded this choice on a Reporting Sheet. Each subject, following the determination of participant numbers, was given a stack of Reporting Sheets, each one of which contained the participant's number, as well as three boxes which could be checked. The first two boxes were for the action choices of Up or Down. The third box was labelled "Round Over" and subjects were instructed to only check this box once any of the participants in their group had chosen Down in a previous period. This was implemented to prevent participants from learning the identity of others in their group by their failure to mark on their Reporting Sheet in a period. The Reporting Sheets were then torn from the stack and collected by the experimenter.

Once all of the Reporting Sheets were collected and sorted, the entire set of actions was written on the board at the front of the room and read out loud twice by the experimenter. Subjects where then instructed to record the actions of the other players in their group on their Record Sheet for that round and proceed to the next period.

One more period than necessary was included in both treatments so that the number of periods would not serve as a cue for the desired behavior. Thus, both treatments were conducted with $n + 1$ periods. Once the game arrived at Period $n + 1$, behavior could no longer be consistent with the theoretical prediction.

At the end of the last period, the experimenter removed the Type Sheet from the display and proceeded to publicly display the entire sheet to all participants. At this time, subjects were instructed to record their own type as well as to verify that the information that they received at the beginning of the round, concerning the types of the other players in their group, was correct. After everyone calculated their earnings for that round, the experiment proceeded to the next round.

Upon the completion of the third round, participants were privately paid, in cash, their earnings in all three rounds plus an $11 participation bonus. This bonus guaranteed that no participant could finish the experiment with a negative total. Thus, the possible earnings in the experiment ranged from $0 (which would only occur if a subject was of type O in all three rounds and picked Down all three times) to $13.20. Each session typically lasted about one hour in both the $n = 2$ and $n = 3$ treatments.

The experiments were conducted in March and April 1996 using as subjects graduate and undergraduate students at the California Institute of Technology with little or no formal training in game theory.[18] These subjects do not necessarily represent a typical population in terms of quantitative and reasoning ability. For this particular study, however, this does not present a drawback since the study is intended to analyze whether seemingly extreme predictions about the rationality within a population can be supported by laboratory results. Failure by subjects from this population to behave according to the theoretical prediction, therefore, would indicate that this failure of the normative prediction would apply equally, if not more strongly, to a majority of other populations.

4.2 Results

4.2.1 Group Behavior

The complete results for both the $n = 2$ and $n = 3$ experimental rounds are given in Tables 3 and 4. There are a total of 13 and 14 groups for the two- and three-player treatments, respectively. The first row for each round indicates the type of each player as determined by the roll of the die. Subsequent rows contain the actions selected by each player, U for Up and D for Down.[19]

[18] Instructions are available in the appendix.

[19] A reporting error was committed by the experimenter in the second round for the second group in Set 4 of the 2 player game. At the end of the second period, the experimenter erroneously reported that one of the two participants had selected Down when, in fact, both subjects had chosen the action Up. Furthermore, neither of the two participants indicated that they realized this error had occurred and instead recorded that the other participant had chosen Down. Although the round was thus prematurely terminated for that group, both subjects had failed to behave according to the equilibrium prediction and, therefore, their third period actions would not have been informative. Furthermore, since the error occurred at the end of the second round and the experiment ended immediately following the conclusion of this round, there is no possibility that the reporting error affected subsequent actions.

Round	Player	Ses. 1 (3 groups)						Ses. 2 (4 groups)							
		A1	A2	B1	B2	C1	C2	A1	A2	B1	B2	C1	C2	D1	D2
I	Type	X	X	X	X	X	O	X	X	X	X	X	X	X	O
	1	U	U	U	U	D	U	U	D	U	U	U	U	D	U
	2	U	D	U	D					U	U	D	U		
	3									U	U				
II	Type	X	O	X	X	X	X	X	O	X	X	X	O	X	X
	1	D	U	U	U	U	U	U	U	U	U	D	U	U	U
	2			D	D	D	D	U	D	U	U			U	D
	3									D	U				

Round	Player	Ses. 3 (2 groups)				Ses. 4 (4 groups)							
		A1	A2	B1	B2	A1	A2	B1	B2	C1	C2	D1	D2
I	Type	X	X	X	O	X	X	X	X	X	X	X	X
	1	U	U	D	U	U	U	D	U	D	U	U	U
	2	U	D			D	D					D	U
	3												
II	Type	X	O	X	X	X	O	X	X	X	X	X	X
	1	D	U	U	U	D	U	U	U	D	U	U	U
	2			U	D			U	U			D	D
	3							*	*				

Table 3: Results for two player game

The first, and perhaps most striking observation is the extent to which subjects overplay the strategy Down when they have no information, beyond their prior, indicating their type.[20] This uniformed action by the participants results in that, for the $n = 2$ condition, only 14 of 18 cases where there are two players of type X reach the second period. Even more surprisingly, none of the three player groups reaches the third period, even though

[20] This is also true in the initial practice round, the results of which are not reported in the tables. In this round, since no announcement was made, expected payoff maximizing subjects should never have chosen Down. However, 12 out of 26 subjects in the $n = 2$ condition and 20 out of 42 subjects in the $n = 3$ condition chose Down in some period. This result, stronger than in later rounds, can be in part attributed to the payoffs being considerably small in the first round and to the fact that subjects were able to end the round by choosing Down.

Round	Player	Ses. 1 (3 groups) A1	A2	A3	B1	B2	B3	C1	C2	C3
I	Type	X	X	X	O	X	X	X	X	X
	1	U	U	U	U	D	U	U	D	U
	2	D	U	U						
	3									
	4									
II	Type	X	O	X	X	O	O	O	O	X
	1	U	U	U	D	U	U	U	U	D
	2	D	U	U						
	3									
	4									

Round	Player	Ses. 2 (3 groups) A1	A2	A3	B1	B2	B3	C1	C2	C3
I	Type	X	X	X	X	X	X	X	X	O
	1	U	D	D	D	D	U	U	U	U
	2							U	D	D
	3									
	4									
II	Type	X	O	X	O	X	X	X	X	O
	1	U	D	U	U	D	U	U	D	D
	2									
	3									
	4									

Round	Player	Ses. 3 (3 groups) A1	A2	A3	B1	B2	B3	C1	C2	C3
I	Type	X	X	X	O	O	X	X	O	X
	1	U	D	D	U	D	D	U	U	U
	2							D	U	D
	3									
	4									
II	Type	X	X	X	X	X	X	X	X	X
	1	U	U	D	U	U	U	U	U	U
	2				U	D	U	D	U	U
	3									
	4									

Round	Player	Ses. 4 (2 groups) A1	A2	A3	B1	B2	B3
I	Type	X	X	X	O	X	X
	1	D	U	U	U	U	U
	2				U	D	U
	3						
	4						
II	Type	O	X	O	X	O	O
	1	D	D	U	D	U	U
	2						
	3						
	4						

Round	Player	Ses. 5 (3 groups) A1	A2	A3	B1	B2	B3	C1	C2	C3
I	Type	X	X	O	X	X	X	X	O	O
	1	U	U	U	D	U	U	D	D	U
	2	D	U	U						
	3									
	4									
II	Type	X	X	X	O	X	X	X	X	X
	1	U	D	U	U	U	U	U	D	U
	2				D	D	D			
	3									
	4									

Table 4: Results for three player game

there are 12 cases where all players are of type X.

