

Globally Inaccurate Stereotypes Can Result from Locally Rational Exploration:

Evidence from a Formal Model and Human Experiments



Xuechunzi Bai



Susan T. Fiske



Thomas L. Griffiths

(invited revision *PsychSci*)

Motivation

Inaccurate stereotypes are prevalent and consequential



Motivation

Inaccurate stereotypes are prevalent and consequential

WHY do people develop inaccurate stereotypes in the first place?

Existing research

Motivational and Cognitive



Existing research

Motivational and Cognitive

- **Motivational (as a group member):**
 - Identity
 - Dominance

Social identity theory (Tajfel & Turner, 1979)
In-group favoritism (Brewer, 1999)
Social dominance theory (Sidanius & Prato, 1999)
System justification theory (Jost & Banaji, 1994)

Existing research

Motivational and Cognitive

- **Motivational (as a group member):**

- Identity
- Dominance

Social identity theory (Tajfel & Turner, 1979)
In-group favoritism (Brewer, 1999)
Social dominance theory (Sidanius & Prato, 1999)
System justification theory (Jost & Banaji, 1994)

- **Cognitive (as an information processor):**

- Categorization
- Selective attention

Error-prone heuristics (Tversky & Kahneman, 1974)
Cognitive miser (Fiske & Taylor, 1984)
Classification in line judgment (Tajfel & Wilkes, 1963)
Illusory correlation (Hamilton & Gifford, 1976)

Our work: A functional minimal-process paradigm

A simple exploratory sampling

Our work: A functional minimal-process paradigm

A simple exploratory sampling

- **Minimal:**
 - Without motivational bias
 - Without cognitive limits

Our work: A functional minimal-process paradigm

A simple exploratory sampling

- **Minimal:**
 - Without motivational bias
 - Without cognitive limits
- **Functional:**
 - Optimal solution in the environment given

Our hypothesis

Globally inaccurate stereotypes can result from locally rational exploration

Our hypothesis

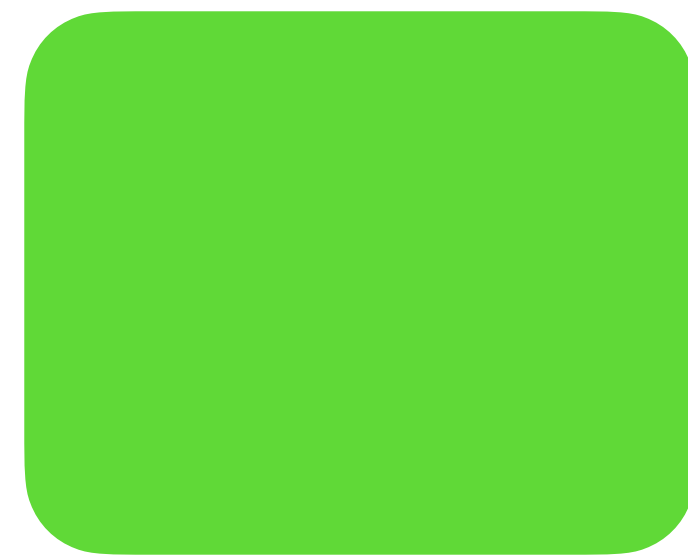
Globally inaccurate stereotypes can result from locally rational exploration

