COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

AT YALE UNIVERSITY

Box 2125, Yale Station
New Haven, Connecticut

COWLES FOUNDATION DISCUSSION PAPER NO. 247

SOME EXPERIMENTAL NON-CONSTANT-SUM GAMES REVISITED

PERCEPTION OF OPPONENT'S PAYOFFS

PART III

Martin Shubik and David Stern

March 11, 1968

SOME EXPERIMENTAL NON-CONSTANT-SUM GAMES REVISITED·

PERCEPTION OF OPPONENT'S PAYOFFS

PART III

by

Martin Shubik and David H. Stern*

1. Introduction

In two previous papers[1/] we analyzed the play of subjects in a series of experiments involving 2 x 2-payoff-matrix two-person nonconstant-sum games in which the players knew only their own payoffs and not those of their opponents. In this paper we explore the effects of the last mentioned feature, the lack of information about the opponent's payoff matrix.

First, however, in order to make this paper self-contained we summarize from the earlier two papers that material which forms essential background for this presentation.

1.1 Form of the Experiments[2/]

Figure 1 shows the six games used in the experiments. In each game the subject labeled "Player 1" chose the row and the subject labeled "Player 2" the column; the labeling of the players was arbitrary. In each cell of a matrix is first the payoff to player 1 and second the payoff to player 2 resulting from the two players' corresponding strategy choices.

---

* Research undertaken by the Cowles Commission for Research in Economics under Contract Nonr-3055(01) with the Office of Naval Research.

Game 1

6, 3	6, 7
10, 3	10, 7

Game 2

1, 3	2, 3
1, 1	2, 1

Game 3

2, 1	-1, -1
-1, -1	1, 2

Game 4

3, 3	-1, -1
-1, -1	2, 2

Game 5

3, 3	-2, 7
7, -2	-1, -1

Game 6

5, 2	-10, -13
4, 1	-20, -23

Figure 1

In our experiments subjects did not see the matrices shown in Figure 1 but instead saw only their own payoffs, shown as the first two columns of matrices in Figure 2, where Player 1 chooses a row and Player 2 a column.

For certain purposes in this discussion it will be convenient to think of both players as choosing rows; for Player 1 this involves no change, but for Player 2 it implies transposition of the matrix (reflection about the main diagonal), as shown in Column (3) of Figure 2. When a player's payoff matrix is displayed so that he chooses rows we say the matrix is in its standard aspect; when it is displayed in the manner in which it is actually played, we say it is in game aspect. For Player 1 standard and game aspects coincide; for Player 2 they generally do not.

	Player 1's Payoff Matrix (1)		Player 2's Payoff Matrix in Game Aspect (2)		Player 2's Payoff Matrix in Standard Aspect (3)	
Game 1	6	6	3	7	3	3
	10	10	3	7	7	7
Game 2	1	2	3	3	3	1
	1	2	1	1	3	1
Game 3	2	-1	1	-1	1	-1
	-1	1	-1	2	-1	2
Game 4	3	-1	3	-1	3	-1
	-1	2	-1	2	-1	2
Game 5	3	-2	3	7	3	-2
	7	-1	-2	-1	7	-2
Game 6	5	-10	2	-13	2	1
	4	-20	1	-23	-13	-23

Figure 2

1.2 Motivation of Subjects to Maximize Their Payoffs

No monetary rewards were used in any of these experiments to motivate the players to equate maximization of their scores with maximization of individual utility. However, most of the subjects were students of game theory in classes taught by one of the authors, who advised that these experiments would teach non-constant-sum game phenomena in an easily assimilable way. Furthermore the students were specifically instructed to maximize their scores.

The experimenters may, however, have introduced a distracting element themselves by their requirement that subjects provide information about the matrix they believed their opponents had faced. This interference with the score-maximization motive resulted if a subject altered his strategy choices in order to gain information about his opponent's matrix. Further discussion of this is given in section 2. 1. 1. below.

1.3 Taxonomy of Standard-Aspect Matrices[3/]

In Part II a rationale was developed for a taxonomy of standard-aspect matrices based on first a decision rule for picking the initial strategy from any matrix, and second, the various motivations to switch strategies that a player might encounter in iterated play. The initial-strategy decision rule used was:

α Dominant Strategy Subrule: If the matrix contains a dominant strategy (in the strong sense), choose it.

β Bayesian Subrule: If α fails to select a strategy, choose the strategy with the highest expected value.

λ Security Maximization Subrule: If β fails to select a strategy, choose the strategy with the highest minimum value; choose the strategy which maximizes the "security level."

δ Randomization Subrule: If γ fails to select a unique strategy, randomize.

The three motivations for switching after using rule $\alpha\ \beta\ \lambda\ \delta$ initially were defined as follows:

<u>Maximizing Incentive</u> - Subject can increase his payoff by switching if opponent continues to use his same strategy.

<u>Investment Incentive</u> - Subject can increase own payoff by switching provided opponent also switches.

<u>Signalling Incentive</u> - Subject can increase own payoff if he can induce opponent to switch without switching himself. In order to do this he switches in order to disturb the status quo; if he succeeds in getting his opponent to switch, he switches back again.

The three motivations can be distinguished by their desired endpoints, as shown in Figure 3. "S" is the starting outcome, "E" the desired ending outcome; the subject in each case is switching rows, though in the last case only temporarily.

S	
E	C

Maximizing

S	
C	E

Investment

S	E
C	C

Signalling

Figure 3

A switch fails if the ending outcome is other than E, specifically, if it is an outcome labeled C. If $C < S$, the switch is called "costly"; if $C \geq S$, it is called "costless".

Table 1 classifies all possible 2 x 2 matrices, with payoffs specified ordinally: $A > B > C > D$. Row 1 is picked initially (except of course where subrule δ must be invoked). The "overall measure of incentive to switch" summarizes the information in columns (4) - (9); it is explained in Part II[4/]

A matrix will be referred to by its identification tag for Column 1 of Table 1; if it appears exactly as in the table, the tag is followed by 11; if columns are interchanged, 12; if rows are interchanged, 21; if both rows and columns are interchanged, 22. Thus $\begin{array}{c|c} C & A \\ \hline D & B \end{array}$ is matrix 4a-12 . Where interchanging makes no difference, the number is replaced by an x : $\begin{array}{c|c} A & A \\ \hline B & B \end{array}$ is matrix 1b-1x . The number of distinct forms of each matrix is shown in column (11).

2. Perceived Matrices of Opponent

In these experiments each subject, after playing each game, reproduced the game-aspect ordinal matrix he believed his opponent had faced. These perceived opponent's matrices provide the raw data for the present paper.

The questions which interest us concern, for example, how well the perceived opponent's matrices (POM's) agree with the actual opponent's matrices (AOM's), what factors are related to good and bad guesses, and what sorts of effects on play follow from players' having had accurate or inaccurate notions of their opponents' payoffs.

There are two matters to be dealt with first: motivation and measurement. Then we examine a variety of questions including those in the preceding paragraph, dividing our efforts into three parts. In section 3 the factors that affect the individual in his specification of a POM are discussed, in section 4 the factors that affect a pair

Table 1 - Taxonomy of 2x2 Ordinal Matrices

Type Number Identification Letter (1)	Matrix (2)	Subrule Selecting Row 1 as Initial Strategy Choice (3)	Incentive to Switch from Initial Strategy Choice: Maximizing Exists (4)	Maximizing Is Costless (5)	Investment Exists (6)	Investment Is Costless (7)	Signalling Exists (8)	Signalling Is Costless (9)	Overall Measure of Incentive to Switch (10)	Number of Distinct Forms of Matrix (11)
1a	AA BC	α	No	No	No	No	No	No	0	4
1b	AA BB	α	No	No	No	No	No	No	0	2
1c	AA AB	α	No	No	No	No	No	No	0	4
2a	AB CD	α	No	No	No	No	Yes	No	1	4
2b	AB DC	α	No	No	No	No	Yes	No	1	4
2c	AB CC	α	No	No	No	No	Yes	No	1	4
2d	AB BC	α	No	No	No	No	Yes	No	1	4
2e	AB CB	α	No	No	No	No	Yes	No	1	4
3	AB BB	α	No	No	No	No	Yes	Yes	2	4
4a	AC BD	α	No	No	(Yes)	No	Yes	No	3	4
4b	AB AC	α	No	No	Yes	No	Yes	No	3	4
5	AC BC	α	No	No	(Yes)	Yes	Yes	Yes	5	4
6a	AC DB	β & λ	(Yes)	No	No	No	Yes	No	5	4
6b	AB CA	β & λ	Yes	No	No	No	Yes	No	5	4
7	BB AC	λ⊛	Yes	No	Yes	No	No	No	6 or 7⊛	4
8	AC CB	β	(Yes)	Yes	No	No	Yes	Yes	7	4
9a	BC AD	λ⊛	Yes	No	Yes	No	(Yes)	No	7 or 8⊛	4
9b	CB AD	λ⊛	Yes	No	Yes	No	(Yes)	No	7 or 8⊛	4
10	AB AB	δ	No	No	Yes*	Yes	Yes*	Yes	8	2
11	AB BA	δ	Yes*	Yes	No	No	Yes*	Yes	9	2
12	AA AA	δ	Yes*	Yes	Yes*	Yes	Yes*	Yes	12 or 6*	1

Notes to Table 1

(*) With ordinal matrix entries subrule β is inconclusive.

