Backward Induction Failures in Bargaining Games

Niels Boissonnet

May 2020

Bargaining Games

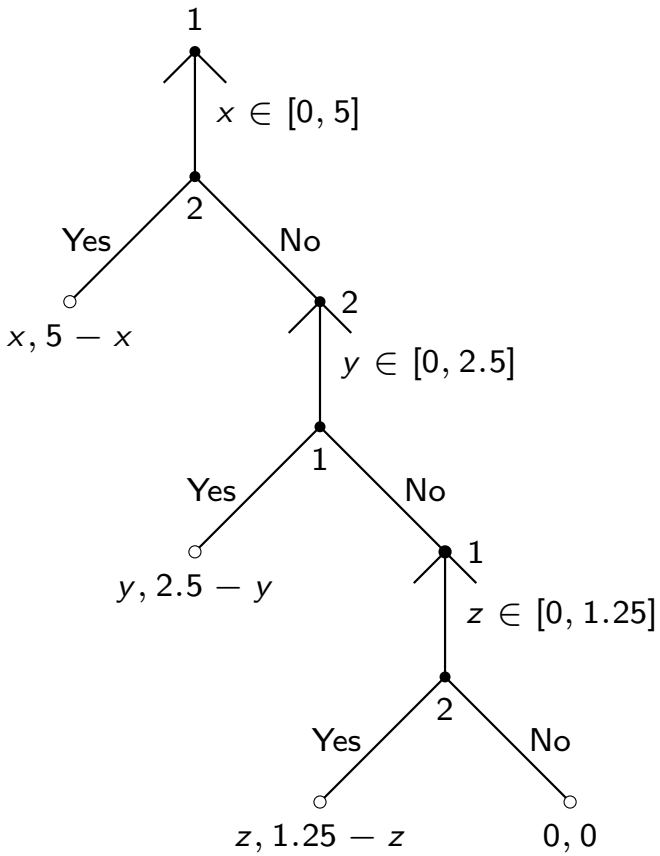


Figure: Bargaining Game

Backward induction yields that player 1 demands 3.75 at the first round, and player 2 accepts the offer.

Nevertheless, subjects in lab demands between 2.5 and 3.75 with an average of 3.

Two competing explanations

Limited-cognition:

- ... players do not intuitively reason game-theoretically; looking ahead is computationally expensive and unnatural because it means reasoning about events that probably will not occur.
- Players offer equal splits or a little less because they are not sure what else to do or are not willing to bet too much on the self-interest and rationality of others.
- Incompatible with backward induction.

Social preferences

- Social preference theories account for departures from perfect equilibrium by positing a "social utility" for others' payoffs or a preference for reciprocity or fairness etc... E.g. : Ultimatum game
- Compatible with backward induction.

Paper 1:
A Backward Induction Experiment

Binmore, McCarthy, Ponti, Samuelson, and Shaked (2002)

Method

They argue that Backward induction is equivalent to the three following statement:

- **Rationality:** Given a choice between two (vectors of) payoffs, a player chooses the most preferred.
- **Subgame Consistency:** Play in a subgame is independent of the subgame's position in a larger game.
- **Truncation consistency:** Replacing a subgame with its equilibrium payoffs does not affect play elsewhere in the game.

Assuming that **Rationality** is correct they show that **Subgame Consistency** and **Truncation consistency** are violated.

