5/1/01 20:11 PM

Why Fairness?:
Facial expressions, evolutionary psychology, and the emergence of fairness in simple bargaining games.

by

Catherine Eckel
Department of Economics
Virginia Polytechnic Institute and State University

and

Rick K. Wilson
Department of Political Science
Rice University

Paper prepared for WOW2 -- the 25th Anniversary Workshop on the Workshop. The authors thank Laura Delgado and Sean Bolks, Rice University, for there help in conducting some of the experiments. Thanks also go to <VPI crew?> Countless nameless, but not faceless, others have listened to different versions of this paper.

Introduction

Fairness is a fundamental normative concept that has long been the focus for political philosophers. It constitutes a fundamental normative category. Positive political theorists have been less concerned with notions of fairness and far more concerned with the outcomes of strategic interaction.[1] Increasingly, however, positive theorists are facing the possibility that fairness is an important part of human strategic calculation. In part this is driven by persistent results from laboratory experiments that call into question standard game theoretic predictions such as Nash equilibrium or sub-game perfect equilibrium (see the discussion by Ostrom (1998; (2000)).

Data generated in the laboratory provide considerable fodder for game theorists who are trying to link observed regularities in human behavior with game theoretic models. This approach has been coined "behavioral game theory" (Camerer (1997)). Two branches of theoretical and experimental research have dominated recent research: the investigation of bounded rationality and learning behavior – how individuals learn to play a new game in an unfamiliar context; and the extension of utility functions to include so-called "exotic" preferences – other-regarding preferences for fairness, altruism, spite, status and sympathy.[2] While both approaches have achieved considerable insight our research focuses on the latter.

Introducing exotic preferences constitutes a marked shift in thinking about strategic play. Most models of strategic play assume that all actors approach the game in the same manner. While actors may differ in their preferences over outcomes, they are assumed to be the same in terms of strategic calculation. But suppose there are different types of people (heterogeneity in the population) and these types play different strategies.

---

[1] Of course, there are many exceptions to this overgeneralization. See, for example, Brams and Taylor (1996) and Bottom et al. (2000).

[2] This is a very active research area, with many recent journal publications and working papers. For bounded rationality and learning see, among others, Anderson (1997), Camerer and Ho (1997), Cooper, Garvin, and Kagel (1997), McKelvey (1995; (1996), Roth (1995), Slonin and Roth (1998), Samuelson (1996). On the topic of "exotic" preferences see Andreoni (1990), Andreoni and Miller (1998), Frank (1988), Rabin (1993), Rabin (1998), Sally (2001) and the references therein. Of course, many other books and papers could be cited, and what we offer here is just a sample.

For example, some individuals refuse to cooperate in bargaining settings, while others conditionally cooperate with other cooperators. In order for such a distinction to matter, there must be some way to credibly signal a "type." At the same time others must be able to read those signals and find them credible. For traditional game theorists, credibility requires costly commitment. Signals about one's type, especially absent *any* reputation, are usually costless and at best constitute "cheap talk." Nonetheless, it appears that people who are engaged in exchange often signal their type and their partners are often very good at inferring the others' type.

This research focuses on the problem of *social signaling* in which agents seek cues about others' intentions and send cues about their own intentions. Rather than accounting for all possible social signals that individuals might display (and that can be inferred), we concentrate on simple, stylized, non-verbal facial expressions. We argue that people are adept at "reading" facial expressions. In the absence of any other information about their partner, they will rely on expressions to build expectations about their partner's behavior. This ability to "read" another individual yields a mix of behavioral strategies that are much richer than most game theoretic models allow. In this paper we focus on a class of social signals that yield consistent behavioral predictions of out-of-equilibrium play in the ultimatum game.

In the next section we sketch the motivation for the consistency of behavioral strategies that deviate from sub-game perfection. In the third section we review some of the extensive literature on facial expressions and simplified icons. The fourth section discusses four experiments that lay the foundation for expectations about behavior and test those expectations using the ultimatum game. The final section discusses these results.

Motivation

An under-utilized empirical model in political science is the “ultimatum “ game. This game is a special case of the Baron and Ferejohn (1989) agenda-setting model, which in turn is related to the Rubinstein (1982) bargaining model. The Baron and Ferejohn model has spawned an enormous amount of theoretical work in political science, but has generated surprisingly few empirical results. The generic model selects one actor, according to some commonly agreed upon rule, who then proposes an allocation of benefits to a group. In turn the group votes on the allocation according to some voting rule and if accepted, the game ends. If the allocation is not accepted then a new proposer is selected and a new allocation is proposed. The key finding holds that if actors are impatient (they heavily discount the future) then under many voting rules the first proposer enjoys a substantial advantage when making allocations. That is, the first proposer can take a large share.

The "ultimatum" game is a simple variant of Baron and Ferejohn’s "closed rule" model. It is equivalent to a one-period unanimity rule in which a randomly selected proposer offers an allocation that must be accepted by all parties. This way of thinking about the problem is closely related to Rubinstein's (1982) model. In turn, that model has given rise to the study of a two-person, sequential play game that has been nicknamed the "ultimatum" game.

The structure of the ultimatum game is relatively simple. It is a two-person, complete information game with sequential play. One actor, the proposer, is given the right to divide some good. The second actor, the responder, is given the right to accept or reject the proposed division. Unlike the Baron and Ferejohn game, this game ends following the second move. Figure 1 provides a simplified version of the ultimatum game in extensive form. In this game Player 1 is the proposer and can choose left or right. By convention the top value at the end node is Player 1’s payoff, while the bottom value is Player 2’s payoff. The numbers are purely arbitrary (think of them as dollars). Once Player 1 has made her choice, then Player 2 makes a choice at the second node. Player 2 can either accept or reject the division. Accepting involves a move to the left or right (depending on the first move), while a rejection involves moving down. A rejection means both parties get nothing. Each branch of the game constitutes a subgame, and in each subgame, accepting the division is the equilibrium. Moreover, the unequal division (player 1’s move to the left) is the subgame perfect equilibrium.

<Figure 1 About Here>

While this game is very simple, political scientists have not paid much attention to this class of games. One exception is Morton and Diermeier (1997) who test aspects of Baron and Ferejohn (1989) and find it wanting. By contrast, in experimental economics the ultimatum game has generated considerable research (see reviews by Thaler (1988); Guth and Tietz (1990); Camerer and Thaler (1995); Roth (1995); Cameron (1999)). While the theoretical properties of this game are well understood, the empirical findings show that equal splits are common and rejects are not unusual. The pattern of outcomes appear to point to ways in which individuals conceive fairness.

In laboratory experiments, where subjects can choose their own allocation, the majority of subjects propose somewhere between a 50/50 and 60/40 split. This finding has been replicated in many settings and cultures (see for example, Heinrich, 2000 AER) and those proposed divisions are invariably accepted. However, when proposers take larger shares, such as an 80/20 split, then over half of the time the division is rejected. It is, Camerer (1997) notes, as if subjects have an implicit notion of fairness that they bring to these games. Traditional equilibrium concepts from game theory predict neither fairness nor rejection.

These findings pose two questions for theorists: First, what motivates relatively equal offers in this game? Second, what motivates a responder to reject any positive offer? The problem for game theorists is to conceptualize this motivation in a simple, but informative manner. Bolton (1998) argues that three explanations of motivation have dominated in discussions of fairness. The first focuses on *altruism*. Here actors have other-regarding preferences in which there is a positive return for giving to someone else. A second explanation focuses on the *distribution* of outcomes. In this sense individuals are not only concerned with their own outcomes, but also their outcomes relative to others. The final explanation focuses on *intentions*. Here people have preferences for their "type" of partner in bargaining setting. If another has been nice to you, then you might want to reply in kind.

These three motivations have been modeled as extensions to standard models of utility. In a simple (and non-technical) manner, consider an individual i's value for outcome x. The usual formulation is given by (1)

$$u_i(x_i) = v(x_i) \tag{1}$$

In this case the utility to i is completely captured by all attributes of the outcome x – it is assumed that individuals are concerned only with characteristics tied directly to the outcome. Theorists conceptualizing altruism assume that it is a trait associated with an individual and adds an additional component to i's utility. Stylized equation (2) gives a sense of this by noting that altruism is independent of the outcome x, and is a specific characteristic of the individual.

$$u_i(x_i) = v(x_i) + a_i \tag{2}$$

A more complicated motivation involves a relative comparison between what i obtains and what her partner j obtains. Equation (3) characterizes this approach.

$$u_i(x_i, x_j) = v(x_i) + v(x_i - x_j) \tag{3}$$

Clearly i's own outcome is paramount. But, at the same time i is also concerned with the perceived difference over outcomes. This effectively allows for positive and negative changes to i's utility, depending on the perceived difference in outcomes. The third motivation is not concerned with inequality aversion, but rather a concern with the partner's intention. Equation (4) illustrates this approach.

$$u_i(x_i, x_j) = v(x_i) + \alpha_i v(x_j) \tag{4}$$

Here too an individual has a separable utility function with preference for consumption not only over one's own outcome, but for the other's outcome as well. The coefficient α_i has specific characteristics related to j's intention. If i thinks j is being "nice" then $\alpha_i > 0$ and i obtains a positive benefit from j's outcome. On the other hand, if $\alpha_i < 0$, then i may regard j as being "mean" and as a consequence, i's utility will decrease. Finally, if I does not care about j, then the equation reduces to (1). The common feature to all of these approaches is that something about the other is incorporated into the utility of the actor and it is done without running afoul of interpersonal comparisons of utility. The evidence for each of these approaches is mixed.

Altruism.

As noted, altruism is modeled as part of an individual's personality. It involves an actor giving up something to another and deriving utility from doing so. Support for altruism was found in experiments using the "dictator game," which is a close relative to the ultimatum game except the responder has no opportunity to reject an allocation. A common finding is that proposers often give responders something, even though responders can do nothing in return. The splits are seldom equal, as in the ultimatum game, but giving any amount to another can be construed as altruism on the part of the proposer (see Forsythe (1994)).

Several experiments, however, call into question whether altruism is driving these results. Altruism, if a personality trait, ought to persist across a variety of strategic environments. An experiment by Hoffman et al. (1994) used a standard dictator game, but systematically manipulated the anonymity of subjects -- not only with respect to respondents, but also with respect to the experimenter. In one version of the experiment half the subjects were given two envelopes; one with ten one-dollar bills and the other with ten blank pieces of paper the size of dollar bills. A form of a mail drop was used, with ID numbers that further masked subject's identities. The other half of the subjects received the allocation. Hoffman et. al. (1994) find that as proposer anonymity increases, so too does the share taken by the proposer. They conclude that other-regarding behavior occurs when proposers are socially proximate to either the experimenter or other subjects. In a sense, proposers are building a social reputation, but only when they can be observed.

In a second experiment Eckel and Grossman (1996) use the Hoffman et al. (1994) double blind procedure, but they change the responder from an anonymous student to a reputable charity (the American Red Cross). Under this treatment Eckel and Grossman observe a dramatic increase in giving. While they conclude "that altruism is a motivating factor in human behavior" (p. 188), it is clear that altruism is not simply showered on everyone. Instead, other regarding actions require a context and a target. These findings do not support the idea that altruism is a persistent character trait, but rather they point to people being conditional in their "other regardingness."

Inequality Aversion.

A second approach to modeling other-regarding preferences has attracted substantial attention. Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) propose that individuals care about relative comparisons. Consequently, individuals pay considerable attention to how they are faring with respect to others. If one member of the relevant comparison group is made much better off, then actors may adopt strategies designed to bring that individual back in line – even if it comes at a cost. Such strategies might be conceptualized as "punishment" and should be familiar to anyone who has violated a social norm and been shamed. Conversely, if another is made much worse off, strategies might be invoked that, at a cost to one's self, makes the other better off (this might be conceptualized as empathy).

The advantage of this approach is that the same individual may use very different strategies, depending on the relevant comparisons. It does not require that individuals play the same strategies in all settings. Several experiments have been designed to test between this motivation and motivation centering on the intention of others. In one test, Bolton, Brandts, and Ockenfels (1998) turn to an experiment in which subjects makes choices in a normal form game using the strategy method. Subjects are required to specify a choice in response to every possible action by their counterpart. In their design the row player has two choices (under three distinct conditions), while the column player has six different choices. In every experimental condition the column player always has the same dominant strategy. Bolton et al. are interested in seeing whether the column player selects different columns contingent on the row choice. Indeed, there are numerous deviations from dominant strategy play. Given their design, Bolton et al. are able to make explicit predictions about whether subjects use inequality aversion or pay attention to the intentions of their counterparts. Their findings are consistent inequality aversion. However, applying the strategy method is a bit bothersome for drawing inferences about intention. Using the strategy method a subject needs on to look inwardly and never has to think too carefully about the counterpart. The strategy choices are only hypothetical until the row player's choice is revealed. Playing alone may be different than playing with someone else.

In dictator games Andreoni and Miller (1998) use a menu of choices in which the amounts that the dictator splits varies as does the payoff to the receiver. While there is considerable heterogeneity among subjects, with some keeping everything and others minimizing the differences, the results do not neatly conform to a model of inequality aversion. Numerous players sacrifice their earnings in order to increase the total payoff, with the bulk of the surplus going to the other player. Andreoni and Miller interpret this as evidence for altruism, although it is not clear whether such a trait would persist if social distance were imposed. Even so, these data question whether inequality aversion fully characterizes other-regarding preferences.

Intention.

It is reasonable that people pay attention to their partners and try very hard to read their intention before committing to a particular action. In economic exchange, as well as in political negotiation, actors often want to "get a read" on their counterpart. It is as if being able to read the other confers an advantage. This concept has gained some currency among game theorists who borrow the concept of "mindreading" from psychologists. Basically the idea has to do with the capacity of individuals to read the intention of their partner, while not drawing false conclusions. Others have fixed on the idea of "Machiavellian intelligence" which holds that humans have an evolved capacity to negotiate complex social spaces (Byrne and Whiten (1988); Whiten and Byrne (1997)). The ability to read hierarchies, know which coalitions to join and how to seek strategic advantage, are all connected with an evolved capacity for higher order cognition.

Frank (1988) has dealt extensively with the problem of reading intention and avoiding the problem of infinite regress. He argues that humans give off numerous cues about intention. Particularly useful are emotions, which he contends are difficult to mask and as such become credible signals. For example, if I can tell that my partner is angry, then Nash best response may not always be best response. Smith (1998) takes a slightly different tack. He argues that Adam Smith's concept of the "invisible hand" and his concept of "natural empathy" are founded on a single behavioral postulate: a "propensity to truck, barter and exchange one thing for another." This gregarious aspect of humans is grounded in an evolutionary past in which reciprocity outside of kinship was crucial for the spread of the species. Reciprocity, however, is a double-edged sword. If one cannot differentiate between those who are likely and unlikely to reciprocate, then reciprocity will not survive. Smith (1998) goes on to argue that the capacity to read intention is fundamental and in fact is an important part of human cognition. He claims:

> "Normal human beings, and even those with various, and substantial, limitations on their general intelligence, have intact mental modules that enable them to be intuitively aware if mental phenomenon in others. This enables me to see not

> only the value to me of possessing certain rights to act but also to know intuitively the value of such rights for others. Hence, my willingness to defend friends against internal foes. This evolved social capacity appears to be a normal part of the development of the human mind: it is as much a part of the natural order as being hungry and requiring that hunger to be satisfied." (p. 8)

As we will note later, Smith's idea of "mindreading" is not so farfetched. There is considerable research on autism that points to the importance of reading the intention of others (see for example Baron-Cohen (1995)).

Most attempts at modeling intentions and the value of reading them make reference to Rabin (1993) who offers an explicit model of a preference for fairness and clearly states it by noting

> "If somebody is being nice to you, fairness dictates that you be nice to him. If somebody is being mean to you, fairness allows -- and vindictiveness dictates -- that you be mean to him." (p. 1281)

This is not only good advice for the playground (and a norm most children learn early on), but has tractable mathematical characteristics.

Falk and Fischbacher (1998) take Rabin's model several steps further. First, they worry about whether the model can be extended to games with sequential play. Second, they worry not just about the consequences of action (outcomes) but the underlying motivation of an actor in taking some action. Falk and Fischbacher (1998) suggest that it is important to model the ways in which actors infer the intentions of their partners based on something observed about a partner. But getting a handle on the dispositions of others is a central conundrum for social science. Falk and Fischbacher rely on the context surrounding a set of decisions and use it to motivate their models. Basically they agree with Rabin that actors will differentiate among actions by others. However, that assessment will be made relative to some marker. For example, in a binary ultimatum game, the marker is an equitable allocation versus an inequitable allocation. However, in a different setting in which the proposer has the choice over two unequal splits, e.g., substitute (3,19) for (11,11) on the right branch of Figure 1, then the proposer can be forgiven for choosing the inequitable allocation that benefits the proposer. As Falk and Fischbacher argue, fairness is no longer a consideration because there is no fair allocation.

While in the spirit of models of "exotic" preference, it seems that the Falk and Fischbacher model requires a huge cognitive investment. Actors need to discern the decision context, compare it with many other contexts that they know, assess the intention of actors in each of those contexts, and finally figure out what is the fair response given the current situation. In our view, suppose people use very simple cues about others to try to discern intentionality?

On a slightly different tack, Levine (1998) proposes a version of the Rabin (1993) model, in which actors place different weights on their partner's payoffs and those weights are a function of their own characteristics (altruistic or spiteful) and beliefs about their partners. As with Falk and Fischbacher, the type of game may yield somewhat different payoffs for the same type of individual. More importantly, as Levine notes, this transforms the game into a signaling game since player's actions will reveal something about their levels of altruism or spite (1998, p. 595).

There is some reasonable evidence supporting these claims. An imaginative experiment by Blount (1995) focuses on responders in a series of ultimatum games. Each responder is asked to reveal the value at which a proposer’s allocation will be rejected. Subjects faced three different types of proposers: the first an interested party who reaps the value of the allocation; the second a neutral party who decides the allocation but has no stake in the outcome; and the third a random device that selects an allocation from a distribution. Under the first condition responders can hold beliefs about the intention of the proposer and have a basis for rationalizing a rejection. Under the second condition intentions also can be inferred, but a rejection harms an individual who had nothing to do with deciding the allocation. Finally, no intentions can be ascribed to the random mechanism. Blount (1995) finds clear differences between the random mechanism and those conditions in which people decide allocations. Subjects are much more willing to reject "unfair" offers when people make them, presumably through ascribing intention.

Experiments by Eckel (1998) and Solnick and Schweitzer (1999) vary information about partners in ultimatum games. Eckel (1998) focus on the sex of partners and design an experiment where subjects observe the group (all male, all female, or mixed) but they do not know the identity of their partner. They find that in female groups, women are less likely to reject and offers from women are less likely to be rejected. These findings are quite striking in that differences in the rejection rates are due to the sex of the actors. Solnick and Schweitzer (1999) look at both the attractiveness and the sex of the partner. In this experiment subject play against photographs of other subjects. The photographed subjects have previously played against others, an independent panel has rated their attractiveness, and the least and most attractive male and female photographs have been culled. In the experiment they find that the highest offers fo the the attractive males, while attractive people, in general, do better. These two experiments are interesting in that differences in offers and rates of rejection are solely a function of observed characteristics of subjects.

Not all of the empirical evidence is supportive. Fahr and Irlenbusch (2000) design an experiment to test Smith (1998) and the "mind reading" concept. Using a one-shot sequential “trust-reciprocity” game, Fahr and Irlenbusch focus on the capacity of a first mover to infer what the second mover will do in the game. In effect the first mover has to decide whether to pass money to a second player. The amount that is passed is tripled and the second mover then decides whether to return anything to the first mover. Fahr and Irlenbusch (2000) provide several distinct hypotheses based on experimental conditions that vary the property rights granted to the two subjects. Their key condition grants the first mover very strong property rights and predicts very high investments. When this does not happen, it is taken as evidence rejecting "mind reading." Their argument is that the first mover is unwilling to depend on the second mover understanding the intention behind investing. The design is a bit complicated and conflates property rights structures and partner characteristics. The design requires too many inferential steps in order to demonstrate either the presence or absence of mind reading. By contrast, Burnham, McCabe, and Smith (2000) show that a slight alteration in the label attached to a subject ("partner" versus "opponent") has a marked effect on behavior in these kinds of trust games. Their claim is that the differences in labels cause people to draw different inferences.

Our sense is that modeling intention is a fruitful endeavor. It appears that humans evolved in an environment where it paid to be attentive to others. Although social cues are culturally derived, the capacity to pick up and process those cues is part of an evolved cognitive structure. The fact that we pay attention to others, that we try to infer intention behind another's action, and that we conditionally respond, provides new insights into strategic behavior.

Inferring Intentions

The musings by game theorists and experimentalists are firmly grounded in work by psychologists. Psychologists suggest that individuals generate a raft of signals -- ranging from the timbre of one's voice to the stance of one's body -- which betray one's intention. Research on the communication of intention has focused on people with autism and on primates (see Baron-Cohen (1995) and O'Connell (1998)). Psychologists are interested in the capacity of individuals to put themselves in the place of others, referred to as having a "theory of mind" (TOM) about another person. This capability requires that a person be able to separate what he knows or understands from what another might.