This result may be due to several reasons. First of all, it may be unlikely that subjects calculate the expected payoff, but instead focus on the prior probability of 0.8. That is, participants in the experiment may view the uniformed decision as having an 8 out of 10 chance of a "good" payoff and a 2 out of 10 chance of a "bad" payoff and, therefore, ignore the sizes of the payoffs. This may be particularly true since, if someone else in their group selects Down before they do, the participant will not be able to realize any payoffs. Additionally, it is likely the case that the subjects view the additional time involved with additional periods as having negative utility. Therefore, as long as at least one person in each group does so, choosing Down guarantees that the round will end more quickly. Moreover, it may be the case that subjects do not view the determination of types for each participant as being independent of the types of the other participants. There is some support for this in that, for the 3 player condition, over-playing Down is more frequent when subjects observe two players of type X as opposed to one player of each type. This provides weak evidence that participants are behaving as though the types within a group are positively correlated. Furthermore, over-playing the action Down is considerably more frequent in the $n = 3$ condition, where subjects play D when they have no information about their type 30 percent of the time, compared to 8 percent of the time in the $n = 2$ case. This perhaps implies that more subjects are randomizing in the 3 player game, where the solution is less transparent. Finally, there is weak evidence that the over-playing decreases across rounds. For $n = 2$ the decrease is from 12 percent in the first round to 4 percent in the second, while for $n = 3$ the decrease is from 33 percent to 26 percent. Similarly to the previous explanation, this observed decrease also implies that the cause of the over-playing may be confusion which leads more players to randomize both in earlier periods and in a more confusing situation (the three player game). While the above are all possible explanations for the observed result, it is most likely that a combination of these is responsible for the subjects' behavior.

Tables 5 and 6 provide some summary results for the two conditions. The left side of each table reports the occurrence of each composition of types within a group. Since the game is symmetric, no distinction is made between participants in a group. Therefore, the case where subject A1 is of type X and subject A2 is of type O is identical to the situation where A1's type is O and A2's type is X. The frequencies of the different draws indicate that the results of the random process approximated the expected frequencies. Furthermore, the modal outcome in both conditions occurred, as expected, where all of the players are of type X. Also note that it was always the case that there was at least one player of type X in a group. Therefore, the announcement was always that there is at least one player of type X in the group for all groups.

Types	n	Predicted Behavior	Actual Behavior	
OO	0			
XO	8	(DU)	7	(0.88)
XX	18	(UU)(DD)	4	(0.22)
Total	26		11	(0.42)

Table 5: Summary of group results for $n = 2$

Types	n	Predicted Behavior	Actual Behavior	
OOO	0			
XOO	6	(DUU)	3	(0.50)
XXO	10	(UUU)(DDU)	1	(0.10)
XXX	12	(UUU)(UUU)(DDD)	0	(0.00)
Total	28		4	(0.14)

Table 6: Summary of group results for $n = 3$

On the right hand side of Tables 5 and 6, a summary of the behavior across groups is provided, aggregating between the two experimental rounds. The column *Predicted Behavior* refers to the expected behavior of the group given its composition. For example, in the XX case for $n = 2$, ((UU)(DD)) indicates that both players should choose Up in the first period and Down in the second. The next column provides the actual number of groups which behaved according to this prediction, as well as the corresponding frequencies. Note that while the number of groups whose behavior corresponds to the theoretical prediction is low in both conditions, this is particularly true for the three player game. In fact, while for the two player case 42 percent of the groups behave according to what is predicted, this is true of only 14 percent of the groups when $n = 3$. Additionally, none of the 12 groups in the XXX situation behave according to the theory. More surprisingly, the predicted behavior occurs in only 50 percent of the groups with types XOO, where only one of the players knows her type and should choose Down. The results in the tables indicate quite convincingly that group behavior does not conform to the predicted play in either of the games.

An interesting question is whether or not the failure of groups to behave according to theory differs systematically between rounds and across the two games. In the two player game, for instance, 31 percent of the groups in the first round behave as predicted, while for the second round, this is true of 54 percent of the groups. Similarly, for the three player game, 7 percent of the groups in Round I and 21 percent in Round II behave accordingly. These differences, however, are not significant. A Fisher exact test of the null hypothesis that behavior is the same across rounds produces p-values of 0.218 and 0.298 for the 2 and 3 player games, respectively. There is a difference, however, between groups in the two conditions. In the two player game, groups behave according to the theory 42 percent of the time, while this is true for only 14 percent of the observations in the 3 player game. This difference is significant in both a Fisher exact test ($p = 0.022$) and a Chi-Square Test ($p < 0.05$).[21] In summary, the results provide weak evidence for effects of both group size and experience on whether or not groups behave according to what is predicted.

Returning to the situation in the three player game where the types are XOO and the player of type X, therefore, knows her type, it is worthwhile to re-examine the result that the groups behaved according to the prediction only one half of the time. Although this is surprising since the solution appears to be obvious, it is true that in each of these cases the player who knew her type did choose Down in the first period. The failure of the groups to behave according to the theory, then, is the result of the previously observed over-playing of the action Down by other members of the group, and does not imply that the player who knew her type behaved irrationally. For this reason, it seems appropriate to further examine behavior on an individual, rather than group, basis.

[21] However, the significance levels are possibly exaggerated since both tests rely on the independence of the observations, which is not present in this sample across two rounds. Despite this limitation, pooling data within groups is necessary for the tests to have power and, furthermore, common in practice.

4.2.2 Individual Behavior

In order to examine individual behavior, and whether it corresponds to predicted behavior, it is necessary to consider the information held by players when they make their choice. Thus, a player choosing an action at any time may or may not be behaving according to the theory, depending on what this player has previously observed with regard to the types and actions of other players. Tables 7 and 8 present the behavior of players conditional upon their information. The second column presents each action, Up or Down, along with the possible information a player received about the types of the other subjects in her group. Hence (D,X) in the two player case corresponds to the situation where a participant observed that the other player was of type X and, subsequently, chose Down. The next set of columns contain the number of participants with the given information who took that action, and the frequencies expected probabilities of that behavior. Behavior for each period is reported separately. The data is provided for both rounds, as well as for the aggregate of the two rounds.

Note that, at the individual level, adherence to the equilibrium prediction is considerably greater than for the groups. For example, looking at the first period in the two player case, 88 percent of subjects behave according to the equilibrium prediction in the first round, while 92 percent do so in the second round. In the second period, 50 percent of the subjects in Round I and 56 percent in Round II behave as predicted.