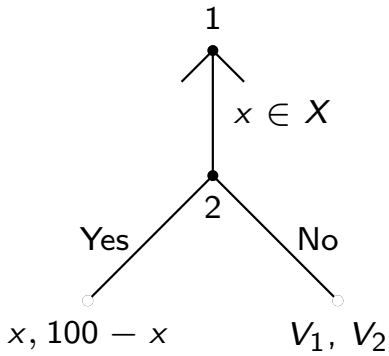


Figure: Game I

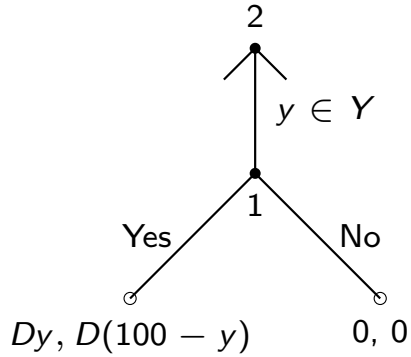


Figure: Game II

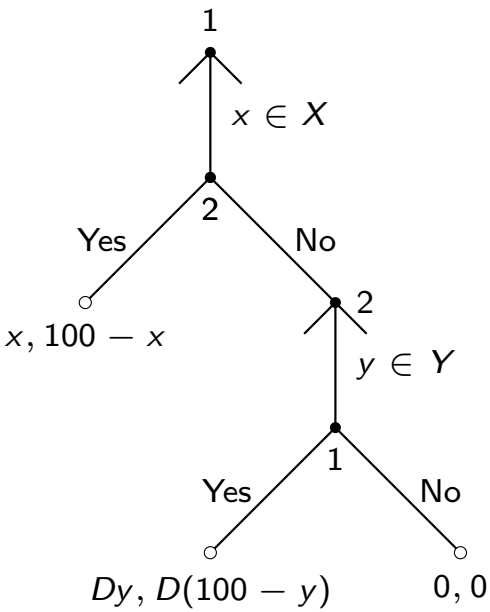


Figure: Game III

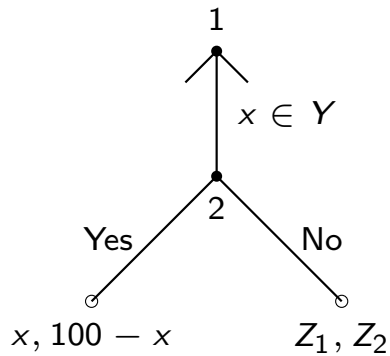


Figure: Game IV

Subgame Consistency

"Play in a subgame is independent of the subgame's position in a larger game."

They compare Game II and the second stage of game III

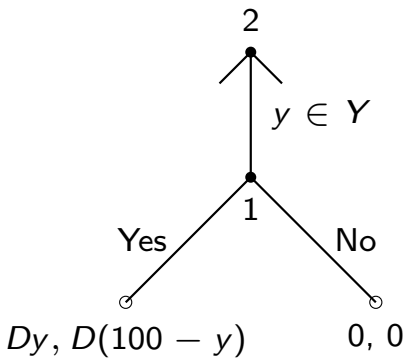


Figure: Game II

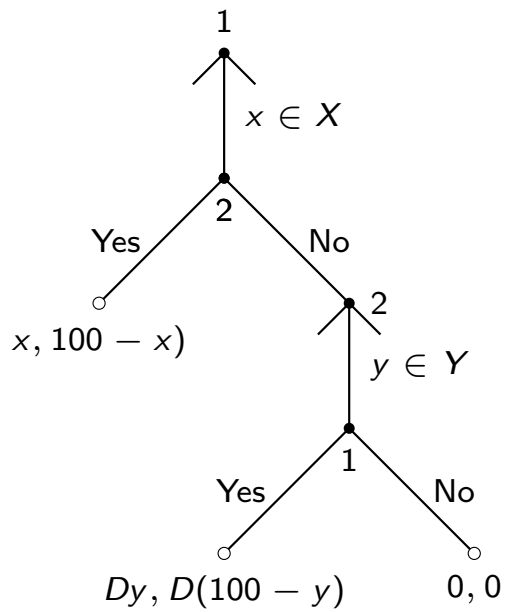


Figure: Game III

Subgame Consistency

"Play in a subgame is independent of the subgame's position in a larger game."