There are three components to TOM (Baron-Cohen (1995), Chapter 4). The first is characterized as an Intentionality Detector (ID). This amounts to an ability to impute purpose and cause to another's actions. In large part this entails recognizing that others have goals and then deriving hypotheses about how actions are related to attaining those goals. Children as young as seven months old can distinguish between events that have a clear cause, such as a hand picking up a doll, and events with no apparent cause, such as a doll being lifted by an invisible wire (see the discussion by O'Connell, 1998, pp. 41-42). The second component involves an Eye-Direction Detector (EDD) that enables an individual to recognize another person's focus of attention and to draw inferences about the intentions of that person. As O'Connell puts it, "The evolutionary reason why you should take very good care to detect eye gaze is because when another animal is looking at you it can mean one of the three 'F's. Either that animal wants to fight you, feed on you, or mate with you." (O'Connell (1998), p. 47). The final component involves a Shared Attention Detector (SAD). Again, the eyes are important for what is communicated. In its most mentally complex form, SAD can be characterized as a relationship between two individuals and an object. The ability to follow the gaze of another and infer its meaning is important. Consider the luck of many a Hollywood hero who manages to avoid being stabbed in the back by reading the eyes of the person facing him. By noticing a shift in eye direction, the hero quickly reads that a third party is sneaking up behind him and usually whirls around in time to be saved from imminent death. The capacity to read gaze occurs in many more mundane settings, while equally important is the capacity to use gaze to send a signal. SAD is a complex mental concept that can only occur with joint mental attention. In an interesting sense, it may guarantee the “common knowledge” assumption crucial for game theory.

Baron-Cohen argues that a general “Theory-of-Mind Mechanism” ties together these volitional, perceptual and epistemic mental-state concepts. Such a mechanism enables people to draw inferences about the intentions of others by going outside themselves. A variety of laboratory experiments finds that subjects with autism and Asperger's syndrome have a difficult time reading and interpreting emotional expressions from still photographs. People with autism differ significantly from "normal" functioning individuals who have little difficulty in reading the emotional state of another. (Loveland et al. (1994) and Baron-Cohen, Wheelwright, and Jolliffe (1997)). The lesson drawn from this research is that the structure of the brain is crucial for enabling individuals to read signals that give insight into the intention of others. While learning to read those signals may be due to socialization, the capacity to do so is an important part of the way in which the brain is designed. As such, reading intentions should not be dismissed and should be incorporated into behavioral modeling of social interaction.

Faces.

We think that the ability to read the intentions of others is important for humans. We also think that the face is an important source for signals about intention. The question is whether facial expressions serve as a credible source of intention. There is an extensive body of research on human faces and what they mean to observers. Much of this literature derives from Darwin (1872/1998), who argued that humans, like animals, have evolved patterns of signaling behavior, including (but not limited to) facial expressions.

Contemporary researchers largely follow the lead of Ekman (1972; (1982) who contends that there is a universal set of evolved human facial expressions. Many are thought to be involuntary and reflect basic emotions. The bulk of the research has turned toward understanding what facial expressions reveal about someone's underlying emotional state. The claim is that facial expressions are emotional leakage, and that the emotional content is obvious to others because they share the same universal repertoire of facial expressions. Learned social behavior works to mask these emotions and cultural differences lead to different forms of masking. Therefore facial expressions can sometime be hard to read (for a general critique of the "universal" recognition of emotion, see Fridlund (1994), Chapter 10).[3]

Researchers for the most part have not taken up the question of the role of facial expressions in *social* signaling. Because the focus has been on the meaning and interpretation of facial expressions, there has been almost no investigation of behavioral responses to expressions (for an exception, see Yik and Russell (1999)). In a challenge to what he calls the "emotions view" of faces, Fridlund (1994) proposes a "behavioral ecology" view of faces, arguing that facial expressions *and their interpretation by others* is crucial. "The balance of signaling and vigilance, countersignaling and countervigilance, produces a signaling 'ecology' that is analogous to the balance of resources and consumers, and predator and prey, that characterize all natural ecosystems." (Fridlund (1994), p. 128). Facial expressions and their interpretation involve a delicate game in which expressions are signals about intention.

It is surprising that little research has focused on expressions as social signals, because there is an extensive literature on what facial expressions mean for children. From the outset, faces are important for child development. Johnson et al. (1991) trace the reaction of new born infants to a variety of paper stimuli the size and shape of a human head. Face-like images range from a human-like face to one with the same parts, but scrambled, to a blank piece of paper. Measuring eye and head movement, they find that newborns pay much closer attention to images resembling a human face than to others images. These findings are all the more impressive in that the infants tested were less than one hour old. These researchers conclude that children are born with a system that orients them toward face-like patterns; only as they mature do they develop a cortical system that allows for sophisticated face-processing activities.

[3] The work on facial images is not limited to psychology. There is a literature in political science that deals with facial expressions and their implications for electoral behavior and leadership (Mullen (1986); Sullivan and Masters (1988); McHugo and Smith (1996)). Much of this literature, like the work in psychology, is concerned with the ways in which facial expressions are interpreted.

The point from work on human faces is that they are an important source of information for social interaction. Humans are pre-attuned to pay attention to faces. From constant exposure they learn to read faces.

Abstract Images of Facial Expressions.

Considerable research shows that particular expressions are difficult to "read," and that the emotional content of an expression is often unclear (see the critique in Fridlund (1994)). Even something as simple as a "smile" can easily be misinterpreted or misrepresented (Ekman, Friesen, and Sullivan (1998), Leonard, Voeller, and Kuldau (1991), Fernandez-Dols and Ruiz-Belda (1997)). Human facial expressions are complex; the muscle groups on the face can easily send a wide spectrum of signals. While most researchers are looking for the six primary emotions -- happiness, sadness, anger, fear, surprise and disgust -- humans are capable of sending subtle blends of expressions. In addition differences in physical attractiveness, slight differences in expression, and unfamiliarity with the posed face all lead to variations in assessing emotions. To correct for these problems a handful of researchers have adopted highly stylized aspects of faces in order to detect the primary elements of facial expressions.

If there are specific components of expressions that signal specific emotional states, then these should be susceptible to systematic evaluation. Taking this insight, McKelvie (1973) designed an experiment in which he used schematic representations of faces. These schematics resemble variations on the ubiquitous "happy face" wishing everyone a nice day. McKelvie used an oval to represent a head and then drew in line segment representations of eyebrows, eyes, nose and mouth. These were systematically varied and then presented as stimuli to subjects.

A total of 128 schematic faces were used and each subject was presented with a sample of 16 faces. Working one at a time, subjects were asked to rate how easy it was to find an adjective to describe the face and then asked to score the appropriateness or inappropriateness of each of 46 adjectives for describing the face. The adjectives reflect four different emotional categories (happy, sad, angry and scheming) and one other category (vacant). His analysis shows that the shape of eyes and the structure of the nose had little effect on evaluations. Instead, eyebrow and mouth shapes have the greatest effect. He cautions that neutral (horizontal) eyebrow or mouth expressions signal little. "However, when brow and mouth move from the horizontal, clear differences in meaning emerge: medially down-turned brows indicate anger or schemingness; medially upturned brows are seen as sad; an upturned mouth denotes happiness; and a down-turned mouth is seen as angry or sad." McKelvie (1973), p. 345. In short, even simple schematic representations of faces can trigger emotional affect. These findings have been replicated in a number of different environments with very different populations (Yamada (1993), MacDonald, Kirkpatrick, and Sullivan (1996), and Katsikitis (1997)).

The lesson to draw from these studies is that humans are very good at recognizing emotional content even in highly stylized schematics. Pictures have meaning and they are readily interpreted. In a subsequent section of this paper we will rely on these results in order to concentrate on relatively clean expressions.

Experiments

Quite simply, we believe that humans use all the information about their partner when contemplating social exchange. In doing so, humans make a calculated choice, struggling to determine what information is credible. We posit that certain facial expressions produce credible information concerning intention. Moreover, we posit that such information systematically affects the behavior of individuals in strategic environments.

In order to be confident that people "read" something about others, we designed an experiment in which subjects were presented with a set of abstract facial icons. We used a battery of items to determine whether subjects drew inferences about intention from those icons. Subsequent experiments were run in which those abstract facial icons served as the stimulus for subjects in a series of ultimatum games. Several hypotheses are offered that tease out whether subjects infer intention by their partners in these games.

Experiment I.

The first experiment, carried out at Virginia Tech, was a survey designed to gauge subjects' impressions of schematic faces. The instrument probes whether subject's expectations about behavior are shaped by facial expressions.

The survey instrument was administered to 524 subjects (324 male, 192 female and 8 who failed to indicate their sex) in Principles of Economics classes at Virginia Polytechnic Institute and State University in January 1998. The classes consisted primarily of college sophomores; about one-third were business majors, one-third engineering majors and the remainder from assorted fields. Subjects were asked to complete a three-page survey during a regular class meeting time, either at the beginning or the end of class, and were not compensated for their participation. On the first page of the survey, each subject was assigned one of nine icons and asked to rate its characteristics. The icons are based on a 3x3 design involving three manipulations of the mouth and three manipulations of the eyebrows. The icons used in the survey are shown in Figure 2.

<Figure 2 About Here>

Subjects were randomly assigned to a particular icon and told that the icon "is supposed to represent a type of person." They then were asked to choose the most appropriate response for their icon on twenty-five word-pair items using a seven-point semantic differential scale. In the scale, a value of (1) means the word on the left is "very" close to matching the meaning of the icon, (2) is "somewhat" close, (3) is "slightly" close and (4) is "neither." The scale is symmetric to the right of (4). Left/right word order was randomly assigned for the word pairs.

The items were ordered in a consistent direction and factor analysis was used to uncover any underlying structure to the data. Eight items, forming two dimensions, are of interest.[4] The first dimension includes the pairings kind/cruel, pleasant/unpleasant, friendly/unfriendly and amiable/hostile. This dimension taps a general assessment of the "niceness" of the icon. The second dimension includes the pairings honest/dishonest, generous/selfish, trustworthy/untrustworthy and considerate/inconsiderate. This dimension taps several behavioral attributes related to social interactions.

Two scales were built based on the items constituting each dimension. For both the "niceness" and the "behavioral" scale four items were added and an average score was calculated for each respondent. Two models were then estimated for each scale; the first

[4] Under principal components all 8 items scale together quite well. However, using a varimax rotation, two dimensions are recovered.

model focuses on the main effects while the second includes all the interaction terms. Four main effects are included in the models. The variable SMILE is a dummy variable for icons with an upturned mouth. Likewise FROWN is a dummy variable for a down-turned mouth. UPBROW is a dummy variable for upturned eyebrows and DOWNBROW does the same for down-turned eyebrows. The neutral position is reflected in the intercept term of the regression. In model 2, the same main effects as well as all interaction terms are estimated. Finally a dummy variable for the SEX of the respondent was added.

Table 1 details the regressions for both scales, using the two models. Both models confirm findings by McKelvie (1973) and others. The positioning of both the mouth and eyebrows makes a difference. With respect to the behavioral intention scale the intercept term reflects the midpoint of the general semantic differential scale and is consistent with what we might expect from a neutral icon. The effect of eyebrows is pronounced. When the eyebrows are upturned, they decrease the evaluation (move it toward the “trustworthy” end of the scale) by almost a full point. Down-turned eyebrows have exactly the opposite effect. The down-turned mouth position has a modest effect on behavioral assessments and the upturned mouth has little effect. Under Model 2 a smile now has a modest independent effect on assessing behavioral intentions. It yields a more favorable evaluation of behavior. There is also a strong interaction between a smile and down-turned brows. The coefficient is positive indicating a less favorable evaluation of behavior. The effect of a smile can be deceiving; its effect is positive or negative depending on the position of the eyebrows. While the main effects alone lead one to believe that the frown/downbrow combination is the most negatively perceived, the interaction terms adjust the evaluations so that the conflicting message of the smile/downbrow icon is perceived with suspicion.

<Table 1 About Here>

Focusing on the "niceness" scale, both the smile and the eyebrow positions are strongly related to the evaluation of the icon. These results are consistent with the behavior scale in that a smile and upturned eyebrows result in a more positive assessment, while a frown and down-turned eyebrows lead to a more negative assessment. The interaction terms have the greatest effect with the combination of a frown and down-turned brows. The negative coefficient indicates some dampening of the strong main effects for the frown and down-brow. It is also the case that female respondents are more likely to evaluate the icons harshly with respect to affect. The effect is not large, but is statistically strong.

In short, the position of eyebrows and mouth both matter for the inferences that respondents draw about the icon. While the direction is the same for judgments about intentions and character, the latter has uniformly stronger main effects. These data strongly support the idea that respondents draw meaning from the icons. The next task is to see if these inferences carry over in the play of strategic games.

Experiment II.

Experiment 2 incorporates a series of two-person bargaining games in extensive form. At the outset, each subject was randomly assigned a facial icon (three of which are drawn from Experiment 1). Subjects then played a series of games in which they were randomly matched with other players each round.

Experimental Design. In this experiment pairs of subjects participated in a series of distinct two-person games. A total of 80 subjects were recruited from the local student population at Rice University. Students were contacted in their dining hall and asked to volunteer for a decision making experiment. Subjects signed up for one of 11 planned experimental sessions.

The laboratory accommodated eight subjects, each seated in a cubicle formed by moveable partitions, facing a computer. Although subjects were in the same room and could hear one another, they could not see one another's computer screen. At the outset of the experiment subjects were cautioned not to speak and told that if they did so, then the experiment would be canceled (none were). All experimental sessions were conducted over a local area network that handled all communication between subjects. In four of the 11 sessions only six subjects participated in the experiment.

Upon arriving subjects chose their seat at a computer. They were given self-paced instructions and shown how choices were made in the experiment. These instructions are attached as Appendix 1. In experimental sessions subjects participated in as many as 30 decision periods and were randomly matched prior to each decision. Subjects could not be uniquely identified, so even though same-pair play often occurred, it was impossible for subjects to know with whom they were paired at each decision.

In this experiment subjects faced as many as 18 distinct games played over 30 periods. These games included simple ultimatum games (analyzed here), a variation of a "trust game" and several simple games with obvious equilibrium. The first five periods had a fixed order for the games, with the third decision an ultimatum game. All remaining periods had games presented in random order.

Procedure. Regardless of the manipulation, the same procedure was used in each period of play. Before each game began subjects were randomly assigned and told their role (in the experiment, they were called “Decision Maker” 1 or 2). The subject assigned as player 1 moved first. In the ultimatum game this meant choosing either a left or right branch. Once a choice was made, player 2 was notified of the move and was given the choice to accept or reject the proposed division. Following Player 2's choice the corresponding payoff box was circled on the computer screen, both players were notified of the outcome and both were asked to record their payoffs for that period. The computer mediated all communication between players. Subjects were only told the moves of their partner and not the play of other subjects in the experiment. Once the period ended, subjects were instructed to wait until all pairs of players completed their decision. At that point the subjects were re-shuffled and re-paired.

At the conclusion of the experiment participants were paid in cash and in private. Subjects were told at the outset that they would be paid only for a single period of play. At the conclusion of the experiment they were asked to draw one card from a deck of 100 electronic cards displayed on the computer screen. Subjects were told that each period had an equal probability of being chosen, and the algorithm for the selection ensured this. When the card was turned over the subject learned which period was drawn and what was earned. Subjects were asked to verify that payoffs for the period drawn matched what they had recorded. Before being paid, subjects filled out an on-line questionnaire that asked them questions about their participation. The session lasted approximately 45 minutes with subjects earning, on average, $13.21. One subject earned the maximum of

$29.00 and four subjects earned $0.00 for their play. These latter were paid a show-up fee of $3.00, but not informed until their debriefing that they would be paid this amount.

Icon assignments and pairings. The first manipulation in this experiment relates to what subjects know about their partner. At the outset each subject was randomly assigned a permanent identity. Figure 3 presents the pairings of icons used in the experiment. In each session an equal number of subjects were assigned to one of the icons.

<Figure 3 About Here>

In each game within an experimental session a subject could be paired either with an individual with the same icon or with a different icon. At the beginning of each period subjects were shown the entire set of icons in the game (for an 8-person group this meant four images of each type of icon). When the subject was ready to begin, the icons were shuffled on the screen and the program randomly selected an icon. The subject's own icon and the counterpart's icon was then displayed.

We were very deliberate in not tying the icon to any personal characteristics of subjects. They were simply told the icon was theirs for the entire experimental session. Our rationale is that this constitutes a very weak stimulus, and that a stronger connection between the icon and the subject would strengthen any observed behavioral effects.

Our primary concern is with icons with human facial characteristics. As detailed below, we offer explicit predictions about out-of-equilibrium play in the ultimatum game. Three icons were used in which the angle of the eyebrows and orientation of the mouth are varied, as shown in Figure 3. A second set of icons were used that have no human facial content: a rectangle and an oval. These icons constitute one control condition for the experiment. It may be that subjects do not rely on human facial content, but instead they view the world as consisting of two types: “us” and "them." “In-group/out-group” effects are common in social psychological experiments (Turner (1978); Tajfel and Turner (1979)) so these icons are introduced as a control treatment.

In a final control condition subjects have no information about their counterparts, their icons are "blank" and they see no screen telling them about their counterpart’s identity. Subjects simply make a series of decisions in which they are randomly re-matched with another participant.

Games. The second manipulation changes the type of game. In this experiment an assortment of games were included, three of which were ultimatum games, several were simple bargaining games with a unique subgame perfect Nash equilibrium, and several were designed to test trust and reciprocity (the full set of these games are given in Appendix 2). The subset of games analyzed in this section are given in Figure 4. On the Figure the subgame perfect equilibrium for each game is circled

<Figure 4 About Here>

Games 1, 2, and 3 on figure 4 tap different aspects of the ultimatum game. Game 1 has a large asymmetry in divisions (a split of {19,3} versus {11,11}). Game 2 is identical except that the payoff for a rejection increases from {0,0} to {1,1}. Finally, Game 3 has a smaller asymmetry in divisions (a split of {15,5} versus {10,10}). In each instance the subgame perfect equilibrium is the asymmetric split.

We call games 4, 5 and 6 "simple" equilibrium games. All share the property that subjects have no incentive to use strategies that yield out-of-equilibrium behavior. Neither equity considerations nor fear of retaliation should enter into a subject's strategic calculation when playing these games. Game 6 is a two-branch game that is a bit more complex and requires more calculation on the part of subjects.

Predictions. Our predictions are derived both from game theoretic expectations and from behavioral expectations linked to the content of the icons. These predictions vary as a function of the type of game subjects play and the icons that they hold and against which they play.

The first prediction is taken directly from game theory. Despite the fact that subjects are assigned to different conditions -- some with and some without icons -- game theory treats these as the same. The icons are uninformative -- they amount to "cheap talk."

Prediction 1 (Game Theoretic): Subjects will choose actions leading to the subgame perfect equilibrium.

In the ultimatum game proposers should choose the left branch and responders should not reject the offer. In the simple equilibrium games, the subgame perfect equilibirum will be chosen.

The second prediction builds on ideas of equality and fairness. As Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) contend, subjects will choose actions that lead to more equal splits. Because equality (or inequality aversion) is a trait held by subjects, there should be no differences across manipulations.

> *Prediction 2 (Equality/Fairness)*: The greater the inequality in division the more likely subjects will choose equal divisions.

In the ultimatum game, proposers will choose the right branch, which involves an equal split, and there should be no rejections by responders. Proposers should be insensitive to the size of the division on the left branch, so the rate of choosing the right branch will be the same across all three ultimatum games. It is also the case that in game 6 the second mover will choose the exit option on the right branch, foregoing a larger gain in order to obtain an equal split.

The third prediction draws on our earlier discussion of intention. We expect subjects to rely on information about their counterparts and then conditionally choose actions. In this experiment we expect that the icons we assigned to subjects trigger specific responses. We expect three patterns of behavior. First, only icons with facial content will be informative. Second, subjects are more likely to choose an equal division when they perceive their counterpart to be "nice." Third, rejections are more likely from "nice" icons that are given unequal offers.

These three patterns are predicated on the idea that the icons, even though a weak stimulus, have informative content -- an idea supported by Experiment 1.

> *Prediction 3.1*. Conditions with no icon or an icon without a facial expression are uninformative. Subjects in these conditions will choose the equal divisions at lower rates than conditions with facial icons.

If subjects use the icon expressions to draw inferences then what the first mover sees in the partner will make a difference. The data from experiment 1 provides explicit orderings across the icons used here. First, the "happy" icon is regarded as the "nicest," followed by "devious" and then "angry." All three icons were statistically different from one another on the "niceness" scale, although the latter two items are closely related.[5] On the behavioral scale, "happy" is regarded as the most trustworthy, followed by "angry" and then "devious." As before, all pairwise comparisons are statistically different and there is greater separation between "happy" and the remaining two icons then between "angry" and devious.[6] This yields two distinct hypotheses:

> *Prediction 3.2a*. If proposers focus on the "niceness" of their counterpart, then the pattern of equal splits will be:
> $\text{Happy}_{\text{respondent}} > \text{Devious}_{\text{respondent}} > \text{Angry}_{\text{respondent}}$
>
> *Prediction 3.2b*. If proposers focus on behavioral inferences about their counterpart, then the pattern of equal splits will be:
> $\text{Happy}_{\text{respondent}} > \text{Angry}_{\text{respondent}} > \text{Devious}_{\text{respondent}}$

In both instances if the responder has a "happy" icon then the proposer should offer an equal split.

Similar hypotheses can be offered for patterns of rejection. The question now turns to how the respondents perceive proposers. If Rabin (1993) or Falk and Fischbacher (1998) are correct, then respondents have already observed the proposers action, they need to consider whether the action was in line with their beliefs and infer the proposers' intention. If a responder expects a proposer to be "nice" and those expectations are violated, then there ought to be higher rates of rejection. If a responder expects that a proposer will be mean, then rejections should occur at higher rates. The same should hold true if subjects only form their beliefs over expected behavior. This yields the following pair of predictions:

> *Prediction 3.3a*. If respondents use the icons to draw inferences about the "niceness" of proposers, then the pattern of rejections will be:
> $\text{Happy}_{\text{proposer}} > \text{Devious}_{\text{proposer}} > \text{Angry}_{\text{proposer}}$
>
> *Prediction 3.3b*. If respondents use the icons to draw inferences about the behavior of proposers, then the pattern of rejections will be:
> $\text{Happy}_{\text{proposer}} > \text{Angry}_{\text{proposer}} > \text{Devious}_{\text{proposer}}$

[5] The mean scale ratings for the icons were: Happy=2.38; Devious=5.20; Angry=6.04. The lower the value, the "nicer" the evaluation.
[6] The mean behavioral scale ratings were: Happy=2.75; Angry=4.83; Devious=5.27.