		Period 1			Period 2*		
		n	freq.	pred.	n	freq.	pred.
Round I	(U,O)	0	0.00	0.00			
	(D,O)	3	1.00	1.00			
	(U,X)	20	0.87	1.00	7	0.50	0.00
	(D,X)	3	0.13	0.00	7	0.50	1.00
Total Agreements		23	0.88		7	0.50	
Total Violations		3	0.12		7	0.50	
Round II	(U,O)	1	0.20	0.00	1	1.00	0.00
	(D,O)	4	0.80	1.00	0	0.00	1.00
	(U,X)	20	0.95	1.00	6	0.40	0.00
	(D,X)	1	0.05	0.00	9	0.60	1.00
Total Agreements		24	0.92		9	0.56	
Total Violations		2	0.08		7	0.44	
Total	(U,O)	1	0.13	0.00	1	1.00	0.00
	(D,O)	7	0.87	1.00	0	0.00	1.00
	(U,X)	40	0.91	1.00	13	0.45	0.00
	(D,X)	4	0.09	0.00	16	0.55	1.00
Total Agreements		47	0.90		16	0.53	
Total Violations		5	0.10		14	0.47	

Table 7: Summary of individual behavior in two player game

* Both players chose U in Period 1

The results are similar for the three player case, where the first period percentages are 71 and 81, respectively, for Rounds I and II. In the second period, subjects behave according to the theory 67 percent of the time in both rounds.

Noticing again that the percentages differ by group size and by round, the tests from the previous section were repeated using individual data. The data used in these tests are reported in Table 9. This table again reports a player's observation of the types of the other players and the predicted behavior. In this case, however, the behavior is not by periods but across periods. Hence, if a participant observed XO, the predicted behavior is that she would select Up in the first period and Down in the second. This table reports the numbers

		Period 1			Period 2*		
		n	freq.	pred.	n	freq.	pred.
Round I	(U,OO)	0	0.00	0.00			
	(D,OO)	2	1.00	1.00			
	(U,XO)	11	0.79	1.00	3	0.38	0.00
	(D,XO)	3	0.21	0.00	5	0.62	1.00
	(U,XX)	17	0.65	1.00	5	0.71	1.00
	(D,XX)	9	0.35	0.00	2	0.29	0.00
Total Agreements		30	0.71		10	0.67	
Total Violations		12	0.29		5	0.33	
Round II	(U,OO)	0	0.00	0.00			
	(D,OO)	4	1.00	1.00			
	(U,XO)	15	0.83	1.00	1	0.25	0.00
	(D,XO)	3	0.17	0.00	3	0.75	1.00
	(U,XX)	15	0.75	1.00	5	0.62	1.00
	(D,XX)	5	0.25	0.00	3	0.38	0.00
Total Agreements		34	0.81		8	0.67	
Total Violations		8	0.19		4	0.33	
Total	(U,OO)	0	0.00	0.00			
	(D,OO)	6	1.00	1.00			
	(U,XO)	26	0.81	1.00	4	0.33	0.00
	(D,XO)	6	0.19	0.00	8	0.67	1.00
	(U,XX)	32	0.70	1.00	10	0.67	1.00
	(D,XX)	14	0.30	0.00	5	0.33	0.00
Total Agreements		64	0.76		18	0.67	
Total Violations		20	0.24		9	0.33	

Table 8: Summary of individual behavior in three player game

* Both players chose U in Period 1

and frequencies of agreements and violations in each case, and for each round. Although the hypothesis that experience reduces the number of violations is supported for both games, that is, the second round produces a higher frequency of agreements for both games, this difference is not significant at any reasonable levels in a Chi-Square Test. The hypothesis that violations are more frequent in the $n = 3$ condition is also not supported by these data; neither of the differences is significant and, in fact, it is in the wrong direction for Round II. Thus, there is no support for this hypothesis and only weak support for the former one.

Number of Players	Round	Obs.	Pred.	Agreements n	freq.
2	I	O	D	3	1.00
		X	UD	13	0.57
		Total		16	0.62
	II	O	D	4	0.80
		X	UD	14	0.67
		Total		18	0.69
3	I	OO	D	2	1.00
		XO	UD	8	0.57
		XX	UUD	15	0.58
		Total		25	0.60
	II	OO	D	4	1.00
		XO	UD	14	0.78
		XX	UUD	12	0.60
		Total		30	0.71

Table 9: Individual Behavior Across Periods

It is also possible to use individual behavior to test the validity of models which give predictions as to the behavior in these games. Tables 10 and 11 provide tests of and comparisons between alternate models for the two and three player games respectively. In both tables, the models are estimated separately for Rounds I and II as well as for the aggregate data of both rounds. For each model, the table presents the predicted probabilities of playing Up, given a player's information about the types of the other player or players. Hence, $P(U_1, X)$ represents the probability of playing Up in the first period after observing that the other player is of type X in the two player game and $P(U_2, XO)$ represents the probability of playing Up in the second period after observing one player of type X and one of type

O in the three player game.[22] The probabilities are only given for the information sets for which there is data and, hence, the models are only compared with respect to their predictions for these probabilities. In addition, the observed frequencies in each case are also given.

Some of the models include a parameter which determines the resulting predicted probabilities from a correspondence of probabilities. This parameter is given, where applicable, in the second to last column. Finally, the log-likelihood for each set of probabilities is also reported.

Round	Model	$P(U_1,O)$	$P(U_1,X)$	$P(U_2,X)$	$P(U_2,O)$	param.	LL
I	Nash*	0.000	1.000	0.000			$-\infty$
	Rand.	0.500	0.500	0.500			-27.73
	NNM	0.250	0.750	0.250		0.500	-22.49
	QRE	0.321	0.820	0.324		3.155	-20.90
	PIM	0.170	0.830	0.394		0.660	-19.62
	Actual	0.00	0.87	0.50			
II	Nash*	0.000	1.000	0.000	0.000		$-\infty$
	Rand.	0.500	0.500	0.500	0.500		-29.11
	NNM	0.214	0.786	0.214	0.214	0.572	-21.82
	QRE	0.340	0.862	0.303	0.070	3.677	-20.77
	PIM	0.142	0.858	0.345	0.142	0.716	-19.72
	Actual	0.20	0.95	0.40	1.00		
Aggregate	Nash*	0.000	1.000	0.000	0.000		$-\infty$
	Rand.	0.500	0.500	0.500	0.500		-56.84
	NNM	0.232	0.768	0.232	0.232	0.536	-44.39
	QRE	0.328	0.838	0.315	0.088	3.360	-41.74
	PIM	0.156	0.844	0.370	0.156	0.689	-39.43
	Actual	0.13	0.91	0.45	1.00		

Table 10: Comparisons between predictions of behavior in two player game

[22] In the latter case, the participant also has the information that both of the other players chose Up in the first period.