They compare Game II and the second stage of game III

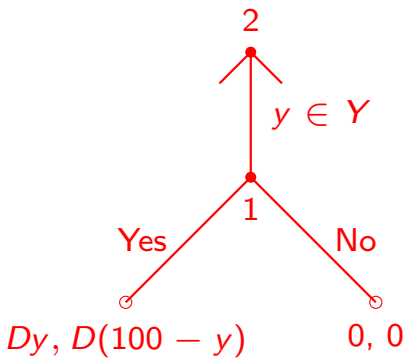


Figure: Game II

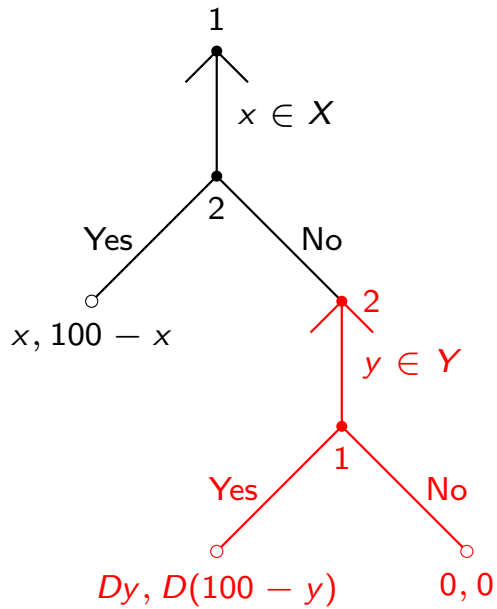


Figure: Game III

Strategy to show that Subgame Consistency is Violated

Isn't the mere fact that the second stage is reached evidence that backward induction fails?

- No, because the players' (interdependent) preferences are not commonly known.
- Not necessarily: it may reflect that player 2 has inferred something about player 1, and hence about her optimal second stage offer, from the demand that 2 rejected to reach the second stage of the two-stage game.

So, the question is whether we can use belief updating in a sufficiently straightforward way to avoid the violation of Backward induction.

- If player 2 can draw inferences about player 1 from the latter's period-1 demand, then we would expect player 2's period-2 offer to vary significantly with the identity of the demand rejected by player 2 to reach the second stage.

No evidence of such a relationship.

Significant Discrepancy between the median offer in Game II and the second stage of game III

Rounds	(V_1, V_2)	Observations	Mean offer	Median	5th %tile	95th %tile
1-10	All	1200	33.4	33	15	50
11-20	All	1200	32.2	31	20	48.5

Figure: Game II

Specification	Rounds	Observations	Mean offer	Median	5th %tile	95th %tile
All	1-10	276	43.3	44	23	60
All	11-20	225	43.8	45	25	60
All	1-20	501	43.5	44	25	60

Figure: Game III

Higher rejection rate for the second stage of Game III

		Offers		
Game	Observations	All	30–40	35–45
II	2400	20%		
III	501	29%		
II	1064		15%	
III	131		47%	
II	733			9%
III	190			36%

Figure: Rejection rates in Game II and in the second stage of Game III

Effect on payoff of this discrepancy

Game	D	II-median, y_m^{II}	Payoff	III-median, y_m^{III}	Payoff
II	.2	31	13	45	11
II	.3	28	19	47	16
II	.4	35	23	48	21
II	.5	33.5	33	44	28
II	.6	35	34	37.5	33
II	.7	31	41	45	39
II	.8	35	52	40.5	48
II	.9	25	68	45	50
III	.2	31	0	45	7
III	.3	28	0	47	13
III	.4	35	0	48	9
III	.5	33.5	0	44	18
III	.6	35	32	37.5	38
III	.7	31	0	45	39
III	.8	35	15	40.5	27
III	.9	25	0	45	41

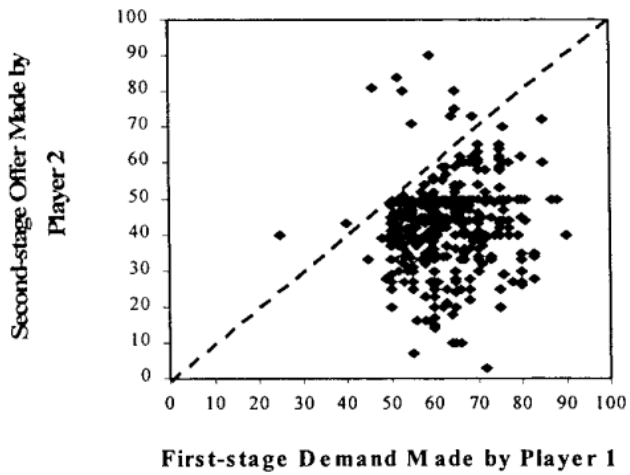
Figure: Effects on payoff

Two possibilities

1. **Belief update about the opponent's preferences:** Could proposers in the second stage of the two-stage game be simply reacting to information gleaned about their opponents' preferences from the demand made by their opponents in the first stage?
2. **Player 2s who rejects in Game III are of special type:** The second stage of game III can only be reached if player 2 rejects player 1's initial demand. Could play in the continuation game differ from that of the second stage of game III because the latter game is not played by a random sample of player 2s?