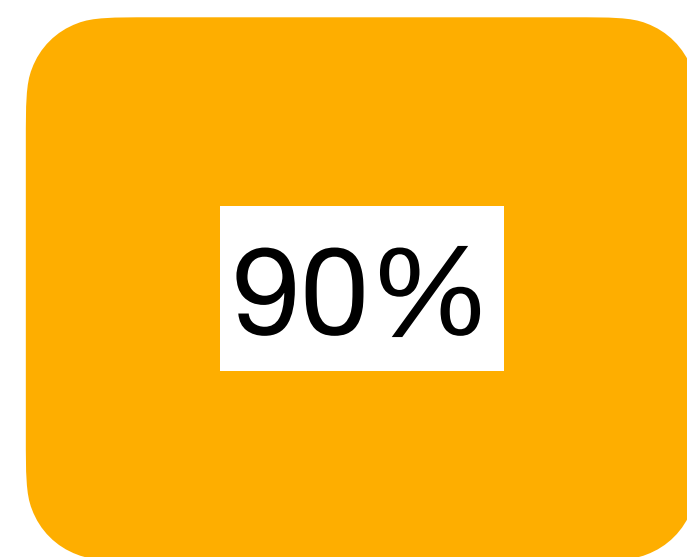
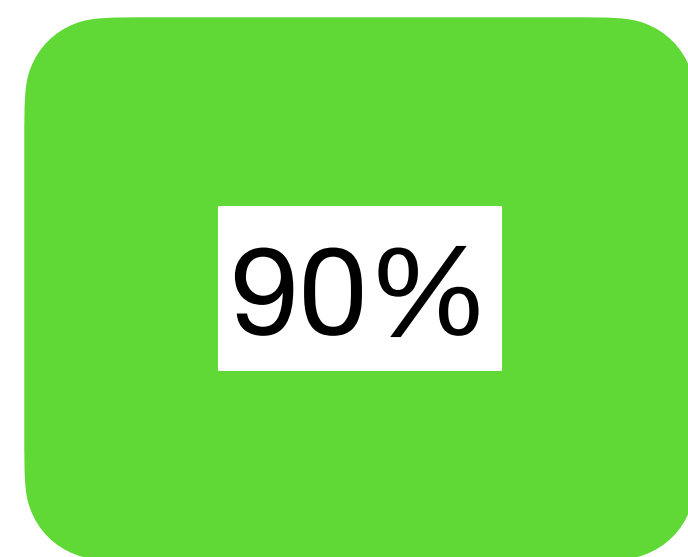
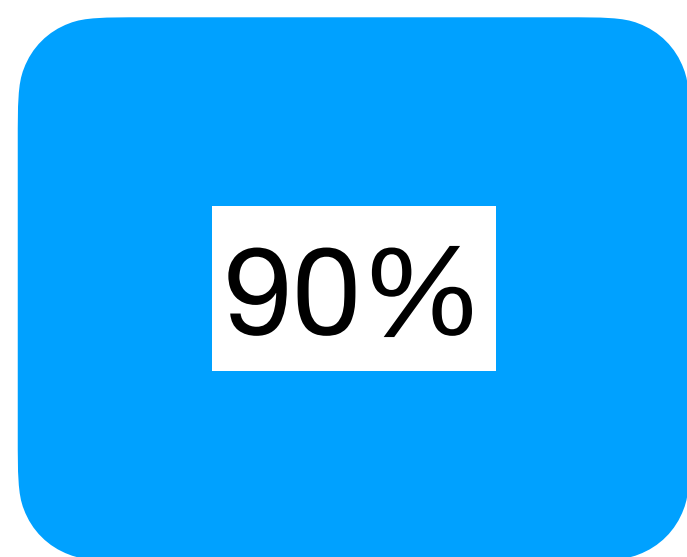
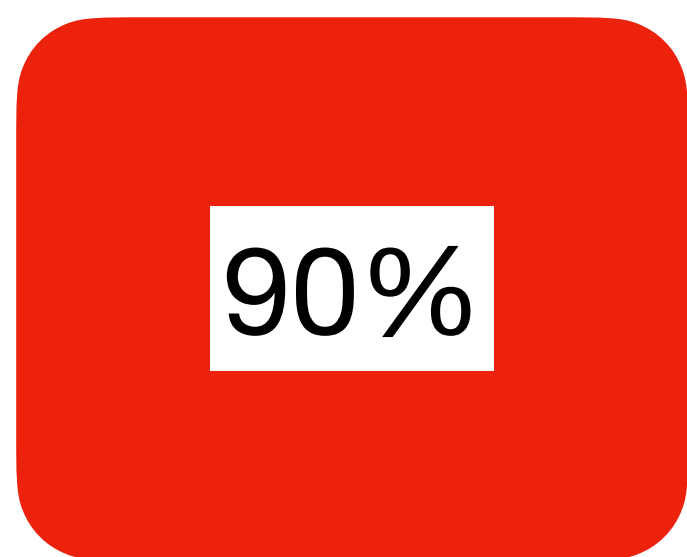
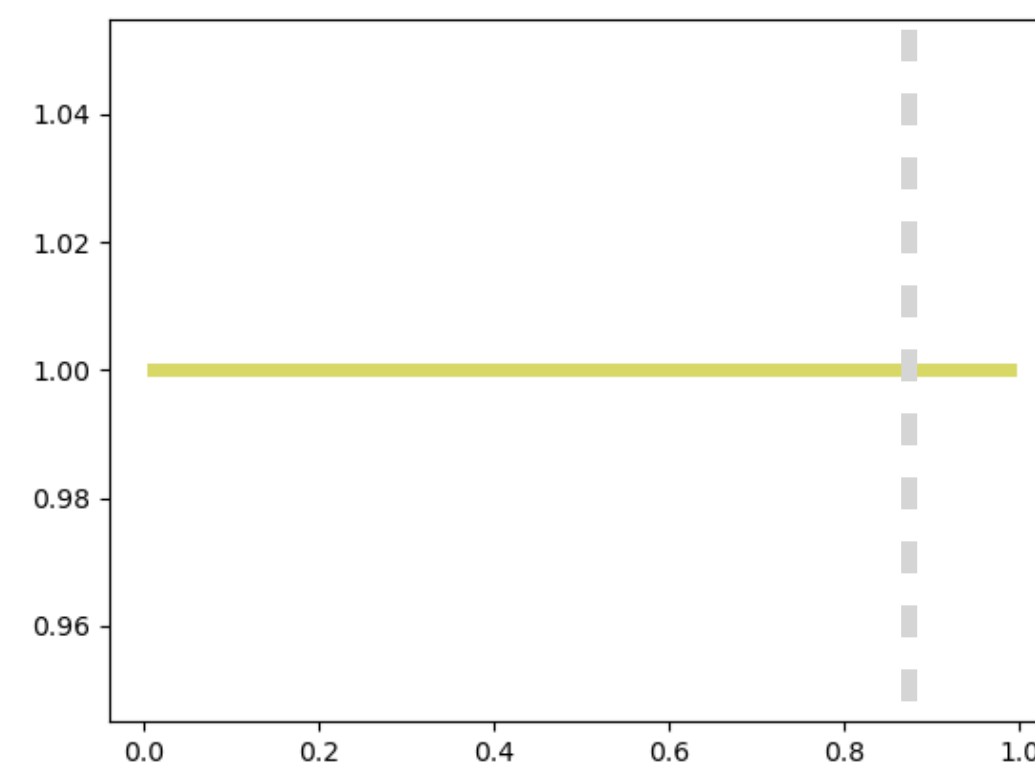
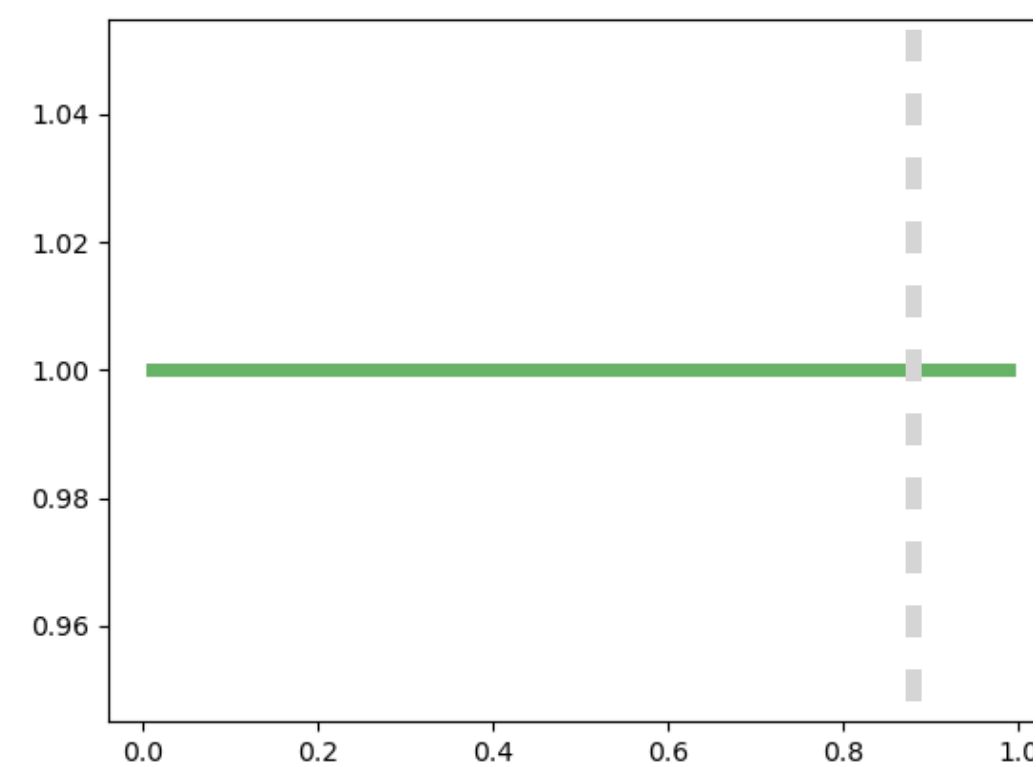
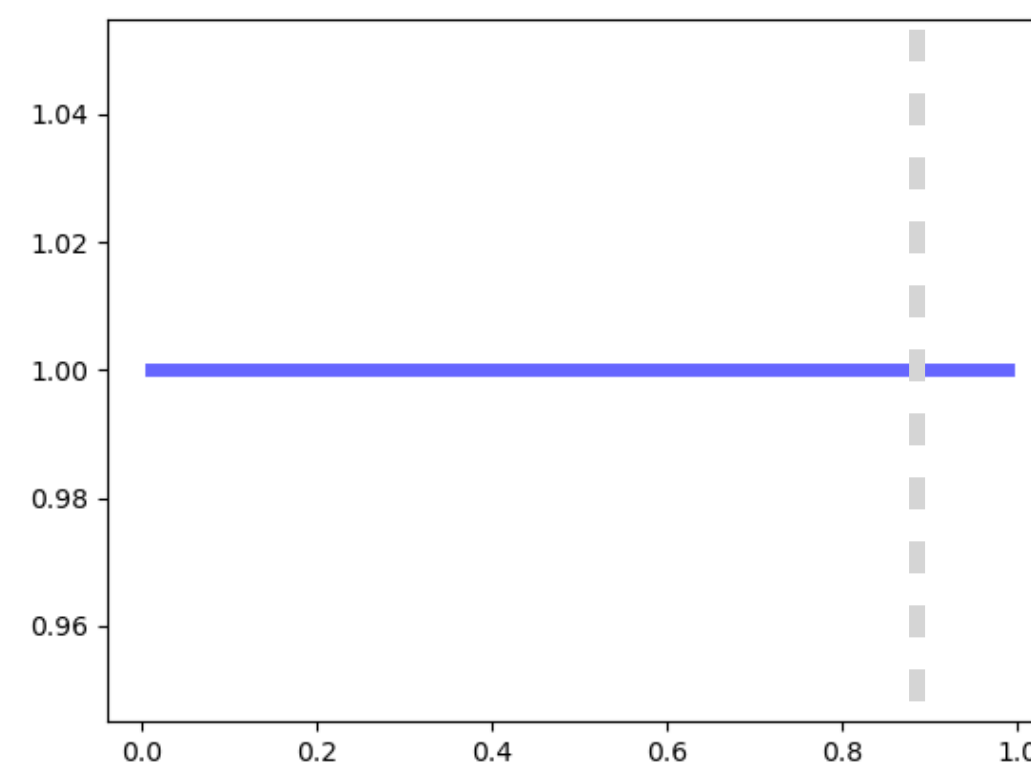
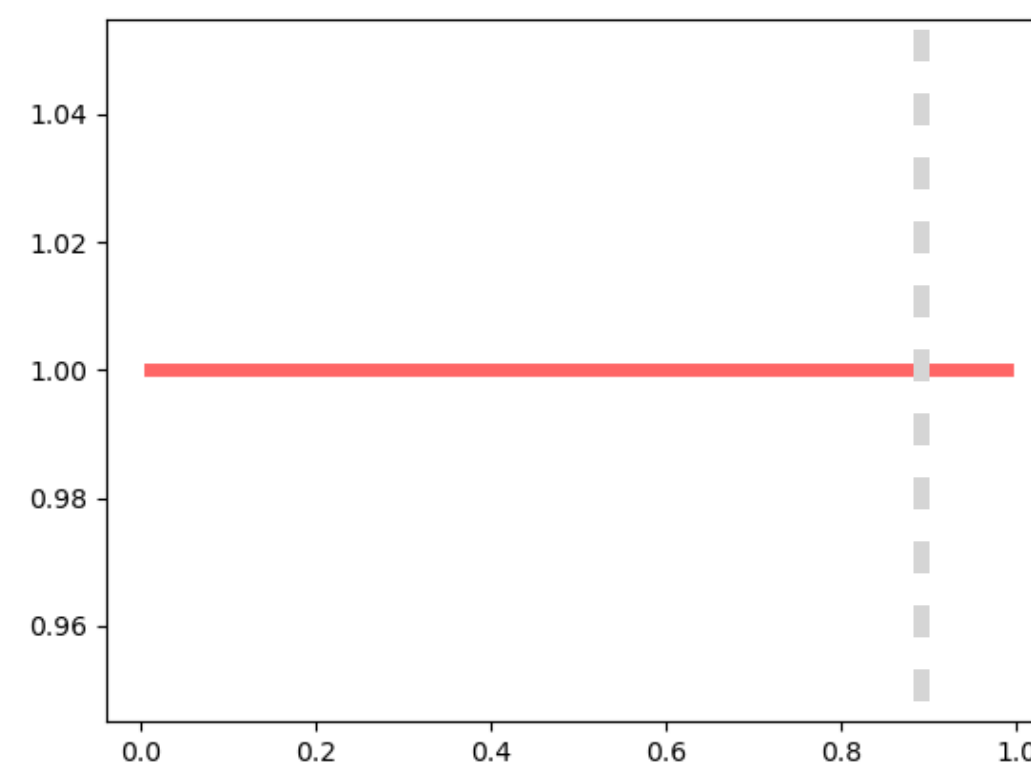
**By adopting a functional/rational analysis,
we are NOT saying social stereotypes are accurate or morally right.**