Round	Model	$P(U_1,OO)$	$P(U_1,XO)$	$P(U_1,XX)$	$P(U_2,XO)$	$P(U_2,XX)$	param.	LL
I	Nash*	0.000	1.000	1.000	0.000	1.000		$-\infty$
	Rand.	0.500	0.500	0.500	0.500	0.500		-39.51
	NNM	0.298	0.702	0.702	0.298	0.702	0.404	-34.73
	QRE	0.220	0.691	0.652	0.324	0.803	2.199	-34.54
	PIM	0.270	0.730	0.730	0.518	0.730	0.460	-34.97
	Actual	0.00	0.79	0.65	0.38	0.71		
II	Nash*	0.000	1.000	1.000	0.000	1.000		$-\infty$
	Rand.	0.500	0.500	0.500	0.500	0.500		-37.43
	NNM	0.222	0.778	0.778	0.222	0.778	0.556	-28.60
	QRE	0.223	0.744	0.728	0.274	0.854	2.553	-29.61
	PIM	0.206	0.794	0.794	0.448	0.794	0.558	-28.96
	Actual	0.00	0.83	0.75	0.25	0.62		
Aggregate	Nash*	0.000	1.000	1.000	0.000	1.000		$-\infty$
	Rand.	0.500	0.500	0.500	0.500	0.500		-76.94
	NNM	0.261	0.739	0.739	0.261	0.739	0.478	-63.76
	QRE	0.220	0.714	0.684	0.301	0.828	2.357	-64.40
	PIM	0.237	0.763	0.763	0.487	0.763	0.525	-64.22
	Actual	0.00	0.81	0.70	0.33	0.67		

Table 11: Comparisons between predictions of behavior in three player game

Five different predictive models are compared with each other. The first two of these are an approximation of the Nash equilibrium and a model where players act entirely in error and, therefore, choose each action with equal probability in all cases.[23] We also estimate a simple model, labelled the Noisy Nash Model (NNM), which determines a choice probability correspondence as a function of a parameter, γ.[24] In this model players make probabilistic errors which are ignored by all players. The predicted probabilities in this model consist of all points which lie along the linear combination between the Nash equilibrium and random play. The parameter γ, therefore, is estimated from the interval [0,1] and, for a given value

[23]An approximation is used rather than the pure strategy equilibrium itself for the reason that, since the predictions are in pure strategies, the log-likelihood at the Nash equilibrium is undefined. Therefore, the reported value is the limit as the probabilities approach the equilibrium.

[24]This approach is similar to that of Smith and Walker (1993).

of γ, a player will choose correctly only with probability $0.5(1-\gamma)+1\gamma = 0.5+0.5\gamma$. Thus, when $\gamma = 0$, the model predicts random play, and for values of γ close to 1, the prediction approximates the Nash equilibrium.

Similarly to the above model, the logit specification of McKelvey and Palfrey's Quantal Response Equilibria (1998) produces, as a correspondence of a parameter λ, a progression from random play to the Nash Equilibrium. The QRE model incorporates error into the best response of players in a game, so that better responses are more likely to be played, but in this case the behavior of players is in equilibrium and they take into account the error in everyone's choices. The parameter λ, which is inversely related to error, measures precision in that when $\lambda = 0$ players are acting entirely in error while high values of λ correspond, in the limit, to the Nash equilibrium.[25].

Finally, a simple model called the Probabilistic Information Model (PIM), which incorporates the possibility that players probabilistically ignore the actions of other players is included in the estimation. This model incorporates the principal aspect of the NNM, that players make probabilistic errors and fail to realize that such errors occur, and develops it one step further. In all cases where only **R**, or rationality, is necessary to make a decision, such as in the first period, behavior is identical to that under the NNM, with the parameter

[25]The calculations for the results reported in these tables were conducted using the GAMBIT Command Language (McKelvey, et al., 1996). For a formal and more detailed description of Quantal Response Equilibria for extensive form games, see McKelvey and Palfrey (1998)

here labelled ρ. However, when $\mathbf{R}\&\mathcal{K}^1(\mathbf{R})$ is necessary, players may also ignore information they have received with probability ρ. Hence, in the two player case where both players are of type X, each player will select Up with probability $0.5 + 0.5\rho$ in the first period. Under the standard assumptions, however, if the game reaches the second period, then each player should now be aware of her own type. Using the same parameter, ρ, in this model, the player probabilistically either fails to realize or ignores this information, and, hence, selects Down in the second period only with probability $\rho(0.5+0.5\rho)+(1-\rho)(0.5-0.5\rho) = 0.5-0.5\rho+\rho^2$. This is similarly extended to the three person case when a player reaches the second period and observes types X and O. Here, the subject will again choose Down with probability $0.5 - 0.5\rho + \rho^2$, rather than 1.[26]

Looking at Tables 10 and 11, it can be seen that both the Nash approximation and randomness predictions do significantly worse than all three one-parameter models with error. Of the three error models, none performs consistently better than the others. In the two-player game, PIM performs best across rounds, while for the three-player game the NNM (which performs worst for the two-player case) does best.[27]

An interesting observation is that, for all three models and in both games, the measure of

[26]This model can be extended to the third period in the three person case, where the probability of misperception by another player in the second period is included in the decision probabilities. Since there is no data for the third period, however, this is irrelevant to this analysis.

[27]The significance of these differences is not performed for two reasons. First, since the relative performanc of the models differs across treatments and rounds, it is unreasonable to test the hypothesis that one outperforms the others. Second, since the models are non-nested, standard likelihood-ratio tests can not be used.

precision (γ, λ, ρ) is always higher in the second round than in the first. The consistency of this observation again provides some support for the idea that learning, or at least a decrease in error, is taking place across rounds.[28]

In sum, while all three models do a better job of predicting behavior than the Nash or randomness predictions, none convincingly outperforms any of the others in both games. In addition, the increase in precision in all three models provides modest support for the hypothesis of learning. While there is some support for the hypothesis that learning occurs across periods, it is impossible to determine whether any learning is taking place without examining behavior over more periods. A second set of experiments was conducted to address this concern.

5 Experiment 2: Learning in the dirty faces game

5.1 Experimental Design

Additional experiments were conducted to address the issue of whether learning would take place with more experience. Group sizes of both 2 and 3 were again used. The experiments were conducted similarly to those above with a few minor exceptions.

[28]However, individual tests of the restrictions $\gamma_I = \gamma_{II}$, $\lambda_I = \lambda_{II}$, and $\rho_I = \rho_{II}$ fail to reject, at any reasonable levels, the constrained models where the parameter is the same for both rounds,

First, the game was played for more rounds to allow any learning to take place. Subjects played the game for 9 rounds. Before the first round, there was again a practice round in which payoffs were smaller and in which no announcement was made. Instead of being paid their earnings in all rounds, subjects were now paid according to a lottery procedure at the end of the experiment that randomly selected two out of the nine experimental rounds. Their earnings for the experiment were determined by their earnings in these two rounds plus a $29 participation bonus.[29]

Second, the game was changed slightly in order to attempt to reduce the extent to which subjects over-played the action "Down" in experiment 1. The value of p was changed from 0.8 to 0.67 ($\frac{2}{3}$) and the payoffs associated with outcomes in each round were changed to the values in Table 12. Therefore, a subject choosing D without any additional information about her type faced an expected value of -$2.33.[30]

Finally, subjects in these experiments consisted of both UCLA and Caltech graduate and undergraduate students with little or no formal training in game theory. While the inclusion of UCLA subjects complicates the possibility of making direct comparisons with the results

[29]The lottery procedure was used because paying subjects for each experimental round would have required a much higher participation bonus since there is the possibility that a subject could lose a substantial amount of money in every round. While paying subjects according to this lottery procedure might have an effect on behavior compared to experiment 1, the purpose of these experiments was not such a comparison, but rather simply to examine whether learning takes place with experience.