Analysis. Prediction 1 contends that subjects will choose the subgame perfect equilibrium. Figures 5a, b present the aggregate results for each game. The figures present the percentage of times each of payoff node was chosen and the numbers in parentheses are the frequencies. Figure 5a presents data for the three ultimatum games, while Figure 5b presents results for the simple bargaining games.

<Figure 5a, b About Here>

With respect to the ultimatum games there is no support for Prediction 1. Almost two-thirds (63.8 percent) of the ultimatum game outcomes result in equal splits. Moreover, 20.3 percent of the proposed unequal divisions were rejected, while proposed equal divisions were never rejected. Although these results are consistent with other ultimatum game experiments, they do not support the game theoretic prediction. Taken as a whole, only 28.8 percent of the outcomes are at the subgame perfect equilibrium in the ultimatum game.

The findings for the ultimatum game do not imply that game theory is meaningless. For the simple equilibrium games depicted on figure 5b the game theoretic prediction fares quite well. In games 4 and 5 over 90 percent of all outcomes are at the equilibrium. These results make it clear that most subjects understood the structure of the game in extensive form. From a game-theoretic standpoint, there are very few "errors" that constitute out-of-equilibrium play. Game 6 presents a greater challenge to the concept of subgame perfection. Here the first mover chose the right branch in 61 out of 62 decisions. However, the subgame perfect equilibrium was chosen only 53.2 percent of the time. Over 45 percent of the time the second mover chose the equal division payoff, even though that actor could have been made better off by passing the move back to the first proposer. It appears that considerations of equity influenced play in this game.[7]

Two findings stand out. First, the game theoretic prediction does not fare well in the ultimatum game, although this is consistent with a large number of ultimatum game experiments (see Camerer (1998)). Second, equilibrium predictions fare very well in simple bargaining games. Certainly the results from Games 4 and 5 lead us to think that subjects understand the structure of the extensive form game. The more complex Game 6

[7] Unreported analysis shows there are no main effects from any of the treatment variables in these simple games.

contains an attractive equal-split outcome that induces equitable play on the part of some subjects. We now turn to Prediction 2 to address questions of equity.

Prediction 2 holds that subjects will seek an equal division irrespective of the manipulations. Moreover, subjects will be sensitive to differences in possible earnings, with larger unequal splits leading to increased equitable behavior. We find that proposers are more likely to offer an equal split in games1 and 2 than in game 3 (where the difference in the unequal offer decreases from $16 to $10). The differences in equal splits are statistically significant (χ^2=4.67, df=1, p=.03) and support Prediction 2.

If subjects are concerned with equity, then equal splits should hold up with respect to the icon manipulations. The data were coded according to whether the manipulation involved an icon with a facial expression, an icon with no facial expression or the control group with no icon. Here we find strong differences across manipulations with equal splits occuring 76.4 percent of the time for subjects assigned icons with facial expressions, 30.3 percent for icons with no expression and 55.9 percent for the control group. Under a simple chisquare test, these groups are significantly different from one another (χ^2=24.59, df=2, p<.001).

Such a finding undermines Prediction 2. If subjects are concerned with equity or fairness, this should hold irrespective of the identity or label assigned to themselves or their counterpart. Conjectures by Fehr and Schmidt (1999) or Bolton and Ockenfels (2000) point to individuals either holding a sense of equity or being concerned with relative comparisons. In either case, introducing an identity should not have an impact on overall rates of choosing an equal split. Our finding leads us to discount Prediction 2 and we move to predictions that deal with inferring a partner's intention.

The results reported above lend support for Prediction 3.1. Subjects choose an equal split at higher rates when viewing an icon with facial content. To further detail this relationship, table 2 produces the equal split offers for different pairings of subjects. The left column illustrates the proposer's icon, while the top row indicates the responder's icon. Each cell in the table displays the percentage of times subjects offered an equal split. Several interesting patterns are apparent, although one should be cautious in over-interpreting these data given the small number of observations in some cells.

<Table 2 About Here>

The table allows us to distinguish between Prediction 3.2a and 3.2b. Focusing on the bottom row of the table and looking only at the three columns with facial icons, these data support the behavioral prediction. We find that $\text{Happy}_{\text{respondent}}$ (85.7%) > $\text{Angry}_{\text{respondent}}$ (74.1%) > $\text{Devious}_{\text{respondent}}$ (59.3%), as predicted under 3.2b. It is as if proposers take into account the respondent's icon and make a decision to settle on an equal split accordingly. Overall these rates of choosing an unequal split are significant across the icons (χ^2=7.17, df=2, p<.03).[8] This finding is very encouraging for hypotheses about reading the intention of others.

The pattern is less clear when turning to rejections. Table 3 details rejections given that the proposer chose an unequal split. Predictions 3.3a and 3.3b focus on the response by the second mover in the game. If intentions matter, then respondents should take into account their partner's icon when deciding whether to reject an unequal split. Generally, we find that the lowest rates of rejections are for non-facial icons, while rejections for the control group are 26.7 percent. The facial icons vary considerably, from no rejections for paired Happy icons to 50 percent rejections for paired Devious icons. However, it is difficult to interpret the rejection rates given these very small numbers. In part this is because of the low rates with which proposers with a facial icon chose an unequal split.

<Table 3 About Here>

Aggregating these data and looking at the last column of the table we find that neither prediction 3.3a nor 3.3b is supported. The pattern of rejections, based on the proposers icon is $\text{Devious}_{\text{proposer}}$ (40%) > $\text{Happy}_{\text{proposer}}$ (25%) = $\text{Angry}_{\text{proposer}}$ (25%). This behavior is incongruent with what is predicted under a model of intention. Based on the results from the first experiment, we predicted that subjects would expect a Devious icon to behave badly (offer an unequal split) and subjects would not punish such an individual. We expected that when responder's expectations were violated (e.g., when a proposer with a Happy icon chose an unequal split) then rejections would occur at higher rates.

The analysis presented so far aggregates the data across many subjects. To double check these findings we turn to an analysis of individual choices. First we

---

[8] Pairwise comparisons between Happy and Angry and Angry and Devious show that these differences are not statistical significant at the .05 level.

examine the choice of an equal split and then turn to rejections. Proposers were constrained to two choices and we estimate a Probit model for the likelihood of choosing the equal split. To control for unique subject characteristics (and unobserved between subject variation) we use a random effects model. The independent variables tap different aspects of the inequality aversion and inferring intentions models. The first variable, DIFF, takes the difference in payoffs between player 1 and 2 under the unequal split. In line with inequality aversion, the greater the difference, the more likely a proposer will choose an equal split. The type of icon with which a subject was paired is also included. Dummy variables are created for the extreme facial icons, happy and devious, or whether a partner had a non-facial icon. Finally, a dummy variable is included for subjects sharing the same icon in the decision (subjects in the control condition with no icons are coded as zero).

The first column of Table 4 presents the estimates. The coefficient for the difference in earnings is strongly correlated with choosing a fair outcome. Consistent with Bolton and Ockenfels (2000), the larger the difference in earnings for an unfair choice, the more likely an equitable split is proposed. However, at the same time the icon of the other subject has an independent effect on whether an equitable split was proposed. A devious icon has no effect while a happy icon has a positive effect. The ordering is consistent with what is predicted by a model of inferring intention. In addition, in conditions where subjects have no facial icon (either a rectangle or oval), unfair splits are more likely to be proposed. Finally, there is little support for the conjecture that an in-group identity has taken root. While the parameter estimate for the "same icon" is positive, it is weak and insignificant. In short, these findings support hypotheses that subjects are paying attention to fairness and are trying to draw inferences about their partner's likely response.

<Table 4 About Here>

The second column in Table 4 estimates a similar model for rejections. In this case the model estimates the likelihood that an unequal split was rejected. In these data there are relatively few rejections (a bit over 20 percent) and the total number of unequal splits is low so some caution should be exercised when interpreting these results. We find that responders pay no attention to the magnitude of the difference in payoffs under the unequal split. One interpretation is that responders who reject are less concerned with inequality aversion and instead only respond to actions directed at them. An unequal proposal is an unequal proposal no matter its magnitude. There is no effect across any of the first mover's icons except for the negative effect of a non-facial icon. This supports a view that the icons signal something about the counterpart and that the non-facial icons carry no meaning. Overall there is little that is systematically related to an individual choosing to reject an offer.

Discussion.

The results in this second experiment are both encouraging and puzzling. First, we find mixed support for models of inequality aversion. While subjects choose equal splits at high rates, those choices vary across icon manipulations. Support for models of reading intention also receive mixed support. Proposers act as if respondents will reject at different rates. However, when we turn to predictions about rejection rates, we do not find support for models of intention.

These mixed results may be due to the experimental design. First, subjects played many different types of games in this experiment and their experiences in other games may have affected their behavior in the ultimatum game. For example, prior to playing the ultimatum game, subjects experienced a version of a trust game that invited a form of cooperation. This might have led to inflated rates of equal splits. Second, subjects experienced the role of both the proposer and responder. It is possible that bouncing between roles led subjects to develop norms of reciprocity. Finally, there are only a small number of observations across manipulations. These small numbers may yield misleading trends in the data.

Experiment III.

Experiment III was designed to address several limitations to Experiment II. The number of decisions and the types of games were limited. As well, the number of subjects was increased. A total of 126 subjects participated in 18 different sessions. All procedures were identical to those in Experiment II.

Once again subjects were randomly assigned an icon identity that was held throughout the course of the experiment. At each decision period they were randomly paired with another subject and that subject's icon was revealed. The icons used in this experiment are the same as in Experiment II, except that now Devious and Angry were paired. Decision makers were randomly assigned to be either the proposer or responder and could be in either position over the course of the experiment.

In this experiment subjects made nine decisions. The first was based on one of the three ultimatum games from experiment II; the second and third decisions were one of the simple equilibrium games from experiment II; the fourth decision was a second ultimatum game; the fifth and sixth decisions were again simple equilibrium games; the seventh decision was the third ultimatum game; the eighth decision was a dictator game and the ninth decision was a simple investment game. The order of the non-ultimatum games was fixed and the order of presentation of the ultimatum games was blocked: in a third of the sessions Game 1 (from figure 4) was the first game subjects encountered, in another third Game 2 was first, and so on. This experiment was designed to control for ordering effects and to ensure that the other games did not have cooperative outcomes that affected behavior in the ultimatum game. The experiment took less than 25 minutes and, on average, subjects earned $13.55.

There were no ordering effects for the blocking and as a consequence the data are pooled(χ^2=4.32, df=2, p=.115). Overall, proposers were less likely to choose an equal split in this experiment. A little over 45 percent of the choices were for an equal split (compared with 63.8 percent in Experiment II). This rate is still well above what is predicted under game theory. Consistent with game theory, subjects continue to pick the subgame perfect equilibrium at high rates in games 4 and 5 (84.9 percent). Exactly two-thirds of the pairs go to the subgame perfect equilibrium in game 6. As such we are confident that the subjects understood the experimental design.[9]

We have exactly the same set of predictions in this experiment as for Experiment II. Prediction 2 holds that the frequency of equal splits will increase when the difference between proposer and responder payoffs from unequal splits increases. Contrary to

[9] When asked, at the conclusion of the experiment, 96.8% (122/126) of the subjects responded that the instructions were clear.

findings in Experiment II, we cannot reject the null hypothesis that the splits are the same across games (Game 1 = 41.3%; Game 2 = 52.4%; Game 3 = 42.9%). If Games 1 and 2 are combined, because the size of the unequal splits are the same, and compared with Game 3, there is still no difference across games (χ^2=.27, df=1, p=.61). Other evidence strengthens this point. The penultimate game played by all subjects was a dictator game involving a decision similar to Game 1 in which the proposer could chose between a division of {19,3} or {11,11}. In this game the second mover could not reject the offer. Although this game occurred late in the set of decisions, 71.4 percent of the proposers chose the unequal split. Overall the experiment provides little support for the inequality aversion model.

At the same time this experiment provides no support for Prediction 3.1. Aggregating the data and then testing for differences across different manipulations -- the control condition, icons without facial expressions and icons with facial expressions -- we find no difference (χ^2=0.81, df=2, p=.67). In fact, the aggregate data points out that the control group, with no icons, chooses an equal split at higher rates (51.1 percent) than the group with facial icons (43.0 percent).

There is considerable variation among specific pairings of icons as can be seen from table 5. Looking at the last row of the table, it appears that proposers pay some attention to their counterpart providing modest support for Prediction 3.2b. Here, the pattern of equal splits is $\text{Happy}_{\text{respondent}}$ (47.1%) > $\text{Angry}_{\text{respondent}}$ (43.8%) > $\text{Devious}_{\text{respondent}}$ (37.0%). While the directions are consistent with Experiment II, the differences between the icons are insignificant.

<Table 5 About Here >

Although the effect by icons for choosing an equal split is weak, the effect for rejections is much stronger. Table 6 details rejections based on icon pairings. While the overall number of rejections is small (20/102) the rates of rejections support prediction 3.3b. When respondents are paired with a happy icon and that proposer offers an unequal split, responders are much more likely to reject the proposal [$\text{Reject}_{\text{Happy}}$ (40.0%) > $\text{Reject}_{\text{Angry}}$ (13.3%) > $\text{Reject}_{\text{Devious}}$ (8.7%)]. This is entirely in line with the idea that expectations are violated. At the same time, rejections under the non-facial icons are also high, although the pattern is more consistent with an in-group hypothesis. There is only a single rejection when ovals and rectangles are paired.

<Table 6 About Here>

Discussion.

This experiment helps separate between inequality aversion and inferring intention models. There is little support for the former model and mixed support for the latter. These findings are further supported by multivariate estimates similar to those in Table 4, but not reported here. Basically, subjects in these experiments paid little attention to their partner's icon when deciding whether to make an equal split. However, when considering whether to reject an unequal split, the proposer's icon mattered a great deal.

These results are still not clearcut. There are a large number of manipulations and very small numbers of outcomes in each cell. There remain other games that subjects play and these might have an effect on subsequent strategic behavior.[10] Finally, some repeated game effects may intrude. Subjects continue to have their role switched between proposer and responder throughout the experiment. This may result in reciprocity emerging across the games. In order to remove these potential confounds we turn to Experiment IV.

Experiment IV.

A fourth experiment was designed to remove history effects and to remove reciprocity due to subjects switching their role as a proposer or responder. This fourth experiment replicates the design of both Experiment II and III. A different subject pool was used -- this time subjects were recruited from Virginia Tech. The laboratory set-up was similar to that used in the previous experiments. A total of 300 subjects participated in the experiment and 37 sessions were run. Subjects were recruited in groups of 6 or 12 and were told that they would never make a decision with another subject more than once (14 of the 37 sessions were with groups of 12).

In this experiment subjects made three decisions. All three decisions were ultimatum game decisions. Subjects were randomly assigned to be either proposers or responders. They retained that role throughout the experimental session. Proposers made their choice and then were asked to predict whether the responder would accept or reject the offer. Proposers were not informed of the responder's choice until the end of the session. Respondents were asked to predict the proposer's action prior to being informed about the proposers choice. Obviously the responders were informed of the proposer's choice prior to making their own choice. Subjects were paid for one decision at the conclusion of the experiment.

The other primary difference between this experiment and the others was that subjects were assigned one of three icons in their session. A 3x3 blocking design of game order and icons was used in the experiment. All subjects knew they were randomly assigned to an icon and they knew other subjects were also assigned to an icon. In groups of 6 at least two subjects had the same icon (one a proposer and another a responder -- although subjects did not know this division). In the groups of 12, four subjects shared the same icon, again evenly split between proposers and responders. Several new icons were added to this experiment, including an icon with a neutral facial expression and a diamond-shaped icon with no facial expression. In this experiment we also mixed icon types such that subjects with and without facial expressions were paired with one another (see figure 6).

<Figure 6 About Here>

The parameters to the games were slightly changed. Game 1 was identical to the games used in experiments two and three. The second game added $1.00 to the proposed divisions under Game 2 and this gave a subject the option of choosing an unequal split {16,6} or an equal split {11,11}. The third game was markedly changed. Here the proposer had a choice between an unequal split of {19,3} and a second unequal split of {8,14}. This game was chosen for three reasons. First, it removed the simple heuristic of selecting an equal split from the choice set. Second, it provides a strong test of inequality aversion in that the left unequal split is much larger than the right unequal split, except

[10] We checked whether the first decision was somehow different from the other decisions. It was not. All of our analysis was replicated for the first period data only and we find no differences from what is reported

that the proposer is put in an awkward position. Third, it forced responders to consider the plight of the proposer who owned the right to the first move.

Subjects in this experimental series chose equal splits at higher rates than in experiment III, but lower than in experiment II. Overall an equal split was selected 53.8 percent of the time, well above what is expected under a game theoretic hypothesis. This experiment provides mixed evidence for Prediction 2. Proposers are sensitive to differences in payoffs between themselves and responders, a finding mirroring that in Experiment II. The highest rate for an equal split is in Game 1 where the difference between the two players is the greatest on the lefthand side. As that difference decreases, so too does the rate of choosing the equal split. Game 3, which requires the proposer to give up a great deal produces a very low rate of equal splitting (Game 1 = 63.4%; Game 2 = 49.0%; Game 3 = 26.1%; χ^2=43.46, df=2, p<.001). If Games 1 and 2, which are directly comparable with previous games, are analyzed separately, the rates of equal splitting remain statistically significant (χ^2=6.43, df=1, p=.011).

Game 3 is particularly interesting because it puts the proposer in a difficult position with respect to relative inequality. The left branch produces a large unequal split favoring the proposer and the right branch produces a smaller unequal split favoring the responder. Equation 3, in an earlier section of this paper, points to a simultaneous concern not only with respect to differences in outcomes between the proposer and responder, but also the proposer's own utility for the outcome. A strong version of inequality aversion would argue for minimizing the difference in payoffs between the proposer and responder. A weaker version, and one consistent with these data, holds that the proposer's own payoffs are paramount. Proposers, it appears, are willing to forego some gains, but not if they end up with less than their counterparts.

The pattern of outcomes is also consistent with subjects trying to draw inferences about their counterparts. One inference is that no reasonable responder will punish a proposer for choosing the left branch in Game 3. After all, the right to move first was randomly granted and the proposer faces a difficult choice. At the same time, assigning different icons should have some effect on the inferences drawn by proposers. However, in the aggregate there is little support for this conjecture. There is no difference in

---

here.

choosing the right branch when comparing partners with a facial, non-facial or blank icon. This leads us to again reject Prediction 3.1.

Despite the fact that there is no relationship across general categories of facial, non-facial and blank icons, when delving deeper into these data a number of relationships that stand out. In the subsequent analysis we focus on Games 1 and 2 which are the most similar to those analyzed in the prior experiments. Table 7 provides the percentage of times subjects chose the equal split in these games by icon pairings. Again the proposer's icon is the leftmost column and the respondent's icon runs across the top row.

<Table 7 About Here>

The same general pattern from Experiment II holds. Proposers are more likely to give equal splits to Happy icons (59.5 percent) then give them to Devious icons (52.4 percent). The rate of equal splits for the Neutral icon lies in between. The differences between the icons, however, are small and not statistically significant. Surprisingly, however, when a facial icon is present with non-facial icons (Neutral mixed with Oval and Rectangle), the Neutral icon is very likely to receive an equal split (85.7 percent of the time). Although the numbers are quite small, these results come from 14 different proposers participating in 7 distinct sessions. As such the results are not due to sessional effects and certainly not due to subject experience. Equally interesting is that when a facial icon is present, the rate of equal splits rises for the Oval/Rectangle pairings (an overall rate of 76.7 percent).

Table 8 presents rates of rejection in Games 1 and 2 for the different icon pairings. These rejection rates are conditional on the proposer choosing an unequal (left) split. Overall the rate of rejection is quite low (9.7 percent compared with around 20 percent in Experiments II and III). There is little pattern to these data -- certainly no pattern consistent with violations of expectations noted in Predictions 3.3a and 3.3b. Proposers with devious and happy icons have substantially higher rates of rejections than do proposers with neutral, rectangular, oval or no icons. However, the highest rate of rejection occurs when the proposer has a diamond shaped icon. That icon was not expected to transmit any information or generate any expectations.

<Table 8 About Here>

In this experiment subjects were asked about their expectations prior to being informed about their counterpart's choice. Before a proposed division, responders were asked to predict what the proposer intended to do. Across all games subjects correctly predicted their counterpart's actions in 54.7 of the cases (only slightly better than chance). In 26.4 percent of the cases responders predicted the proposer would go left, although the proposer took the equal split. This would constitute a violation of expectations, but would not result in a rejection. The more interesting case is that in which the responder expected an equal split (or right branch choice) but the proposer chose left. This constitutes a miss-matched expectation and should trigger a higher rate of rejections. While there is a difference between met and (negatively) unmet expectations (11.5 percent as compared with 8.4 percent) that difference is slight. At the same time there is no relationship between these miss-matched expectations and the icon manipulations.

Discussion.