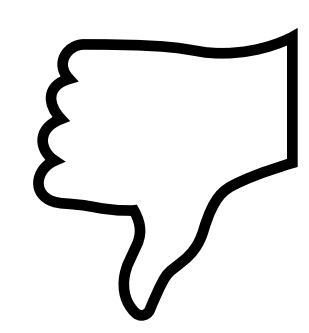
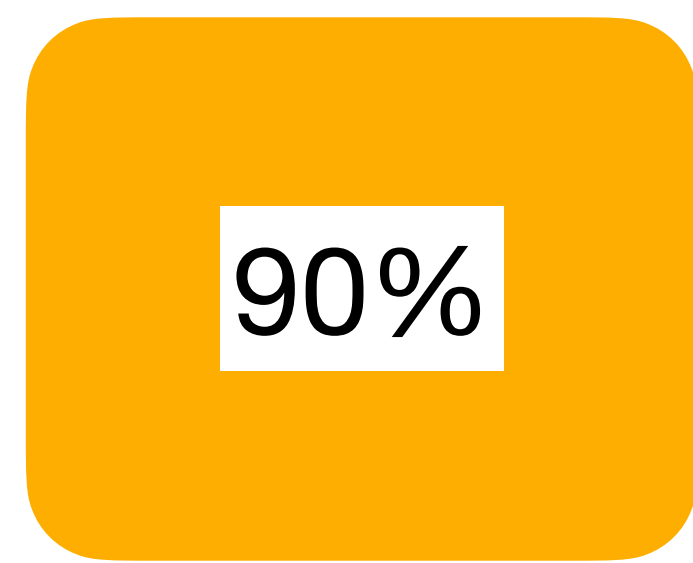
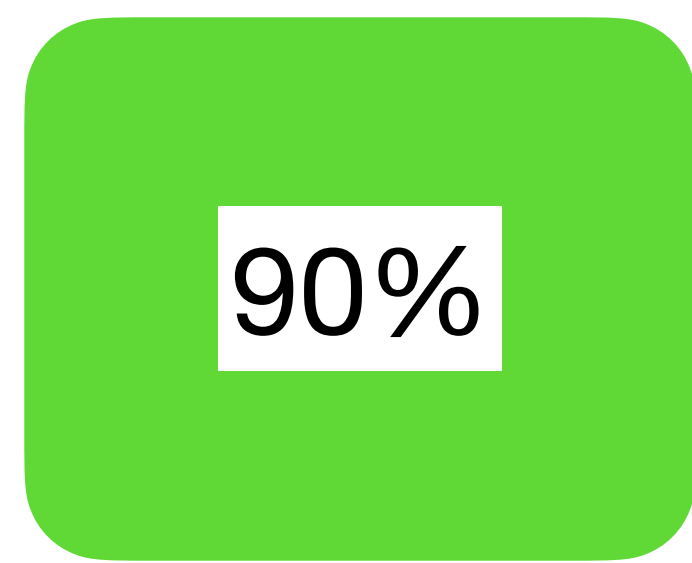
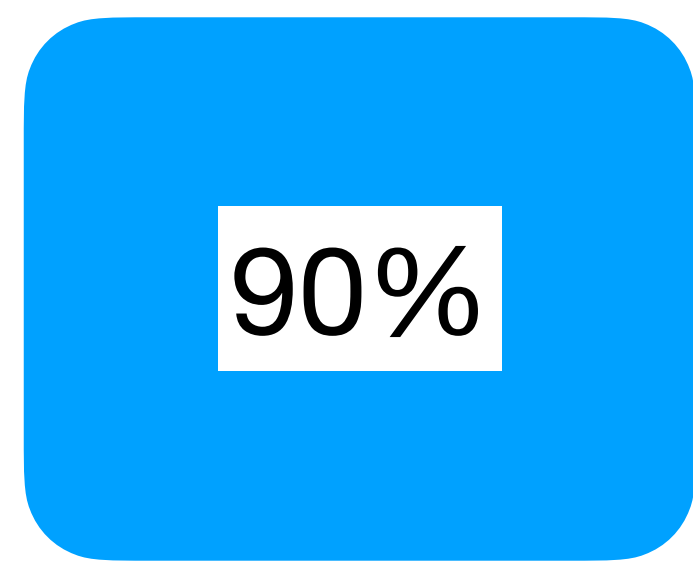
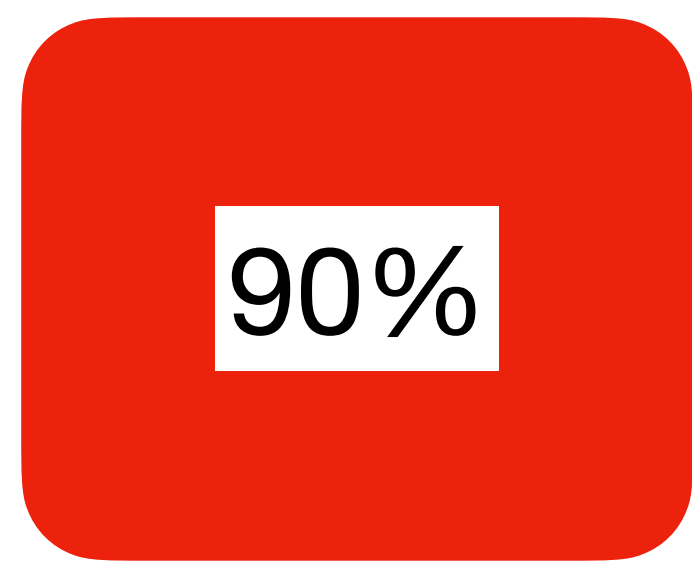