Belief update about the opponent's preferences:

- There appears to be little relationship between the rejected demand and the subsequent offer.
- An examination of player 1's behavior in the second stage of game III suggests that there is little for player 2 to learn from observing player 1's first-period demand.



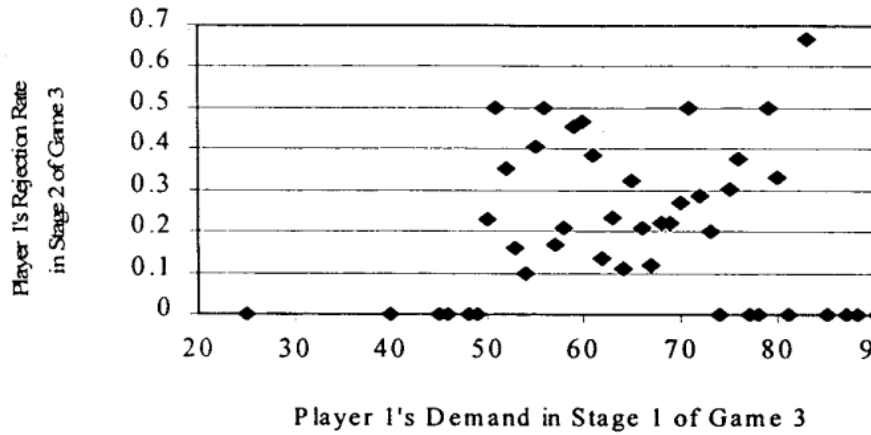


FIG. 14. Player 1's rejection rate in the second stage of game III, as a function of eriod demand that player 1 had rejected in order to reach the second stage.

Not a random sample

Mean offer in Game II of player 2s who rejected at least one demand in game III

Rounds	(V_1, V_2)	Observations	Mean offer	Median	5th %tile	95th %tile
1-10	All	777	33.5	33	16	50
11-20	All	780	32.4	31	20	49

Mean offer in Game II

Rounds	(V_1, V_2)	Observations	Mean offer	Median	5th %tile	95th %tile
1-10	All	1200	33.4	33	15	50
11-20	All	1200	32.2	31	20	48.5

No evidence that the second stage of game III is played by a sufficiently atypical group of player 2s.

Truncation consistency

"*Replacing a subgame with its equilibrium payoffs does not affect play elsewhere in the game.*"

They compare Game II and the second stage of game III

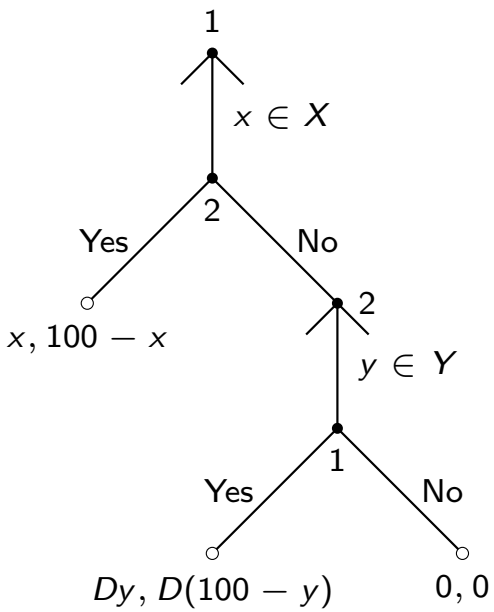


Figure: Game III

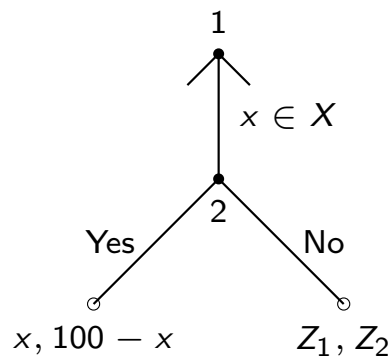


Figure: Game IV

Truncation consistency

" *Replacing a subgame with its equilibrium payoffs does not affect play elsewhere in the game.*"