One point is added to the measure of incentive to switch when β selects a different initial strategy than γ .

✡ The starred incentives and tactics exist in both rows of the matrix.

* The incentives exist, technically, but they are not really attractive because nothing is to be gained by switching (nothing is lost either). The measure of incentive to switch is 12 if the three asterisked yeses are counted, 6 if they are not.

of subjects playing against each other, and in section 5 the elements that arise when the POM's are considered as the output of a group of players of the same game.

2.1 A Motivation Conflict

A player of a game of this sort forms a picture of his opponent's matrix not as an end in itself but as a means to assist his strategic reasoning ("If I do this, he'll probably do that," etc.).

However, occasionally subjects were tempted to pursue identification of the opponent's matrix as a final rather than as an intermediate goal. They would use nonoptimal strategies (nonoptimal for the purpose of maximizing their own scores) in the hope of deducing from the opponent's reaction the appearance of his matrix. Not having formulated clearly the problem of how to identify which moves by an opponent provide useful information about his matrix and which reflect his behavioral idiosyncrosies, these subjects often ended up with poorer guesses than their more conventional fellows--in Game 1 of our six games, for example, as we shall see. On the other hand, in some games--Game 1 not being one of them--rational play, i.e., play aimed at long-run score maximization, probably requires a certain number of "exploratory" moves at first in order to determine the configuration of the opponents' payoffs.

2.2 Measuring How Well The Opponent's Matrix is Perceived

2.2.1 Definition of τ

In order to answer the questions we shall be asking, we need to be able to measure the degree of agreement between two ordinal payoff matrices; for the purpose we use Kendall's Rank Correlation Coefficient, τ . The method of calculating and using this statistic in testing hypotheses has been described by Siegel,[5] so that here we need only conceptualize the formula for computing τ in a way that applies to our purposes:

$$\tau = \frac{c - w}{n}$$

where

c = the number of pairs of payoffs in the POM that are ranked in the correct order as compared with the AOM

w = the number of pairs of payoffs in the POM that are ranked in the wrong order as compared with the AOM

n = the number of possible comparisons (n = 6 when neither the AOM nor the POM contains tied rankings; Siegel explains how n is to allow for ties.[6])

A few examples are given in Figure 4 using the ordinal version of the AOM from Game 5 (see Figure 2 for the cardinal matrix actually used)

W	X
Y	Z

Generic

B	D
A	C

AOM

C	A
B	D

POM_1 $\tau = 0$

B	D
A	C

POM_2 $\tau = 1$

C	A
D	B

POM_3 $\tau = -1$

D	B
A	C

POM_4 $\tau = 0$

C	D
A	B

POM_5 $\tau = .667$

B	C
A	B

POM_6 $\tau = .913$

Figure 4

Using the "generic" matrix let us refer to the payoff in the upper left corner of any matrix as W, the payoff in the upper right corner as X, etc. Thus in the AOM the six payoff pairs rank as follows:

$$W > X , \quad W < Y , \quad W > Z , \quad X < Y , \quad X < Z , \quad Y > Z$$

POM_1 has the payoffs ranked:

$$W < X , \quad W < Y , \quad W > Z , \quad X > Y , \quad X > Z , \quad Y > Z$$

It is easily seen that three of these rankings are correct and three are wrong. Thus $c = 3$, $w = 3$, $n = 6$, and $\tau = \frac{3-3}{6} = 0$.

POM_2 is identical with the AOM and is thus perfectly correct; hence

$$c = 6 , \; w = 0 , \text{ and } \tau = \frac{6-0}{6} = 1 .$$

POM_3 is perfectly wrong --the rankings are reversed--so that

$$c = 0 , \; w = 6 , \text{ and } \tau = \frac{0-6}{6} = -1 .$$

POM_4, though different from POM_1, also receives $\tau = 0$ because it too has three rankings right and 3 wrong. POM_5 is nearly correct-- only one ranking is reversed $(W < Z)$, so that $c = 5$, $w = 1$, and $\tau = .667$. POM_6 is identical to POM_5 except that $w = Z$. This is counted as neither correct nor incorrect, so that $c = 5$ and $w = 0$. In this case n is adjusted to allow for a tied ranking;

Thus $$\tau = \frac{5-0}{\sqrt{6}\sqrt{5}} = .913 .$$

2.2.2 Is τ a Satisfactory Measure?

Defining a measure does not guarantee that it will do its job well. Let us see how it works in practice. Table 2 presents data for all the POM's in Game 1. The AOM in standard aspect was

B	B
A	A

for both players. The POM's are arranged in order of decreasing τ, as can be seen from Column (4).

Table 2 - Perceived Opponent's Matrices in Game 1

POM Number	POM in Standard Aspect	Taxonomic Type of POM	τ of POM	Number of Subjects Guessing This POM	Cumulated Percent of Subjects Guessing This and Preceding POM's	Number of Subjects Guessing This POM: Whose Opponents Used Strategy 1 At Least Once During Play	Whose Opponents Used Strategy 1 Initially	Who Themselves Used Strategy 1 At Least Once During Play	Who Themselves Used Strategy 1 Initially
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	BB AA	1b-2x	1.00	36	54.5	1	0	5	1
2	BC AA	1a-21	.89	1	56.1	0	0	1	1
3	CB AA	1a-22	.89	1	57.6	0	0	1	1
4	CC BA	2c-22	.89	1	59.1	0	0	0	0
5	CD AB	2a-21	.82	3	63.6	2	0	3	2
6	DC AB	2b-21	.82	1	65.2	1	1	0	0
7	CB BA	2d-22	.67	2	68.2	0	0	0	0
8	BB BA	3-22	.58	1	69.7	0	0	0	0
9	DB AC	6a-21	.41	1	71.2	1	0	0	0
10	BC CA	8-22	.22	1	72.7	1	0	0	0
11	BA BA	10-x2	0	6	81.8	1	0	1	0
12	CB DA	9a-12	0	1	83.3	1	1	0	0
13	AB AB	10-x1	0	1	84.8	1	0	1	0
14	AD BC	9a-?1	0	1	86.4	1	0	0	0
15	AC CB	8-11	-.22	1	87.9	1	1	0	0
16	BA CA	46-12	-.22	1	89.4	0	0	1	0
17	AA AB	1c-11	-.50	1	90.9	1	0	0	0
18	BA DC	2a-12	-.82	1	92.4	0	0	0	0
19	AB CD	2a-11	-.82	1	93.9	1	0	0	0
20	-B -A	incom-plete	unde-fined	1	95.5	1	0	0	0
21	-A -B	incom-plete	unde-fined	1	97.0	1	1	1	0
22	A- A-	incom-plete	unde-fined	1	98.5	1	1	1	0
23	-- --	incom-plete	unde-fined	1	100.0	0	0	1	0
Total	—	—	—	66	—	16	5	16	5

Comparison of the taxonomic types of the POM's in Column (3) with their τ 's reveals that τ is a good measure of the degree of agreement between POM's and the AOM, at least in Game 1. All the POM's with $\tau > .5$ are matrices in which the initial strategy chosen by rule $\alpha\beta\lambda\delta$ is strategy 2 and it is chosen by subrule α (that is, row 2 dominates). Intuitively we regard all such matrices as at least fairly good approximations of the AOM. The POM's with $0 < \tau < .5$ have row 2 chosen by subrule β , those with $\tau = 0$ are inconclusive about which row is chosen at first; those with negative τ's have row 1 chosen at first by rule $\alpha\beta\lambda\delta$, and the very lowest τ's are awarded to POM's with row 1 the dominant strategy--which is exactly what we mean in this game by a "bad guess".