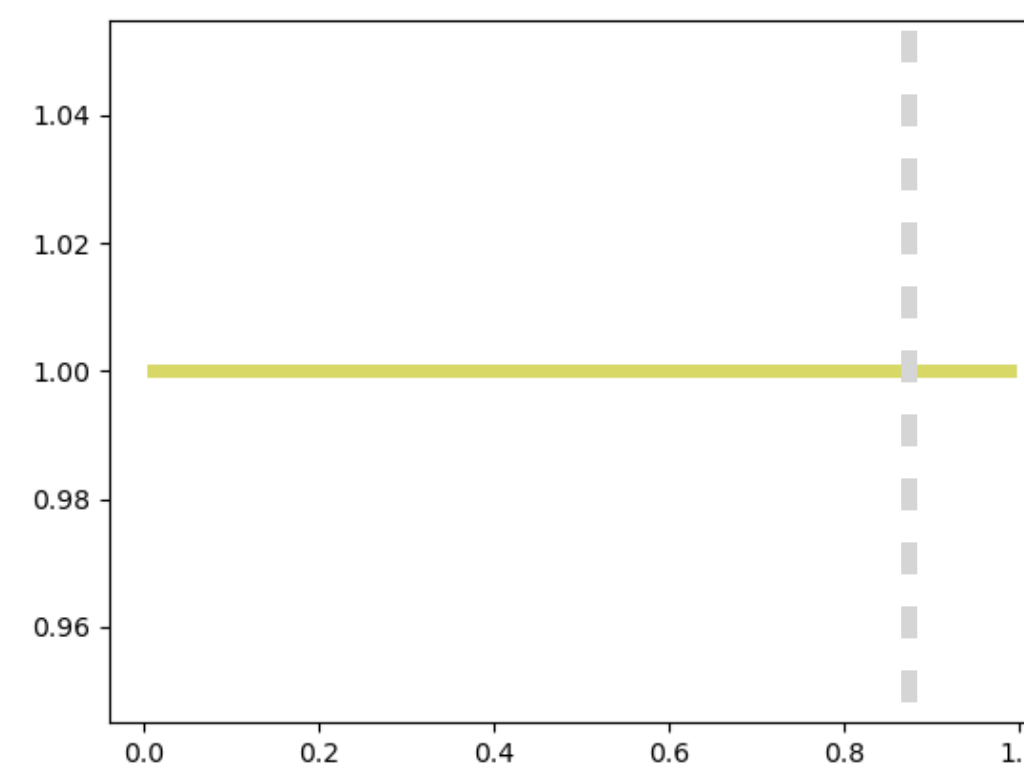
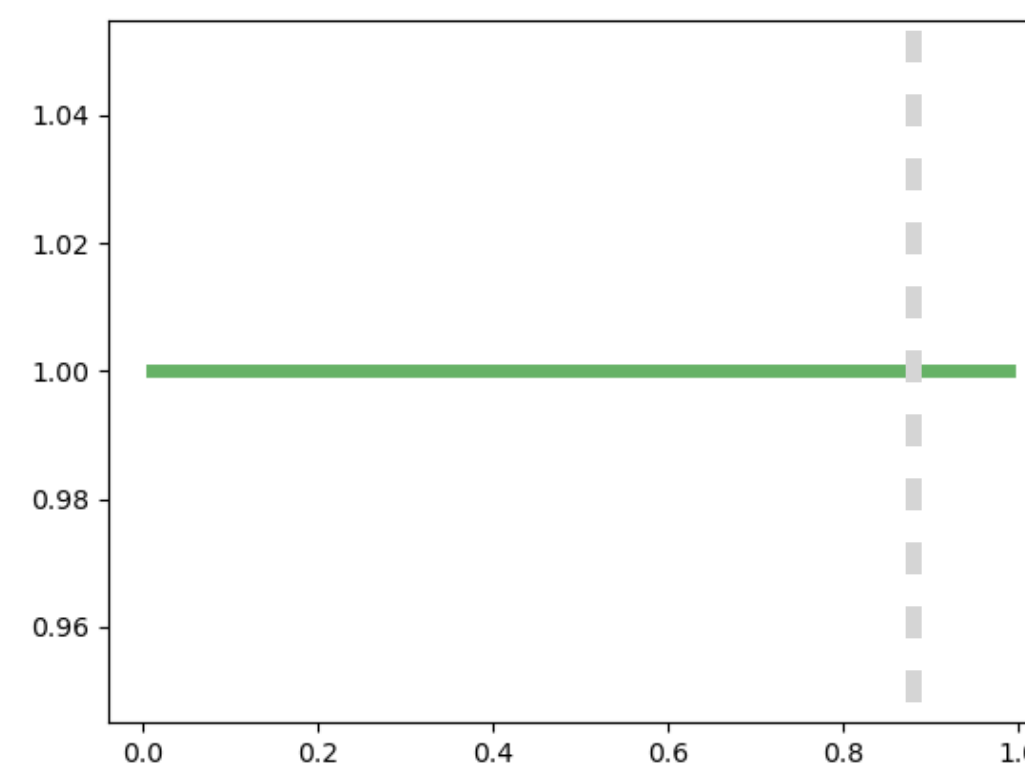
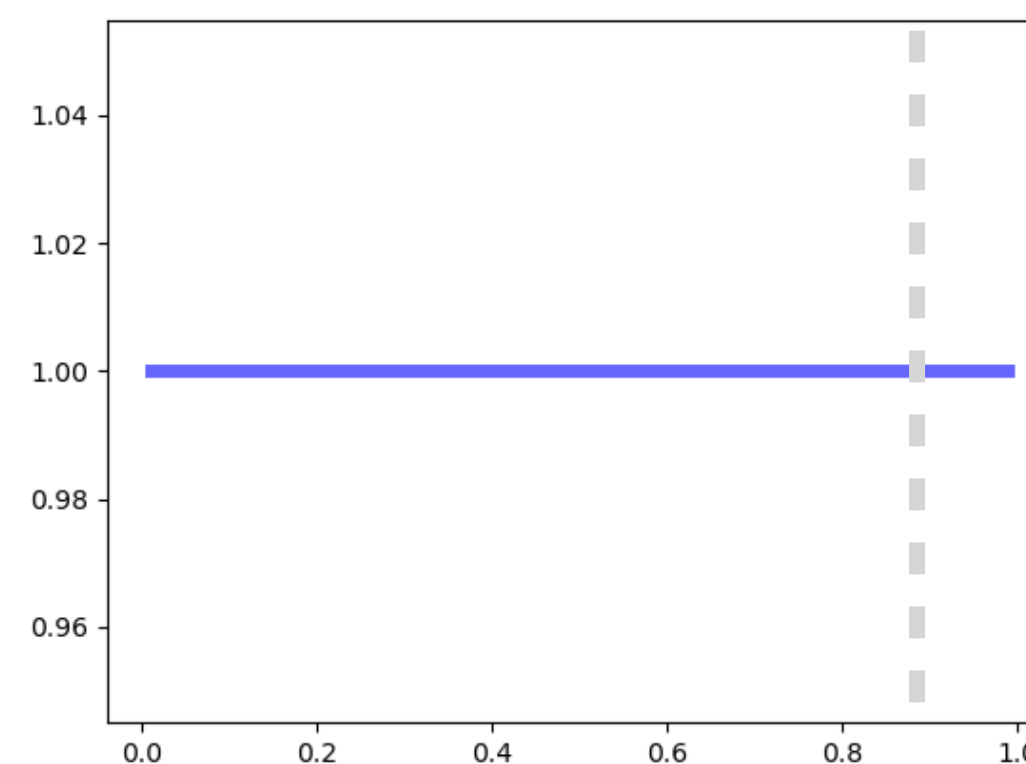
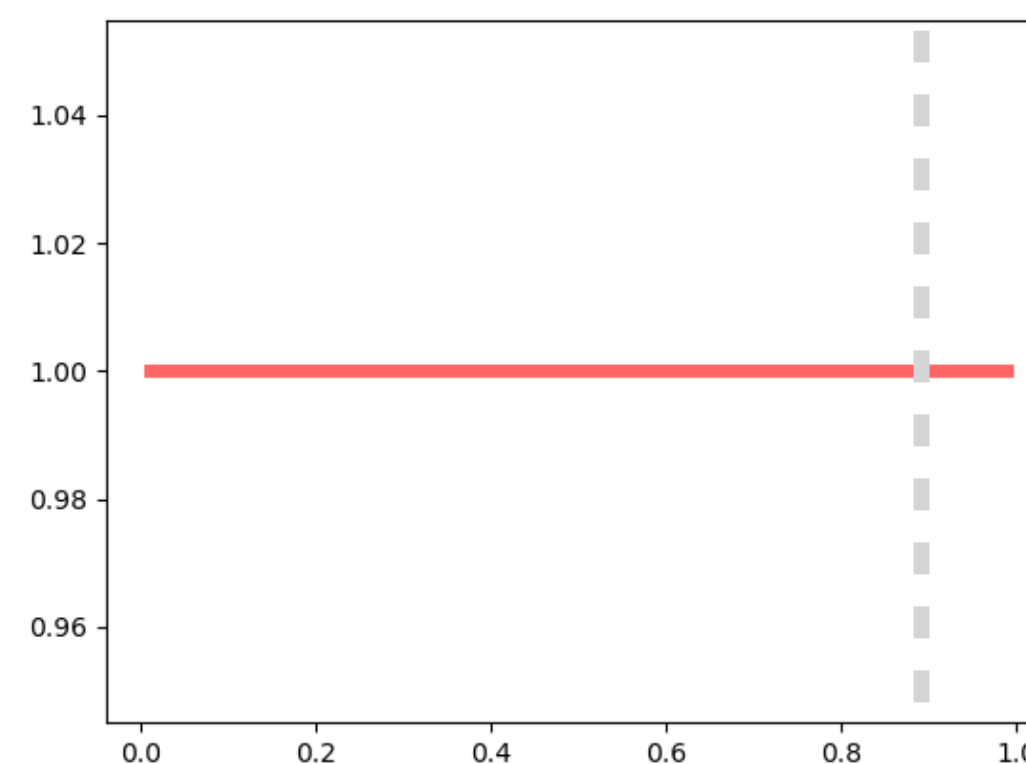
Our hypothesis