They compare Game IV and the first stage of game III

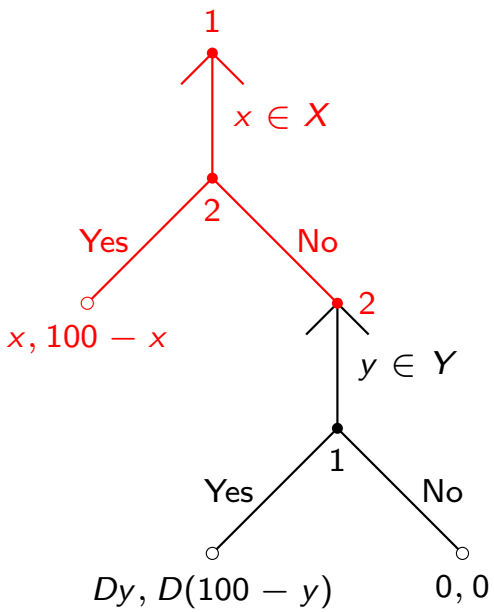


Figure: Game III

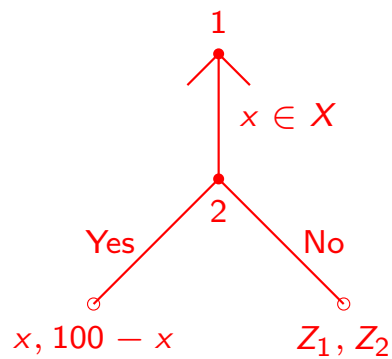


Figure: Game IV

How (Z_1, Z_2) was determined

Z_1 and Z_2 varied across subjects so we denote them by $(Z_1(i), Z_2(i))$

In the nutshell, they chose $(Z_1(i), Z_2(i))$ as the estimated expected payoff of both Players in the second stage of Game III

- $Z_1(i)$ to be the average of the player-1 payoffs in those second-stage games in which player i participated,
- $Z_2(i)$ to be the average of the player-2 payoffs in those second-stage games in which player i participated,

Result

- Player-1 initial demands are significantly less sensitive to rejection payoffs in game III than to rejection payoffs in game IV.
- In contrast to the prediction of truncation consistency, players react more sharply to variations in fixed terminal payoffs than they do to equivalent variations in the expected value of a continuation game.

Conclusion of the Paper

Our experimental results provide several indications that payoff-interdependent preferences and backward induction, in the form of subgame and truncation consistency, are inconsistent.

Failure of subgame consistency: players regard the second stage of the two-stage game and a seemingly stand-alone equivalent as different strategic situations.

Truncation consistency failure: Players are more responsive to variations in future prospects when a rejection leads to a fixed pair of rejection payoffs, as opposed to the case in which a rejection leads to a game involving another offer and response.

Paper 2:
Detecting the Failures of Backward Induction

Johnson, Camerer, Sen, and Rymon (2002)