Inspection of the POM's , their taxonomic types and their τ's for the other games reveals a similar correspondence between intuition and our measure, but in Game 2 the correspondence is perhaps less clear. In this game Player 1's AOM in standard aspect is $\begin{array}{|c|c|}\hline A & B \\ \hline A & B \\ \hline\end{array}$. If he submits $\begin{array}{|c|c|}\hline B & A \\ \hline B & A \\ \hline\end{array}$ as his P O M , it gets a τ of -1 , the lowest possible, because he has reversed the entries. Yet out of the 75 possible matrices he has picked one of the 5 that require the Randomization Subrule, δ , to choose an initial strategy. On the other hand, with this POM , Player 1 will believe Player 2 would like him to play Strategy 2, whereas in fact exactly the opposite is true. One might use the absolute value of τ rather than the signed value as a measure in this game--but this too has flaws when applied to particular matrices. Though it is not perfect we will stick with our present measure.

3. Factors Affecting The Individual's POM

3.1 Are POM's Chosen Randomly?

The fact that one requires of subjects that they submit a POM does not prove that they have given the matter any thought. Our first task will be to test the null hypothesis that the POM's submitted are simply chosen at random from the 75 possible distinct ordinal matrix forms.

Table 3 classifies the POM's into five groups according to what subrule is needed to pick an initial strategy and whether the strategy so picked is right or wrong. The observed frequencies of POM's in these five classes and the frequencies expected under H_0 are shown for each game. The group into which the AOM itself falls is indicated by a circle around the observed frequency; and always the encircled observed frequency exceeds the corresponding expected frequency. With four exceptions (all underlined), all other observed frequencies are less than expected. This suggests that POM's are chosen in an organized way rather than at random and tend toward being approximately correct. But we shall apply a more impartial test.

Table 4 shows the results of testing H_0 by a test.[7] H_0 is clearly rejected in Games 1, 2, 3, 4 and 6 but not in Game 5. The Prisoner's Dilemma evidently is very difficult to apprehend correctly when the opponent's payoff is not known.

Table 3 - Observed and Expected Frequencies of POM's of Various Types in Six Games

Game	Data Specification	P O M Matrix Class: Correct Initial Strategy Chosen by Subrule α	Correct Initial Strategy Chosen by Subrules β or λ	Initial Strategy Equally Likely to be Correct or Incorrect (Subrule δ)	Incorrect Initial Strategy Chosen by Subrules β or λ	Incorrect Initial Strategy Chosen by Subrule α	Total Number of Different POM's	Incomplete POM
1	Observed Frequency	46	3	7	2	4	62	4
	Expected Frequency	19.0	9.9	4.1	9.9	19.0	62	
2*	Observed Frequency	28	15	21			64	0
	Expected Frequency	38.2	20.5	8.3			64	
3	Observed Frequency	11	28	4	16	5	64	2
	Expected Frequency	19.6	10.3	4.3	10.3	19.6	64	
4	Observed Frequency	20	23	1	10	7	61	5
	Expected Frequency	18.7	9.8	4.1	9.8	18.7	61	
5	Observed Frequency	24	8	2	15	15	64	2
	Expected Frequency	19.6	10.3	4.3	10.3	19.6	64	
6	Observed Frequency	48	3	2	4	3	60	4
	Expected Frequency	18.4	9.6	4.0	9.6	18.4	60	

* In Game 2 the first move is chosen randomly (by Subrule δ); hence the POM's with initial strategies chosen by subrules α , β and λ cannot be separated into "correct" and "incorrect."

Table 4 - χ^2 Test of No 11 Hypothesis that POM's are Chosen Randomly

Game	χ^2	Degrees of Freedom	Probability That H_0 is True
1	63.1	4	< .0000001
2	23.6	2	< .0001
3	46.6	4	< .0000001
4	27.5	4	< .00001
5	6.9	4	> .15
6	68.9	4	< .0000001

3.2 Relationship between Accuracy of POM's and Irrational Strategy Choices by Opponent and by Self

What factors are likely to influence the quality of agreement between POM and AOM? We think it likely in Game 1 that poor guesses will result when the opponent uses his first strategy. The "first order of inference" rationale in "solving simultaneously" the information problem posed by an astute subject is that it is natural for a player in guessing his opponent's matrix to assume his opponent behaves rationally and attempts to maximize his score. (This sets aside for the moment the possibility of immediate deliberate false signalling.) While in some games this assumption does not by itself imply a unique pattern of moves, it does in Game 1. Therefore we propose these two hypotheses for Game 1:

1. τ will be lower if the opponent has used strategy one at any time during the course of play.

2. τ will be lower if the opponent has used strategy one initially.

The first hypothesis is the stronger, since it implies the second. We also may test hypotheses that one's own irrational behavior will be associated with poor guessing--but we have no reason to expect them to be true.

3. τ will be lower if the subject himself has used strategy one at any time during the course of play.

4. τ will be lower if the subject himself has used strategy one initially.

As before, Hypothesis 3 implies Hypothesis 4.

Columns (7) - (10) of Table 2 display the raw data, and Table 6 below organizes the results of the four tests. The information contained therein can be used to generate a 2x2 contingency table for each hypothesis.

Thus for hypothesis 1:

Table 5: Contingency Table for Hypothesis 1

	Number of Subjects Whose Opponent Used Strategy 1 At Least Once During Play	Number of Subjects Whose Opponent Used Only Strategy 2	Totals
Number of Subjects Whose POM had $\tau > .5$	4 [11.2]	42 [34.8]	46
Number of Subjects Whose POM had $\tau < .5$	12 [4.8]	8 [15.2]	20
Totals	16	50	66

If there is no association between opponent's use of strategy 1 and the subject's achievement of a POM with $\tau > .5$, we expect to observe the four categories with the frequencies shown in the triangles in each cell. The actual values from samples, if H_0 is true, will be distributed about these means. To test H_0, χ_2 is calculated for each of the four hypotheses by Siegel's formula allowing for continuity of the underlying variable (which is reasonable in the case of degree of agreement).[8/] The value of ^{2} and the associated probabilities that the null hypotheses are true are given in Table 6.

Table 6 - Factors Affecting Agreement of POM and AOM in Game 1

Level of τ	Total Number of Subjects	Number of Subjects Whose Opponent Used Strategy 1: At Least Once During Play	Whose Opponent Used Strategy 1: Initially	Who Themselves Used Strategy 1: At Least Once During Play	Who Themselves Used Strategy 1: Initially
Larger than .5	46	4	1	10	5
Less than .5	20	12	4	6	0
Total	66	16	5	16	5
χ^2		14.56	3.92	.14	.16
p		$<.0001$	$<.025$	$>.3$	$>.3$

These figures confirm our expectations: opponent's irrational behavior does reduce the quality of POM's, but one's own irrational behavior does not, in Game 1.

3.3 Accuracy of POM's as a function of Game Structure

Next we wish to investigate which games produce, on the average well-perceived opponent's matrices and which produce poorly-perceived ones. Tables 7 and 8 and Figure 3 present the data.

Table 7 shows for a given percent of subjects the level of τ achieved. Example: in Game 3, 50% of the subjects achieved $\tau = .40$ or better, 75% achieved $\tau = 0$ or better. The 50% line gives the median τ.

Table 8 shows the percent of subjects achieving a given τ or better. Example: in Game 2, 37.8% of the subjects achieved $\tau = .18$ or better.

The data from this table are graphed in Figure 9. In this figure a game in which the subjects do no better than might be expected from random choice of POM would be represented by a curve close to the diagonal--as is the case with Game 5. The farther above the diagonal a game's curve bulges, the more that game encourages good perception of the opponent's matrix. (These curves are like the Lorenz curves used to measure income inequality.)

From these tables and the figure emerges a pattern resembling that found when we examined the quality of prediction of the standard game-theoretic solutions in Part I.[9] The best POM's are submitted in games in which the Joint Maximum and Non-Cooperative Solutions coincide and are unique, and both players have dominant strategies (Games 1 and 6). Next best is Game 4, in which there is a unique JM-NCE outcome but neither player has a dominant strategy. Next is Game 3, in which the JM and NCE Solutions coincide but include two outcomes. And the poorest POM's come from the two games (2 and 5) in which the JM and NCE Solutions diverge.

