Measuring Higher-Order Rationality with Belief Control

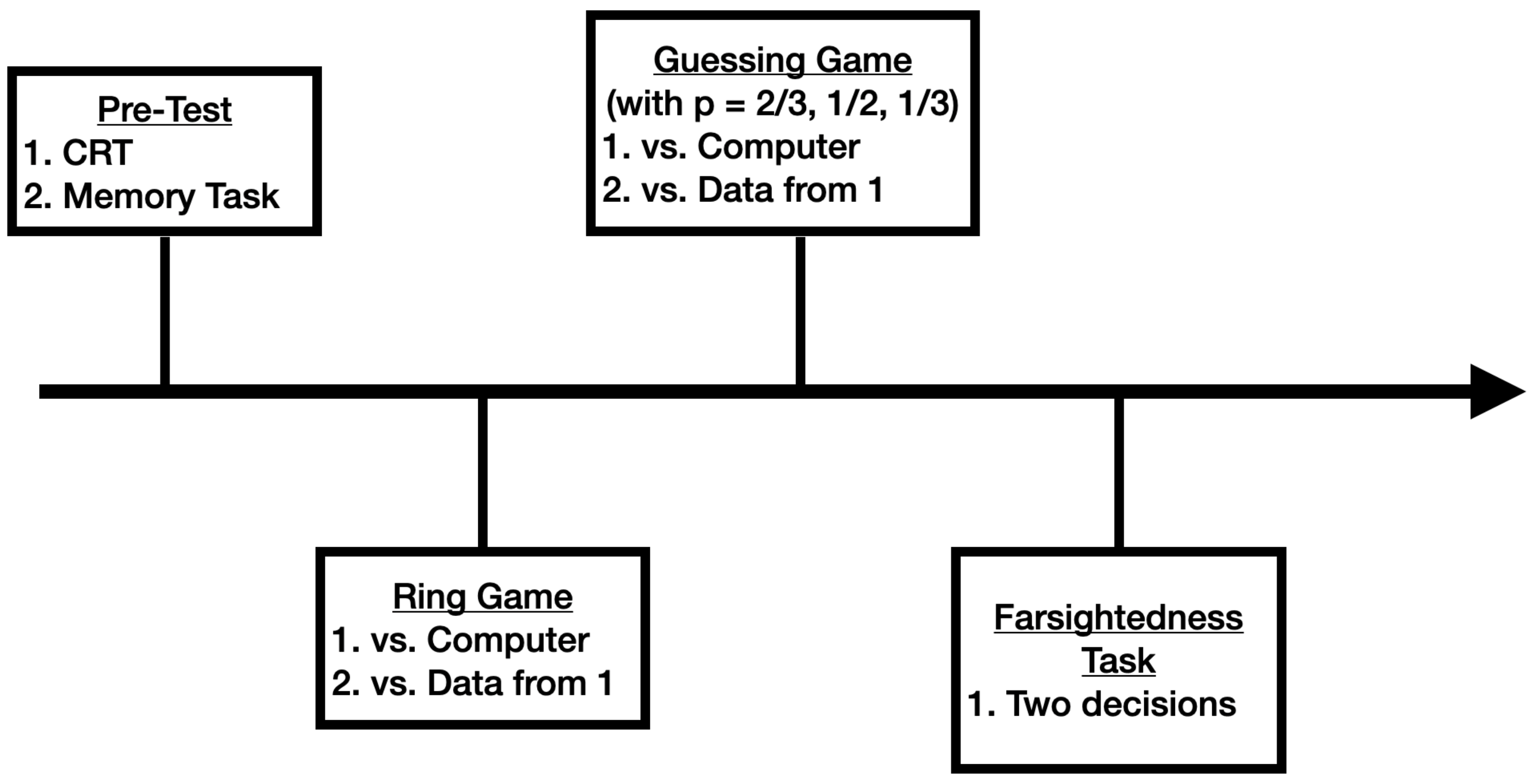
Wei James Chen[1], Meng-Jhang Fong[2], Po-Hsuan Lin[2]