Results from Experiment IV are instructive. In this experiment the history of play was eliminated for proposers. Likewise responders had little incentive to reject proposals because they knew they would not be paired with the same player more than once. In this setting, with rather minimal levels of information, the rate of even splits was high, although nowhere near the level reported for Experiment II. At the same time the rate of rejection was very low. The icons were unconnected with increasing either rates of equal splits or rejections of unequal splits. At best it appears that subjects were reacting to a model of inequality aversion. Clearly when past history is eliminated the icons assigned to players lose much of their meaning. Why, then do these icons have so great an impact in Experiment II?

Experiment II Redux.

Results from experiment IV indicate that once subject feedback is curtailed, then the signaling value of a facial icon disappears. It appears that subjects do not respond directly to the icon, but rather they use the icon as a labeling device. With sufficient repetition the icons take on meaning -- meanings that are consistent with the directions predicted under models by Rabin (1993) and others. Starting with this idea we return to the results from experiment II.

Testing a conjecture that labels take on meaning requires that the history of play by subjects be carefully attended to. We use a variation of discrete event analysis (see Beck, Katz, and Tucker (1998)). Our observations are of two types. First, we are interested in whether the proposer made an equal allocation. Second, we are interested in whether the responder rejected the proposer's allocation. We treat each model separately. Our primary interest concerns the icons and their effect on choice. The facial icons, devious, angry and happy, were coded as dummy variables and we looked at both the icon held by the counterpart as well as the decision maker's own icon. The icons for the non-facial icons (oval and rectangle) were treated as dummy variables. Finally, the omitted category was the "blank" control condition in which subjects observed no signal. To control for impact of repeated play, we included a simple counter measuring the period of play. While Beck, Katz, and Tucker (1998) recommend creating a series of dummy variables for each time period and incorporating them into an estimate, Box-Steffensmeier and Jones (2001) argue that directly incorporating the time period efficiently corrects for any time dependent trends. Finally we build in two measures that tap different aspects of being a first or a second mover (and the fact that a subject's role changes across the session). The first calculates the percentage of time that a subject was treated non-cooperatively. In the ultimatum game this involves instances when the subject was given the unequal allocation. Subjects played several other games in which they could have been treated in a non-cooperative manner and these instances were recorded, accumulated and calculated for each period. The aim behind this variable is to account for the subject's experiences over the course of the experiment. The more a subject experiences cooperative behavior, the more likely that subject will propose an equal allocation. The second variable tests the flipside of this idea. It measures the percentage of time that a subject makes a non-cooperative choice. Again, this is calculated for each game and the actors position over the course of the experiment. The aim behind this variable is to account for the subject's actions over the course of the experiment.[11]

The first model estimates the likelihood that a subject makes a fair choice contingent on observing the identity of their counterpart and their own experiences across the course of the experiment. Estimates are derived from PROBIT. Several different models were estimated, including random effects models (to account for individual variation) and random effects models for each experimental session (to account for sessional variation). Using likelihood ratio tests between the initial model and the different random effects models, we cannot reject the null hypothesis that any of the alternative models improve the fit.[12] Consequently we report coefficients for a standard PROBIT estimate.

The estimates are presented in table 9. Not surprisingly, the ordering across the facial icons is preserved. Subjects are least likely to offer an equal split to counterparts with a devious icon and the most likely to do so with a happy icon. The standard errors are reasonably large, and only the happy icon approaches statistical significance. Over time a pattern of splitting equally takes hold. It should be no surprise that repeated interaction favors cooperation (choosing an equal split). This effect is quite strong, but is tempered by the experiences of the subject. The more often the subject has been treated unfairly, the less likely that subject is to offer an equal split. The negative (but insignificant) coefficient points out that subjects are sensitive to how they have been treated in the past. A much stronger effect relates to the subject's own past behavior. The more a subject treats others unfairly, the more likely the subject is to offer an equal split.

<Table 9 About Here>

When turning to the second model, in which the icons are those of the proposer, little stands out from the estimates. All three facial icons have positive effects, but there

[11] In looking at the final period only, the percentage of cooperative experiences averaged 31.7% (with a standard deviation of 34.9). The percentage of non-cooperative actions averaged 33.4% (with a standard deviation of 39.0). A variety of other variables were explored in different models. We examined interactions between the time period and these two variables. We included the sex of the subject. We examined the subject's most recent experience (whether treated cooperatively or non-cooperatively) and we looked at various measures of how a cooperative and a non-cooperative experience were discounted with the passage of time. The equation presented here provides the best fit of all the different estimations that were tried.

[12] There are no changes in the significance levels of the coefficients. The initial model is fit with robust estimators using the Huber/White sandwich estimator of variances.

is considerable noise as shown by the standard errors. At the margin, an angry icon is slightly more likely to offer an equal split than a happy icon. However, none of these parameters approach statistical significance. The period and deceitfulness variables continue to show strong effects. In general these estimates imply that proposers are paying more attention to their counterpart then seeking to enhance their own reputation. This idea is reinforced by the strong coefficient for an individual's own deceitfulness. Once launched on a non-cooperative path, subjects stick with that strategy.

To give a bit better sense of this, Figure 7 shows the estimated probability of a proposer choosing an equal split while viewing either a happy icon or a nonfacial icon. Along the bottom axes are values for the period and percentage of time the proposer was non-cooperative. The top, curved plane is for the receivers with a happy icon, while the lower plane is for the nonfacial icons. When a subject begins by being non-cooperative and is at the beginning of the experiment, such proposers are very unlikely to choose an equal split, regardless of their counterpart's icon (although there are clear differences between the two groups). Those differences increase as non-cooperation goes to zero (toward the front, bottom corner of the graph). With additional experience, the likelihood of choosing an equal split increases, regardless of the propensity of the proposer to behave non-cooperatively. Throughout the experiment the Happy icon is more likely to receive an equal split then are subjects with nonfacial icons.

<Figure 7 About Here>

The last column of Table 9 estimates a similar model, only for rejections. Most of our discussion of rejections has focused on what happens once an unequal allocation has been proposed. However, estimating rejections only on the basis of unequal proposals runs the risk of selection bias (Heckman (1979); Reed (2000)). There are many unobserved variables underlying the initial choice of an equal or unequal split (although we have tried to capture these in the models noted above). While we expect that the icon manipulations should account for any differences, we use a selection model for our estimates. Although we expect that the first mover's icon accounts for any differences in rejection rates by the second mover, we should not ignore the face that rejections only take place when the first mover chooses and unequal split. Therefore we need to account for characteristics of the second mover (their own icon) as it systematically affects the first mover's choice. Table 9 reports the coefficients for rejection conditional on estimating the effects for the initial proposed division.

There is little to show with the rejection model. Once history is accounted for, none of the icons make any difference. Only the dummy variable for non-facial icons is significant and it indicates that responders are less likely to reject an unequal offer made by someone with a non-facial icon. This is consistent with earlier findings that non-facial icons have little signaling value. What is surprising is that neither the time period nor subject's past experiences or dispositions matter. All of these coefficients are zero.

Discussion.

It should not be surprising, but the history of play matters a great deal in these games. In experiments where we limit information about the history of play and where we eliminate the possibility of reputation effects, we find very different patterns of behavior then where subjects can see a complete history and where labels evolve to take on a meaning. This latter point seems crucial. Subjects are assigned a series of cheap talk signals that provide non-credible information. Those labels are treated in exactly that way from the beginning of the experiment. However, fairly quickly those labels evolve meanings within the sessions and they take on a value when people make decisions. Contrary to what we conjectured, subjects do not appear to be making inferences about the play of their counterpart by reading information about their partner's icon. But, they are building beliefs about the play of others with similar icons over the course of the game. As a consequence, the icons provide a signal value that arises through endogenous play. In a sense, subjects are not using rote models of inequality aversion nor trying to tap into the intention of their counterpart. Instead, they are learning about their population and evolving meaning about commonly observed labels.

Conclusion

These experiments allow us to reach three conclusions. First, subjects engage in a good deal of out-of-equilibrium behavior. Second, subjects behave conditionally, taking something about their partner into account. Third, repeated play is crucial, even when identities are anonymous. Taken together these conclusions point to the importance for conceptualizing human decision processes differently from what is usually done in political science.

We first find that there is a strong propensity for subjects to engage in out-of-equilibrium behavior. In the ultimatum game this means subjects choose to forego significant gains for themselves in order to give a counterpart more. Subjects often choose a fair outcome. At the same time, subjects are also willing to reject offers they consider unfair, even when such action is costly. As surprising as fair divisions may be in these games, the fact that subjects punish others is even more surprising. Yet the behavior is common. We find that subjects have a propensity for seeking fair outcomes, although this is not true across all conditions or in all experiments. We are surprised by the fact that fair outcomes occur with such high frequency in experiment IV under the no-feedback, no history condition. We also find a propensity for punishing others, but the pattern of rejections happens more often in settings in which there is a signal value to punishment. In experiment IV there is little rationale for rejecting an offer because the first mover never sees that rejection until the end of the experiment, and the rate of rejection is much lower in this experiment. Rejections make a good deal of sense when they can have a signal value -- a form of reputation is certainly possible in experiments II and III and here we observe higher rates of rejection. Basically, the out-of-equilibrium behavior that we observe is too consistent to be nothing more than noise. This tells us that we should heed scholars like Ostrom (1998) and Camerer (1997) when they call for the integration of behavior and game theory.

Our second finding makes it clear that subjects take their partner into account when making strategic choices. However, they are not considering their partner in the way in which standard models of game theory might predict. We do not find that actors simply consider their partner's payoffs and choose the appropriate subgame perfect equilibrium. Instead, they pay attention to characteristics of their counterpart. Our results do not allow us to determine whether subjects are driven by inequality aversion or intention detection. Subjects do pay attention to issues of fairness, although not consistently. By the same token, subjects drawn inferences about their counterpart based on very stylized representations. However, the response to the facial icons is not consistent across all experiments. We have sufficient evidence to conclude that subjects respond to some information about their counterpart. We find patterns of behavior in indicating that subjects do not lock themselves into a single strategy. Instead, they "conditionally" cooperate. Their conditional choice is grounded in signals sent by their counterpart. What is interesting for us is that the signals are no more than cheap-talk in that they are not costly and, as a consequence, not credible. Even so, they are interpreted as carrying some signal value. We find those signals most valuable when they evolve a meaning through repetition. This should be no surprise and is consistent with models of reputation formation (see for example, Fudenberg and Levine (1992), Chong (1992), Celentani (1996) and Cripps, Schmidt, and Thomas (1996)). Even though the different icons take on the meaning, the signals are not costly.

Our third finding is that repeated play in a population is important. This should be no surprise to anyone who has followed the extensive literature on repeated games. However, our findings are not in the vein of Evolutionary Stable Strategies (ESS) -- see for example, Axelrod and Hamilton (1981). That is we do not find that strategies evolve over the course of play (or that strategies held by actors are displaced over time). Instead the information content of a cheap talk signal evolves over time and this leads people to conditionally adjust their strategies. The labels that are randomly assigned to subjects evolve meaning. Such a finding lends credence to modeling efforts by Crawford (1995), Van Huyck, Battalio, and Rankin (1997) and Bednar and Page (2001) to characterize population models. One interpretation is that the labels become cultural artifacts that have meaning within the context of the group. A useful direction for future modeling efforts may be to focus on the evolution and transmission of cultural meaning (see Boyd and Richerson (1985)).

The broader point ought to be that political scientists should pay attention to the work going on outside the discipline. It appears that there are interesting constraints on human cognition (for one approach by political scientists, see the recent paper by Lubell and Scholz (2001)). This manuscript began with the notion that actors in their initial encounters use information about one another to begin strategic play. The central problem is that the available information is likely to be non-credible. However, extensive work in neuropsychology has demonstrated the importance of facial expressions for communicating information while the literature on emotion points to facial expressions as exhibiting credible signals of intention. Rather than a constraint on cognition, facial expressions may serve as simple, credible heuristic that allows for a quick adjustment in the play of a strategy. This is speculative, as only a tiny amount of research has focused on facial expressions in strategic settings (see, Solnick and Schweitzer (1999); Mulford et al. (1998); Eckel and Wilson ((forthcoming)).) We expect that the face is an important source for signaling intention and expectations.

What can be said about fairness? Fairness is a core concept for political scientists. These results indicate that people pay attention to issues of fairness, even when there is little reason to do so. Our experiments indicate that subjects pay attention to relative inequality. But, by the same token, subjects are discriminating in the cues that they use when making decisions. Fairness is a social concept that is invoked conditionally rather than a trait that is embedded in an individual. The question of whether to be fair goes back to whether one thinks that others are trying to be fair. Laboratory experiments, in which attributes of fairness can be systematically manipulated, are quite useful. Rather than probing attitudes about fairness, individuals should be pressed, under a variety of conditions, to see how they behave when being fair is costly. The value of the ultimatum game (and its close relative, the dictator game) is that it taps an individual's sense of fairness. The parameters of the game can easily be manipulated in order to gauge when subjects will forego earnings to give to another. By systematically manipulating the characteristics of subjects, the social conditions under which subjects will behave fairly can also be probed. There is a huge literature on the ultimatum game and it provides a rich environment for thinking about strategic behavior. We think that political scientists should be interested in the efforts that are underway in economics and in psychology to understand when fairness will be invoked.

Bibliography

Anderson, Simon, Jacob Goeree and Charles Holt. 1997. "Stochastic Game Theory: Adustment to Equilibrium Under Bounded Rationality". *University of Virginia Thomas Jefferson Center Discussion Paper* 304:304.

Andreoni, James. 1990. Impure Altruism and Donations to Public Goods: A Theory of Warm Glow Giving. *Economic* Journal 100:464-477.

Andreoni, James, and John H. Miller. 1998. "Giving According to GARP: An Experimental Study of Rationality and Altruism.". *Unpublished manuscript, University of Wisconsin*.

Axelrod, Robert M., and William D. Hamilton. 1981. The Evolution of Cooperation. *Science* 211:1390-1396.

Baron, David P., and John A. Ferejohn. 1989. Bargaining in Legislatures. *American Political Science Review* 83 (4):1181-1206.

Baron-Cohen, Simon. 1995. *Mindblindness: An Essay on Autism and Theory of Mind.* Boston: MIT Press.

Baron-Cohen, Simon, Sally Wheelwright, and Therese Jolliffe. 1997. Is There a 'Language of the Eyes'? Evidence from Normal Adults, and Adults with Autism or Asperger Syndrome. *Visual Cognition* 4 (3):311-331.

Beck, Nathaniel, Johnathan N. Katz, and Richard Tucker. 1998. Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable. *American Journal of Political Science* 42 (4):1260-1288.

Bednar, Jenna, and Scott Page. 2001. Explaining Behavioral Diversity through Strategic Choice in Game Ensembles: A Formal Analysis of Culture. Paper read at Midwest Political Science Association, April 19-22, at Chicago, IL.

Blount, Sally. 1995. When Social Outcomes Aren't Fair: The Effects of Causal Attributions on Preferences. *Organizational Behavior and Human Decision Processes* 63 (2):131-144.

Bolton, Gary E. 1998. Bargaining and Dilemma Games: From Laboratory Data Towards Theoretical Synthesis. *Experimental Economics* 1 (3):257-281.

Bolton, Gary E. , Jordi Brandts, and Axel Ockenfels. 1998. Measuring Motivations for the Reciprocal Responses Observed in a Simple Dilemma Game. *Experimental Economics* 1 (3):207-219.

Bolton, Gary E., and Axel Ockenfels. 2000. ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review* 90 (1 (March)):166-193.

Bottom, William, Cheryl L. Eavey, Gary Miller, and Jennifer Nicoll Victor. 2000. The institutional effect on majority rule instability: Bicameralism in spatial policy decisions. *American Journal of Political Science* 44 (3):523-540.

Box-Steffensmeier, Janet, and Bradford Jones. 2001. Timing and political change: Event history modeling in political science.

Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process.* Chicago:: University of Chicago Press.

Brams, Steven J. , and Alan D. Taylor. 1996. *Fair Division: From Cake-Cutting to Dispute Resolution.* Cambridge: Cambridge University Press.

Burnham, Terence, Kevin McCabe, and Vernon L. Smith. 2000. Friend-or-foe intentionality priming in an extensive form trust game. *Journal of Economic Behavior & Organization* 43 (1):57-73.

Byrne, Richard W., and Andrew Whiten, eds. 1988. *Machiavellian intelligence : social expertise and the evolution of intellect in monkeys, apes, and humans*. New York: Oxford University Press.

Camerer, Colin. 1998. Simple Bargaining and Social Utility: Dictator, Ultimatum and Trust Games. In *Experiments in Strategic Interaction*: Unpublished Manuscript.

Camerer, Colin F. 1997. Progress in Behavioral Game Theory. *Journal of Economic Perspectives*, 11 (4):167-88.

Camerer, Colin F., and Teck Hua Ho. 1997. Experience-Weighted Attraction Learning in Normal Form Games. In *California Institute of Technology, Division of Humanities and Social Sciences, Working Paper 1003*.

Camerer, Colin F., and Richard H. Thaler. 1995. Ultimatums, Dictators and Manners. *Journal of Economic Perspectives* 9 (2):209-219.

Cameron, Lisa A. 1999. Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia. *Economic Inquiry* 37 (1):47-59.

Celentani, Marco. 1996. Reputation with Observed Actions. *Economic Theory* 7 (3):407-419.

Chong, Dennis C. 1992. Reputation and Cooperative Behavior. *Social Science Information* 31 (4):683-709.

Cooper, David J., Susan Garvin, and John H. Kagel. 1997. Adaptive Learning vs. Equilibrium Refinements in an Entry Pricing Game. *Economic Journal* 107 (442):553-75.

Crawford, Vincent P. 1995. Adaptive Dynamics in Coordination Games. *Econometrica* 63 (1):January.

Cripps, Martin W., Klaus M. Schmidt, and Jonathan P. Thomas. 1996. Reputation in Perturbed Repeated Games. *Journal of Economic Theory* 69 (2):387-410.

Darwin, Charles, ed. 1872/1998. *The Expression of the Emotions in Man and Animals.* Edited by P. E. (Editor), *Series in Affective Science*. Oxford: Oxford University Press.

Eckel, Catherine C. and Philip Grossman. 1998. Chivalry and Solidarity: Gender Differences in Ultimatum Game Experiments.". *Working Paper, Department of Economics, Virginia Polytechnic Institute and State University*.

Eckel, Catherine C., and Philip Grossman. 1996. Altruism in Anonymous Dictator Games. *Games and Economic Behavior* 16:181-191.

Eckel, Catherine C., and Rick K. Wilson. (forthcoming). The Human Face of Game Theory: Trust and Reciprocity in Sequential Games. In *Trust and Reciprocity: Interdisciplinary Conceptual and Empirical Lessons.*, edited by E. Ostrom and J. Walker. New York: Sage Foundation.

Ekman, P. , W. V. Friesen, and M. O. Sullivan. 1998. Smiles when lying. *Journal of Personality and Social Psychology* 54:414-420.

Ekman, Paul. 1972. *Emotion in the Human Face: Guide-Lines for Research and an Integration of Findings.* Vol. 11, *Emotion in the Human Face Guide-Lines for Research and an Integration of Findings.* New York: Pergamon Press.

Ekman, Paul. 1982. *Emotion in the human face, 2nd ed.* Cambridge: Cambridge University Press.

Fahr, René, and Bernd Irlenbusch. 2000. Fairness as a constrint on trust in reciprocity: earned property rights in a reciprocal exchange experiment. *Economic Letters* 66:275-282.

Falk, Armin, and Urs Fischbacher. 1998. Kindness is the Parent of Kindness: Modeling Reciprocity.

Fehr, Ernst, and Klaus M. Schmidt. 1999. A Theory of Fairness, Competition and Cooperation. *Quarterly Journal of Economics* 114 (3):817-868.

Fernandez-Dols, J. M., and M. A. Ruiz-Belda. 1997. Spontaneous Facial Behavior During Intense Emotional Episodes: Artistic Truth and Optical Truth. In *The Psychology of Facial Expression.*, edited by J. A. Russell and J. M. Fernandez-Dols. Cambridge: Cambridge University Press.

Forsythe, R., Horowitz, J. L., Savin, N. E. and Sefton M. 1994. "Fairness in simple bargaining experiments.". *Games and Economic Behavior* 6:347-369.

Frank, Robert. 1988. *Passions Within Reason: The Strategic Role of the Emotions*. New York: W. W. Norton & Co.

Fridlund, Alan J. 1994. *Human Facial Expression: An Evolutionary View*. San Diego: Academic Press.

Fudenberg, Drew, and David K. Levine. 1992. Maintaining a Reputation When Strategies Are Imperfectly Observed. *Review of Economic Studies* 59 (3):561-579.

Guth, Werner, and Reinhard. Tietz. 1990. Ultimatum Bargaining Behavior: A Survey of Comparison of Experimental Results. *Journal of Economic Psychology* 11 (3):417-449.

Heckman, James D. 1979. Sample Selection Bias as a Specification Error. *Econometrica* 47:153-162.

Hoffman, E., K. McCabe, K. Shachat, and V. L. Smith. 1994. Preference, property rights and anonymity in bargaining games. *Games and Economic Behavior* 7:346-380.

Johnson, Mark H., Suzanne Dziurawiec, Haydn Ellis, and John Morton. 1991. Newborns preferential tracking of face-like stimuli and its subsequent decline. *Cognition* 40:1-19.

Katsikitis, Mary. 1997. The classification of facial expressions of emotion: a multidimensional scaling approach. *Perception.* 26:613-626.

Leonard, C. M. K. , K. S. Voeller, and J. M. Kuldau. 1991. When's a smile a smile? Or how to detect a message by digitizing the signal. *Psychological Science.* 2:166-172.

Levine, David. 1998. Modeling Altruism and Spitefulness in Experiments. *Review* of Economic Dynamics. 1:593-622.