Two Experimental studies

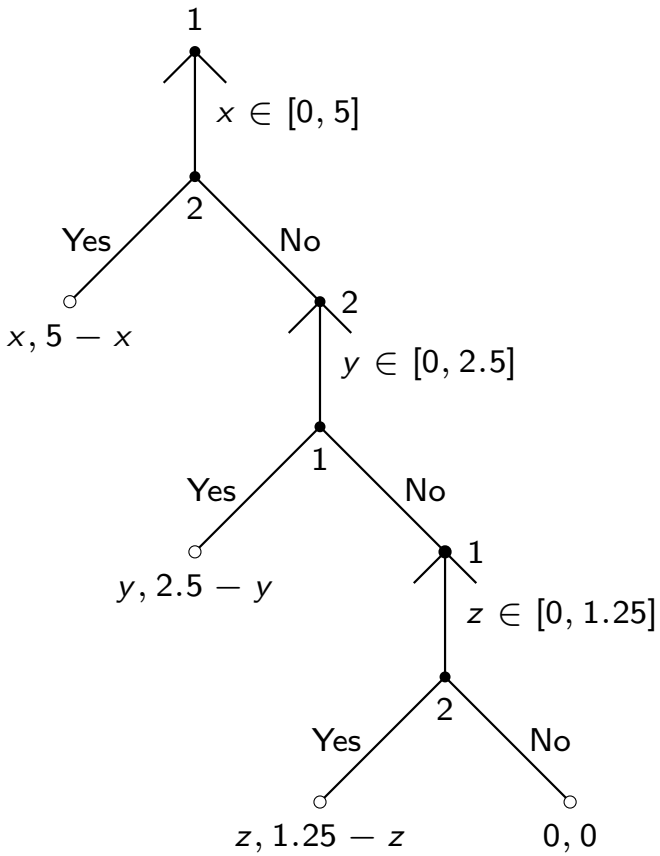


Figure: Bargaining Game

1. They collected measures of information search with a "Mouselab" system to study individual information seeking process.
2. Players' opponents were computers.

First Study

- To see what is in a box, subjects used a mouse to move a cursor into the box.
- By recording which box is open every 16 milliseconds, we collect a fine-grained measure of the time when each box is entered and exited, allowing us to study the order in which boxes were opened and for how long each box was opened.

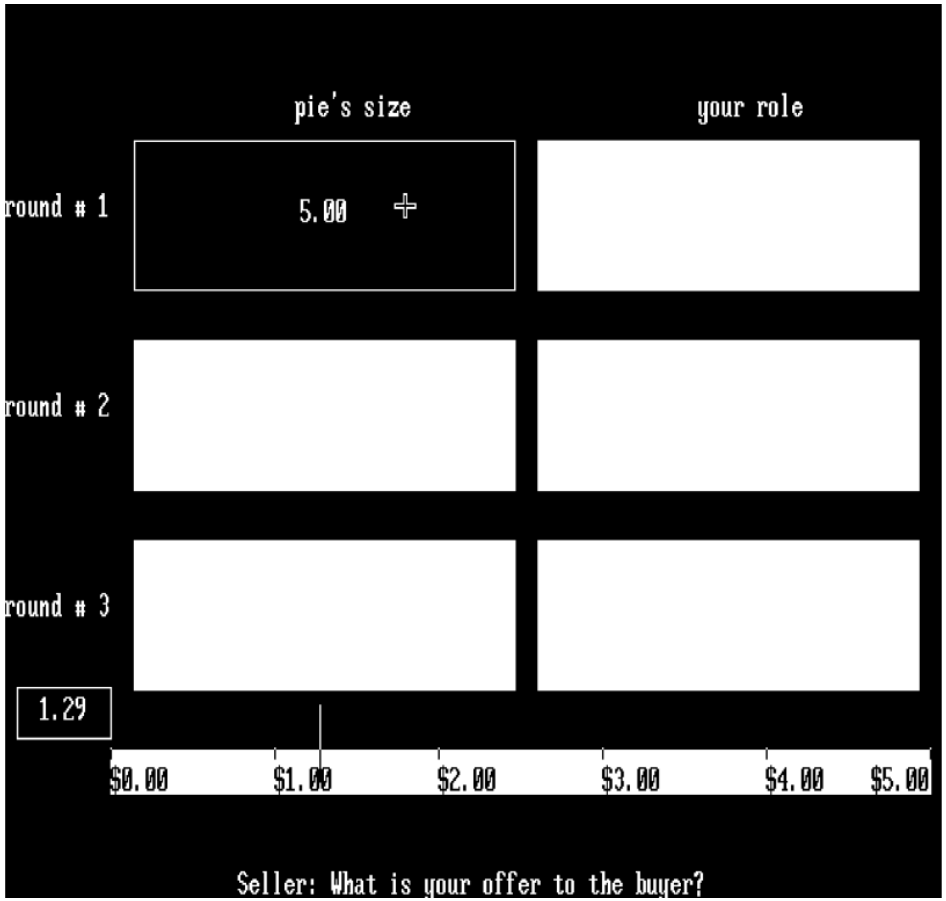


Figure: MOUSELAB computer display.

First Study

- Players make an offer by moving the cursor to the dollar scale at the bottom of the box.