[1]National Taiwan University [2]California Institute of Technology

Research Idea

- We conduct an experiment to study an individual's strategic reasoning levels across games by matching subjects w/ robot players
- Motivation: establishing an approach to measure a subject's strategic reasoning depth in the lab is important
- Challenge: unstable individual strategic reasoning levels across games (E.g., Georganas et al., 2015; Cerigioni et al., 2019)
  - Possible reason: heterogeneous beliefs about human opponents
- Previous studies: using computer players for studying non-equilibrium behavior (E.g., Johnson et al., 2002)
  - Focusing on one family of games in one study

Experiment Protocol



Pre-Test
1. CRT
2. Memory Task

Guessing Game
(with p = 2/3, 1/2, 1/3)
1. vs. Computer
2. vs. Data from 1

Ring Game
1. vs. Computer
2. vs. Data from 1

Farsightedness Task
1. Two decisions

Ring Game (Kneeland, 2015)

- The only difference between G1 and G2 is P4's payoff matrix

G1

Player 1

Player 2's actions

Player 1's actions	a	b	c
a	8	20	12
b	0	8	16
c	18	12	6

Player 2

Player 3's actions

Player 2's actions	a	b	c
a	14	18	4
b	20	8	14
c	0	16	18

Player 3

Player 4's actions

Player 3's actions	a	b	c
a	20	14	8
b	16	2	18
c	0	16	16

Player 4

Player 1's actions

Player 4's actions	a	b	c
a	6	10	8
b	12	16	14
c	8	12	10

G2

Player 1

Player 2's actions

Player 1's actions	a	b	c
a	8	20	12
b	0	8	16
c	18	12	6

Player 2

Player 3's actions

Player 2's actions	a	b	c
a	14	18	4
b	20	8	14
c	0	16	18

Player 3

Player 4's actions

Player 3's actions	a	b	c
a	20	14	8
b	16	2	18
c	0	16	16

Player 4

Player 1's actions

Player 4's actions	a	b	c
a	8	12	10
b	6	10	8
c	12	16	14

2-Person Guessing Game (Costa-Gomes and Crawford, 2006)

- $U_i = 0.2(100 - |Guess_i - p \cdot Guess_{-i}|)$
  - $Guess_i = \{1, 2, \dots, 100\}$ for $i = 1, 2$
  - Dominance solvable given a single-peaked payoff structure

Identification by Revealed Rationality (Lim and Xiong, 2016)

- (First-order) Rationality: the ability to best respond to some belief
- K_{th}-order rationality: the ability to anticipate that the opponents are $(K - 1)_{th}$-order rational and to best respond to such belief
- One is **k_{th}-order revealed rational** if his strategy survives k rounds of iterated elimination of dominated strategies (IEDS)
- A subject is assigned to the lowest type he exhibits across games

Treatments: Robot and History

- Play the games in two different scenarios (without feedback)

1. Robot Treatment: against fully rational computer players
2. History Treatment: against the data drawn from the first scenario