[30]If we consider that each round was selected to be one in which payoffs mattered with probability $\frac{2}{9}$, then the expected value is -$0.52. This is still greater in magnitude than the associated expected value (-$0.20) in experiment 1.

		Type	
		X	O
Action	U	0.00	0.00
	D	3.50	-14.00

Table 12: Payoffs for Experiment 2

of experiment 1, such a comparison is not the goal of these experiments. Instead, the goal is simply to examine whether or not learning takes place when the game is played repeatedly.

5.2 Results

The experiments were conducted in May and June of 1997 at UCLA and Caltech. Three sessions were conducted for each treatment resulting in 10 groups for the two-player treatment and 9 groups for the three-player treatment. The aggregate data for these experiments is reported in the appendix. Table 13 reports individual level data for each round similar to the data reported in Table 9 for experiment 1. Each entry in the table presents the number of times (and the corresponding frequency) that a subject who observed a particular set of other players' types behaved consistently with the theoretical prediction. The total number of observations for each round is not constant for either treatment because the table omits all cases in which all players were of type O.[31]

Notice first the high frequency of compliance across treatments with the predicted be-

[31] All players selected U for all periods in these cases, as expected.

	Number of players						
	2			3			
Obs.	O	X	Total	OO	XO	XX	Total
Pred.	D	UD		D	UD	UUD	
Round	Agreements: n (freq.)						
I	3 (1.00)	5 (0.39)	8 (0.50)	2 (1.00)	4 (0.50)	2 (0.12)	8 (0.30)
II	5 (1.00)	10 (0.77)	15 (0.83)	4 (1.00)	9 (0.56)	4 (0.57)	17 (0.63)
III	3 (1.00)	10 (0.67)	13 (0.72)	1 (1.00)	1 (0.17)	6 (0.35)	8 (0.33)
IV	4 (1.00)	10 (0.71)	14 (0.78)	2 (1.00)	6 (0.75)	7 (0.50)	15 (0.63)
V	4 (1.00)	9 (0.75)	13 (0.81)	1 (1.00)	5 (0.63)	5 (0.28)	11 (0.41)
VI	4 (1.00)	10 (0.83)	14 (0.88)	2 (1.00)	5 (0.63)	10 (0.59)	17 (0.63)
VII	5 (1.00)	9 (0.69)	14 (0.78)	4 (1.00)	13 (0.81)	5 (0.71)	22 (0.82)
VIII	3 (0.75)	11 (0.79)	14 (0.78)	4 (1.00)	12 (0.86)	5 (0.56)	21 (0.78)
IX	3 (0.75)	6 (0.75)	9 (0.75)	2 (1.00)	7 (0.58)	8 (0.62)	17 (0.63)
Agg.	34 (0.94)	80 (0.70)	114 (0.76)	22 (1.00)	62 (0.65)	52 (0.44)	136 (0.574)

Table 13: Individual Behavior Accross Periods in Experiment 2

havior in the simplest case when there is only one player of type X (i.e., a player observes O or OO and knows there is at least one player of type X). In every instance of this but two, players correctly inferred their own type and selected D in the first period. Moreover, both of the exceptions can be accounted for by one group in which social utilities – rather than a lack of rationality – appears to be the source of the "irrational" behavior.[32] Therefore, the corresponding frequencies are high for both treatments (0.94 and 1.00).

As before, the frequency of behavior consistent with the theoretical prediction decreases

[32]In the last two rounds of one session in the two player treatment (Group A in the UCLA 5/9/97 experiment reported in Appendix B), a player observed O but selected U, even though she had selected D both previous times she had encountered the same observation. This caused the other player to select D incorrectly in the next period and sustain an (expected) loss. That this subject was behaving spitefully and not merely making a mistake can be supported by two pieces of evidence. First, as mentioned before the subject had behaved "rationally" in the same situation previously. Second, in the following period, she observed X and, instead of selecting U and then D (as she had done before), selected U in the first two periods. It seems she correctly inferred that the other player was attempting to extract revenge.

as the problem becomes more complicated. In the two-player treatment, the pooled frequency for all nine rounds is 0.94 for the simplest case and 0.70 for the case that requires one step of iterated rationality. In the three-player treatment, the frequency is 1.00 when only rationality is required, 0.65 when one step of iterated rationality is required, and 0.44 when two steps are required. Also as in experiment 1, the frequency of agreements seems to be consistent across group sizes, holding the difficulty of the problem fixed. In both treatments, the frequency of agreements in the simplest case (O or OO) is close to 1. When players observe one other player of type X, the frequency of agreements is 0.70 in the two-player treatment and 0.65 in the three-player treatment.

Finally, note that there is clear evidence of learning across rounds. The frequency of total agreements in the $n = 2$ treatment rises from 0.50 to 0.75 between the first and last rounds. This frequency similarly rises from 0.30 to 0.63 in the $n = 3$ treatment. This is in spite of the fact that there is one group in the $n = 2$ treatment (as noted above) in which both subjects' behavior in the last round and one subject's behavior in the second to last round is counted as "irrational" even though they both behaved consistently with the theory in previous rounds an appear instead to be motivated by social utilities in Rounds 8 and 9. If this group is eliminated from the analysis, the frequencies for the last two rounds become 0.82 and 0.90. Comparing the difference in total number of agreements and disagreements in the first four rounds (1 through 4) with the last four rounds (6 through 9),[33] results in

[33]Choices in the above group for the last two rounds were omitted from the analysis.

a difference significant at the $p < 0.08$ level for $n = 2$ and at the $p < 0.01$ level for $n = 3$, using a one-tailed Fisher exact test.

The results of this experiment support the hypothesis that learning takes place across rounds. Taken together with the consistent – though not statistically significant – results from experiment 1, there is considerable evidence that repeated play leads to greater consistency with the theoretical prediction. However, the frequency of this consistency does not reach one.

6 Discussion

The equilibrium in the Dirty Faces Game makes extreme predictions concerning both players' behavior and their assumptions about other players. As the above results and analysis indicate, the behavior of actual players in the game does not approximate the theoretical prediction. The results may be somewhat compromised by the fact that subjects routinely over-play the action Down, which, in order for the equilibrium to arise, should be played only when a participant is aware of her own type. Nonetheless, that a significant part of individual behavior is consistent, or at least fails to violate, the predicted behavior indicates that play is not entirely random. Hence, models of play which specify and introduce some error perform much better than either the Nash prediction or randomness in explaining the data. Nonetheless, the predictive ability of these models does not approach perfection, indicating that a better model can still be found.

