

Globally Inaccurate Stereotypes Can Result From Locally Adaptive Exploration

Paper by: Xuechunzi Bai, Susan T. Fiske, Thomas L. Griffiths

Presentation by: Gianluca Bencomo

Outline

- Motivation
- Preliminaries
 - Multi-armed Bandits (Beta-Bernoulli)
 - Thompson Sampling
- Simulations
 - Setup + Results
- Experiments 1 + 2
 - Setup + Results
- Conclusion

Motivation

How do stereotypes emerge?

- Humans consistently exhibit inaccurate generalizations and stereotypes in the absence of real group differences. Why?
- Previous research explains stereotypes through motivational biases, cognitive limitations, and information deficits.
- These explanations assume:
 - A social identity, dominance, or threat (motivational)
 - Selective attention or prior biases (cognitive)
 - Majority or minority representation (informational)
- Many, if not most, stereotypes do not carry any of these assumptions.

Motivation

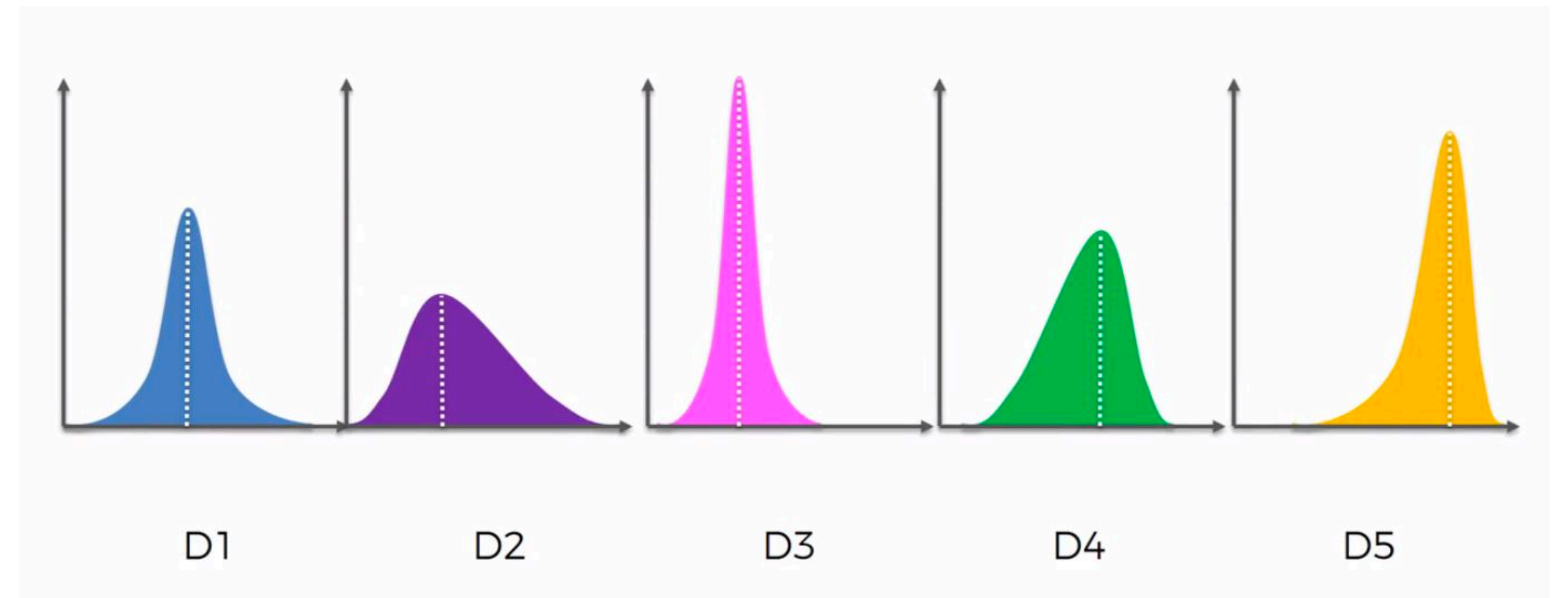
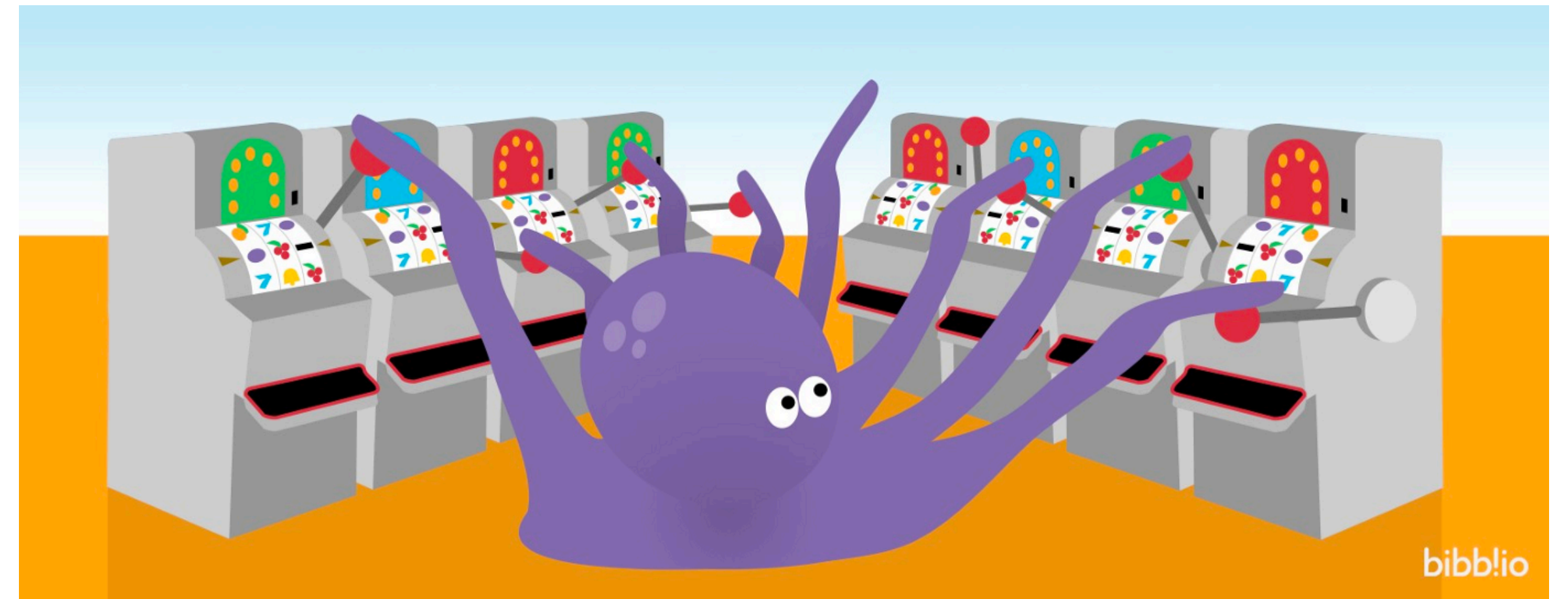
What is the most minimal, yet functional, explanation of stereotype emergence?

- Consider the following:
 - Humans are probabilistic agents in an environment plagued with uncertainty. **Goal:** maximize the long-term rewards of our actions.
- This is a valid assumption across nearly all of the psychological landscape. Can we explain stereotype emergence with just this?
- Using a multi-armed bandits model, the authors demonstrate that stereotypes can emerge in environments with no group differences due to inherent adaptive exploration strategies.