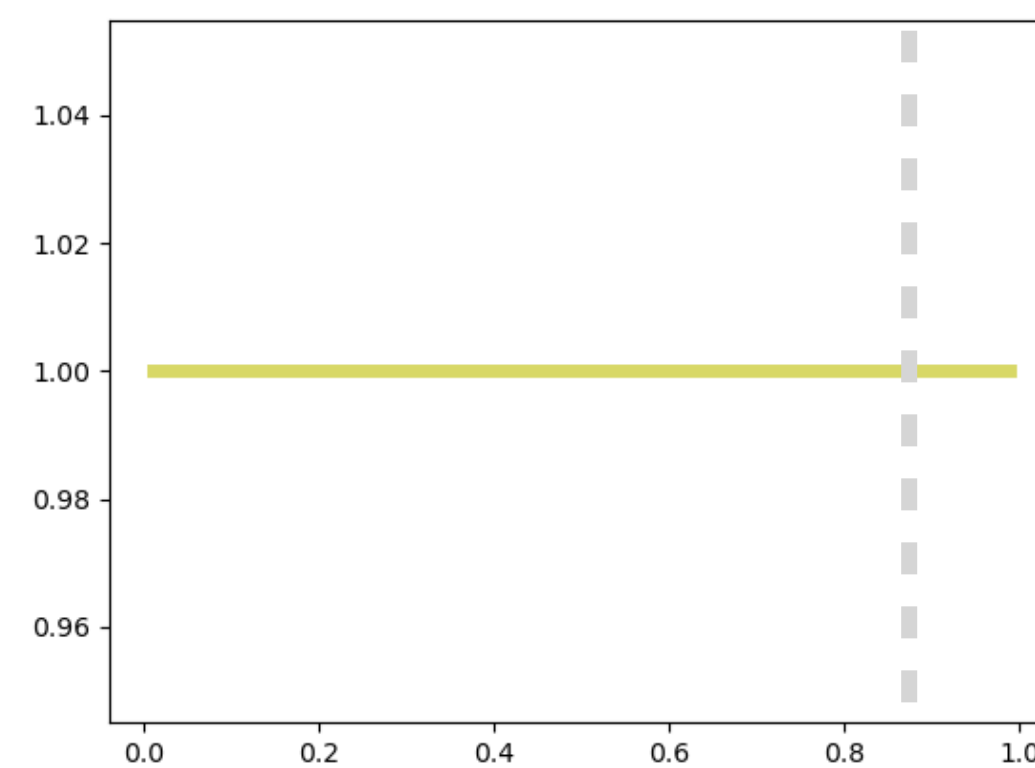
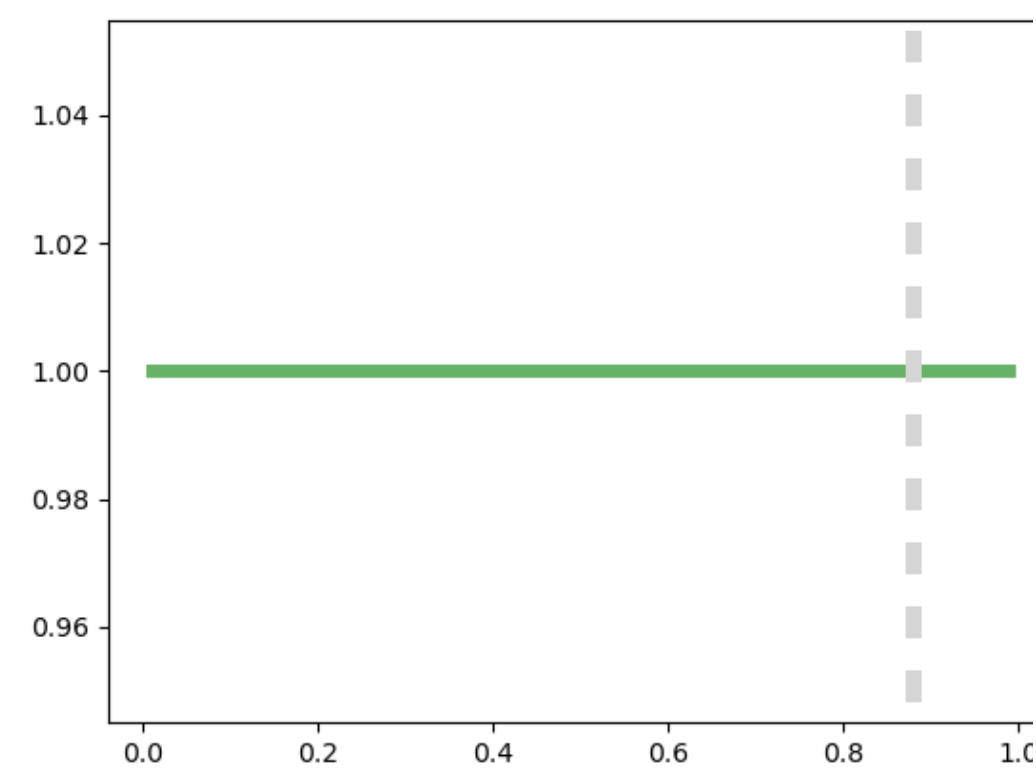
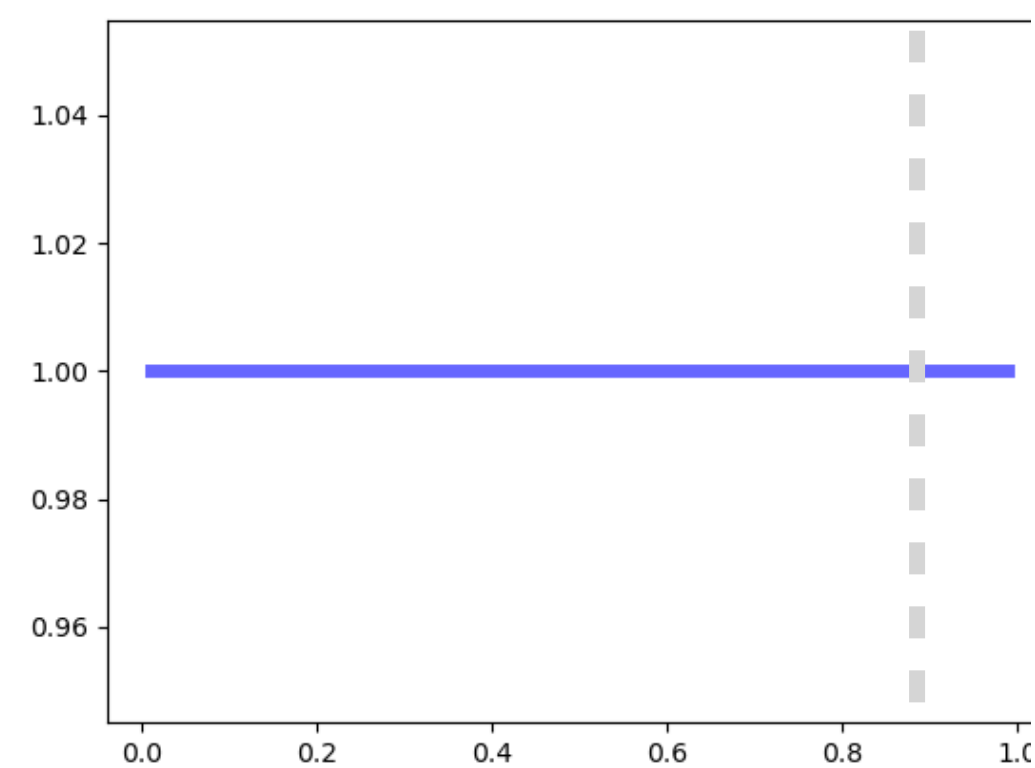
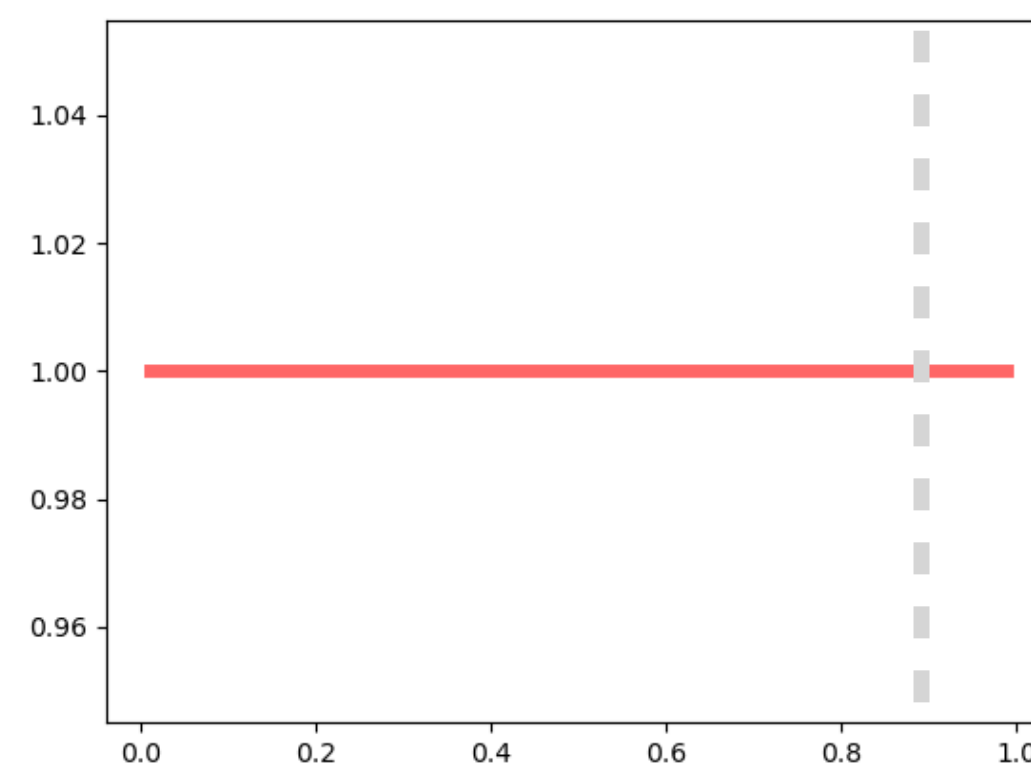
Globally inaccurate stereotypes can result from locally rational exploration

