Erik Brockbank
ebrockbank@ucsd.edu
Edward Vul
evul@ucsd.edu

# Humans fail to outwit adaptive rock, paper, scissors opponents

UC San Diego
computational cognition lab

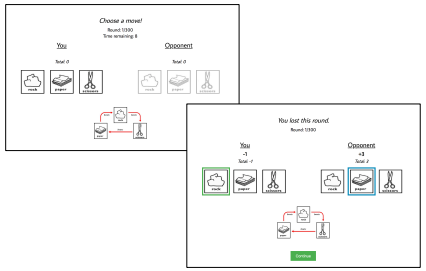
## MOTIVATION


*How do people adapt to exploitation by others?*

- Prior work has studied exploitability through the robust sequential patterns people generate when they are trying to behave randomly[1, 2, 3]
- These exploitable patterns can be detected and reduced through feedback,[4] expertise,[5] and adversarial dynamics[6]
- **Earlier work leaves unanswered whether and how people can detect and respond to exploitation of more complex patterns in their own behavior**

***In the current study***, people play rock, paper, scissors (RPS) against bots that exploit patterns in participant behavior.[7] We explore whether people are able to reduce their exploitability over many rounds of play.

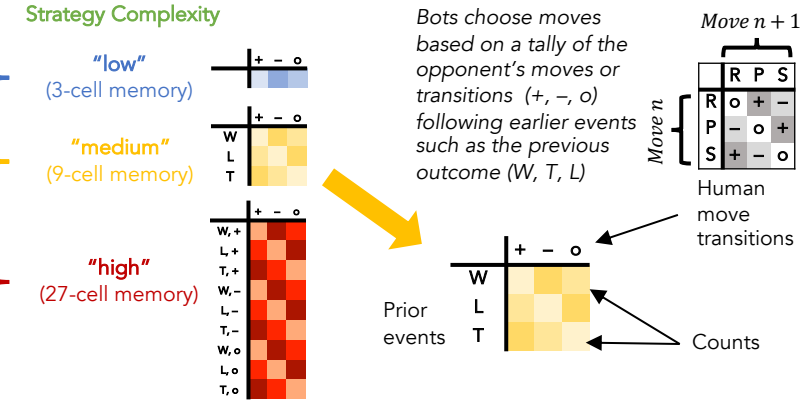
## METHODS



Participants (N = 192) played 300 rounds of rock, paper, scissors against a *strategic bot*

| Bot Strategy | N (Total = 192) |
|---|---|
| *Transition baserate* (+ / – / O) | 21 |
| *Opponent transition baserate* (+ / – / O) | 26 |
| *Transition given player's prior choice* | 24 |
| *Transition given opponent's prior choice* | 28 |
| *Transition given prior outcome* (W / L / T) | 22 |
| *Choice given player's prior choice & opponent's prior choice* | 25 |
| *Choice given player's prior two choices* | 20 |
| *Transition given prior transition & prior outcome* | 26 |

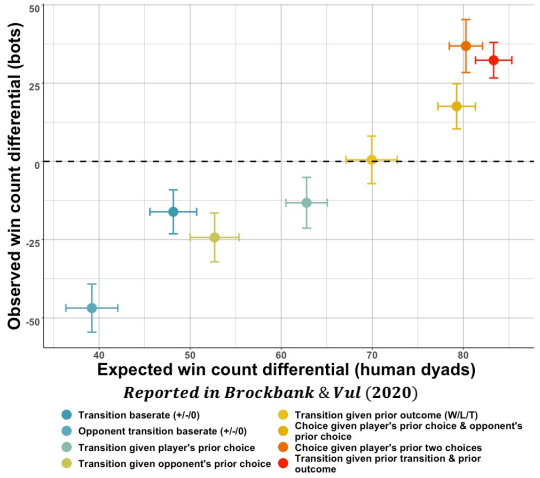
The 8 bots each chose their moves by exploiting a different sequential pattern in their opponent's behavior




Bot strategies varied in *complexity* based on the bot's memory of previous human moves and events

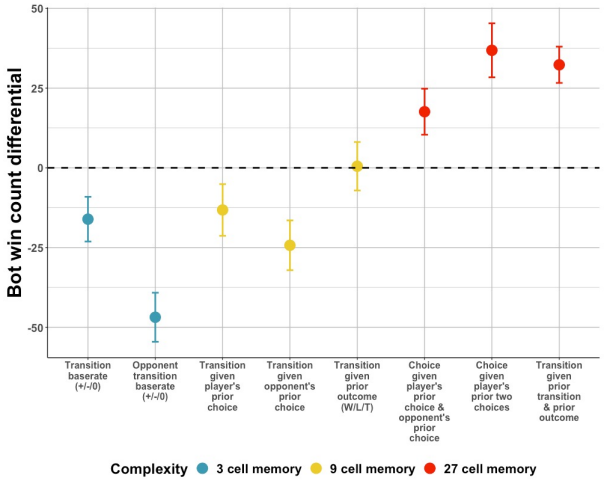
## RESULTS

### Successful exploitation by bot agents aligns with patterns people exhibit in dyad play




Average bot win count differentials (bot wins – human wins) for each bot strategy are highly correlated with the *expected* win count differentials observed in Brockbank & Vul (2020) for the same sequential dependencies ($r = .96$, $p < .001$).

### More complex sequential dependencies are more reliably exploited by bot agents




Average bot win count differentials vary with *memory complexity* of the bot strategies. The most complex bot strategies reliably beat human opponents, but participants *counter-exploited* the simplest bot strategies.

## SUMMARY

***We find that*** across a range of behavioral patterns exhibited in rock, paper, scissors games, people counter-exploit the simplest dependencies but can be reliably exploited using complex patterns in their move choices.

- How well bots exploit patterns in human behavior aligns with how much people exhibit these patterns in dyad play[7]
- Successful exploitation of patterns in people's behavior varies with the ***memory complexity*** of the pattern itself
- People are reliably exploited by complex patterns in their own behavior and show little ability to adapt
- For simpler patterns, people successfully ***counter-exploit*** their bot opponents

## REFERENCES

[1] Kahneman & Tversky, 1972
[2] Lopes, 1982
[3] Bar-Hillel & Wagenaar, 1991
[4] Neuringer, 1986
[5] Walker & Wooders, 2001
[6] Rapoport & Budescu, 1992
[7] Brockbank & Vul, 2020