Preliminaries

beta-Bernoulli multi-armed bandits (MAB)

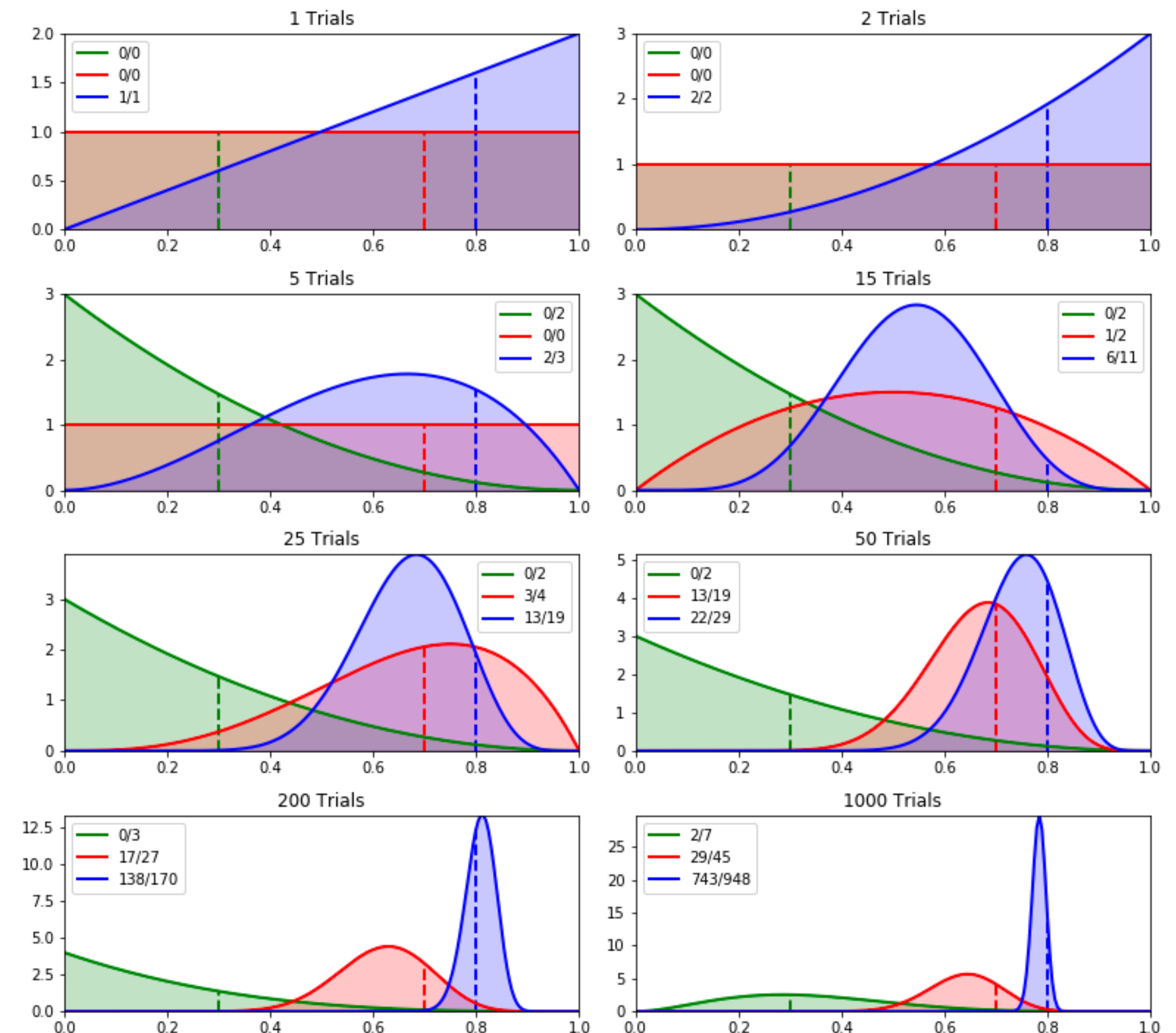
- We have a set of discrete actions we can sample (i.e. slot machines at a casino)
- Every time we sample, we are given a reward (or not), and we can update our beliefs regarding the reward structure of that action
- Goal: discover the best action to take (so we can exploit it) whilst minimizing the number of times we sample suboptimal actions



Preliminaries

Thompson Sampling

- Simple approach to MAB: choose an action with probability equal to the probability of it being the optimal choice.
- Steps:
 1. Maintain a distribution over each action that represents our beliefs over the reward structure of that action
 2. Randomly sample from each choice distribution
 3. Select the action with the highest random sample
 4. Update beliefs about reward structure, given the returned reward. Repeat step 2.