The observation that error appears to be decreasing across rounds in both treatments indicates that subjects may be able to "learn" to play the correct equilibrium. Experiment 2 tests and finds support for this hypothesis. The results of these experiments indicate that including more rounds leads to improved play relative to the prediction. However, after nine rounds, there was still a considerable amount of behavior inconsistent with the standard prediction, indicating that learning may not be sufficient to arrive at the equilibrium. In fact, an examination of the experiment 2 data reveals several instances in which subjects who had previously behaved according to the prediction stopped doing so in subsequent rounds.[34]

Finally, there is also evidence that behavior is the same across group sizes, holding the difficulty of the problem consant. Therefore, when players observe one other player of type X, the extent to which they conform to the theoretical prediction is independent of whether they are in a two- or three-player game.

The goal of this paper is to address the question of whether predicted play is consistent

[34] Some of this appears to be due to the introduction of social utility. However, another part of it appears in groups where players behaved according to theory initially but stopped doing so after observing that other players in their group were not behaving according to theory. Since in order for the equilibrium to arise it is necessary that players not only behave rationally but also that they believe others will, there exists the possibility that subjects may believe that the actions of others are random and therefore uninformative, possibly resulting in a myopic "equilibrium" in which everyone behaves according to the prediction only when they don't have to trust the behavior of other participants (i.e., when there is only one player of type X).

with actual behavior in the Dirty Faces Game. The results from the experiments indicate that subjects do not satisfy the common knowledge of rationality assumption necessary for equilibrium behavior to occur. With repeated play, behavior converges towards – but still falls short of – the theoretical prediction.

References

[1] Aumann, R. J. 1976. Agreeing to Disagree. *Annals of Statistics* 4: 1236-1239.

[2] Aumann, R. J. 1995. Backward Induction and Common Knowledge of Rationality. *Games and Economic Behavior* 8: 6-19.

[3] Bacharach, M. and M. Bernasconi. 1997. The Variable Frame Theory of Focal Points: An Experimental Study. *Games and Economic Behavior* 19(1):1-45.

[4] Bernheim, D. 1984. Rationalizable Strategic Behavior. *Econometrica* 52: 1007-1028.

[5] Binmore, K. and A. Brandenburger. 1988. Common Knowledge and Game Theory. Unpublished Manuscript.

[6] Camerer, C. F., E. Johnson, T. Rymon, and S. Sen. 1993. Cognitions and Framing in Sequential Bargaining for Gains and Losses. In *Frontiers of Game Theory*, ed. K. Binmore, A. Kirman and P. Tani. Cambridge: MIT Press.

[7] Camerer, C. F., M. Knez and R. A. Weber. 1996. Timing and Virtual Observability in Ultimatum Bargaining and 'Weak Link' Coordination Games. California Institute of Technology Social Science Working Paper 9.

[8] Camerer, C. and R. H. Thaler. 1995. Ultimatums, Dictators and Manners. *Journal of Economic Perspectives* 9: 209-219.

[9] Cooper, R., D. V. DeJong, R. Forsythe and T. W. Ross. 1993. Forward Induction in the Battle-of-the-Sexes Games. *The American Economic Review* 83: 1303-1316.

[10] Costa-Gomes, M., V. Crawford and B. Broseta. 1998. Cognition and Behavior in Normal-Form Games: An Experimental Study. Unpublished Manuscript.

[11] Fudenberg, D. and J. Tirole. 1993. *Game Theory.* MIT Press.

[12] Geanakoplos, J. 1993. Common Knowledge. Cowles Foundation Discussion Paper.

[13] Geanakoplos, J. and H. Polemarchakis. 1982. We Can't Disagree Forever. *Journal of Economic Theory* 28: 192-200.

[14] Güth, W., S. Huck and A. Rapoport. 1995a. On Silent Threats: An Experimental Study of the Timing of Decisions. Unpublished manuscript.

[15] Güth, W., S. Huck and A. Rapoport. 1995b. The Positional Order Protocol: Can it Support Non-Equilibrium Outcomes. Unpublished manuscript.

[16] Hanson, R. 1996. Correction to McKelvey and Page, 'Public and Private Information: An Experimental Study of Information Pooling'. Forthcoming in *Econometrica.*

[17] Ho, T. H., K. Weigelt and C. Camerer. 1998. Iterated Dominance and Learning in Experimental 'Beauty Contest' Games. Forthcoming in *The American Economic Review.*

[18] Kahneman, D. and A. Tversky. 1979. Prospect Theory: An Analysis of Decision Under Risk. *Econometrica* 47: 263-291.

[19] Kahneman, D. and A. Tversky. 1992. Advances in Prospect Theory: Cumulative Representation of Uncertainty. *Journal of Risk and Uncertainty* 5: 297-323.

[20] Littlewood, J. E. 1953. *A Mathematician's Miscellany.* Meuthen & Co. Ltd.

[21] McKelvey, R. D. and T. Page. 1986. Common Knowledge, Consensus, and Aggregate Information. *Econometrica* 54: 109-127.

[22] McKelvey, R. D. and T. Page. 1990. Public and Private Information: An Experimental Study of Information Pooling. *Econometrica* 58: 1321-1339.

[23] McKelvey, R. D. and T. R. Palfrey. 1992. An Experimental Study of the Centipede Game. *Econometrica* 60: 803-836.

[24] McKelvey, R. D. and T. R. Palfrey. 1998. Quantal Response Equilibria for Extensive Form Games. *Experimental Economics* 1: 9-41.

[25] McKelvey, R. D., A. McLennan, T. Turocy and G. Wu. 1996. GAMBIT COMMAND LANGUAGE. Program available at *http://www.hss.caltech.edu/ gambit/Gambit.html.*

[26] Moulin, H. 1986. *Game Theory for the Social Sciences.* New York University Press.

[27] Nagel, R. 1995. Experimental Results on Interactive Competitive Guessing. *The American Economic Review* 85: 1313-1326.

[28] Nielsen, L. T., A. Brandenburger, J. Geanakoplos, R. McKelvey, and T. Page. 1990. Common Knowledge of an Aggregate of Expectations. *Econometrica* 58: 1235-1239.

[29] Pearce, D. 1984. Rationalizable Strategic Behavior and the Problem of Perfection. *Econometrica* 52: 1029-1050.

[30] Rapoport, A. and M. A. Fuller. 1995. Coordination in Non-cooperative Three-Person Games under Different Information Structures. Unpublished manuscript.

[31] Rubinstein, A. 1989. The Electronic Mail Game: Strategic Behavior Under 'Almost Common Knowledge'. *The American Economic Review* 79: 385-391.