Table 7: Quartiles Achieving τ at least as Large as Shown

Percent of Subjects	Game					
	1	2	3	4	5	6
25	1.00	.67	.91	1.00	.33	.91
50	1.00	0	.40	.60	0	.67
75	.41	-.22	0	0	-.55	.41
100	-.82	-1.00	-1.00	-1.00	-1.00	-1.00

Table 8 - Percent of Subjects Submitting POM's With τ At Least As Large As Given

Game

τ	1 N=62	2 N=63	3 N=64	4 N=61	5 N=63	6 N=60
1.00	58.1	13.1	18.8	27.9	11.1	15.0
.91	63.9	13.1	25.0	31.2	12.7	26.7
.89	63.9	13.1	28.1	31.2	12.7	26.7
.82	69.3	22.9	28.1	31.2	12.7	33.3
.80	69.3	22.9	31.3	36.1	12.7	33.3
.78	69.3	22.9	32.8	49.2	12.7	33.3
.71	69.3	22.9	32.8	49.2	12.7	41.7
.67	72.6	26.1	32.8	49.2	15.9	65.0
.60	72.6	26.1	43.8	54.2	15.9	65.0
.58	74.2	26.1	43.8	54.2	15.9	65.0
.55	74.2	26.1	48.4	64.0	19.0	68.3
.41	75.8	32.8	48.4	64.[illegible]	20.6	80.0
.40	75.8	32.8	51.6	67.[illegible]	20.6	80.0
.33	75.8	32.8	51.6	67.[illegible]	34.9	88.3
.22	77.4	37.8	56.3	70.5	34.9	88.3
.20	77.4	37.8	60.9	70.5	34.9	88.3
.18	77.4	37.8	70.3	80.3	44.4	91.7
0	91.9	73.8	76.6	85.3	60.4	91.7
-.18	91.9	73.8	79.6	90.2	63.5	93.3
-.20	91.9	73.8	79.6	91.8	63.5	93.3
-.22	95.2	75.3	79.6	91.8	63.5	93.3
-.26	95.2	75.3	79.6	93.4	63.5	93.3
-.33	95.2	75.3	79.6	93.4	71.5	95.0
-.40	95.2	75.3	81.3	95.1	71.5	95.0
-.41	95.2	83.5	81.3	95.1	74.6	95.0
-.50	96.8	83.5	81.3	95.1	74.6	95.0
-.55	96.8	83.5	90.6	95.1	82.6	96.7
-.67	96.8	83.5	90.6	95.1	96.8	98.3
-.80	96.8	83.5	93.8	95.1	96.8	98.3
-.82	100.00	88.6	93.8	95.1	96.8	98.3
-.89	100.00	88.6	95.3	96.7	96.8	98.3
-.91	100.00	88.6	98.4	98.3	96.8	98.3
-1.00	100.00	100.0	100.0	100.0	100.0	100.0

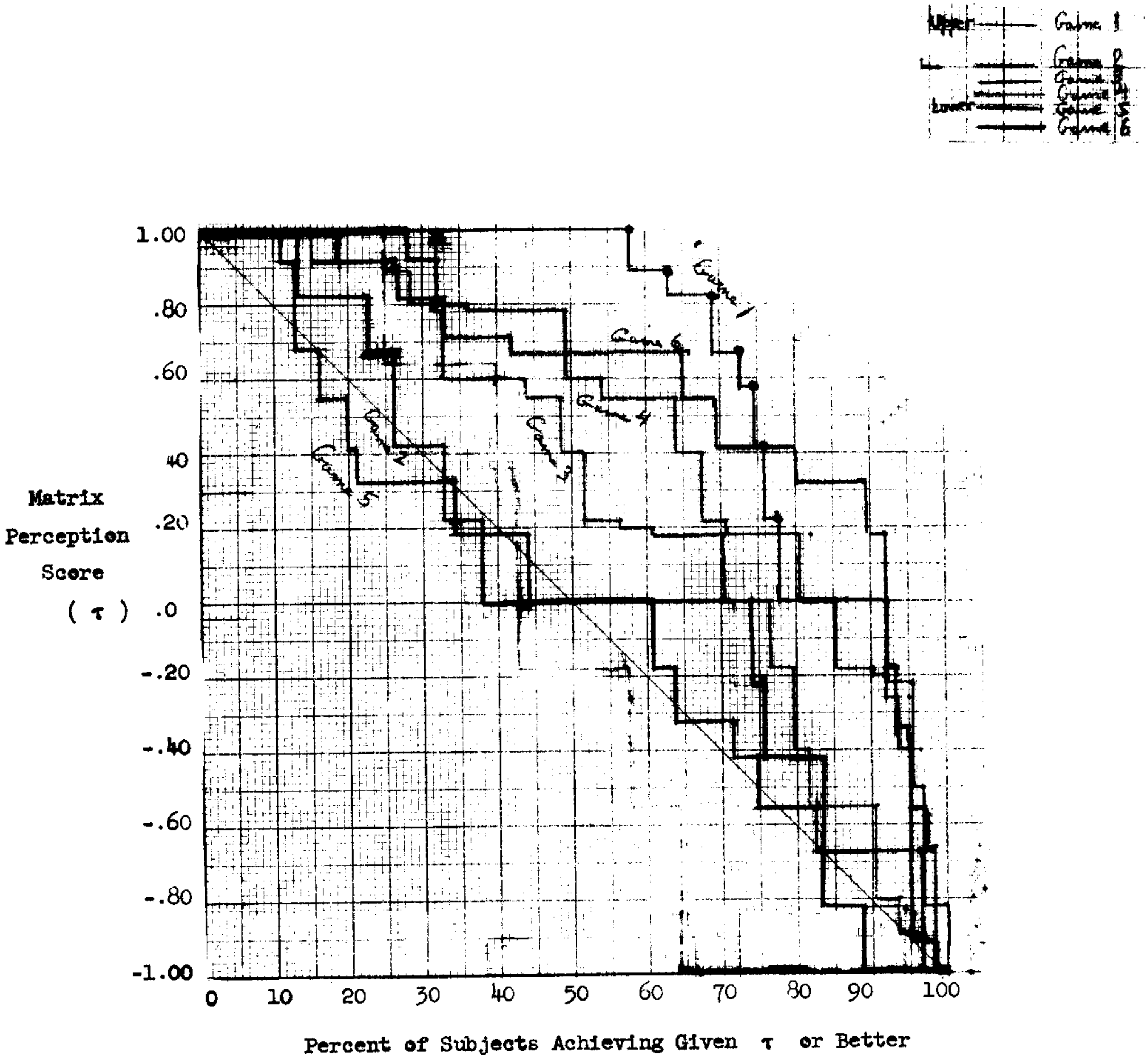
Upper — Game 1
Game 2
Game 3
Game 4
Lower Game 5
Game 6
1.00
.80
.60
40
.20
.0
-.20
-.40
-.60
-.80
-1.00
Matrix Perception Score (τ)
Game 1
Game 6
Game 4
Game 2
Game 3
Game 5
0
10
20
30
40
50
60
70
80
90
100
Percent of Subjects Achieving Given τ or Better

Figure 3

4. Player Pairs and the POM's

4.1 Do Paired Players Influence Each Other's POM's?

Another approach is to sidestep identifying the factors that affect τ and ask merely whether whatever affects one member of a playing pair will affect the other member also. In other words we hypothesize that if Player 1's τ is high (low), his opponent's τ is likely to be higher (lower) than random selection from the population of Player 2 τ's would produce. The test of this hypothesis requires us to rank the Player 1 τ's and to associate with each Player 1 τ the τ of that player's opponent. This has been done for Game 4 in Table 9. Data for 28 pairs are shown; in the remaining 5 pairs at least one POM was incomplete and could not be assigned a τ.

Table 9 τ's and Their Rankings for Paired Players in Game 4

Player 1's τ (1)	Player 2's τ (2)	Rank of Player 1's τ (3)	Rank of Player 2's τ (4)
1.00	1.00	1	1
1.00	1.00	1	1
1.00	1.00	1	1
1.00	.55	1	6
.91	.60	2	5
.80	.78	3	4
.78	1.00	4	1
.78	1.00	4	1
.78	.91	4	2
.78	.78	4	4
.60	-.18	5	11
.55	.18	6	9
.55	-.40	6	14
.40	1.00	7	1
.40	.80	7	3
.22	1.00	8	1
.22	.55	8	6
.18	1.00	9	1
.18	1.00	9	1
.18	.55	9	6
0	1.00	10	1
0	1.00	10	1
0	.60	10	5

Table 9 - τ's and Their Rankings for Paired Players in Game 4 (Continued)

Players 1's τ (1)	Player 2's τ (2)	Rank of Player 1's τ (3)	Rank of Player 2's τ (4)
-.18	1.00	11	1
-.18	.18	11	9
-.20	.80	12	3
-.26	.55	13	6
-.89	-.91	15	16

To test whether the order of the Player 2 τ's is sufficiently like the order of the Player 1 τ's so that we can reject H_0, we calculate the Kendall Rank Correlation Coefficient for the pair of rankings and determine whether it is sufficiently different from, and in this case larger than, zero.[10/]

Table 10 gives the results of this test for all six games. Continuing to use Game 4 as our example, this table tells us that the Kendall Rank Correlation Coefficient K for the data of Table 9 is .094.