Loveland, Katherine A., Belgin Tunali-Kotoski, Deborah A Pearson, and Kristin A Brelsford. 1994. Imitation and expression of facial affect in autism. *Development & Psychopathology* 6 (3):433-444.

Lubell, Mark, and John T. Scholz. 2001. Cooperation, Reciprocity, and the Collective-Action Heuristic. *American Journal of Political Science* 45 (1):160-178.

MacDonald, P. M. , S. W. Kirkpatrick, and L. A. Sullivan. 1996. Schematic Drawings of Facial Expressions for Emotion Recognition and Interpretation by Preschool-

Aged Children. *Genetic, Social and General Psychology Monographs.* 122:375-388.

McHugo, Gregory J., and Craig A. Smith. 1996. The power of faces: A review of John T. Lanzetta's research on facial expression and emotion. *Motivation & Emotion* 20 (2):85-120.

McKelvey, Richard D, Thomas R. Palfrey. 1995. "Quantal Response Equilibria for Normal Form Games.". *Games and Economic Behavior* 10 (1):6-38.

McKelvey, Richard D, Thomas R. Palfrey. 1996. "A Statistical Theory of Equilibrium in Games.". *Japanese Economic Review* 47 (2):186-209.

McKelvie, Stuart J. 1973. The meaningfulness and meaning of schematic faces. *Perception and Psychophysics* 14 (2):343-348.

Morton, Rebecca, and Daniel Diermeier. 1997. Proportionality versus Perfectness: Experiments in Majoritarian Bargaining. Paper read at Paper presented at the annual meeting of the Economic Science Association., Sept 18-21, 1997, at Tucson, AZ.

Mulford, Matthew, John Orbell, Catherine Shatto, and Jean Stockard. 1998. Physical Attractiveness, Opportunity, and Success in Everyday Exchange. *American Journal of Sociology* 103 (6):1565-1592.

Mullen, Brian; et al. 1986. Newcasters' facial expressions and voting behavior of viewers: Can a smile elect a president? *Journal of Personality & Social Psychology* 51 (2):291-295.

O'Connell, Sanjida. 1998. *Mindreading: An Investigation into How We Learn to Love and Lie.* New York: Doubleday.

Ostrom, Elinor. 1998. A Behavioral Approach to the Rational Choice Theory of Collective Action. *American Political Science Review* 92 (1):1-22.

Ostrom, Elinor. 2000. Collective Action and the Evolution of Social Norms. *Journal of Economic Perspectives* 14 (3):137-158.

Rabin, Matthew. 1993. Incorporating Fairness into Game Theory and Economics. *American Economic Review* 83 (5):1281-1302.

Rabin, Matthew. 1998. Psychology and Economics. *Journal* of Economic Literature (36):11-46.

Reed, William. 2000. A Unified Statistical Model of Conflict Onset and Escalation. *American Journal of Political Science* 44 (1):84-93.

Roth, Alvin E., and Ido Erev. 1995. "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term.". *Games and Economic Behavior* 8 (1):164-212.

Rubinstein, Ariel. 1982. Perfect Equilibrium in a Bargaining Model. *Econometrica* 50:97-109.

Sally, David. 2001. On sympathy and games. *Journal of Economic Behavior & Organization* 44 (1):1-30.

Smith, Vernon L. 1998. The Two Faces of Adam Smith. *Southern Economic Journal* 65 (1):1-19.

Solnick, Sara J., and Maurice E. Schweitzer. 1999. The Influence of Physical Attractiveness and Gender on Ultimatum Game Decisions. *Organizational Behavior and Human Decision Processes* 79 (3):199-215.

Sullivan, Denis G., and Roger Masters, D. 1988. 'Happy Warriors': Leaders' Facial Displays, Viewers' Emotions, and Political Support. *American Journal of Political Science* 32:345-368.

Tajfel, Henri, and John Turner. 1979. An integrative theory of intergroup conflict. In *The Social Psychology of Intergroup Relations.*, edited by W. G. Austin and S. Worchel. Monterey, CA: Brooks/Cole.

Thaler, Richard H. 1988. The Ultimatum Game. *Journal of Economic Perspectives* 2 (4):195-206.

Turner, John. 1978. Social Categorization and Social Discrimination in the Minimal Group Paradigm. In *Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations.*, edited by H. Tajfel. London: Academic Press Inc.

Van Huyck, John B., Raymond C. Battalio, and Frederick W. Rankin. 1997. On the Origin of Convention: Evidence from Coordination Games. *Economic Journal* 107:576-596.

Whiten, Andrew , and Richard W. Byrne, eds. 1997. *Machiavellian intelligence II : extensions and evaluations*. Cambridge: Cambridge University Press.

Yamada, Hiroshi. 1993. Visual Information for Categorizing Facial Expression of Emotion. *Applied Cognitive Psychology* 7:257-270.

Yik, Michelle S. M., and James A. Russell. 1999. Interpretation of Faces: A Cross-cultural Study of a Prediction from Fridlund's Theory. *Cognition and Emotion* 13 (1):93-104.

Table 1
Effect of Facial Characteristics on Perceptions of Icons
Standard Errors in Parentheses

	Model 1		**Model 2**	
	Behavior	**Niceness**	**Behavior**	**Niceness**
Intercept	4.89*** (.11)	4.21*** (.12)	3.90*** (.14)	4.08*** (.14)
 **SMILE**	-.13 (.11)	-.82*** (.12)	-.38* (.20)	-.91*** (.22)
 **UPBROW**	-.71*** (.11)	-.97*** (.13)	-.56*** (.19)	-.75*** (.22)
 **FROWN**	.22** (.11)	.67*** (.10)	.45** (.19)	1.12*** (.22)
 **DOWNBROW**	1.03*** (.11)	1.41*** (.12)	.94*** (.17)	1.54*** (.19)
SEX (1=Female)	-.002 (.09)	.21** (.10)	-.01 (.09)	.21** (.10)
 **Interaction**			-.16 (.27)	-.13 (.30)
 **Interaction**			-.47* (.25)	-.79*** (.28)
 **Interaction**			.84*** (.24)	.39** (.28)
 **Interaction**			-.29 (.28)	-.54* (.30)
r2	.37	.54	.41	.56
	***p<.01	**p<.05	*p < .10	

Table 2
Percentage of Proposers Choosing the Equal Split in the Ultimatum Game
(Experiment II; numbers in italics)

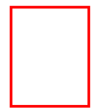
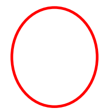
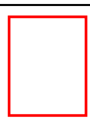
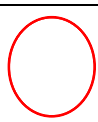
Proposer's Icon	**Respondent's Icon**						
						**Blank**	**Total**
	57.1 *8/14*	--	75.0 *12/16*	--	--	--	66.7 *20/30*
	--	72.7 *8/11*	92.3 *12/13*	--	--	--	83.3 *20/24*
	61.5 *8/13*	75.0 *12/16*	88.9 *24/27*	--	--	--	78.6 *44/56*
	--	--	--	100.0 *4/4*	14.3 *1/7*	--	45.5 *5/11*
	--	--	--	23.1 *3/13*	22.2 *2/19*	--	22.7 *5/22*
Blank	--	--	--	--	--	55.9 *19/34*	55.9 *19/34*
Total	59.3 *16/27*	74.1 *20/27*	85.7 *48/56*	41.2 *7/17*	18.8 *3/16*	55.9 *19/34*	63.8 *113/177*

Table 3
Rejections by Respondents for Unequal Proposed Splits in the Ultimatum Game (Experiment II; numbers in italics)

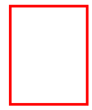
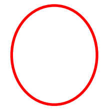
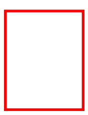
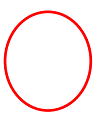
Proposer's Icon	**Respondent's Icon**						
						**Blank**	**Total**
	50.0 *3/6*	-	25.0 *1/4*	-	-	-	40.0 *4/10*
	-	33.3 *1/3*	0.0 *0/1*	-	-	-	25.0 *1/4*
	40.0 *2/5*	25.0 *1/4*	0.0 *0/3*	-	-	-	25.0 *3/12*
	-	-	-	-	0.0 *0/6*	-	0.0 *0/6*
	-	-	-	0.0 *0/10*	14.3 *1/7*	-	5.9 *1/17*
Blank	-	-	-	-	-	26.7 *4/15*	26.7 *4/15*
Total	45.5 *5/11*	28.6 *2/7*	12.5 *1/8*	0.0 *0/10*	7.7 *1/13*	26.7 *4/15*	20.3 *13/64*

Table 4
PROBIT Estimates of Individual Choices.
(Standard Errors in Parentheses and Significance Levels in Italics)

	Equal Split	**Rejection**
Constant	-1.195 (.789) *p=.13*	-1.160 (.884) *p=.19*
Difference	**.135** **(.058)** ***p=.02***	.038 (.066) *p=.563*
Devious Icon	.052 (.585) *p=.93*	.240 (.551) *p=.664*
Happy Icon	.832 (.485) *p=.09*	-.030 (.504) *p=.95*
Non-Facial Icon	**-1.526** **(.689)** ***p=.03***	**-1.165** **(.577)** ***p=.04***
Same Icon	.424 (.360) *p=.24*	.311 (.438) *p=.48*
	ll=-86.692, n=177	ll=-28.13, n=64

Table 5
Percentage of Proposers Choosing the Equal Split in the Ultimatum Game (Experiment III; numbers in italics)

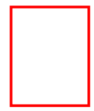
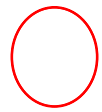
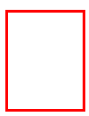
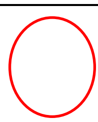
Proposer's Icon	**Respondent's Icon**						
						**Blank**	**Total**
	31.3 *5/16*	53.9 *7/13*	14.3 *1/7*	--	--	--	36.1 *13/36*
	66.7 *4/6*	50.0 *7/14*	45.5 *5/11*	--	--	--	51.6 *16/31*
	20.0 *1/5*	0.0 *0/5*	62.5 *10/16*	--	--	--	42.3 *11/26*
	--	--	--	30.0 *3/10*	38.5 *5/13*	--	34.8 *8/23*
	--	--	--	55.6 *10/18*	50.0 *5/10*	--	53.6 *15/28*
Blank	--	--	--	--	--	51.1 *23/45*	51.1 *23/45*
Total	37.0 *10/27*	43.8 *14/32*	47.1 *16/34*	46.4 *13/28*	43.5 *10/23*	51.1 *23/45*	45.5 *86/189*

Table 6
Rejections by Respondents for Unequal Proposed Splits in the Ultimatum Game (Experiment III; numbers in italics)

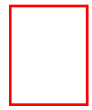
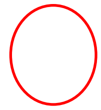
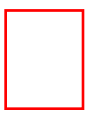
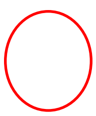
Proposer's Icon	**Respondent's Icon**						
						**Blank**	**Total**
	9.1 *1/11*	0.0 *0/6*	16.7 *1/6*	-	-	-	8.7 *2/23*
	0.0 *0/2*	0.0 *0/7*	33.3 *2/6*	-	-	-	13.3 *2/15*
	50.0 *2/4*	20.0 *1/5*	50.0 *3/6*	-	-	-	40.0 *6/15*
	-	-	-	42.9 *3/7*	12.5 *1/8*	-	26.7 *4/15*
	-	-	-	0.0 *0/8*	40.0 *2/5*	-	15.4 *2/13*
Blank	-	-	-	-	-	18.2 *4/22*	18.2 *4/22*
Total	17.6 *3/17*	5.6 *1/18*	33.3 *6/18*	20.0 *3/15*	23.1 *3/13*	18.2 *4/22*	19.4 *20/103*

Table 7
Percentage of Proposers Choosing the Equal Split in the Ultimatum Game (Experiment IV; Games 1 and 2; numbers in italics)

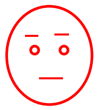
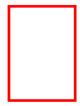
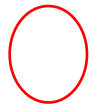
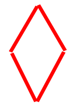
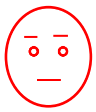
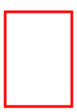
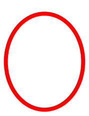
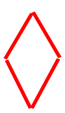
							Blank	Total
	53.3 *8/15*	53.8 *7/13*	78.6 *11/14*	-	-	-	-	61.0 *26/42*
	57.1 *8/14*	39.1 *9/23*	53.8 *7/13*	71.4 *5/7*	57.1 *4/7*	-	-	51.6 *33/64*
	46.1 *6/13*	57.1 *7/15*	46.7 *7/15*	-	-	-	-	50.0 *21/42*
	-	71.4 *5/7*	-	64.3 *9/14*	38.5 *5/13*	33.3 *2/6*	-	52.5 *21/40*
	-	100.0 *7/7*	-	76.9 *10/13*	71.4 *10/14*	50.0 *3/6*	-	75.0 *30/40*
	-	-	-	50.0 *3/6*	16.7 *1/6*	33.3 *2/6*	-	33.3 *6/18*
Blank	-	-	-	-	-	-	58.3 *35/60*	58.3 *35/60*
Total	52.4 *22/42*	56.2 *36/64*	59.5 *25/42*	67.5 *27/40*	50.0 *20/40*	38.9 *7/18*	58.3 *35/60*	56.2 *172/306*

Table 8
Rejections by Respondents for Unequal Proposed Splits in the Ultimatum Game (Experiment IV; Games 1 and 2; numbers in italics)

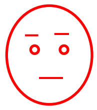
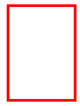
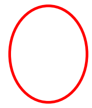
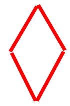
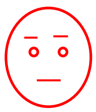
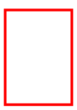
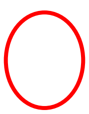
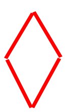
							Blank	Total
	28.6 *2/7*	16.7 *1/6*	0.0 *0/3*	-	-	-	-	18.7 *3/16*
	0.0 *0/6*	0.0 *0/14*	0.0 *0/6*	0.0 *0/2*	33.3 *1/3*	-	-	3.2 *1/31*
	14.3 *1/7*	16.7 *1/6*	12.5 *1/8*	-	-	-	-	14.3 *3/21*
	-	0.0 *0/2*	-	20.0 *1/5*	0.0 *0/8*	0.0 *0/4*	-	5.3 *1/19*
	-	-	-	0.0 *0/3*	0.0 *0/4*	0.0 *0/3*	-	0.0 *0/10*
	-	-	-	0.0 *0/3*	20.0 *1/5*	50.0 *2/4*	-	25.0 *3/12*
Blank	-	-	-	-	-	-	8.0 *2/25*	8.0 *2/25*
Total	15.0 *3/20*	7.1 *2/28*	5.9 *1/17*	7.7 *1/13*	10.0 *2/20*	18.2 *2/11*	8.0 *2/25*	9.7 *13/121*

Table 9
Probit Estimates for Fair Offer in Ultimatum Games -- Experiment II
(Standard Errors in Parentheses)

	Fair Choice		Rejection
	Other	Self	Other
Constant	.209 (.288)	.205 (.287)	.141 (.512)
Devious	-.110 (.317)	.002 (.312)	.214 (.390)
Angry	.156 (.365)	.512 (.417)	.145 (.615)
Happy	.548# (.320)	.251 (.307)	.047 (.453)
Non Face	-.534 (.327)	-.521 (.329)	-1.151* (.490)
Period	.049** (.016)	.052** (.015)	.024 (.021)
%Cooperative	-.001 (.004)	-.002 (.004)	-.001 (.004)
%Non Cooperative	-.016** (.003)	-.016** (.003)	-.006 (.004)
n	190	190	71
pseudo r^2	.25	.24	.11
	#p<.10	**p<.05	**p<.01