- **An intuitive example**
- A formal model and simulations
- Two human experiments
- Implications

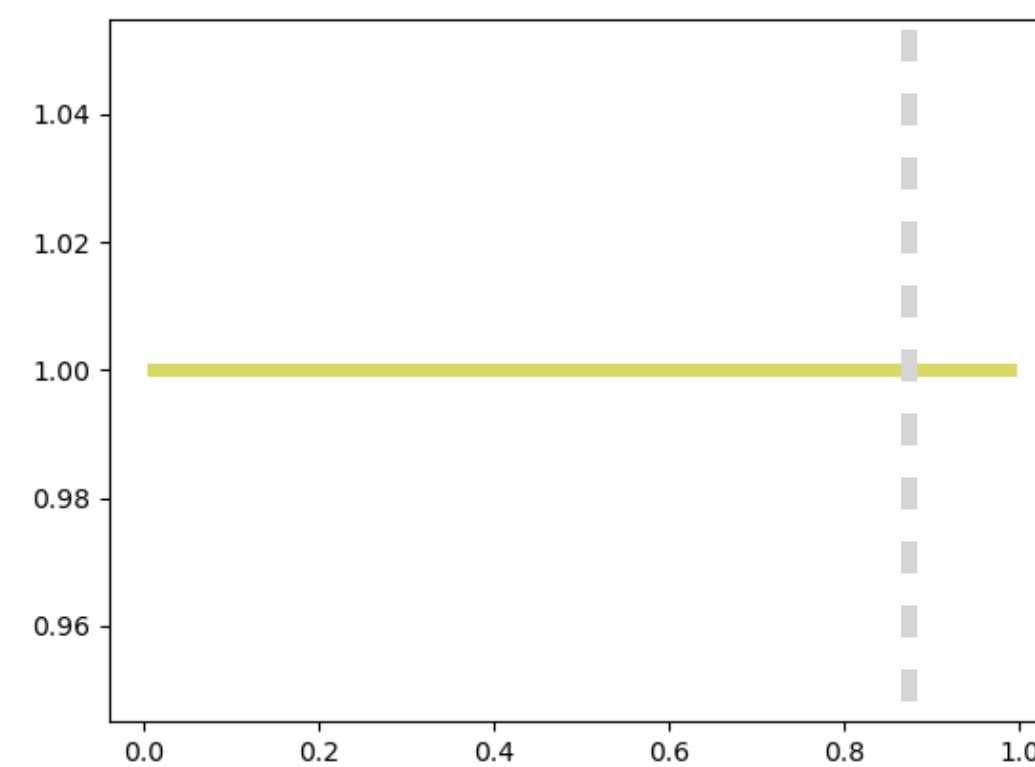
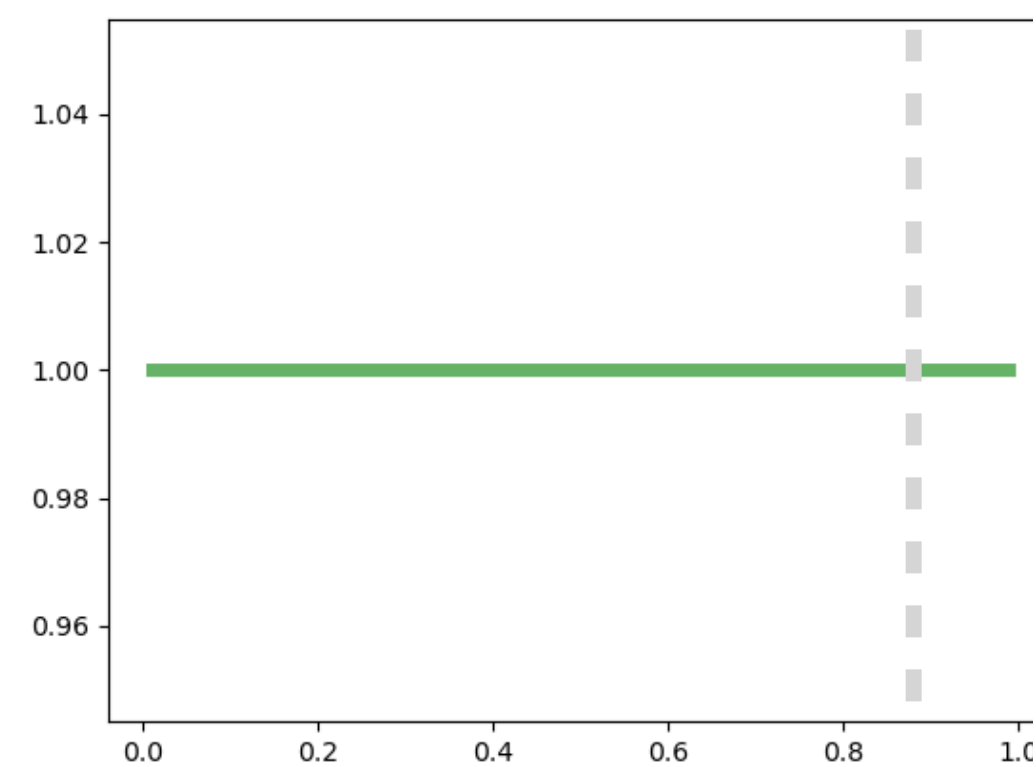
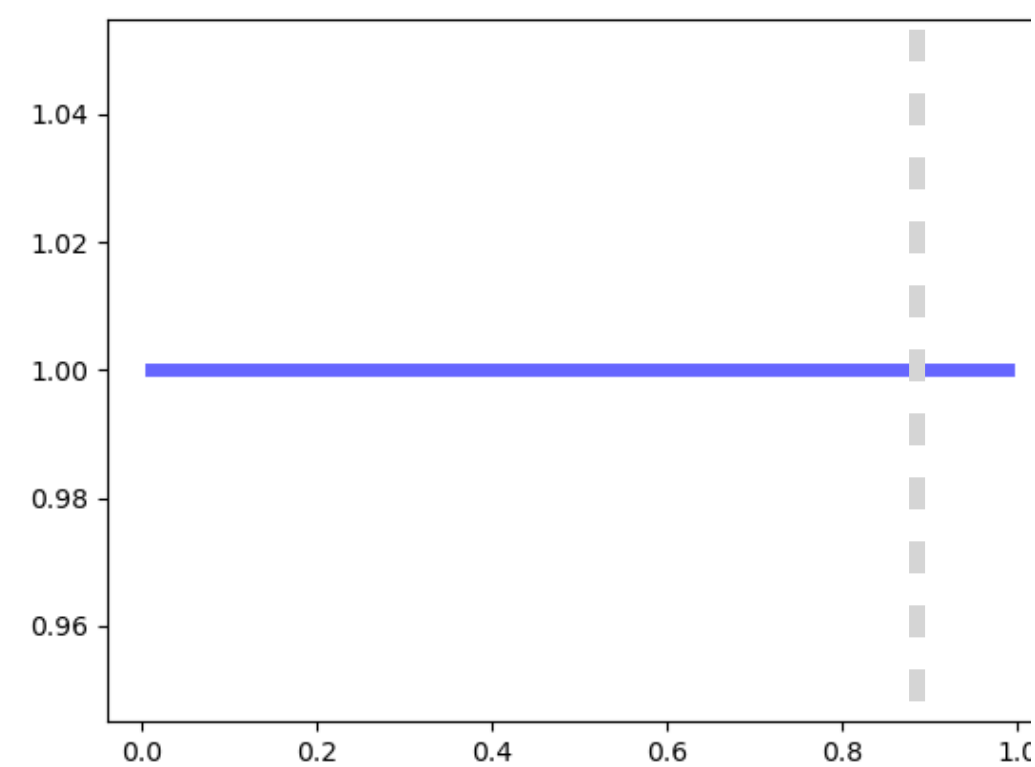
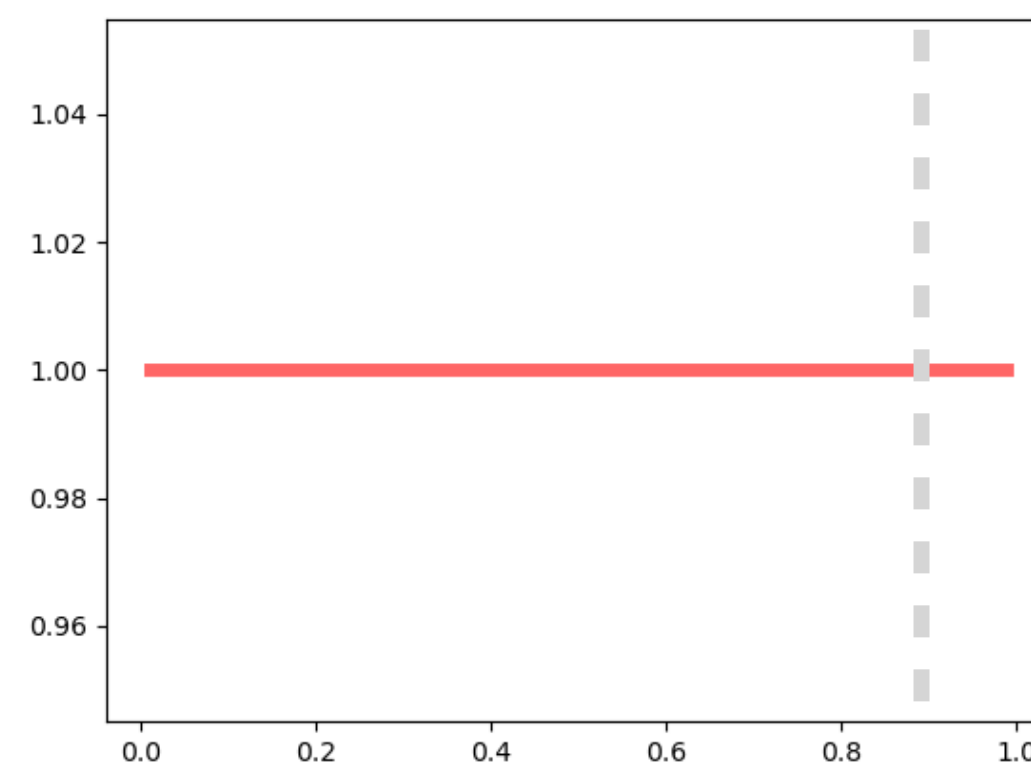