Table 10 - Test of Hypothesis That Paired Subjects Produce POM's With Similar τ's

	Kendall Rank Correlation Coefficient K	z	One-tailed test: Probability That H_0 is true
1	+ .220	1.40	$\approx$ .08
2	+ .174	.89	$\approx$ .19
3	+ .048	.37	$\approx$ .36
4	+ .094	.68	$\approx$ .25
5	+ .077	.55	$\approx$ .29
6	+ .110	.79	$\approx$ .21

K can be mapped into the familiar z of the normal distribution: $z = \frac{K-\mu}{\sigma}$, where $\mu = 0$ = the expected mean of K under the null hypothesis, and σ under H_0 depends on the number of items ranked and the number of ties at each level of each ranking. For Game 4, $z = .68$, which means that we could expect $K \geq 094$ could occur under H_0 about 25% of the time. Therefore we cannot reject H_0 at any reasonably high significance level; we must conclude that there is no significant relationship between one player's τ and that of his opponent. And the same conclusion must be reached for the other five games as well, if we use the 5% significance level for rejection of H_0.

On the other hand, in every one of the games the correlation between the two rankings is positive. Such association as there is between the rankings is in the expected direction.

4.2 Correlation Between Paired Players' POM's and Degree of Involvement

We may also ask whether the magnitudes of the correlation that does exist between Player 1 τ's and Player 2 τ's is positively correlated with the degree of involvement fostered by a game. The degree of involvement, discussed at length in Part II,[11] measures the extent to which strategy changes by one player induce changes by the other. The rationale for there being such an association is that in games with a high degree of involvement a player might experience greater difficulty in separating inferences about the opponent's matrix from inferences about the latter's reactions to the player's own moves.

Table 11-Comparison of Games Ranked by Degree of (Strategic) Involvement and Degree of Information Involvement

Game	Ranked by Degree of Information Involvement	Ranked by Degree of (Strategic) Involvement
1	1	5
2	2	2
3	6	1
4	4	4
5	5	3
6	3	6

Table 11 compares the ranking of the games according to degree of involvement with their ranking according to K,[12] where K can be considered a measure of degree of information involvement. We can test the hypothesis that these two rankings are related by calculating once again the Kendall Rank Correlation Coefficient; it turns out to be -.467, which actually has the wrong sign! Degree of Information Involvement is inversely related to Degree of (Strategic) Involvement--though not significantly so (an unlated pair of rankings of six objects will produce a correlation coefficient as different from zero as $\pm .467$ with probability .272).[13]

We find it surprising that the conjectures tendered in this section are not borne out by the data. Perhaps a larger sample would confirm them; possibly they would be more confirmable in 2x2 games other than the six used by us. In any case, we cannot give any reasonable explanation for the negative correlation observed.

4.3 Consequences for Play of Poor Perception of Opponent's Matrix: An Example

There is an old saying to the effect that when A and B meet there are actually four people present: A as A really is B as B is, A as B thinks he is, and B as A thinks he is. The folk wisdom points out that if the perceived others fail to correspond to the actual others, one can expect compounded confusion.

Can "compounded confusion" be expected if POM's and AOM's diverge widely? An extreme instance from Game 4 (Table 9, last line) suggests that it can. Player 1's POM had a τ of -.89; Player 2's was -.91. Thus each player lived and reacted in a world of his own far from "reality", by which we mean the AOM's. Figure 4 illustrates the situation. We call the game player 1 thinks he is playing Game $4I_1$ ("I" for "imaginary"); Game $4I_2$ is defined analogously.

Game 4 itself has already been analyzed in Part II[14] -- outcome (1, 1) is expected as a steady state; but if one player uses his second strategy, the other can maximize by switching to his second strategy too, so that (2,2) is an equilibrium point; however each has an investment incentive to switch back to his first strategy; and if one does, the other maximizes by doing likewise. Thus the structure of Game 4 encourages stability.

But Game $4I_1$ is altogether different. In this game we see a <u>maximizers' cycle</u>: starting, say, from outcome (1,1) player 2 (that is, the imaginary player of Player 1's POM) can maximize by switching to strategy 2, in which case Player 1 maximizes by switching to strategy 2, after which Player 2 maximizes by returning to strategy 1, following which Player 1 does the same ... and so on.

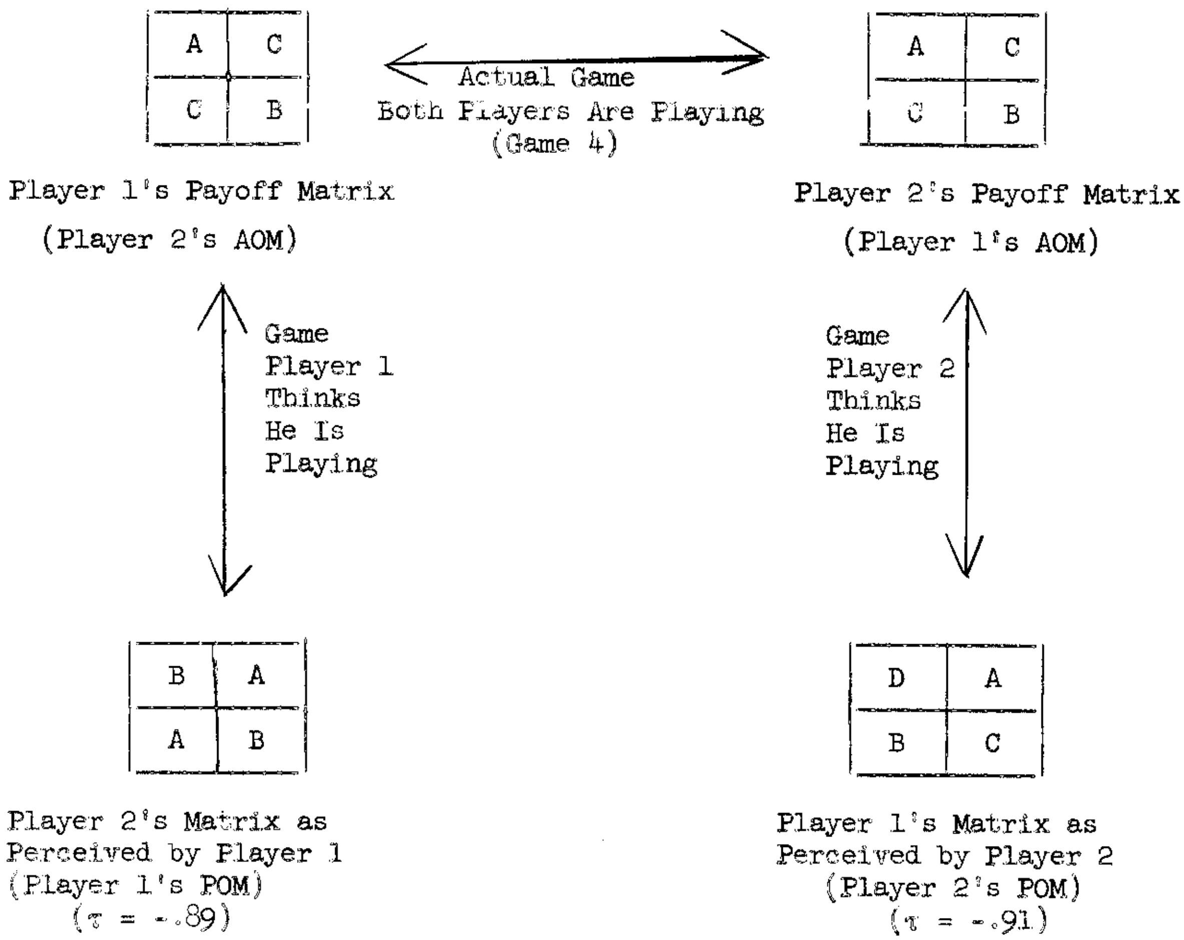

Figure 4

The cycle of outcomes, which we label MM, is

$$(1, 1) \rightarrow (1, 2) \rightarrow (2, 2) \rightarrow (2,1) \rightarrow (1, 1) \ldots .$$

If Player 1 understands that the game he is playing sets up a maximizers' cycle he may try to beat it by anticipating his opponent's move. Instead of suffering the low payoff from outcome (1, 2) that results when his opponent maximizes with respect to Player 1's first strategy, he will shift simultaneously. The Player-1-anticipates Player-2- maximizes cycle (AM) is:

$$(1, 1) \rightarrow (2, 2) \rightarrow (1, 1) \ldots .$$

Conversely, the MA cycle is

$$(1, 2) \rightarrow (2, 1) \rightarrow (2, 2) \ldots .$$

And if each player anticipates the other's efforts to maximize we still get a determinate cycle (AA) which is the reverse of the MM cycle:

$$(1, 1) \rightarrow (2, 1) \rightarrow (2, 2) \rightarrow (1, 2) \rightarrow (1, 1) \ldots .$$

For an outcome to be repeated in Game $4I_1$, one of the players must be an anticipator, and the other must reason (using outcome (1, 1) for the example),

"My opponent could maximize with respect to my first strategy by switching to his second. But he's too smart to do that: he's caught on. He knows I will switch to my second strategy in anticipation of his switching to his, and therefore he will stick with his first. Therefore I can stay a step ahead only if I stick with my first strategy, too." This player we call a "double anticipator" (D). Either DA or AD can account for the repeated observation of an outcome.