Figure 1

A Simple Ultimatum Game

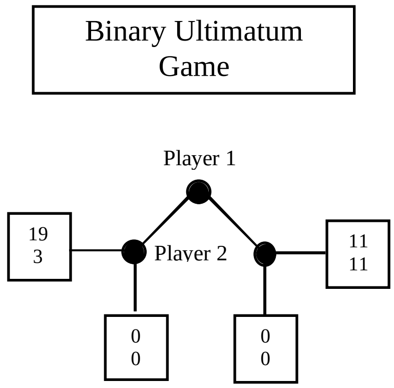


Figure 2
Icons Used in Survey

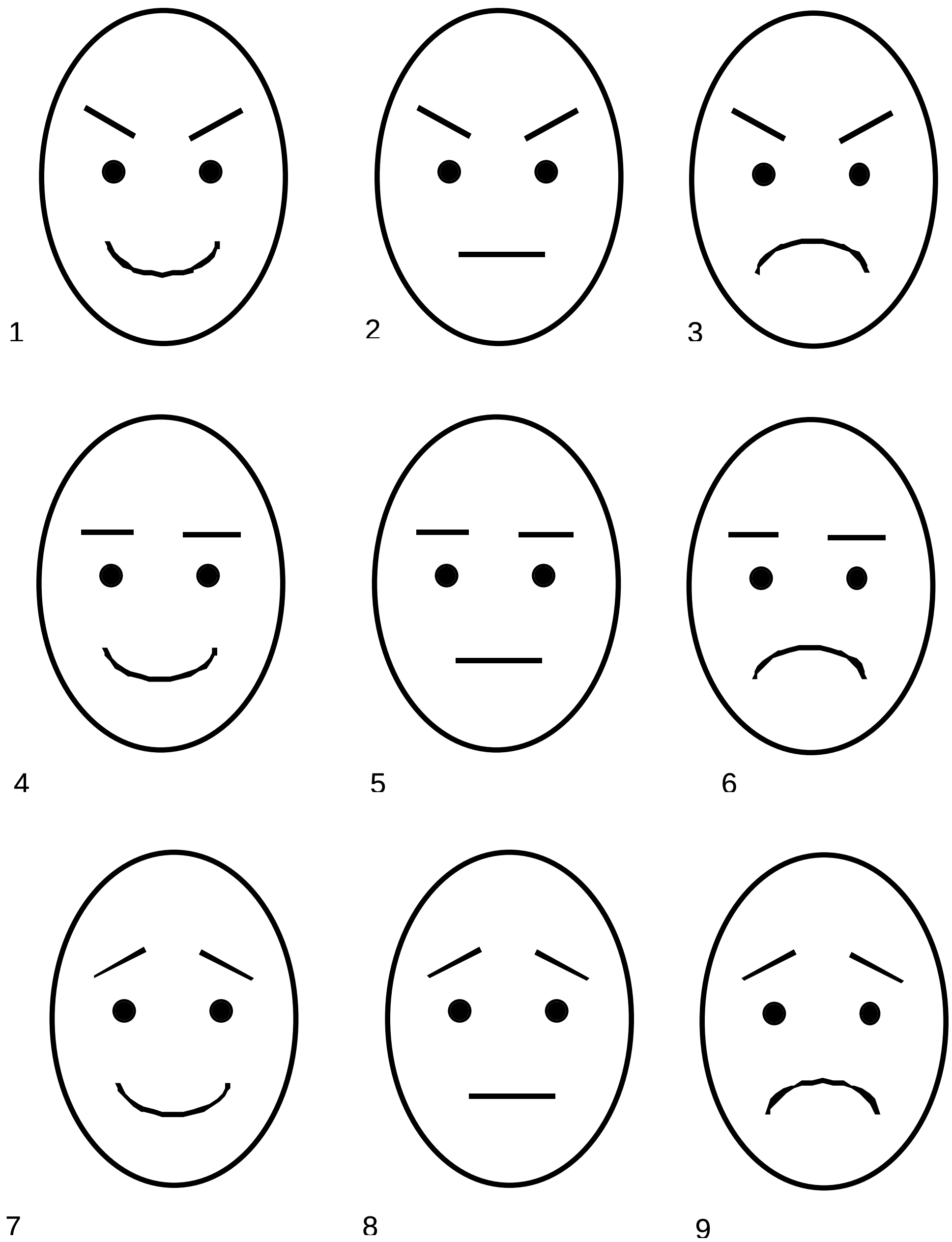


Figure 3
Icon Pairs Used in Experiment

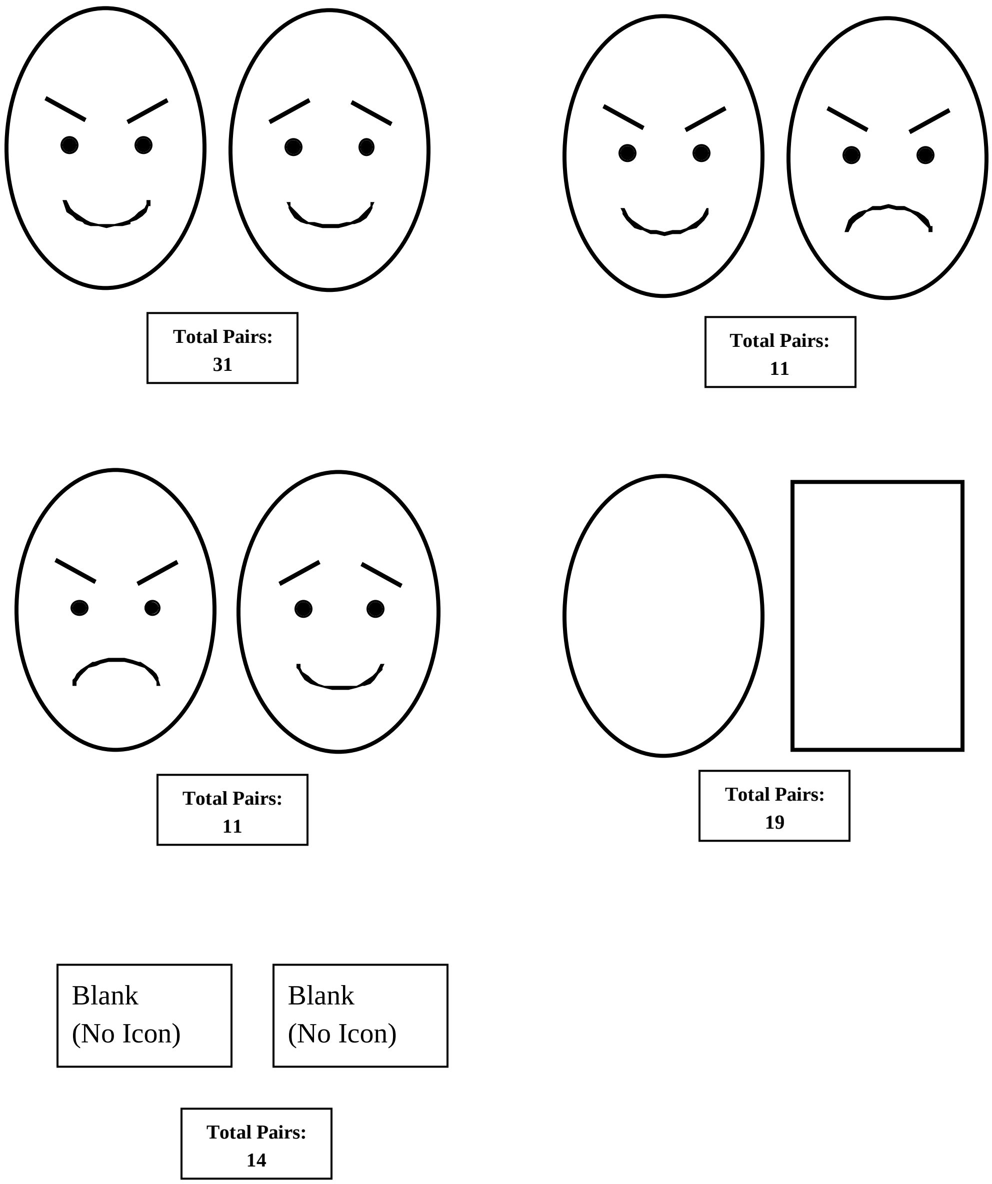


Figure 4
A Subset of Games and Parameters used in Experiments II and III

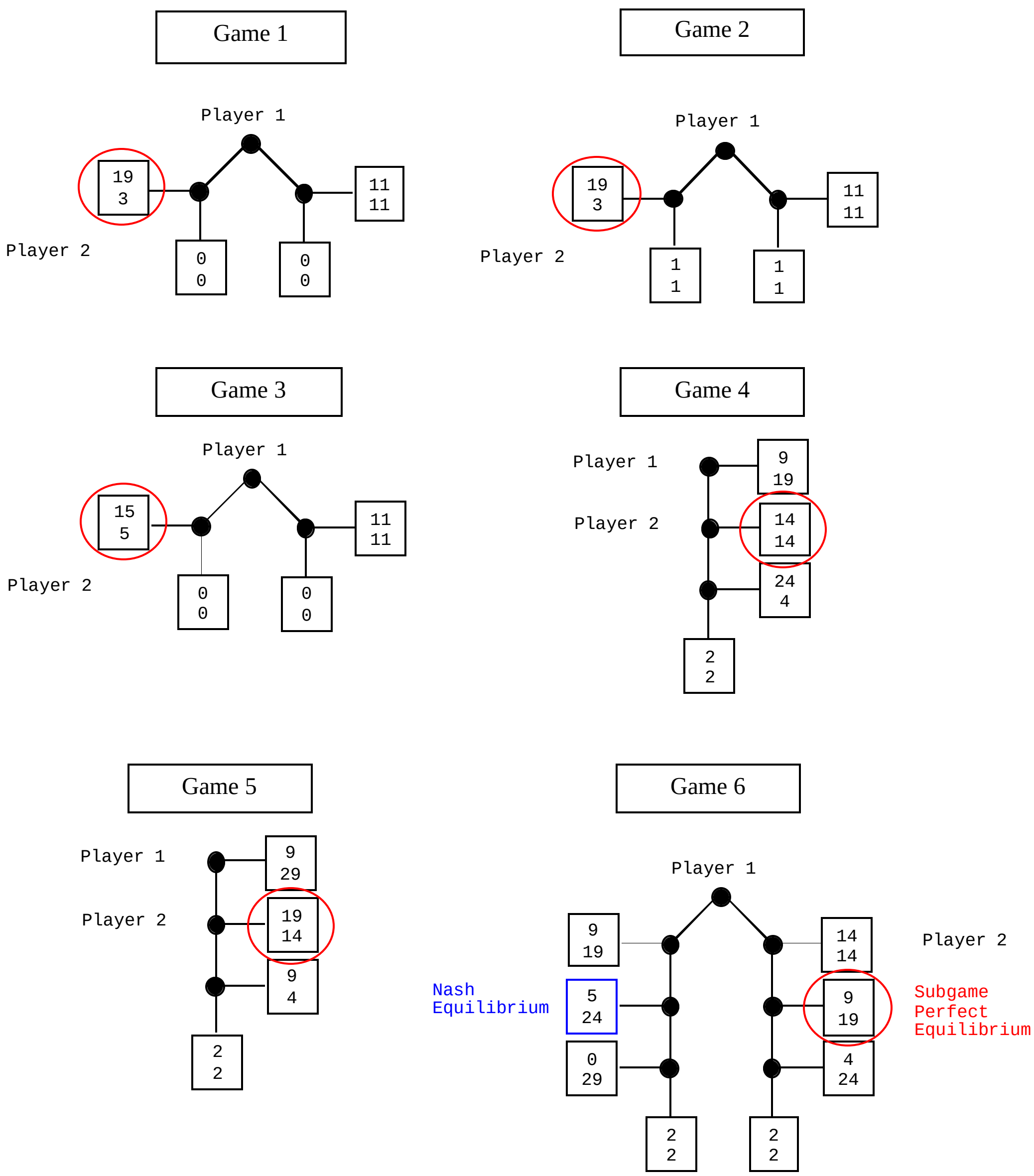


Figure 5a
Aggregate Outcomes for Ultimatum Games, Experiment II

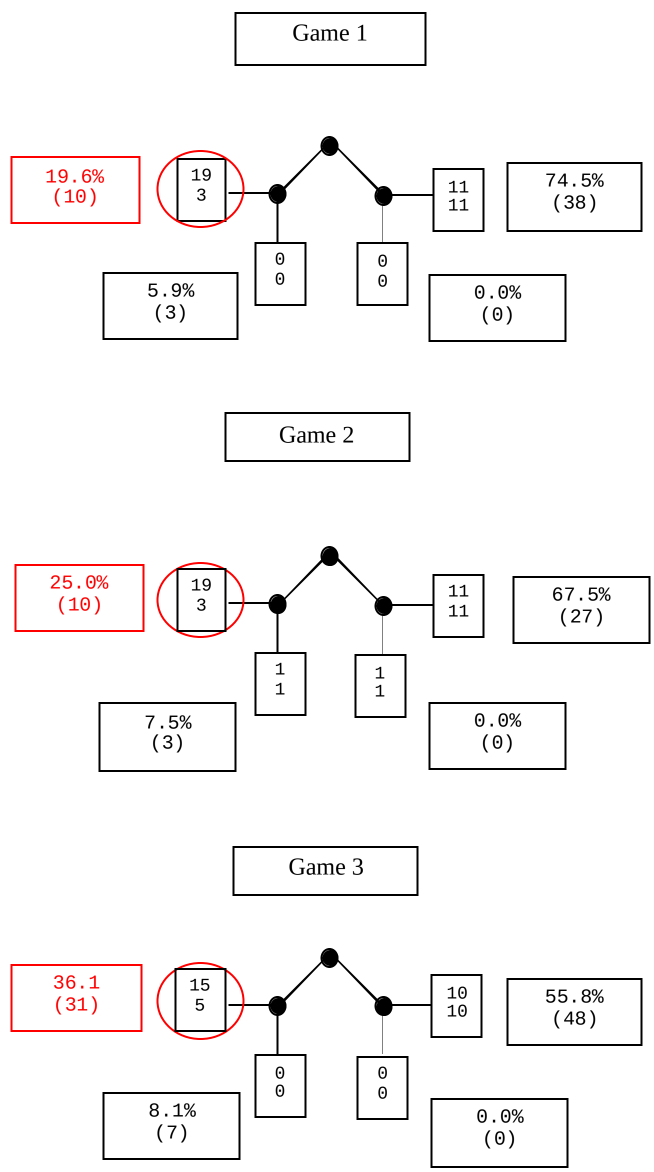


Figure 5b
Aggregate Outcomes for Simple Bargaining Games, Experiment II

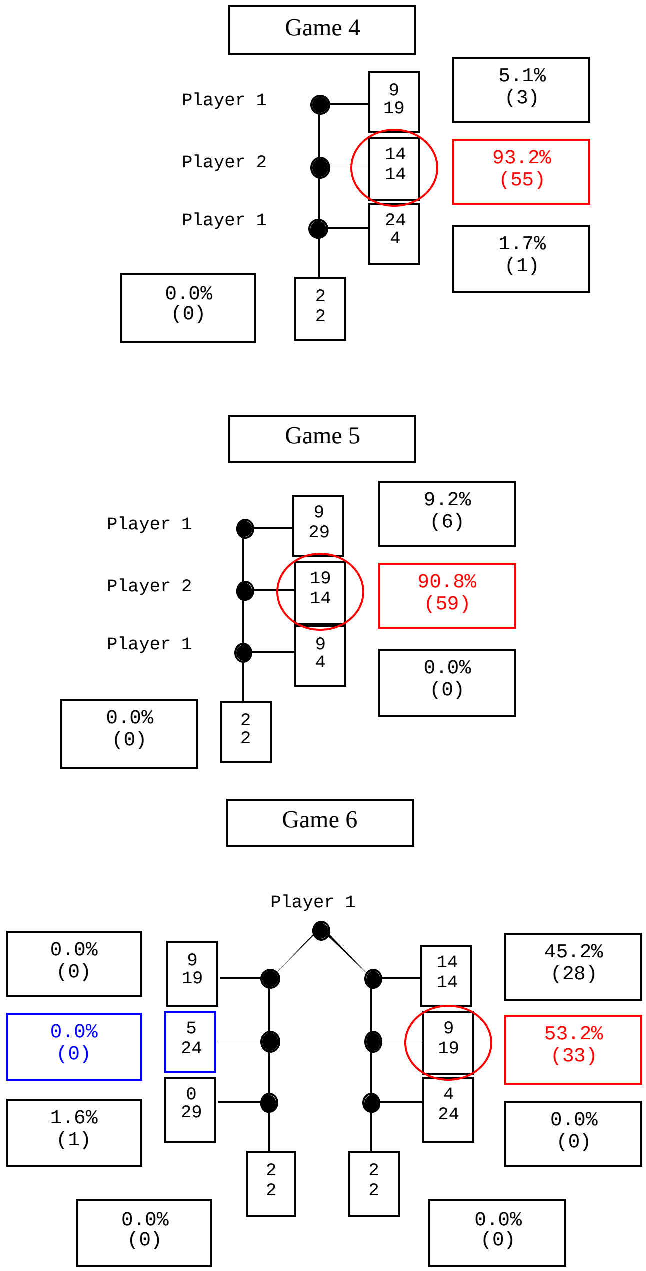
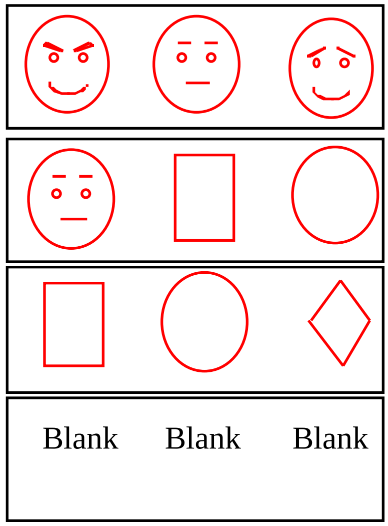


Figure 6
Icons Triples and Payoffs Used in Experiment IV



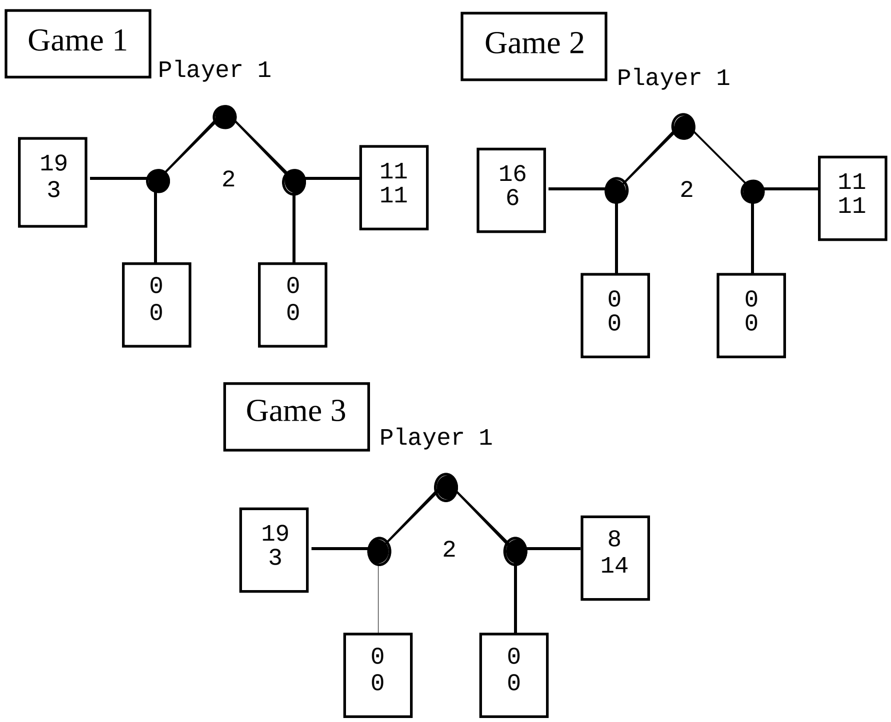
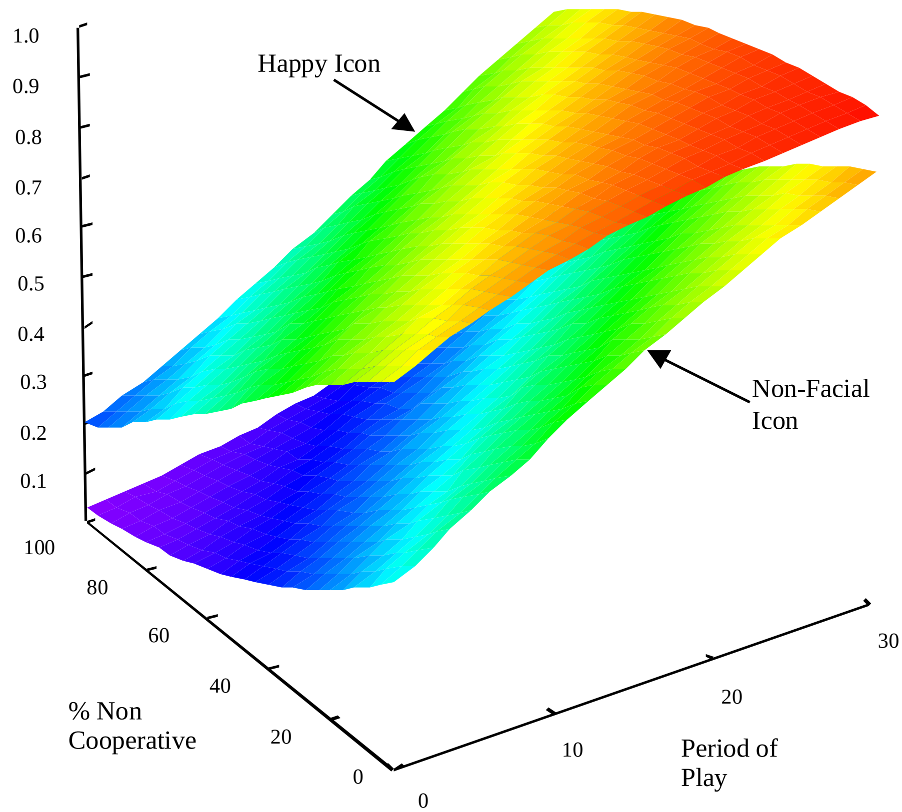


Figure 7
Predicted Probability of an Equal Split Given Past History



Appendix 1
Instruction Set

Screen 1

In this experiment you will participate in several two person decision problems. At each decision you will be randomly paired with another individual in this room: your counterpart.

The joint decisions made by you and your counterpart will determine how much money you will earn for this decision problem.

Your earnings for this decision will be paid to you in cash at the end of this experiment. I will not tell anyone else your earnings and I ask you not to discuss your earnings.

Click OK when you are ready to continue.

OK

Screen 2

You will not be paid for every decision in the experiment. You will make many decisions with the other participants in this experiment.

At the conclusion of the experiment, ONE of the decisions will be randomly selected. You will be paid for that decision.

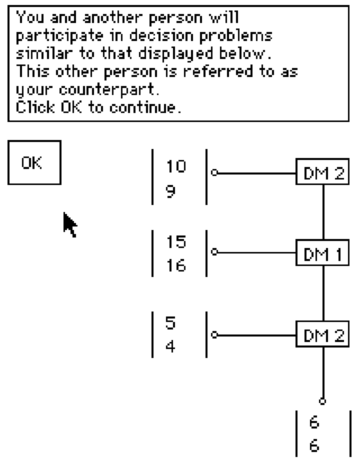
On the sheet of paper I have provided, please record your potential earnings for each decision. This will help you keep track of what you earn at the end.

Click OK when you are ready to continue.