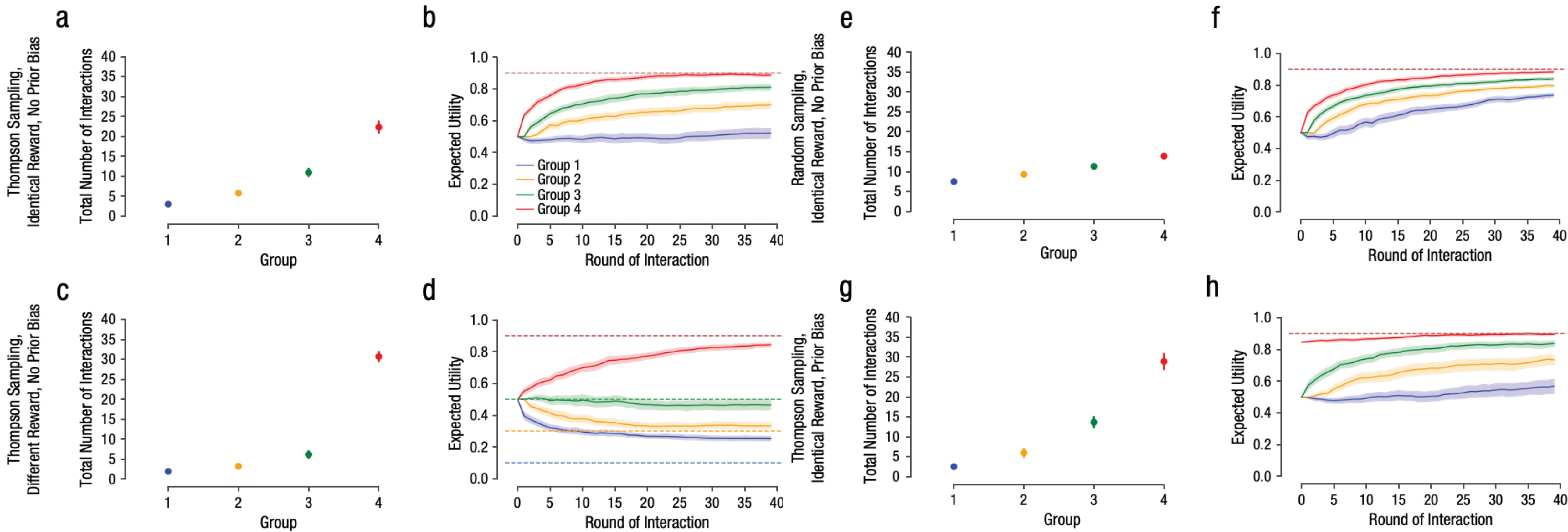
Simulations

Setup

- To understand how inaccurate impressions can emerge from locally adaptive exploration, the authors simulated three MAB variants:
 - **Variant 1:** Thompson sampling (locally adaptive) **versus** random sampling (non-adaptive)
 - **Variant 2:** Identical rewards **versus** non-identical rewards
 - **Variant 3:** Strong priors **versus** no priors