It is of interest that in Game $4I_1$ it is <u>not</u> possible to explain the repeated observation of (1, 1) as simply the result of both players' being satisfied--which the players are in the game actually being played! Truly Player 1 is living in a fantasy world, the complexities of which totally block out his ability to enjoy the simplicity of reality!

The structure of Game $4I_2$ is a mirror of Game $4I_1$'s. Here "Player 1" is the imaginary player of Player 2's IOM. The MM cycle of Game $4I_2$ is the same as the AA cycle of Game $4I_1$:

$$(1, 1) \rightarrow (2, 1) \rightarrow (2, 2) \rightarrow (1, 2) \rightarrow (1,1) \ldots .$$

Game $4I_2$'s AM cycle is Game $4I_1$'s MA cycle and vice versa. Game $4I_2$'s AA cycle is Game $4I_1$'s MM cycle. A repeated outcome explained by DA in Game $4I_2$ would be explained by AD in Game $4I_1$, and vice versa.

If we adduce the structure of Game 4 itself to explain the observed strategy switching in terms of responses in period t to opponent's move in period t-1 we get the following:

Opponent's move, Period t-1	Subject's move, Period t	Explanation
1	1	Subject is maximizing (M)
1	2	Subject is irrational (X)
2	1	Subject is investing (I)
2	2	Subject is maximizing (M)

That is, the structure of Game 4 can explain any sequence except opponent's strategy 1 followed by subject's strategy 2; this sequence is simply irrational--which is what we mean when we say outcome (1,1) is expected as a steady state.

Table 12 offers three alternative "explanations" of the sequence of outcomes observed in the 19 periods of play of Game 4 by these two subjects.

In this context Game $4I_1$ can be interpreted as the conceptual framework Player 1 uses to rationalize his own and his opponent's behavior: it is his Weltanschauung. Game $4I_2$ performs the same function for Player 2. Significantly the structures of these games are such that every strategy choice, both of one's own and of one's opponent--except one's own initial strategy, can be "rationally explained".

Table 12- "Explanations" of Play of Game 4 by A Pair of Subjects With Poor POM's

Period	Player 1's Strategy	Player 2's Strategy	Explanation of Outcome in Terms of The Structure of		
			Game 4	Game $4I_1$	Game $4I_2$
1	2	2	XX	Xδ	γX
2	2	2	MM	AD or DA	DA or AD
3	1	1	II	AM	MA
4	2	1	XM	AA	MM
5	1	2	MM	MA	AM
6	2	2	MX	MM	AA
7	2	2	MM	AD or DA	DA or AD
8	1	2	IM	AA	MM
9	1	2	IX	AD or DA	DA or AD
10	2	2	MX	MM	AA
11	2	1	MI	MM	AA
12	1	2	MM	MA	AM
13	1	1	IM	AA	MM
14	2	2	XX	AM	MA
15	2	2	MM	AD or DA	DA or AD
16	1	1	II	AM	MA
17	2	2	XX	AM	MA
18	1	1	II	AM	MA
19	1	2	MX	MM	AA

Table 12 Notes

Note: In the explanation columns the first letter explains the strategy chosen by Player 1 in period t as a consequence of Player 2's strategy choice in period $t - 1$; the second letter does the same for Player 2's period t strategy with respect to Player 1's period $t - 1$ strategy. The initial strategies are explained in terms of rule $\alpha\beta\lambda\delta$. The symbols used have these meanings:

- M Maximizes
- I Invests
- A Anticipates opponent's maximizing
- D Anticipates opponent's anticipating own maximizing
- X Behaves irrationally (choice cannot be explained in terms of the structure of this game)
- γ Chooses initial strategy by the Security Maximization Subrule, γ .
- δ Chooses initial strategy by the Randomization Subrule, δ .

This conforms to the belief, subscribed to by the authors, that people in general have a need to be able to explain both their own behavior and that of others, no matter how foolish or irresponsible it may "really" be, as reasonable in the face of circumstances; and they will develop models of the world which serve no other purpose.

It is equally significant that the structure of Game 4 is not equipped to perform this function for these players, for it dismisses a number of moves as simply "irrational." Mere importantly, the structure of Game 4 is also insufficient to explain these moves to an impartial observer or game theorist. A scientist possessing the structure of Game 4 may pride himself on "seeing things as they really are;" yet his clear vision avails him nought in explaining the outcomes observed in periods 4, 6, 9, 10, 14, 17 and 19. A truly comprehensive explanation would have to include both the real world of Game 4 and the fantasy worlds of Games $4I_1$ and $4I_2$ in some still broader framework of thought.

5. Are Many Heads Better Than One? - Analysis of POM's as the Output of a Group

Still another approach to the analysis of the POM's considers the players neither as individuals nor as pairs but as a group. Do the players independently submitting POM's develop a consensus?

Table 13 presents the raw data.

We start with the frequencies with which A, B, C and D are assigned to each outcome by all the Player 1's and by all the Player 2's of a game. We assign these ranks the following numerical values : A = 4, B = 3, C = 2, D = 1 . Then the average

Table 13 -Raw Data for Analysis of Group's Perception of Opponent's Matrix

Game	Data Specification	Player 1's POM Outcome (1,1)	(1,2)	(2,1)	(2,2)	Total or Mean	Player 2's POM Outcome (1,1)	(1,2)	(2,1)	(2,2)	Total or Mean
1	Frequency of A(=4)	4	23	6	26	59	2	4	23	26	55
	" " B(=3)	23	6	23	4	56	24	22	5	4	55
	" " C(=2)	2	1	2	2	7	6	3	3	1	13
	" " D(=1)	2	1	1	0	4	0	3	1	1	5
	Total observations	31	31	32	32	126	32	32	32	32	128
	Frequency of blanks	2	2	1	1	6	1	1	1	1	4
	Average of POM Outcome ranks	2.935	3.645	3.063	3.750	3.349	2.875	2.844	3.656	3.719	3.250
	Consensus POM	D	B	C	A		C	D	B	A	
	AOM	B	A	B	A		B	B	A	A	
2	Frequency of A(=4)	19	16	11	12	58	10	11	17	16	54
	" " B(=3)	8	9	12	16	45	15	15	7	13	50
	" " C(=2)	3	4	7	3	17	5	5	5	0	15
	" " D(=1)	3	4	3	2	12	2	1	3	3	9
	Total observations	33	33	33	33	132	32	32	32	32	128
	Frequency of blanks	0	0	0	0	0	1	1	1	1	4
	Average of POM Outcome ranks	3.303	3.121	2.939	3.152	3.129	3.031	3.125	3.183	3.313	3.164
	Consensus POM	A	C	D	B		D	C	B	A	
	AOM	A	A	B	B		A	B	A	B	
3	Frequency of A(=4)	12	10	2	17	41	12	8	8	8	36
	" " B(=3)	12	10	10	9	41	12	5	5	14	36
	" " C(=2)	7	12	14	4	37	6	16	17	6	45
	" " D(=1)	2	1	5	2	10	3	4	3	5	15
	Total observations	33	33	31	32	129	33	33	33	33	132
	Frequency of blanks	0	0	2	1	3	0	0	0	0	0
	Average of POM Outcome ranks	3.030	2.879	2.290	3.281	2.876	3.000	2.515	2.545	2.758	2.705
	Consensus POM	B	C	D	A		A	D	C	B	
	AOM	A	C	C	B		B	C	C	A	

Table 13 (Continued)