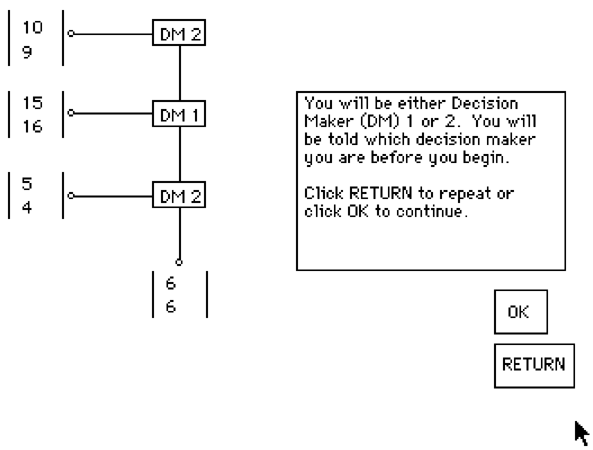
OK

-

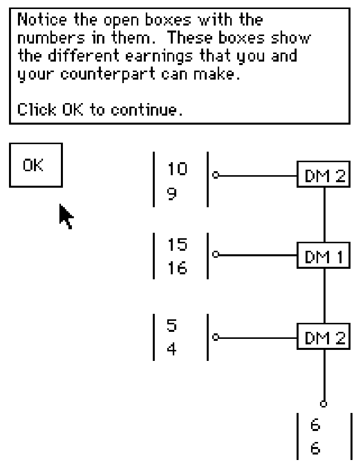
Screen 3



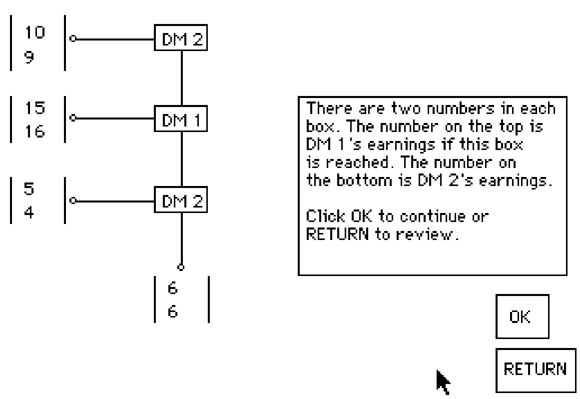
Screen 4



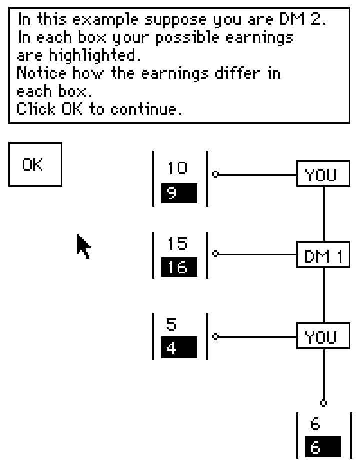
Screen 5



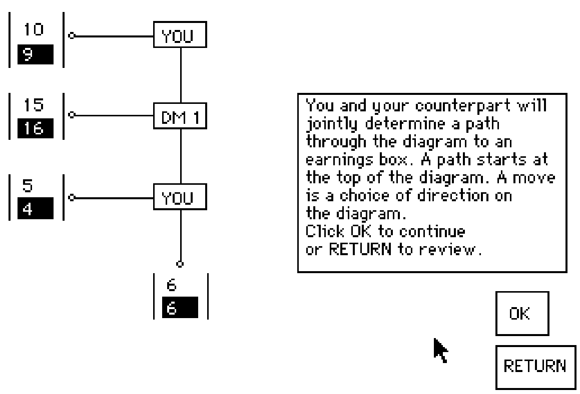
Screen 6



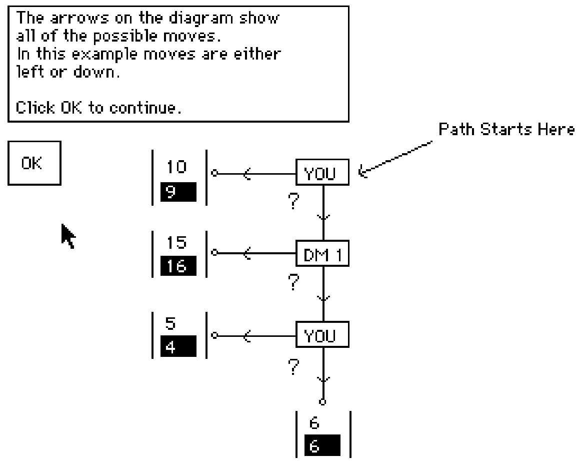
Screen 7



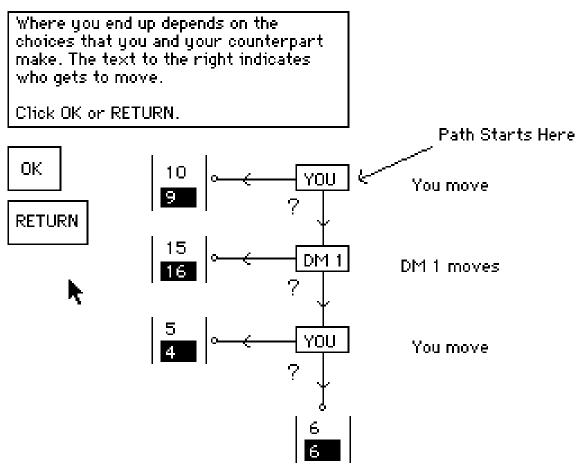
Screen 8



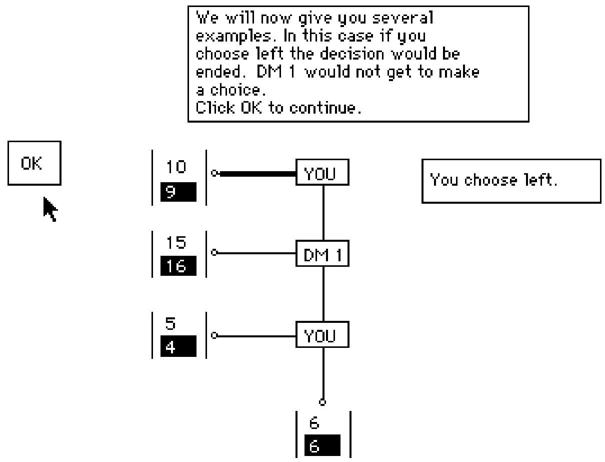
Screen 9



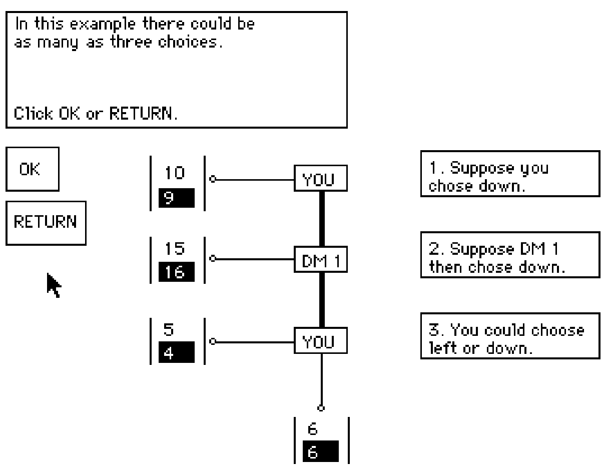
Screen 10



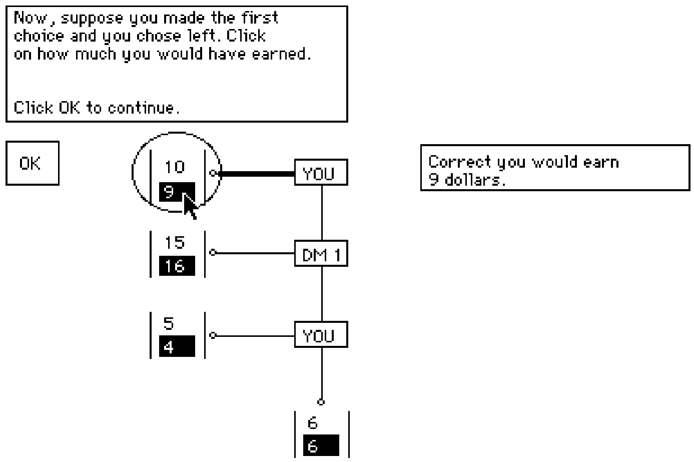
Screen 11



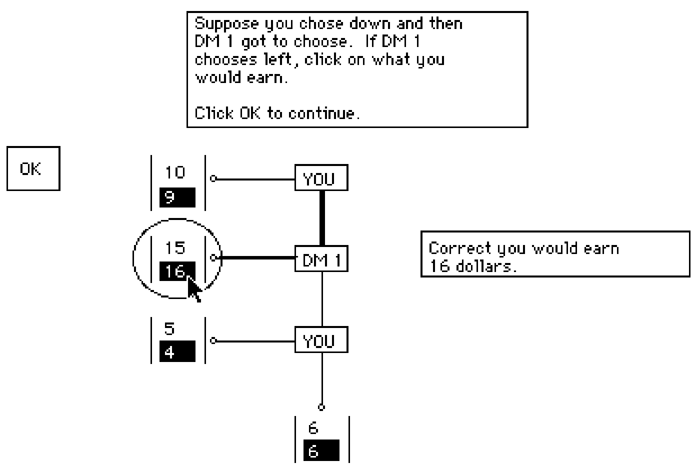
Screen 12



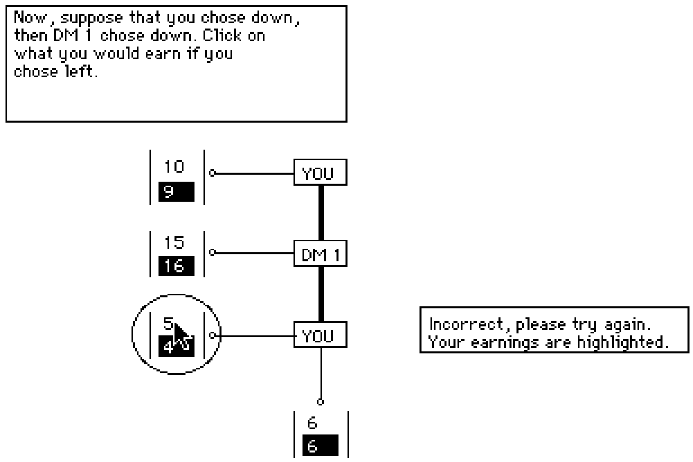
Screen 13



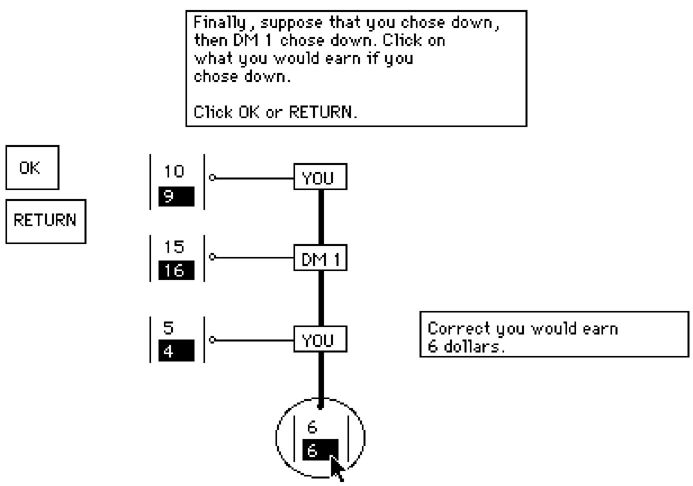
Screen 14



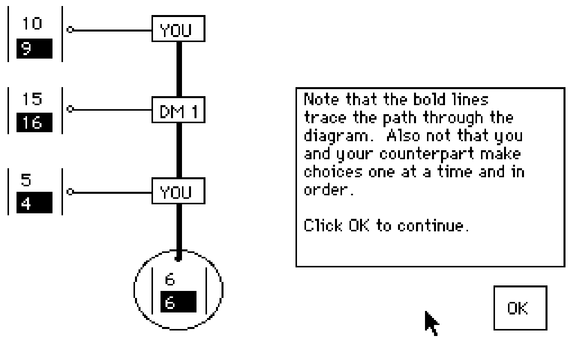
Screen 15



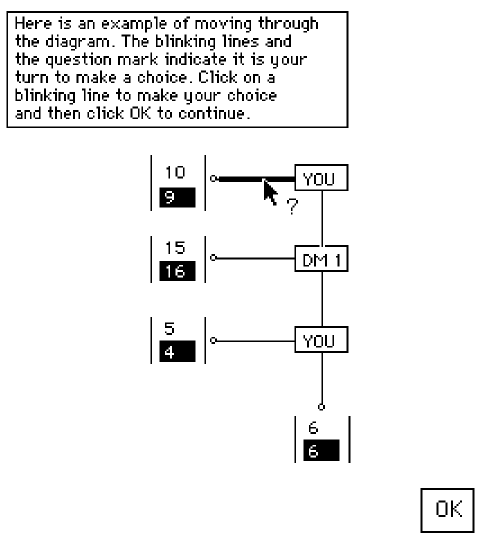
Screen 16



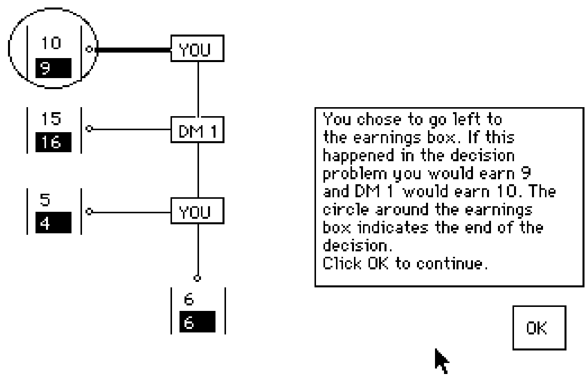
Screen 17



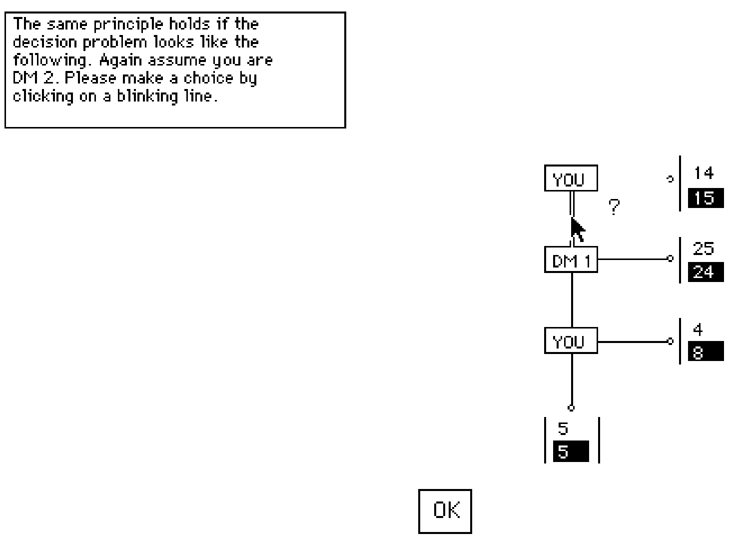
Screen 18



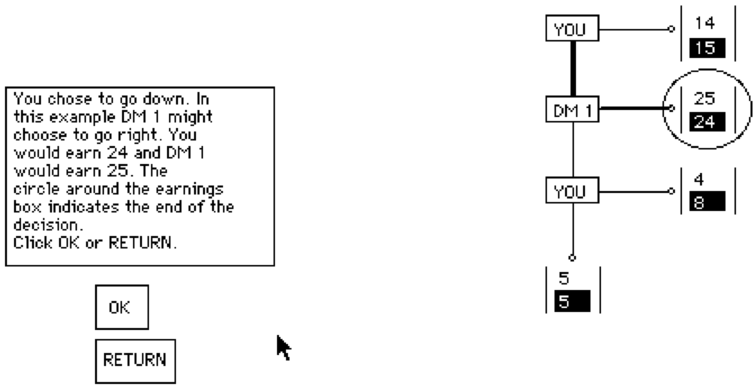
Screen 19



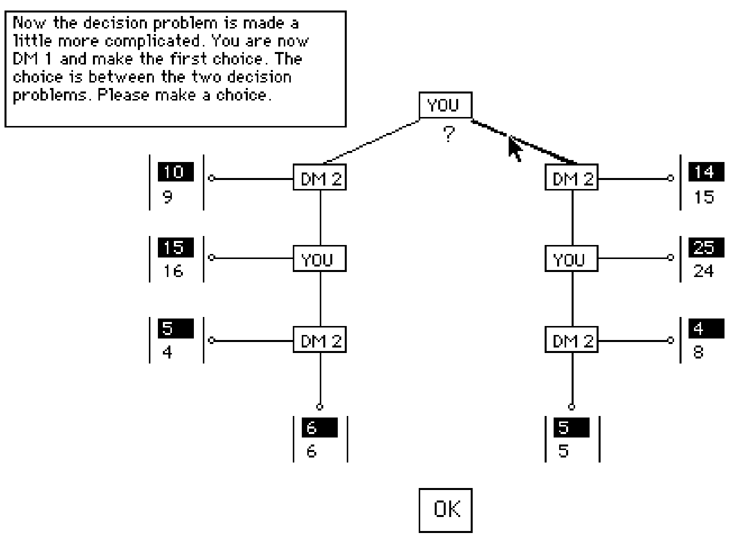
Screen 20



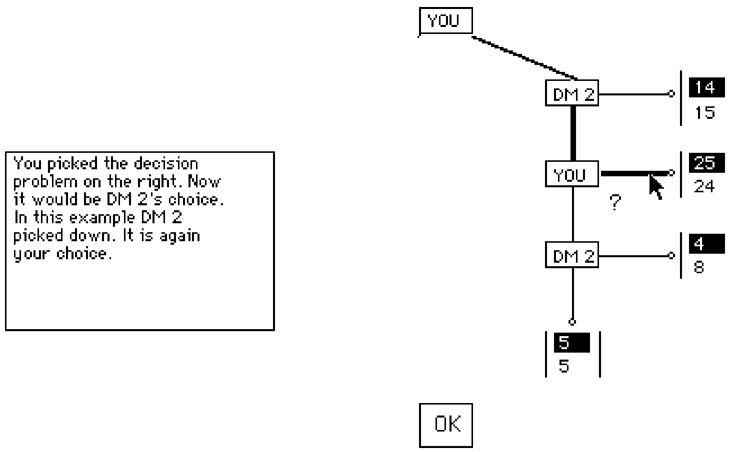
Screen 21



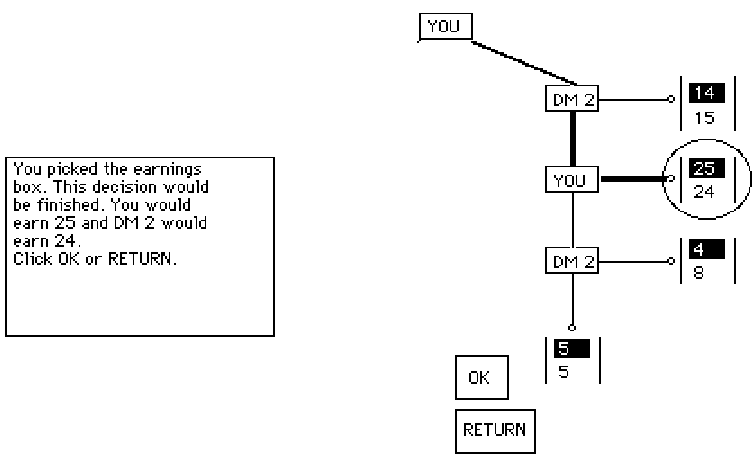
Screen 22



Screen 23



Screen 24



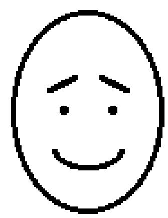
Screen 25

You are about ready to begin. You will make 30 different decisions. However, you will only be paid for one of the decisions that you and your counterpart make. At the end of all of your decisions, you will get to randomly pick one of decisions for which you will be paid. You have a sheet of paper and a pencil to mark your earnings from each decision. Please keep track of how much you could make following each decision. If you have any questions, please ask them now. Otherwise click OK to continue.

OK

Screen 26

Finally, during this experiment you will be represented by the icon illustrated to the right. This is what your counterpart will see before beginning a decision problem. Likewise you will see the icon for your counterpart.

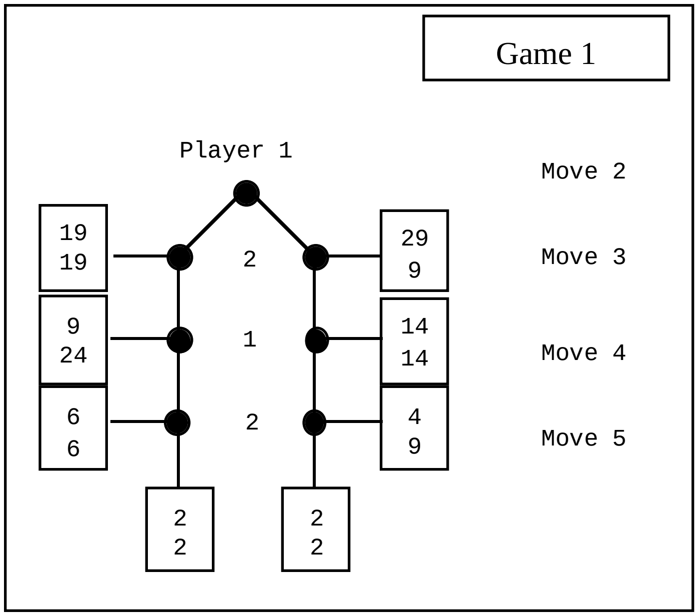


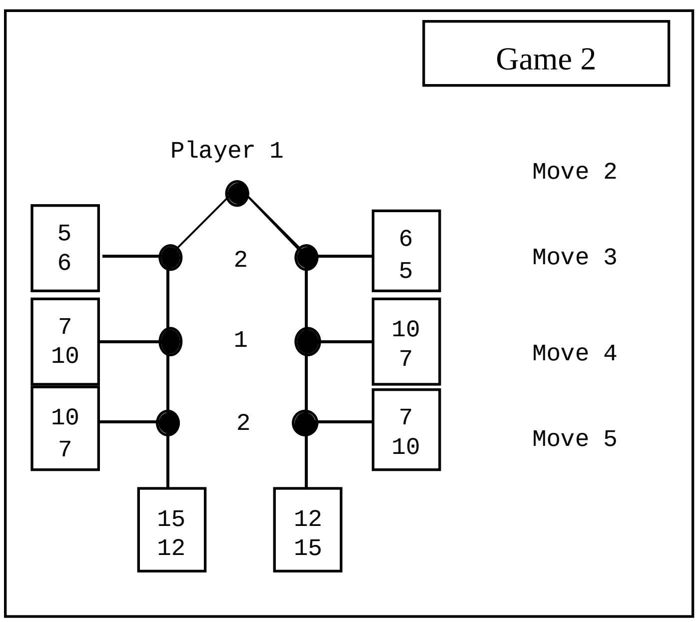
Click OK to continue or RETURN to review.