Simulations

Results



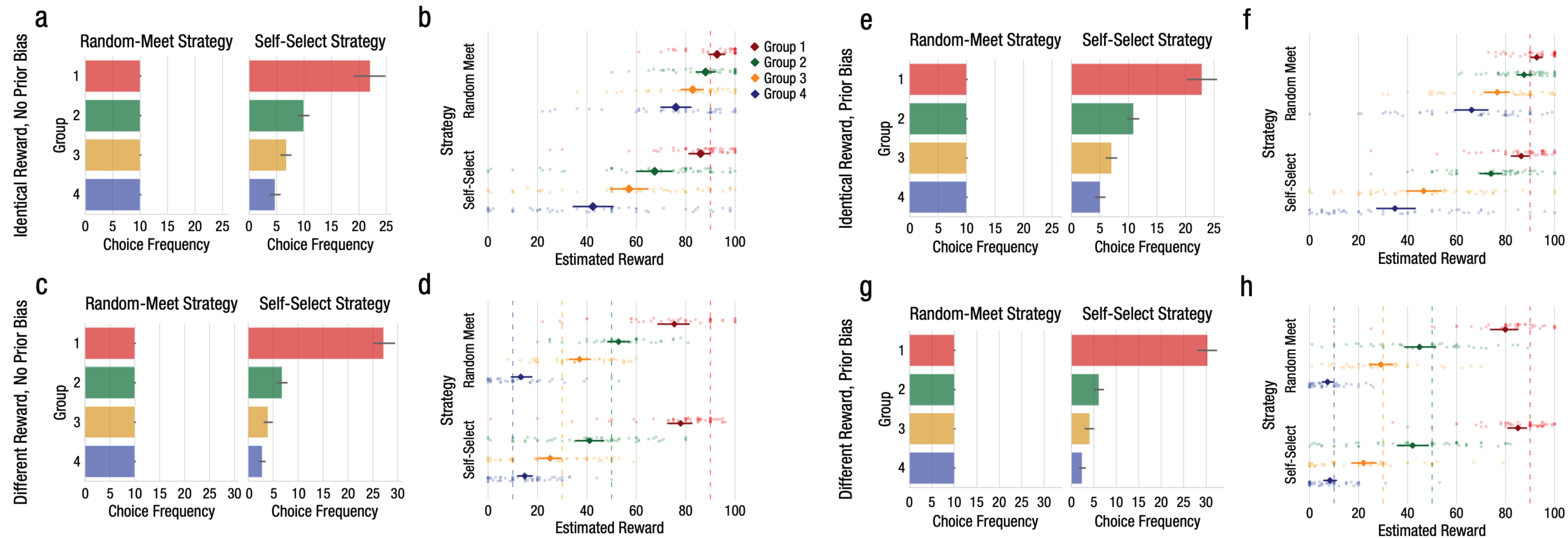
Experiments

Experiment 1: Explore Toma City (setup)

- Human participants are visiting a fictional city, starting a business, and are interacting with **four** groups of people, who can help them with their business. Over 40 rounds, participants were allowed to select one group to help. After 40 rounds, impressions of the four groups were collected.
- Three variants were tested:
 - **Variant 1:** Random-meet **versus** Self-select
 - **Variant 2:** Identical rewards **versus** Non-identical rewards
 - **Variant 3:** Strong priors **versus** No priors

Experiments

Explore Toma City (results)



(Bai et. al., 2022)