[32] Schotter, A., K. Weigelt and C. Wilson. 1994. A Laboratory Investigation of Multi-person Rationality and Presentation Effects. *Games and Economic Behavior* 6: 445-468.

[33] Smith, V. and J. Walker. 1993. Monetary Rewards and Decision Cost in Experimental Economics. *Economic Inquiry* 31: 245-261.

[34] 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.

[35] 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.

7 Appendix A: Instructions

Initial Instructions

This is an experiment in decision making, and you will be paid for your participation in cash. Different participants may earn different amounts. What you earn depends partly on your decisions and partly on the decisions of others.

During the course of the experiment, all interaction between participants will take place through the experimenter. It is important that you not talk or in any way try to communicate with other participants during the experiment. If you disobey the rules, we will have to ask you to leave the experiment.

If you have any questions during the instruction period, raise your hand and your question will be answered so that everyone can hear. If you have any further questions after the experiment has begun, raise your hand, and an experimenter will come and assist you.

At this time, one of the participants will be randomly selected as the monitor for the experiment. The monitor will assist in conducting the experiment and will be paid a fixed sum. Each participant will roll a die at the front of the room and the participant who rolls the highest number will be the monitor for the experiment. Ties will be resolved by another roll. Please note that each participant has an equally likely chance of being selected to be the monitor for the experiment.

The remaining participants will each be randomly assigned a participant number by selecting an envelope from the experimenter. Please select an envelope now as the experimenter passes around the room. Inside the envelope you have selected, there is an index card with your participant number on it. This will be your participant number for the entire experiment.

There will be four groups, each containing two participants. Each participant will interact only with the other participant in his or her group. The groups will be labeled A, B, C, and D. The participants in each group will also have a number which will identify which participant they are within the group. This number will be either 1 or 2. Therefore, the following will be the participant numbers for the experiment: A1, A2, B1, B2, C1, C2, D1, and D2. Please note that the participant numbers are private and should not be shared with anyone during the experiment.

The experiment will consist of three rounds, Round I , Round II, and Round III. Each round will consist of up to 3 periods. At the beginning of Round I, the monitor will roll the die to determine the type of each participant, beginning with A1. The type of each participant will be either "X" or "O". If the monitor rolls a value of 1 or 2, then this participant will be of type "O". If the monitor rolls a value of 3, 4, 5, 6, 7, 8, 9, or 10, then the participant will be of type "X".

When this process is done, all of the participants will be of either type "X" or type "O". Note that each participant has an 8 out of 10 chance of being a type "X" and a 2 out of 10 chance of being a type "O", and that the type of each participant is independent of the types of all other participants. Each participant's type will remain the same for the entire round.

While the monitor is determining the type of each participant, he or she will not be visible to any of the other participants. The outcome of the rolls will therefore not be known by any of the participants. The monitor will record the type of each participant on a sheet identical to the one at the front of the room labeled Type Sheet. Once the monitor has recorded each participant's type on the sheet, it will be placed inside the cardboard display at the front of the room.

The experimenter will then show each participant the type of the other participant in his or her group by lifting the corresponding flap on the cardboard display. When the experimenter comes to you, please record the type of the other participant in your group on the sheet labeled "Round I Record Sheet". The flap corresponding to the participant who is currently viewing the display, however, will remain closed. Thus, every participant will know only the type of the other participant in their group. No participant, however, will be aware of his or her own type. Note that it is the same sheet which is being shown to everyone.

For example, participant A1 will now know the type of participant A2, but will not know the types of participants A1 (him or herself), B1, B2, C1, C2, D1, and D2. Likewise, participant B2 will now know the type of participant B1, but will not know the types of participants B2, A1, A2, C1, C2, D1, or D2.

At the end of the round, the experimenter will hold up the Type Sheet for Round I at the front of the room so that each participant can observe the types of all the participants and verify that the information received at the beginning of the round was correct.

Once every participant has recorded the type of the other participant in their group, then the first period of Round I will begin. In each period, you will have the opportunity to earn or lose money. Please look at Table 1 now, it describes how your payoffs for each period will be determined. In each period, you will choose between one of two actions: Up or Down. Your earnings in each period will be determined by the action you choose and by your own type. Looking at Table 1, you can see that if you choose Up, then you will earn 0 cents, regardless of whether your type is "X" or "O". If you choose Down and your type is "X", then you earn 20 cents. Finally, if you choose Down and your type is "O", then you lose 1 dollar. Please note that the type of the other person in your group does not affect your earnings in each period.

Round I will continue for each group until either participant selects Down in any period, or until three periods have passed. That is, if in any period one of the two participants in a group selects Down, then Round I will end for that group after that period.

I will now describe what happens in each period. At the beginning of each period, each participant will place a mark in one of the three boxes on their Reporting Sheet for that period. If you wish to choose an action of Up in that period, then place a check in the box corresponding to that choice. Similarly, if you would like to choose Down, then you should place a check in that box. Finally, if you or the other participant in your group has previously chosen Down, then you should place a check in the box labeled "Round Over". This box should only be checked in periods where the experimenter has instructed participants in that group to do so. Participants should also record their choice of action for each period on their Round I Record Sheet.

Once you have placed a check in one of the three boxes, then tear the Reporting Sheet for that period off and place it face down on you desk. The experimenter will come by to collect the Reporting Sheets for all participants once they have all done so.

Once the experimenter has collected all of the Reporting Sheets, he will write the actions selected by each participant, referring to them only by their participant number, on the board at the front of the room. Once this is done, all of the participants should record the action selected by the other participant in their group on their Round I Record Sheet. After all participants have done so, the experiment will proceed to the next period. Once three periods have passed, Round I will end and we will proceed to Round II. After Round III, you will be paid, in private, the total you have earned in all three rounds plus an $11 participation bonus. No other person will be told how much cash you earned in the experiment. You need not tell any other participants how much you earned.

Are there any questions before we begin Round I?

If there are no further questions, we will now begin with the experiment by selecting the monitor. If there are any problems or questions from this point on, raise your hand and an experimenter will come and assist you.

Inter-Round Instructions (After Round I)

Round I is now completed. I will now place the Type Sheet for Round I at the front of the room. Please record your type on your Round I Record Sheet. Using your type for Round I, please calculate your earnings for this round and record this amount at the bottom of your Round I Record Sheet. Once everyone has done that, then we will begin Round II.

Rounds II and III will proceed in exactly the same manner as in Round I except for two differences. First, the payoffs will be different from those used in Round I. Please look at Table 2 which the experimenter handed to you with these instructions. Notice that the payoffs are different from those in Table 1. Your earnings in each period will again be determined by the action you choose and by your own type. Looking at Table 2, you can see that if you choose Up, then you will earn 0 cents, regardless of whether your type is "X" or "O". If you choose Down and your type is "X", then you earn 1 dollar. Finally, if you choose Down and your type is "O", then you lose 5 dollars. Please note that the type of the other person in your group does not affect your earnings in each period.