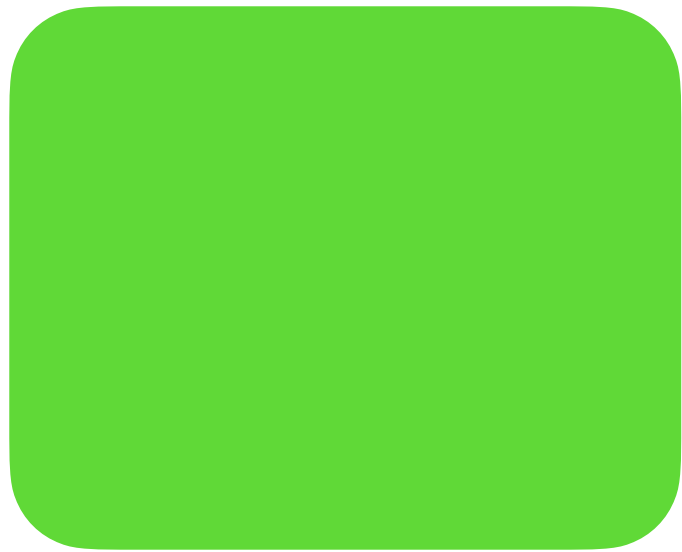
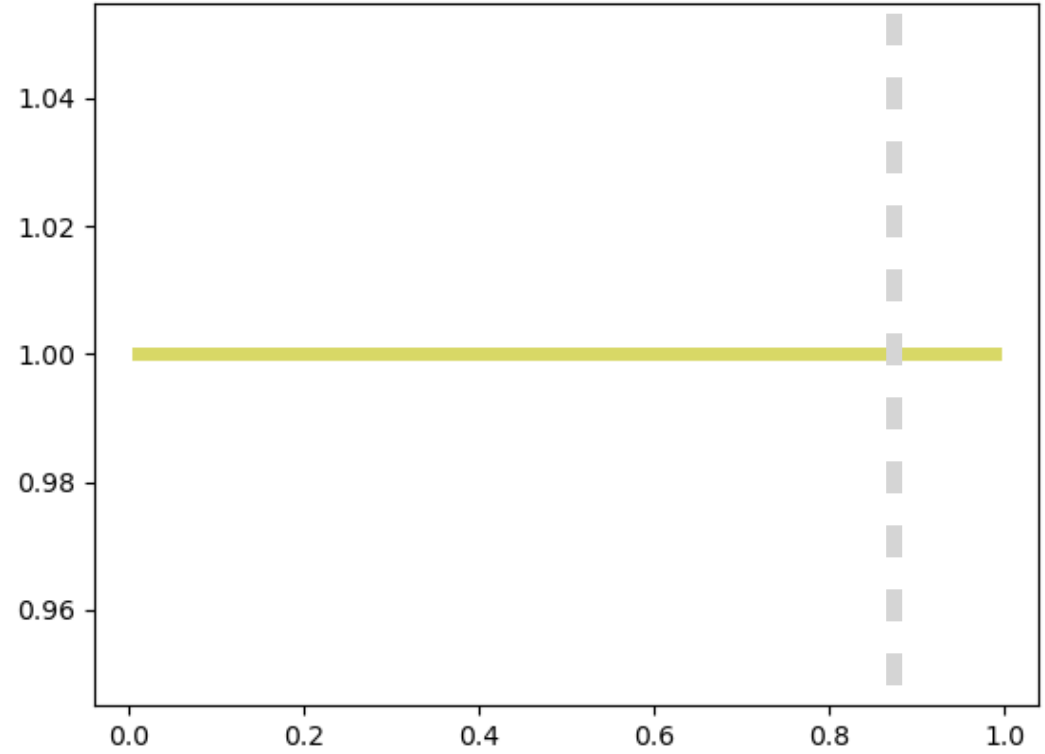
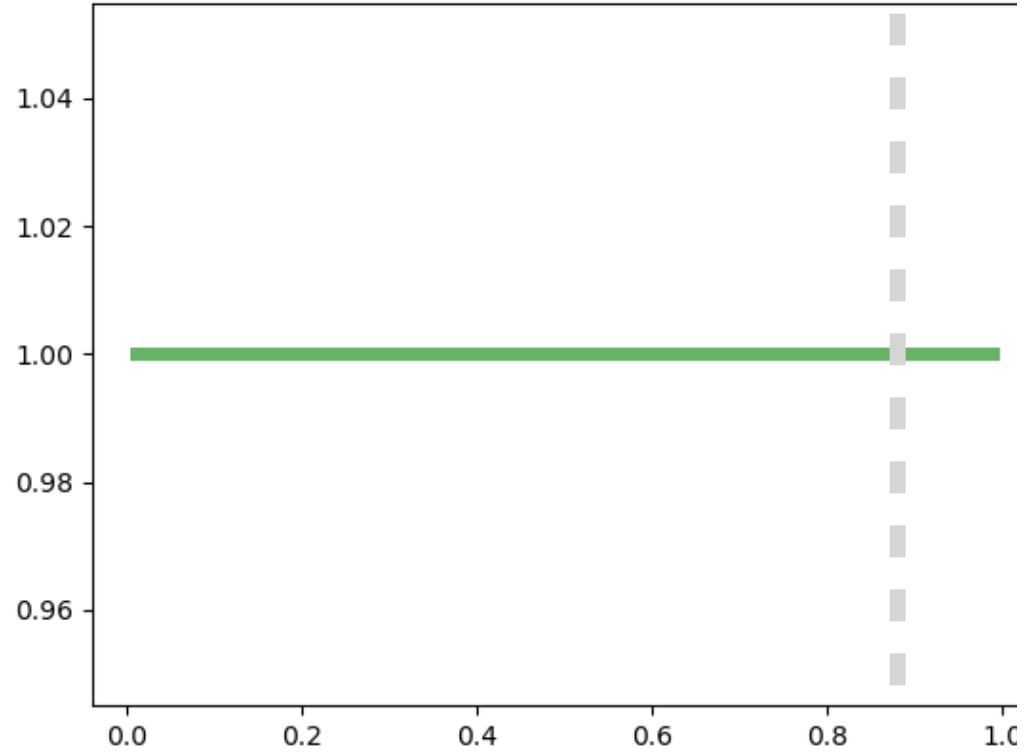
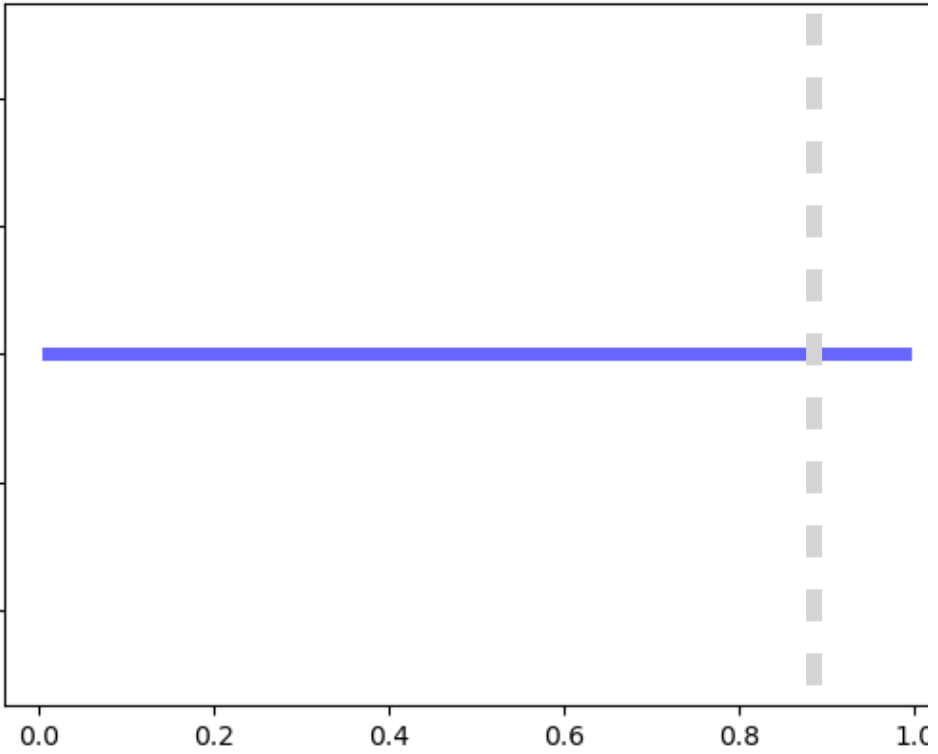
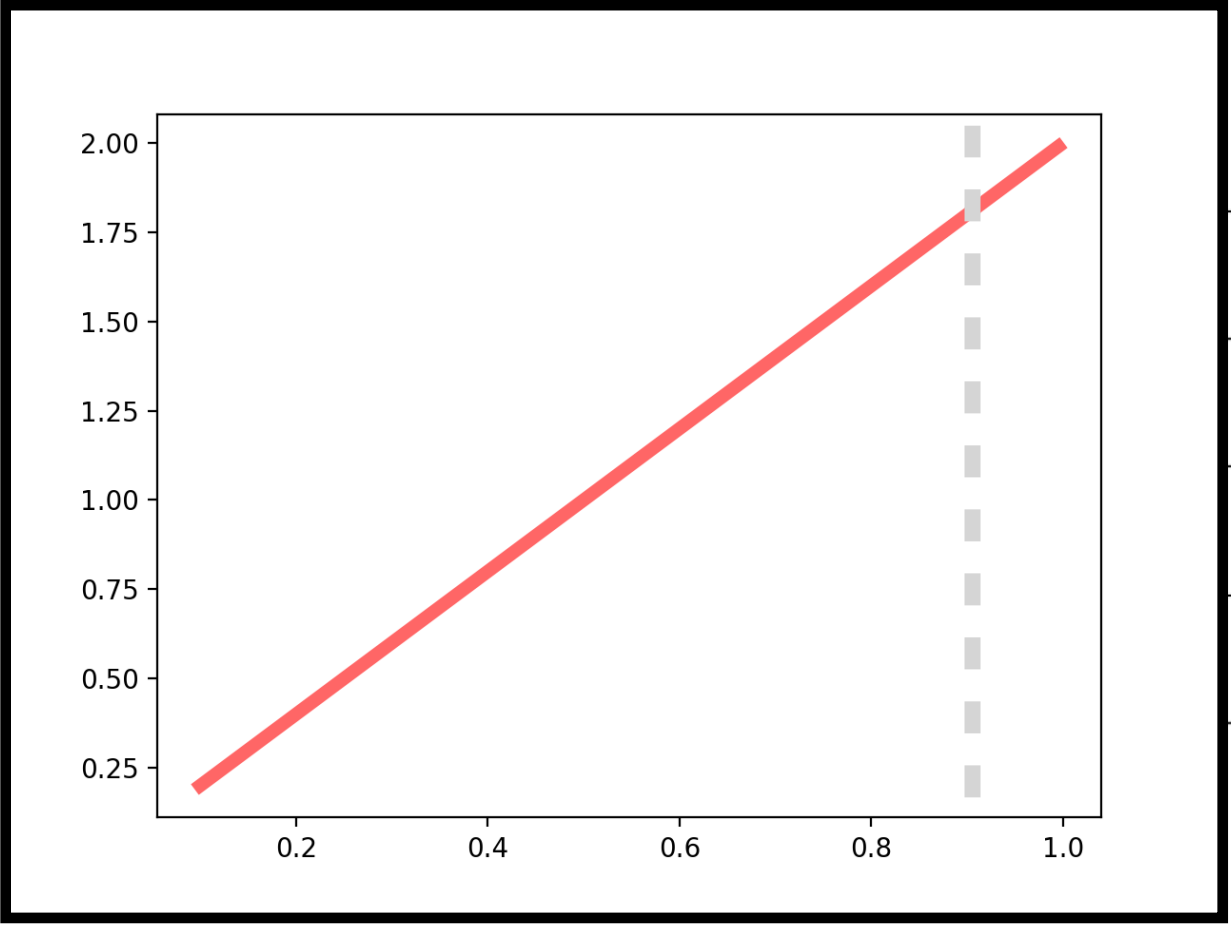