Game	Data Specification	Player 1's POM Outcome (1,1)	(1,2)	(2,1)	(2,2)	Total or Mean	Player 2's POM Outcome (1,1)	(1,2)	(2,1)	(2,2)	Total or Mean
4	Frequency of A(=4)	23	5	6	5	39	25	2	2	6	35
	" " B(=3)	7	13	10	15	45	5	6	9	20	40
	" " C(=2)	1	13	10	6	30	2	22	16	3	43
	" " D(=1)	1	0	4	4	9	1	1	5	2	9
	Total observations	32	31	30	30	123	33	31	32	31	127
	Frequency of blanks	1	2	3	3	9	0	2	1	2	5
	Average of POM Outcome ranks	3.625	2.742	2.600	2.700	2.927	3.636	2.290	2.250	2.968	2.795
	Consensus POM	A	B	D	C		A	C	D	B	
	AOM	A	C	C	B		A	C	C	B	
5	Frequency of A(=4)	7	11	9	11	38	11	7	10	12	40
	" " B(=3)	10	9	7	10	36	13	10	3	10	36
	" " C(=2)	12	7	6	8	33	7	10	10	5	32
	" " D(=1)	3	5	10	3	21	2	5	10	5	32
	Total observations	32	32	32	32	128	33	32	33	32	130
	Frequency of blanks	1	1	1	1	4	0	1	0	1	2
	Average of POM Outcome ranks	2.656	2.813	2.469	2.906	2.711	3.000	2.594	2.394	2.906	2.723
	Consensus POM	C	B	D	A		A	C	D	B	
	AOM	B	A	D	C		B	D	A	C	
6	Frequency of A(=4)	31	3	4	0	38	24	9	3	4	40
	" " B(=3)	1	16	19	9	45	5	14	12	9	40
	" " C(=2)	0	11	6	9	26	1	6	13	6	26
	" " D(=1)	0	1	1	12	14	2	1	3	11	17
	Total observations	32	31	30	30	123	32	30	31	30	123
	Frequency of blanks	0	1	2	2	5	0	2	1	2	5
	Average of POM Outcome ranks	3.968	2.677	2.867	1.900	2.879	3.594	3.033	2.484	2.200	2.837
	Consensus POM	A	C	B	D		A	B	C	D	
	AOM	A	C	B	D		A	C	B	D	

Table 14 - Analysis of Variance for Consensus POM

Game (1)	Source of Variation (2)	Player 1's POM: Degrees of Freedom (3)	Player 1's POM: Sum of Squares (4)	Player 1's POM: Estimate of Variance (5)	Player 2's POM: Degrees of Freedom (6)	Player 2's POM: Sum of Squares (7)	Player 2's POM: Estimate of Variance (8)
1	Between Ranks	3	15.791	5.264	3	19.000	6.333
	Within Ranks	122	50.844	.416	124	61.000	.492
	Total	125	66.635		127	80.000	
2	Between Ranks	3	2.205	.735	3	1.336	.445
	Within Ranks	128	118.606	.928	124	100.219	.810
	Total	131	120.811		127	101.554	
3	Between Ranks	3	16.675	5.558	3	4.993	1.664
	Within Ranks	125	99.341	.793	128	124.484	.968
	Total	128	116.016		131	129.477	
4	Between Ranks	3	21.406	7.135	3	41.686	13.895
	Within Ranks	119	82.935	.696	123	66.991	.544
	Total	122	104.341		126	108.677	
5	Between Ranks	3	3.523	1.174	3	7.714	2.571
	Within Ranks	124	140.782	1.132	126	142.317	1.129
	Total	127	144.305		129	150.031	
6	Between Ranks	3	67.978	22.659	3	35.520	11.840
	Within Ranks	119	49.941	.419	119	95.228	.799
	Total	122	117.919		122	130.748	

Table 15 - Test of Hypothesis That Consensus POM is Significant

Game	Player	F-Ratio	Probability That H_0 is True	Rank of Game In Order of Degree to Which It Encourages Consensus In Perceiving Opponent's Matrix
1	1	12.68	$< .0001$	3
	2	12.89	$< .0001$	
2	1	.79	$> .05$	6
	2	.55	$> .05$	
3	1	7.01	$< .001$	4
	2	1.72	$> .05$	
4	1	10.27	$< .0001$	2
	2	25.51	$< .0001$	
5	1	1.04	$> .05$	5
	2	2.28	$> .05$	
6	1	28.13	$< .0001$	1
	2	14.81	$< .0001$	

Table 16 - Accuracy of Consensus POM

Game	τ for the Consensus POM of All The Player 1's	τ for the Consensus POM of All The Player 2's	Mean Consensus POM τ	Median τ of Individually Submitted POM's	Degree to Which Game Encourages Accuracy of Perception of Opponent's Matrix, As Ranked by Column (4)	Degree to Which Game Encourages Accuracy of Perception of Opponent's Matrix, As Ranked by Column (5)
(1)	(2)	(3)	(4)	(5)	(6)	(7)
1	.82	.82	.82	1.00	2	1
2	.41	-.41	0	0	6	5.5
3	.55	.55	.55	.40	4	4
4	.55	.91	.73	.60	3	3
5	.33	0	.17	0	5	5.5
6	1.00	.67	.83	.67	1	2

Table 17 - Variability of Guesses for Each Outcome of Each Game

Game	Player 1's POM: σ's for Outcome (1,1)	(1,2)	(2,1)	(2,2)	σ's for Whole Matrix	Player 2's POM: σ's for Outcome (1,1)	(1,2)	(2,1)	(2,2)	σ's for Whole Matrix	Average σ, Both Players, Whole Matrix	Rank of Game by Degree of Consensus Encouraged, as Measured by Average σ
1	.680	.709	.619	.519	.646	.572	.768	.806	.635	.702	.674	1
2	.984	1.053	.967	.834	.965	.861	.793	1.029	.896	.900	.932	4
3	.918	.893	.824	.924	.891	.968	1.003	.971	1.001	.984	.938	5
4	.707	.729	.969	.915	.834	.742	.643	.803	.752	.738	.786	3
5	.937	1.091	1.218	.996	1.063	.901	1.011	1.223	1.088	1.062	1.062	6
6	.180	.702	.681	.845	.648	.837	.809	.811	1.095	.895	.772	2

rank is calculated for each outcome. Example:

In game 1, of the 33 players, 2 failed to assign a rank to outcome (1,2); of the remaining 31, twenty-three assigned A , six B, one each C and D. The average of the ranks for outcome (1,2) is

$$\frac{1}{31} (23 \times 4 + 6 \times 3 + 1 \times 2 + 1 \times 1) = 3.645 .$$

The four averages are themselves ranked to give the "consensus POM" (CPOM) seen by the Player 1's in Game 1; that matrix is

D	B
C	A

By construction we get two CPOM's for each game, but in some games the consensus is more pronounced than in others. Speaking heuristically, if the averages of ranks are very close to each other (and to the mean of all four), the group has not reached a strong consensus about how the four outcomes of the opponent's matrix should be ranked. To determine which CPOM's are based on averages of ranks which do in fact summarize a group opinion we have performed an analysis of variance [15/] to determine whether the variation <u>between</u> the outcome ranks is significantly larger than the variation among the players estimating the outcomes ("<u>within</u> ranks"). Table 14 organizes the data for this analysis of variance and Table 15 presents the results. H_0 is that the consensus ranks do not differ from each other sufficiently to allow us to assert they are not random products. If H_0 is rejected we conclude the CPOM is not a random production but does result from some systematic process affecting to a greater or lesser degree all the members of the relevant group (all the Player 1's or all the Player 2's).

The F-ratio in Games 1, 4, and 6 are so large that we can reject H_0 at the .01% significance level. In games 2 and 5 we cannot reject

H_0 at the 5% significance level and must conclude that the CPOM's may be meaningless. In game 3 we have an anomalous result: for the Player 1's H_0 can be rejected at the .1% significance level but for the Player 2's it cannot be rejected even at the 5% level. Since the roles of the players in this game are symmetrical we are at a loss to explain the difference between their abilities to reach a consensus and must accept it as a statistical freak.

Column (5) of Table 15 makes use of the F-ratios as a measure of the degree to which a game encourages the players to arrive at a consensus in perceiving the opponent's matrix. The two F-ratios for each game are averaged; the average is ranked.

We have determined whether the CPOM's reflect any real consensus; now we analyze the CPOM's themselves. Table 16 gives the τ's for the 12 CPOM's. These numbers can be interpreted as measuring for the game as a whole the degree to which its structure encourages accuracy of perception of opponent's matrix by the group; to get a single number for each game the τ's for the two CPOM's are averaged in Column (4). An alternative measure of degree to which a game's structure encourages accuracy of perception is the median τ of the individual POM's; this is taken from Table 7 and presented in Column (5). The two measures are in close agreement in the way they rank the games, as can be seen from Columns (6) and (7), but the mean CPOM τ is higher in four out of six games and lower in only one. While not conclusive it calls to mind the thought that the group may have a "consensual intelligence" which is better (with respect to this problem) than the average individual intelligence--in short, many heads are [a little] better than [an average] one.