Second, once the type of each participant has been determined, the experimenter will make an announcement for each group, indicating whether it is the case that there is at least one player of type "X" in that particular group. For example, if it is the case that no participants in group D are of type "X", then the experimenter will announce, "There are no participants of type 'X' in group D". Otherwise, the experimenter will announce, "There is at least one participant of type 'X' in group D".

Are there any questions before we begin Round II?

If there are no further questions, we will now begin with Round II. If there are any problems or questions from this point on, raise your hand and an experimenter will come and assist you.

Inter-Round Instructions (After Round II)

Round II is now completed. I will now place the Type Sheet for Round II at the front of the room. Please record your type on your Round II Record Sheet. Using your type for Round II, please calculate your earnings for this round and record this amount at the bottom of your Round II Record Sheet. Once everyone has done that, then we will begin Round III.

Final Instructions

Round III is now completed. I will now place the Type Sheet for Round III at the front of the room. Please record your type on your Round III Record Sheet. Using your type for Round III, please calculate your earnings for this round and record this amount at the bottom of your Round III Record Sheet.

After your are done, add together your Round I total, Round II total, Round III total, and the $11 participation bonus to get your final payoff for the experiment. Please record this on your Experiment Earnings sheet along with your name, social security number, and today's date. Please wait until after you have received payment to write your signature. If there are any problems or questions, please raise your hand.

After you are done calculating your earnings for the experiment, please remain seated. You will be paid at the front of the room one at a time in the order indicated by the experimenter. Please bring all of your things with you when you come to the front of the room.

You can leave the experiment as soon as you are paid.

Please refrain from discussing this experiment while you are waiting to receive payment so that privacy regarding individual choices and payoffs may be maintained. Please place all of your experiment materials, except for the Experiment Earnings sheet, inside of the envelope which you were given at the beginning of the experiment.

Thank you all very much for participating in this experiment.

8 Appendix B: Data for Experiment 2

n = 2

	UCLA 5/2/97						UCLA 5/9/97						Caltech 6/10/97							
	A1	A2	B1	B2	C1	C2	A1	A2	B1	B2	C1	C2	A1	A2	B1	B2	C1	C2	D1	D2
Prac.	X D	X D	O U D	X U U	O D	X U	X D	O U	O U U	X U D	X U U D	X U U U	X D	X U	O U D	X U U	X D	X U	O U	X D
I	O D	X D	O D	X D	X U U U	X U U U	X D	X U	X U U	X U D	O U U U	O U U U	O D	X D	X U D	X U D	X U U	X U D	O U U U	O U U U
II	X D	X U	X U D	X U D	O U U U	O U U U	O U	X D	O U	X D	X U D	X U U	X D	X U	O U	X D	X D	O U	X D	O U
III	X D	O U	O U	X D	X D	X U	X D	X U	X D	O U	X U D	X U D	X D	X D	X U	X D	O U U U	O U U U	X U D	X U D
IV	O D	X D	X U D	X U D	O U	X D	X D	O U	X U U	X U D	O U U U	O U U U	X D	X D	O U	X D	X U D	X U D	X U D	X U D
V	O U U U	O U U U	X D	O U	X D	X U	O U	X D	O U	X D	X U D	X U D	X D	X D	X D	O U	O U U U	O U U U	X U D	X U D
VI	X D	O U	O U U U	O U U U	X D	X D	O U	X D	X U D	X U D	X D	O U	O U U U	O U U U	X U D	X U D	X U D	X U D	O U	X D
VII	O U	X D	X D	O U	X D	X D	X U D	X U D	O U	X D	O U	X D	X D	X D	X U D	X U D	O U	X D	O U U U	O U U U
VIII	X D	O U	X D	O U	X U U U	X U U U	X U D	O U D	X U D	X U D	O U U U	O U U U	X D	X U	X U D	X U D	X U D	X U D	X D	O U
IX	O U U U	O U U U	X U D	X U D	O U U U	O U U U	O U U	X U D	O U	X D	X U D	X U D	X D	O D	X D	O U	O U U U	O U U U	O U U U	O U U U

n = 3

	UCLA 5/2/97									UCLA 5/9/97									Caltech 6/12/97								
	A1	A2	A3	B1	B2	B3	C1	C2	C3	A1	A2	A3	B1	B2	B3	C1	C2	C3	A1	A2	A3	B1	B2	B3	C1	C2	C3
Prac.	O U	X D	X U	X D	X U	O D	X U	O D	X U	X U U	O U U	X U D	O U U U U	X U U U U	X U U U U	O D	X U	X U	X U	X D	X U	X U U U U	O U U U U	X U U U D	X U U U	O U U D	X U U U
I	X U U U U	O U U U U	X U U U U	X U U D	O U U U	X U U U	O U	X D	O U	O U	O U	X D	X U U U U	X U U U U	X U U U U	X U U U U	X U U U U	X U U U U	X U	X D	X U	X U U U U	X U U U U	X U U U D	X U U U U	X U U U U	X U U U U
II	X D	O U	O U	O U U U	X U U D	X U U U	O U	X D	O U	X D	O U	O U	X U U	X U D	X U U	X D	X U	O U	O U	X D	X D	X U U U U	X U U U D	O U U U D	O U	O U	X D
III	X U U U U	X U U U U	X U U U U	X U U U U	X U U U U	X U U U D	O U U U	X U U D	X U U U	O U U U U	O U U U U	O U U U U	X U	X D	X U	O D	X D	O U	X U	X D	X U	X U U U U	O U U U U	X U U U D	X U U U	X U U D	X U U D
IV	O U	O U	X D	X U U D	O U U U	X U U U	X U D	X U D	O U U	X D	O U	O U	X U	X D	X U	X U U	X U U	X U D	O U U U U	O U U U U	O U U U U	X U U U U	X U U U U	X U U U D	X U	X D	X U
V	O U	X D	O U	X U U U U	X U U U U	X U U U D	X U U U D	O U U U U	X U U U U	X U	X D	X U	X U	O D	X U	X D	X U	X U	X U	O U	X D	X U U U U	X U U U U	X U U U D	X U U U U	X U U U U	X U U U U
VI	X U	X D	X U	O U U U	X U U D	X U U U	O U	O U	X D	X U	X D	X U	X U	X D	O U	X U D	X U U	X U U	X U U	X U D	X U D	X D	O U	O U	X U U U U	X U U U U	X U U U U
VII	X D	O U	O U	X U U	X U D	O U U	X D	O U	O U	X U	X D	X U	X U	O D	X U	X D	O U	O U	X U	X D	O U	O U	X D	O U	X U D	O U U	X U U
VIII	O U	O U	X D	X U U U U	X U U U U	X U U U D	X U	X U	X D	O U	X D	O U	X U	X D	O U	X U D	O U U	X U D	O U	X D	O D	O U	X D	O U	O U U	X U D	X U D
IX	X U	X D	X D	X U U	X U D	O U U	O U	O U	X D	X U	X D	X U	X U	X D	X U	O U U U	X U U D	X U U D	O U	X D	X D	O U	O U	X D	X U D	O U U	X U D