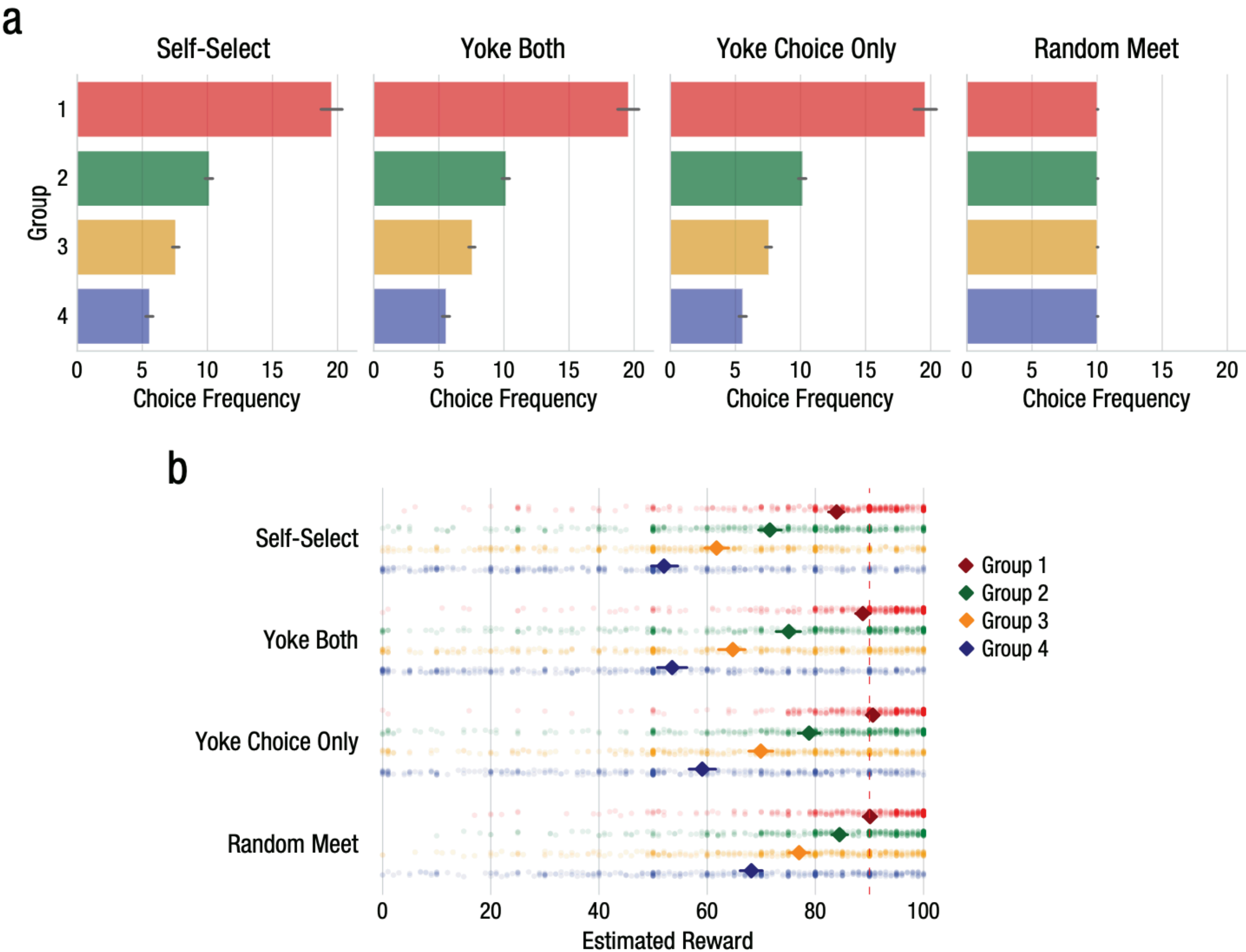
Experiments

Experiment 2 (setup)

- This experiment tested two more sampling strategies to explore the role of active versus passive learning
- Do the selective samples make participants biased, or does the sense of control make them biased?
- Two **new** conditions:
 - yoke-both (another participant selects choice)
 - yoke-choice only (another participant selects choice, rewards are permuted)

Experiments

Experiment 2 (results)



Conclusion

- In the absence of any priors, motivational biases, cognitive limitations, or information deficits, simulations showed that **Thompson sampling** can estimate one group as being better than other groups despite all groups being equally good
- Experiment 1 showed that human behavior, under the same assumptions, reflects a Thompson Sampling procedure
 - A strong empirical argument for inaccurate impressions, or stereotypes, emerging from locally adaptive exploration
- Experiment 2 demonstrated a lower bias in two passive learning scenarios

Conclusion

(Continued)

- Adaptive exploration is functional, but may cause collateral damage to the unexplored groups
- Under this paradigm, we only require two conditions to create biased impressions
 - Choices are adaptive
 - All groups are equal

References

- Bai, X., Fiske, S. T., & Griffiths, T. L. (2022). Globally Inaccurate Stereotypes Can Result from Locally Adaptive Exploration. *Psychological Science*, 09567976211045929.
- Slivkins, A. (2019). Introduction to multi-armed bandits. *Foundations and Trends® in Machine Learning*, 12(1-2), 1-286.
- *I wrote some very rudimentary code that could be instructional for better understanding MABs, Thompson sampling, and two other explore/exploit policies. Feel free to check it out:*
 - <https://github.com/gianlucabencomo/454MAB-supplemental>