Finally in Table 17 are shown the differences in variability displayed by guesses for each outcome of each game. The figures are standard deviations from the average of POM outcome ranks presented in Table 13; thus for Game 1, Player 1, outcome (1,1),

$$\sigma = [4\ (4-2.935)^2 + 23(3-2.935)^2 + 2(2-2.935)^2 + 2(1-2.935)^2]^{\frac{1}{2}}\ .$$

The σ for the "whole matrix" is the square root of the within-ranks variance shown in Table 14.

We conjecture that the σ's for outcome (2,2) in Game 1 and outcome (1,1) in Games 4 and 6 will be lower than the σ's for the other outcomes in these games. The rationale is that the players can easily infer from the course of play that A should be assigned these outcomes, while they remain uncertain about the payoffs from other outcomes. A look at the table does not give our conjecture unequivocal support; it is confirmed for three of the six POM's (Game 1, Player 1, with σ (2,2) = .519 $< \sigma$ for the other three outcomes; Game 4, Player 1; and Game 6, Player 1). The Player 2's managed to arrive at a better consensus (lower σ) for at least one other outcome in each of these three games.

The highest σ was reported for both players at outcome (2,1) in Game 5. A glance back at Table 13 reveals for both players a bimodal distribution, with the greatest frequencies being for A and D; the high σ is thus explained, but the bimodal distribution is not. Perhaps the players expected some sort of symmetry between their own and their opponents' matrices--in which case they ranked their opponents' payoffs the same as or opposite to their own. Since the payoffs were A's and D's we look for A's and D's among the POM's in outcomes (1,2) and (2,1) ...and we find them--that is, we find a bimodal distribution with modes A and D.

The average r for the whole matrix (next-to-last column, Table 17) can be taken as another measure of degree of consensus encouraged by the game; the last column ranks the games accordingly. The ranking is not identical with that in Table 15, but the top three games and the bottom three are the same in both.

6. Conclusion

6.1 Tying It All Together: A Measure of Structural Transparency

Our focus in all three parts of our study has been primarily on the way game structure affects behavior, rather than on such variables as players' personalities and intelligence or on idiosyncratic play. This is because meaningful comparisons between individuals or pairs reacting to an environment are of limited value when the environment itself is not understood. The games, of course, are the environments being studied.

We developed a number of ways of measuring various aspects of these environments (games). Table 18 brings together thirteen such measures or criteria and applies them in ranking our six games. Though it is evident merely by inspection that they give rather similar ranks to the games, we can measure the degree of association among the rankings by the Kendall Coefficient of Concordance, W;[16/] in this case W = .65. So high a W makes the probability that the rankings are unrelated less than .00001 .

If we average the thirteen ranks for each game we can make a fourteenth ranking based on these averages.[17/] A name suggestive of the element common to the sources which have contributed to Ranking Number 14 is measure of structural transparency. According to this measure, the games in order of decreasing structural transparency are 6, 1, 4, 3, 5, 2. Intuition would put Game 1 first; we suspect that prior game-playing experience by the subjects (none in the case of Game 1, five games in the case of Game 6) affected their play and perception enough to account for the reversed order

Table 18 - Six Games Ranked by Fourteen Criteria

Ranking Number	Source of Ranking: Part	Table	Column	Criterion of Ranking	Rank of 1 is assigned to the game with	Rank of Game: 1	2	3	4	5	6
1	I	4	2	H_0 : Frequency of outcomes are equally likely (Random Solution)	χ^2 highest, probability that H_0 is true lowest	3	6	5	2	4	1
2	I	5	4	Quality of prediction of outcomes by Non-Cooperative Solution	Q highest	3	4	5	2	6	1
3	I	6	8	Steady-state frequency	Frequency highest	3	6	4	2	5	1
4	I	7	3	Quality of prediction of steady-state outcomes by Non-Cooperative Solution	Q highest	2.5	6	5	2.5	2.5	2.5
5	II	1	7	H_0: Players do not use initial strategy picked by applying rule αβλδ	Probability that H_0 is true lowest	3	-	5	4	2	1
6	II	3	8	Overall measure of incentive to switch from initial strategy picked by rule αβλδ	Mean M lowest	1	6	4.5	4.5	3	2
7	II	4	2	Degree of (strategic) involvement	I_T lowest	2	5	6	3	4	1
8	III	4	2	H_0: POM's are chosen randomly	χ^2 highest, probability that H_0 is true lowest	2	5	3	4	6	1
9	III	7	row 2	Median accuracy of individuals' POM's	τ highest	1	5.5	4	3	5.5	2
10	III	10	2	Degree of information involvement	K lowest	6	5	1	3	2	4
11	III	15	9	Degree of encouraging consensus (H_0: consensus POM's are based on average outcome ranks sufficiently like each other to have arisen randomly)	Mean of F-ratios highest; probability that H_0 is true lowest	3	6	4	2	5	1

Table 18 - (Continued)

Ranking Number	Source of Ranking: Part	Table	Column	Criterion of Ranking	Rank of 1 is assigned to the game with	Rank of Game: 1	2	3	4	5	6
12	III	16	4	Accuracy of Consensus POM's	Mean τ highest	2	6	4	3	5	1
13	III	17	13	Degree of encouraging consensus (variability of POM guesses)	Mean σ lowest	1	4	5	3	6	2
				Average of 13 Rankings		2.50	5.38	4.26	2.92	4.30	1.58
14	III	18	Cols. 7-12, last row	"Measure of structural transparency" based on average of 13 rankings	Average of 13 rankings lowest	2	6	4	3	5	1

of these two. Otherwise the sequence of games arranged in order of structural transparency--empirically determined--accords well with our a priori expectations.

6.2 Remarks to Experimental Gamesters

We wish to encourage other workers to run experiments like these. The importance of replication of experiments and reconfirmation of results is acknowledged in all the experimental sciences, which a fortiori emphasizes its importance in the young science of experimental economics. Frequently in this field experiments have been done, described and forgotten. The aversion to replication of another's work is to be regretted; sometimes it can be explained by poor documentation on the part of the original experimenters, while in other instances it stems from a feeling that following in another's tracks is a less prestigious kind of scholarly endeavor. We hope our presentation is clear enough to acquit us of the first charge, and we also hope that interested parties will free themselves of the second.

Our experiments are simple, easy to duplicate, cheap, and valuable to the subjects--who learn something about non-zero-sum game theory--as well as the experimenters. While in the larger business-game-type "experiments" until much more core in design and analysis has been done [18/] there can be doubt whether the activity is valuable for either research or teaching, we believe our very simple games develop a synergistic relationship between research and teaching which recommends their more widespread use.

FOOTNOTES

1/ Martin Shubik and David H. Stern, "Some Experimental Non-Constant-Sum Games Revisited," Parts I and II, CFDP 236 and 240.

2/ Subsections 1.1 and 1.2 are lifted verbatim from ibid., Part II, pp.2-4. [Special note: The pagination of Part II as originally issued is in error. All references to Part II in this paper are to the correct pagination; there are altogether 42 pages in Part II.]

3/ Subsection 1.3 summarizes pages 4-7 and 10-25 of ibid., Part II.

4/ Pages 19-20.

5/ Sidney Siegel, Nonparametric Statistics for the Behavioral Sciences, Mcgraw-Hill Book Company, New York, 1956, pp. 213-223.

6/ Ibid., pp. 217-219.

7/ Ibid., pp. 42-47, 249.

8/ Ibid., pp. 104-111, 249.

9/ Shubik, M., and D. H. Stern, op. cit., Part I, pp. 18-23.

10/ Siegel, S., op. cit., pp. 213-223, 247.

11/ Shubik, M., and D. H. Stern, op. cit., Part II, pp. 27-33.

12/ Ibid., p. 33.

13/ Siegel, S., op. cit., pp. 213-223, 285.

14/ Shubik, M., and D. H. Stern, op. cit., Part II, pp. 28-29.

15/ The rationale and method for this often used statistical techniques are found in many texts, for example, Quinn McNemar, Psychological Statistics. (John Wiley and Sons, New York, 1962), pp. 246-267, 433.

16/ Siegel, S., op. cit., pp. 229-238.

FOOTNOTES

17/ Siegel (ibid, p. 238) says, "Kendall (1948a, p. 87) suggests that the best estimate of the 'true' ranking of the N objects is provided, when W is significant, by the order of the various sums of ranks..." The reference is to M. G. Kendall, Rank Correlation Methods (Griffin, London, 1948).

18/ For further discussion, see Levitan, R. and Shubik, M. "A Business Game for Teaching and Research Purposes " Mimeographed M. S. Yale 1968.