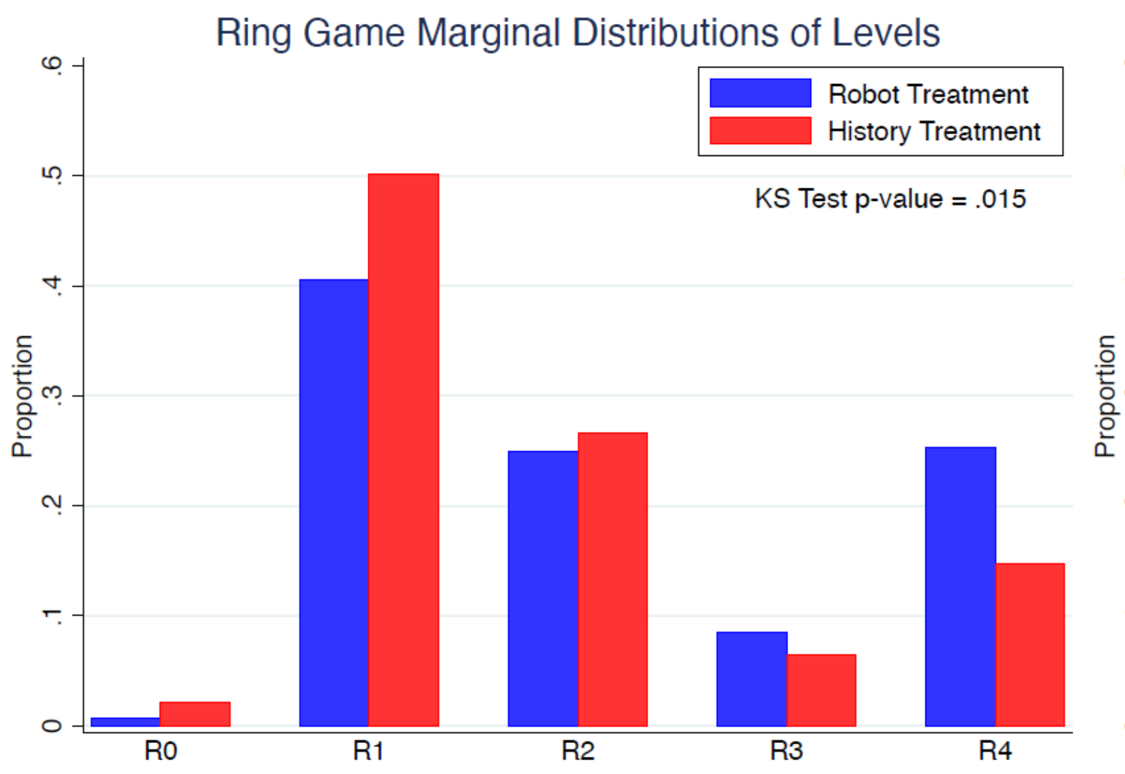
Instructions for Robot Treatment

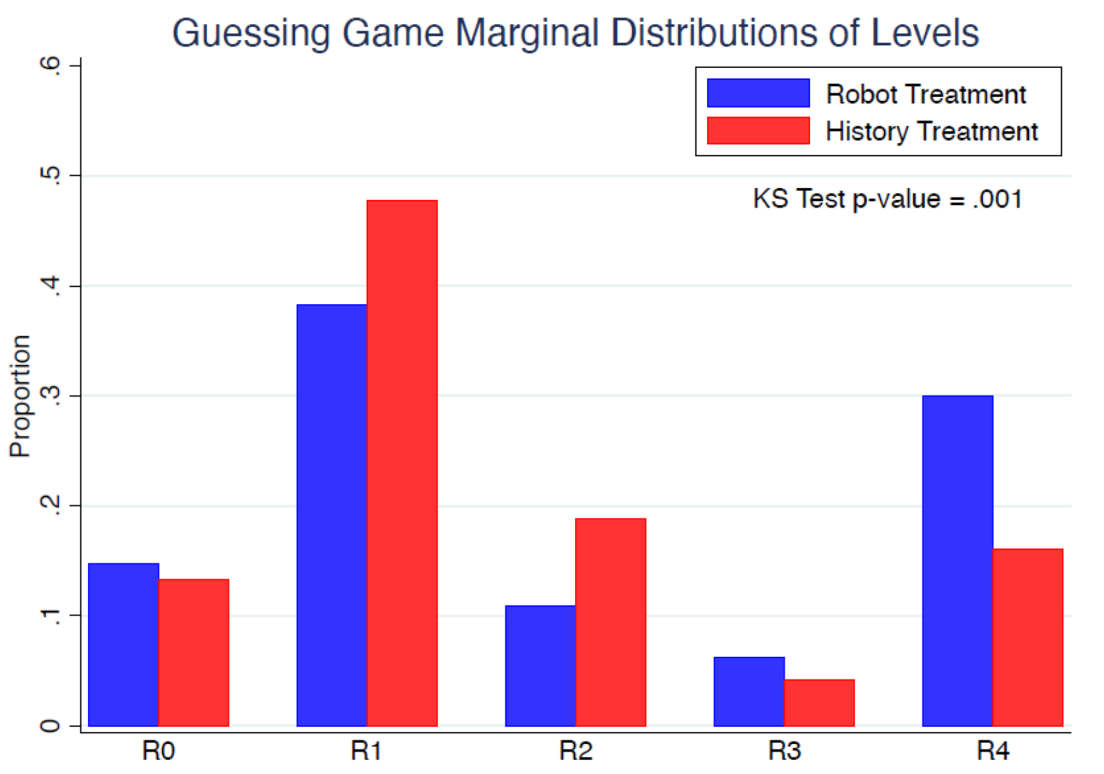
The other participants will be computers that are programmed to take the following strategy:

1. *The computers aim to earn as much payoff as possible for themselves.*
2. *A computer believes that every participant will try to earn as much payoff as one can.*
3. *A computer believes that every participant believes "the computers aim to earn as much payoff as possible for themselves."*
   - Adapted from the instruction used in Johnson et al. (2002)

Result: Type Distribution

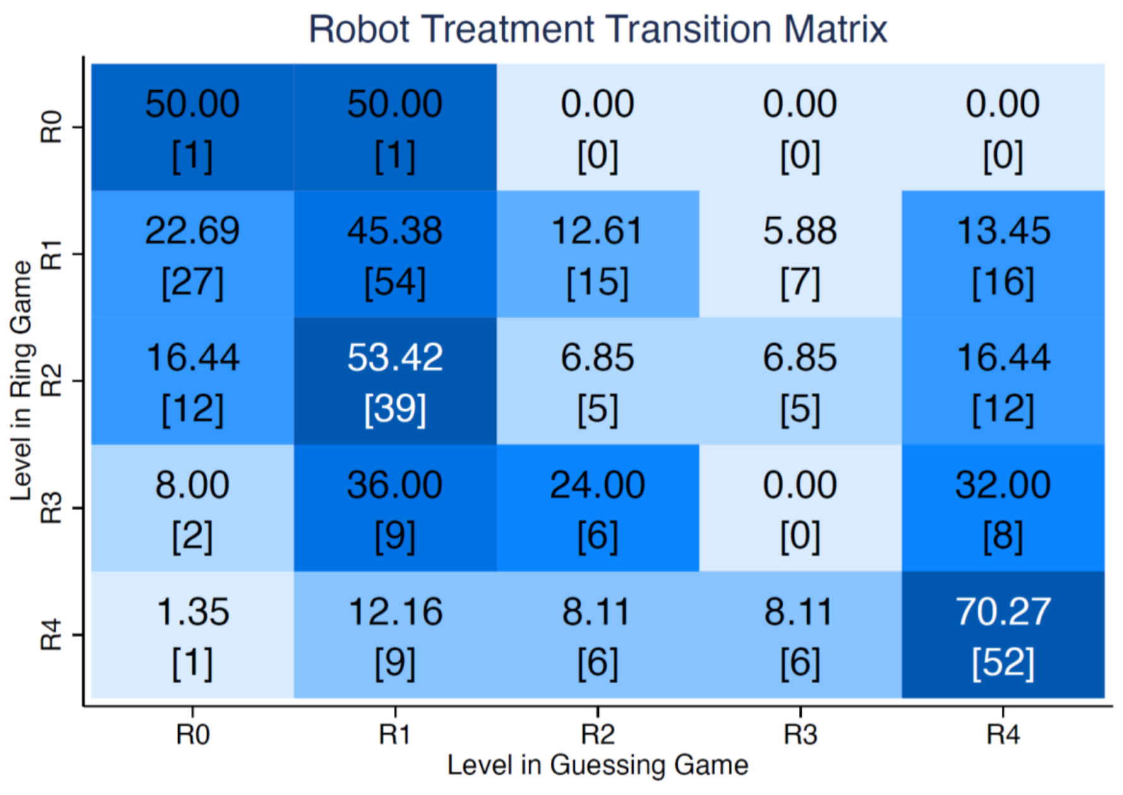
- Does rationality levels against robots capture individual strategic reasoning capacity? ($n = 293$)
  - Within-subject analysis: signed-rank test ($p < 0.001$)