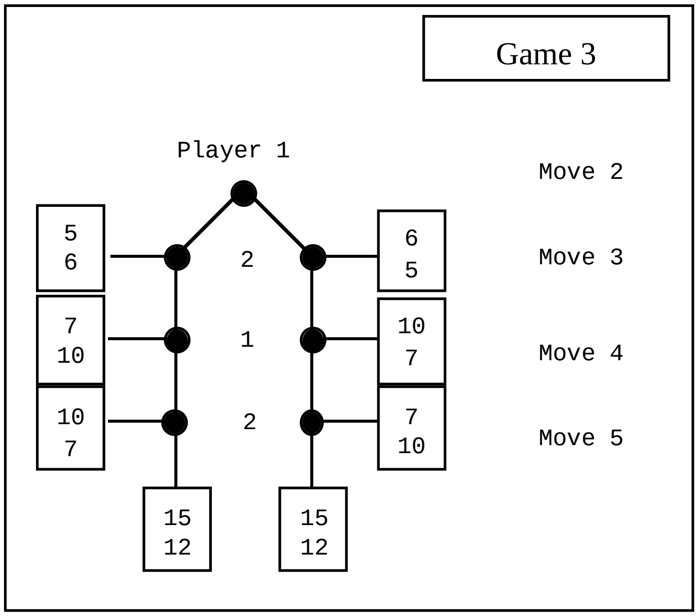
OK

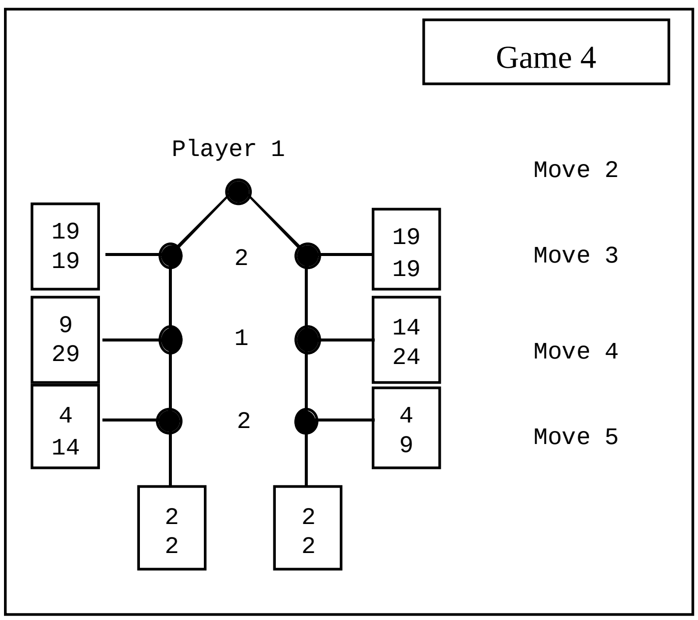
RETURN

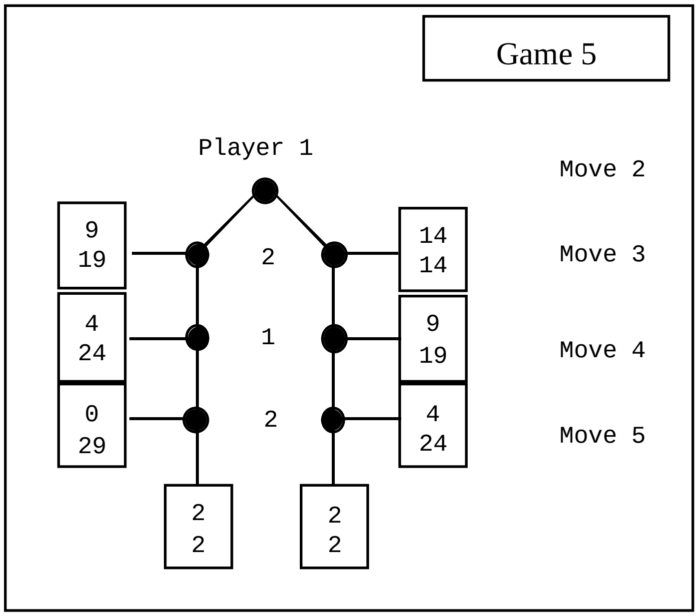
Appendix 2
Complete Set of Games Used in Experiment II

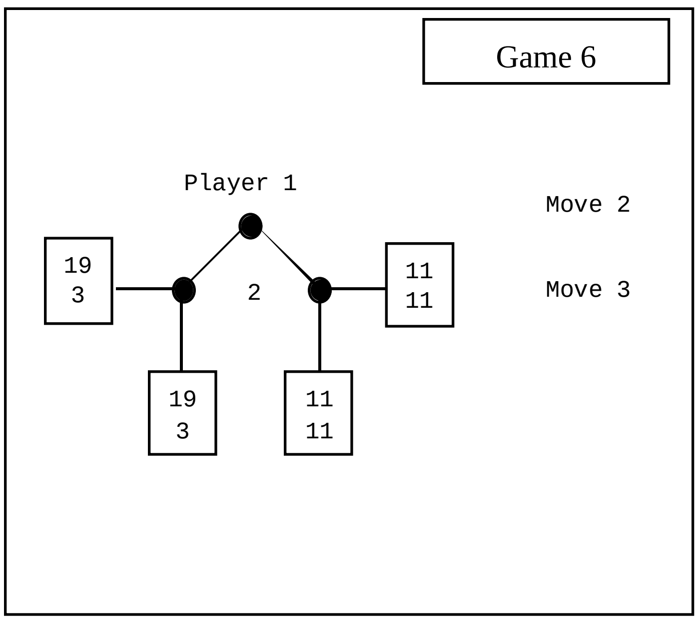


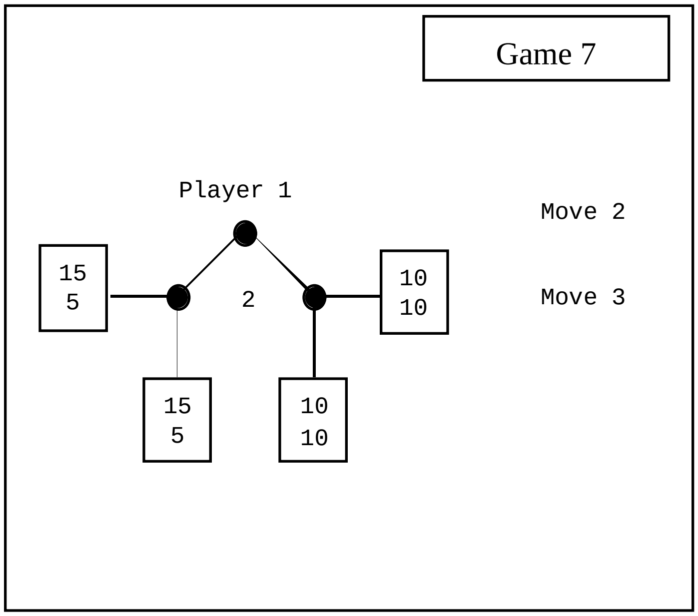


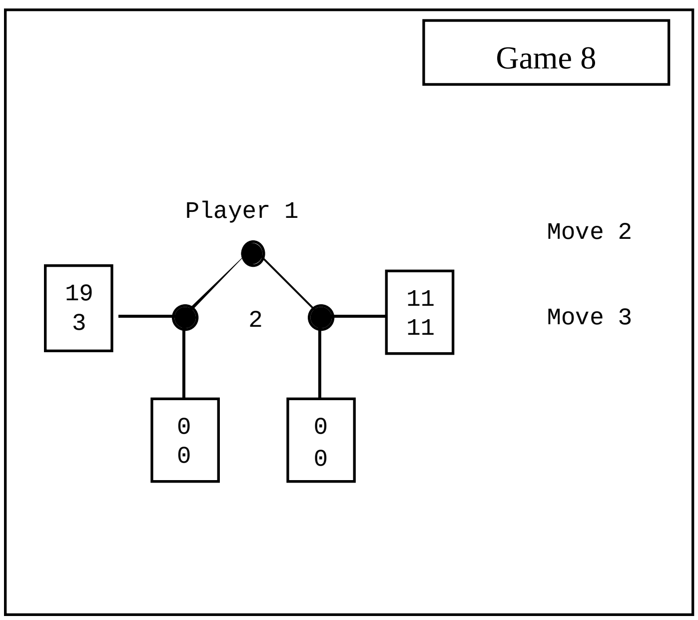

Game 3
Player 1
Move 2
Move 3
Move 4
Move 5
5
6
6
5
2
7
10
10
7
1
10
7
7
10
2
15
12
15
12

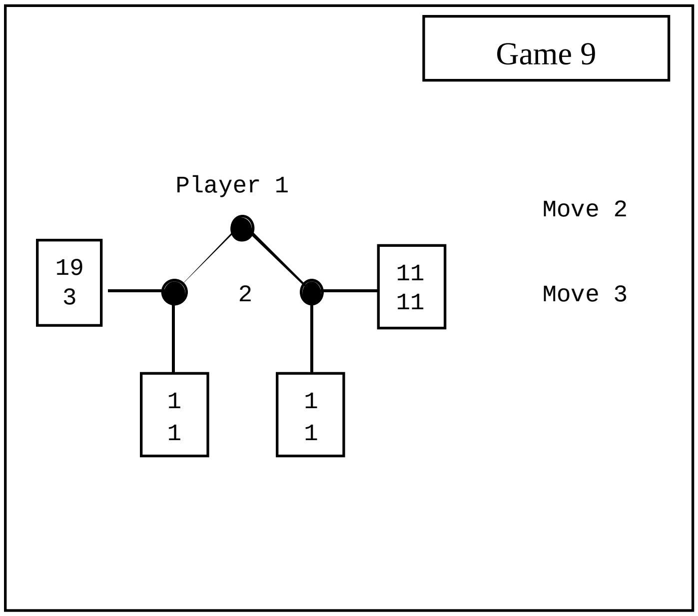

Game 4
Player 1
Move 2
Move 3
Move 4
Move 5
19
19
19
19
2
9
29
14
24
1
4
14
4
9
2
2
2
2
2

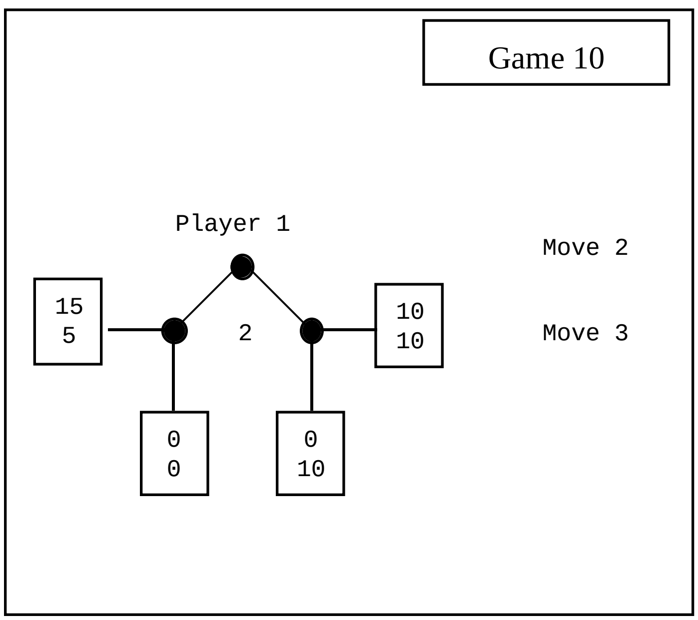

Game 5
Player 1
Move 2
Move 3
Move 4
Move 5
9
19
2
14
14
4
24
1
9
19
0
29
2
4
24
2
2
2
2

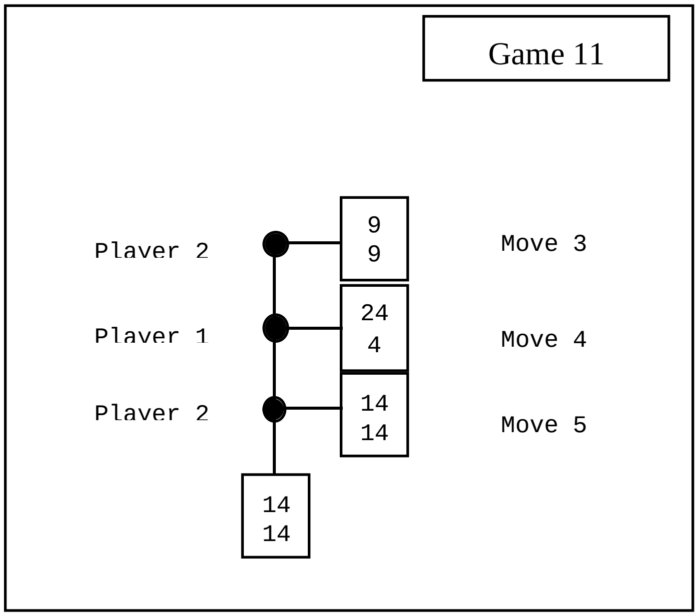

Game 6
Player 1
Move 2
Move 3
19
3
2
11
11
19
3
11
11

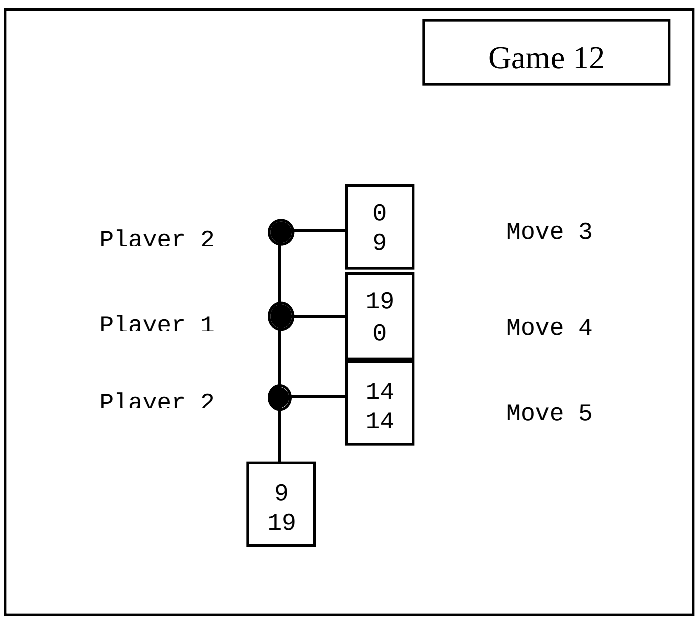

Game 7
Player 1
Move 2
15
5
2
10
10
Move 3
15
5
10
10

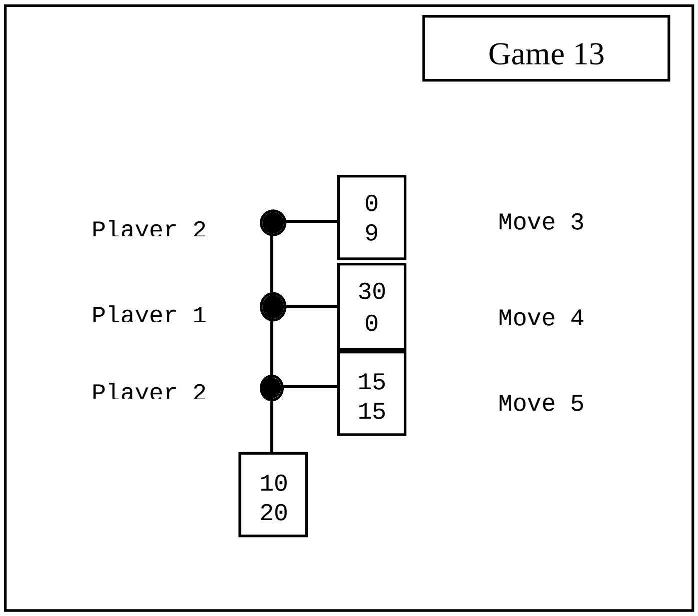

Game 8
Player 1
Move 2
19
3
2
11
11
Move 3
0
0
0
0

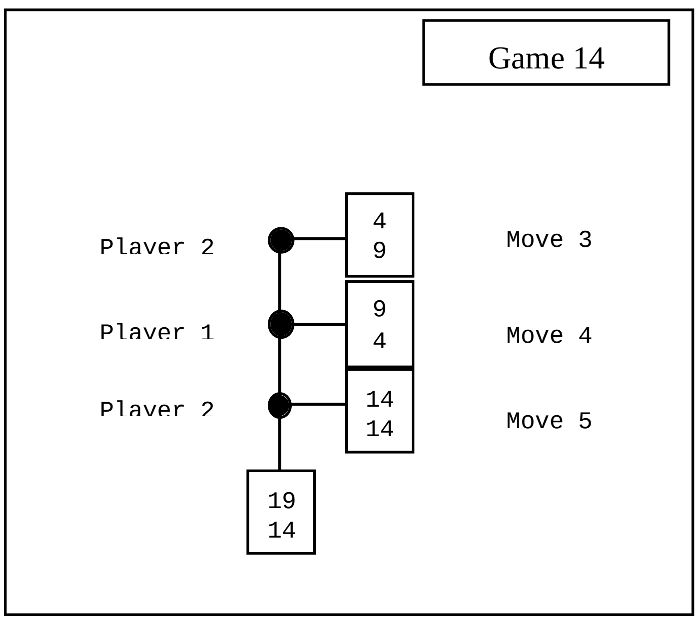

Game 9
Player 1
Move 2
19
3
2
11
11
Move 3
1
1
1
1

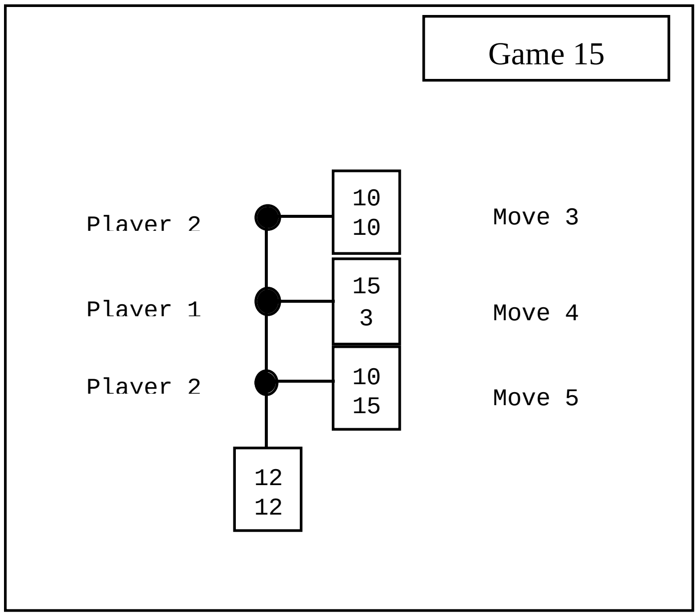

Game 10
Player 1
Move 2
15
5
2
10
10
Move 3
0
0
0
10

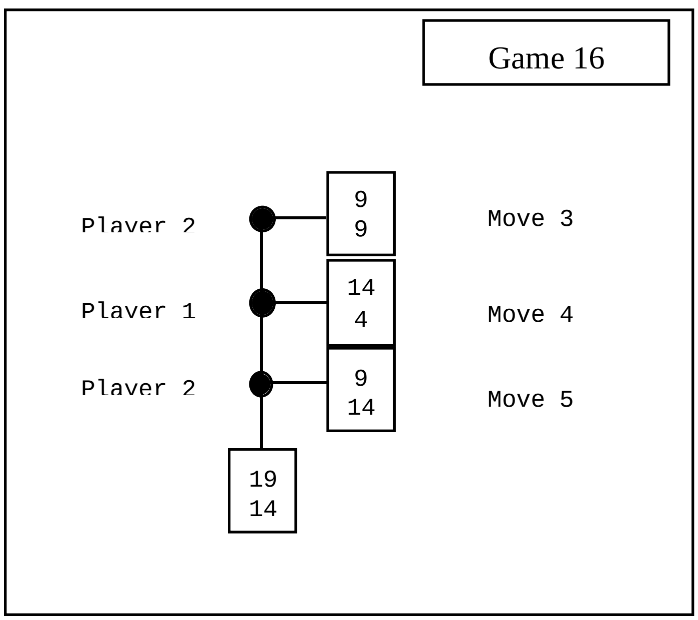

Game 11
Player 2
Player 1
Player 2
9
9
24
4
14
14
14
14
Move 3
Move 4
Move 5

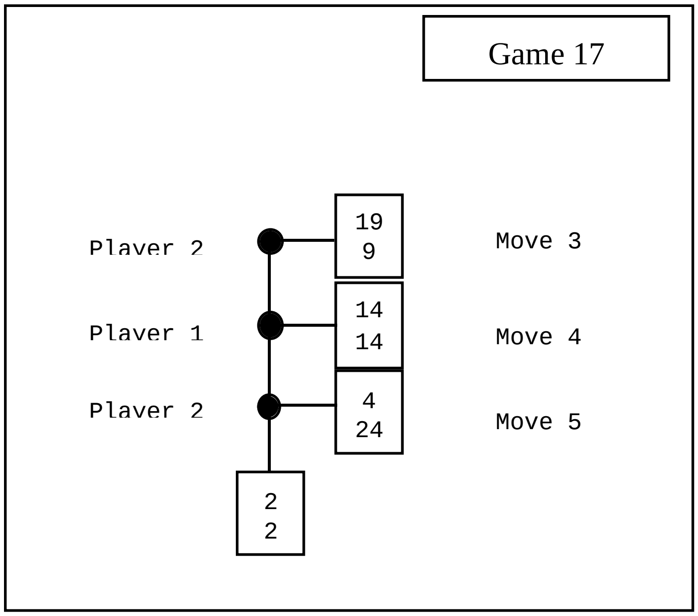

Game 12
Player 2
Player 1
Player 2
0
9
19
0
14
14
9
19
Move 3
Move 4
Move 5

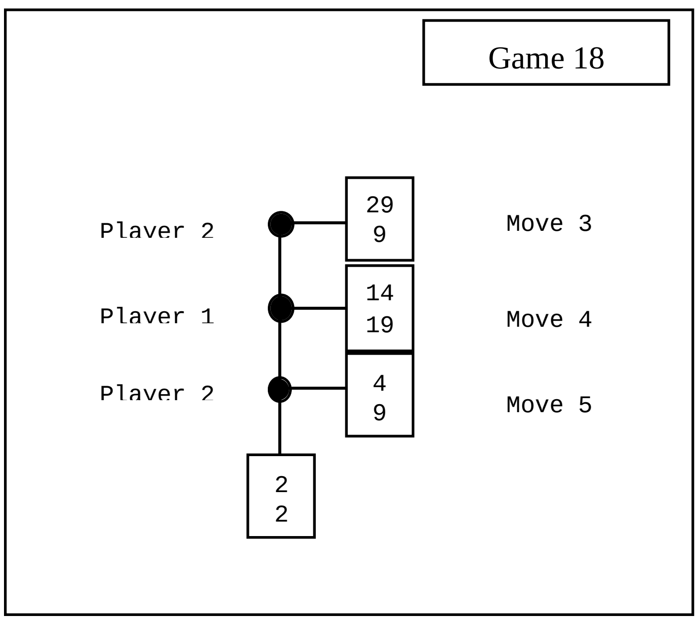

Game 13
0
9
Player 2
Move 3
30
0
Player 1
Move 4
15
15
Player 2
Move 5
10
20


Game 14
4
9
Player 2
Move 3
9
4
Player 1
Move 4
14
14
Player 2
Move 5
19
14


Game 15
Player 2
Player 1
Player 2
10
10
15
3
10
15
12
12
Move 3
Move 4
Move 5


Game 16
Player 2
Player 1
Player 2
9
9
14
4
9
14
19
14
Move 3
Move 4
Move 5


Game 17
Player 2
Player 1
Player 2
19
9
14
14
4
24
2
2
Move 3
Move 4
Move 5


Game 18
Player 2
Player 1
Player 2
29
9
14
19
4
9
2
2
Move 3
Move 4
Move 5