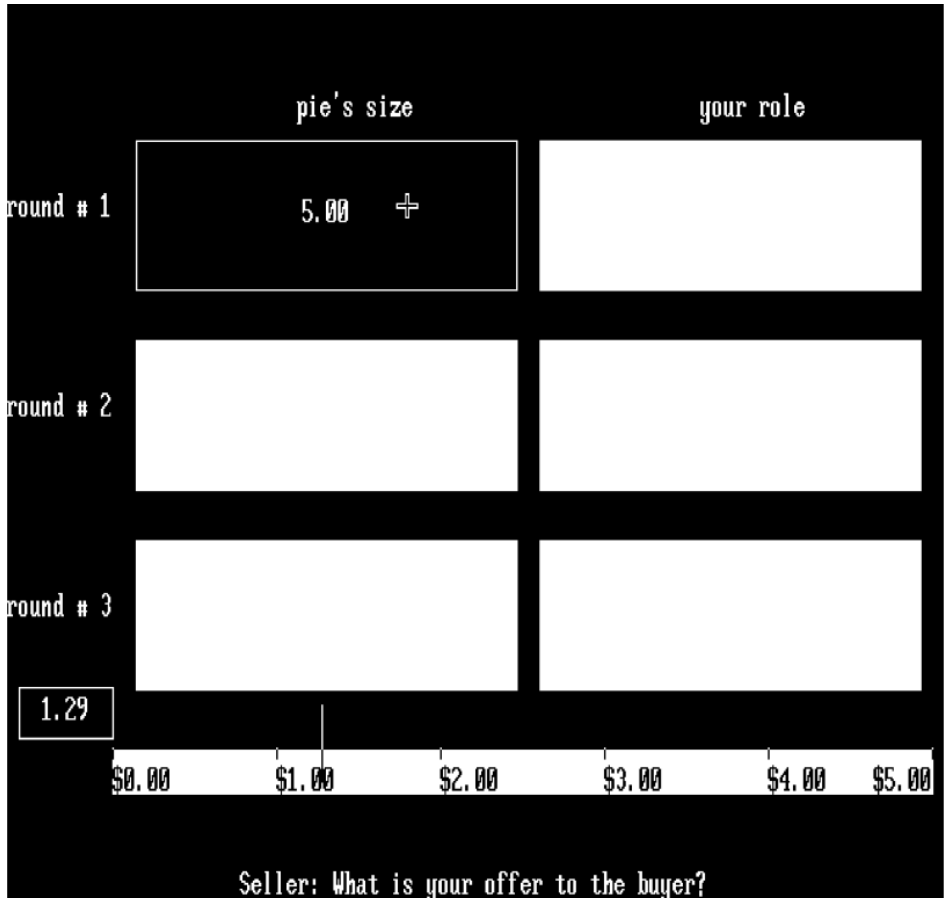


Figure: MOUSELAB computer display.

Players' categorization by cognitive limitations

- **Level 0 players** see the first round as an ultimatum game. They ignore future rounds and offer 40% of the first-round pie ($2), the average offer in ultimatum games.
- **Level 1 players** look one round ahead and act like they are playing a truncated two-stage game. They anticipate that player 2 will make an ultimatum offer of the round two pie (40% of $2.50, leaving $1.50 for player 2), so planning ahead one step, they offer $1.50 to player 2 in the first round.
- **Equilibrium Players** open the second - and third - boxes most often and look at them longest, make many transitions between those boxes, and look least at the first-round box are call equilibrium player.

Result: Information Search

Players look most often at the first-round pie size and make forward transitions more than backward ones.

TABLE I

Round	Number of acquisitions	Total time examining payoff	Transitions		
			1	2	3
Round 1 ($5.00)	4.38	12.91	—	2.55	0.65
Round 2 ($2.50)	3.80	6.67	2.10	—	1.24
Round 3 ($1.25)	2.12	1.24	0.50	0.88	—

Study 2

In study 1, subjects may know how to reason game-theoretically but not do so because they care about fairness or believe other subjects are con- cerned with fairness. In study 2, they control for these possibilities by matching players with robot players.

Results

If concerns for fairness are all that is preventing subgame perfect play among people, offers to robots should be near the subgame perfect equilibrium of \$1.25. If generous offers to people come from limited cognition, we should see generous offers to robots too.

	No instruction, equilibrium = \$1.26			
Round	1–2	3–4	5–6	7–8
Average offer	1.83	2.02	1.67	1.82
Number of rejections (out of 30)	25	11	12	13