Result: Constant Absolute Rationality Levels

- ➢ Is a player's reasoning depth constant across games?
  - 112 (38.23%) exhibit the same rationality levels across games
- Does the seemingly high proportion of constant-level players actually result from two independent type distributions?
- Null hypothesis: the subjects' rationality depths are independently distributed across families of games
- Monte Carlo simulation: 10,000 random samples of 293 pairs of levels (Georganas et al., 2015)
  - Independently drawn from the empirical distribution



Robot Treatment Transition Matrix

Level in Ring Game \ Level in Guessing Game	R0	R1	R2	R3	R4
R0	50.00 [1]	50.00 [1]	0.00 [0]	0.00 [0]	0.00 [0]
R1	22.69 [27]	45.38 [54]	12.61 [15]	5.88 [7]	13.45 [16]
R2	16.44 [12]	53.42 [39]	6.85 [5]	6.85 [5]	16.44 [12]
R3	8.00 [2]	36.00 [9]	24.00 [6]	0.00 [0]	32.00 [8]
R4	1.35 [1]	12.16 [9]	8.11 [6]	8.11 [6]	70.27 [52]

Constant Level Frequency	Pool Data
Robot Treatment	
Simulation mean:	32.9%
Simulation 95% CI:	[27.6%, 38.2%]
Empirical mean:	38.2%
p-value:	0.057
History Treatment	
Simulation mean:	40.3%
Simulation 95% CI:	[34.8%, 45.7%]
Empirical mean:	41.3%
p-value:	0.768

Result: Constant Ordering of Rationality Levels

- ➢ Does the ranking of players (in terms of rationality levels) remain the same across games?
- Define switch ratio = switch frequency/non-switch frequency
  - Under the null hypothesis, the (expected) switch ratio = 1

✓ Switch:

	Player *i*	Player *j*
Ring	2	4
Guessing	4	3

✓ Non-switch:

	Player *i*	Player *j*
Ring	2	4
Guessing	1	2

Ring Game vs. Guessing Game	Empirical Data	Null Hypothesis
Robot Treatment		
Switch frequency:	12.3%	22.5%
Non-switch frequency:	41.3%	22.5%
Switch ratio:	0.30	1.01
p-value:	< 0.0001	
History Treatment		
Switch frequency:	12.9%	17.9%
Non-switch frequency:	34.5%	17.8%
Switch ratio:	0.37	1.02
p-value:	< 0.0001	

Conclusion

- We find some consistency in subjects' rationality depths across games in terms of both absolute and relative levels
- This result suggests that strategic reasoning ability may be a persistent personal trait
- Furthermore, after controlling for a subject's beliefs about his/her opponent's rationality, we may be able to gauge the subject's strategic thinking ability using his/her choice data

References

Cerigioni, F., Germano, F., Rey-Biel, P., & Zuazo-Garin, P. (2019). Higher orders of rationality and the structure of games. *Mimeo.*

Costa-Gomes, M. A. & Crawford, V. P. (2006). Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review*, 96(5), 1737–1768.

Georganas, S., Healy, P. J., & Weber, R. A. (2015). On the persistence of strategic sophistication. *Journal of Economic Theory*, 159, 369–400.

Johnson, E. J., Camerer, C., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory*, 104(1), 16–47.

Kneeland, T. (2015). Identifying higher-order rationality. *Econometrica*, 83(5), 2065–2079.

Lim, W. & Xiong, S. (2016). On identifying higher-order rationality. *Mimeo.*

Meng-Jhang Fong
Email: mjfong@caltech.edu Website: https://mjfong.github.io/