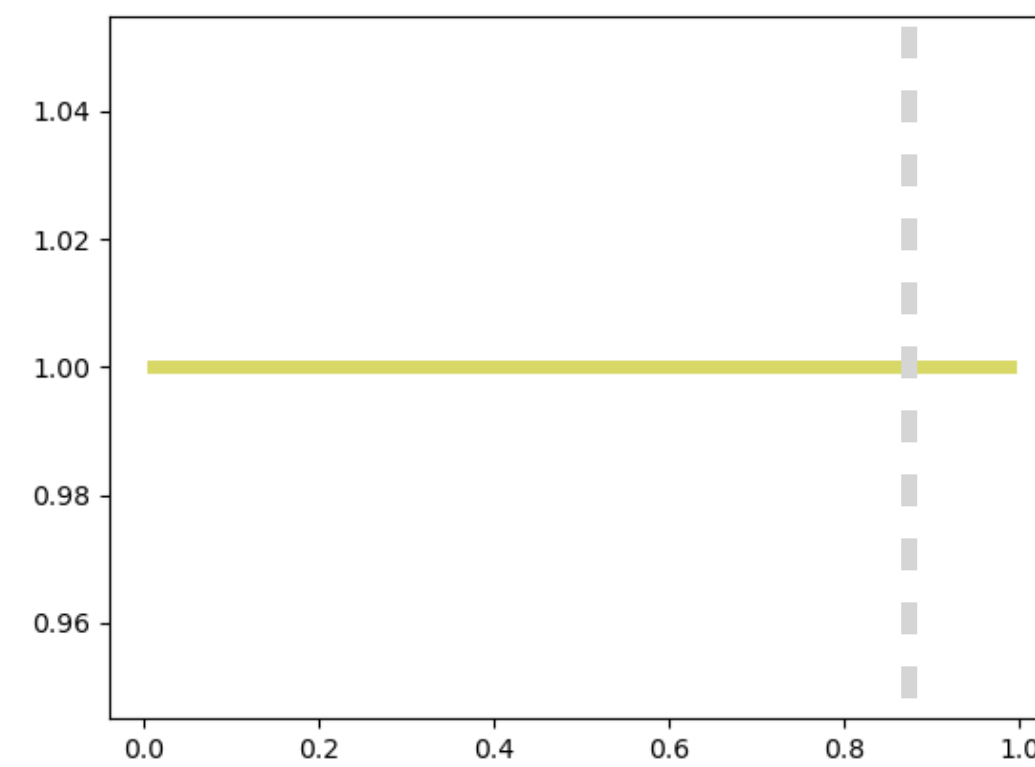
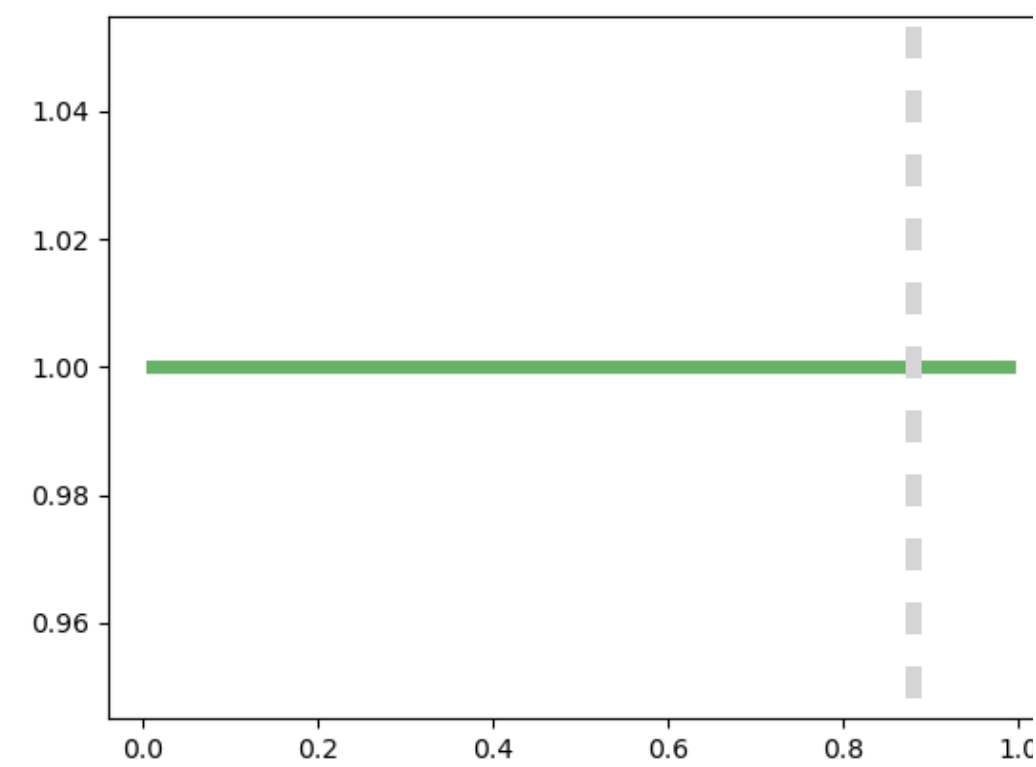
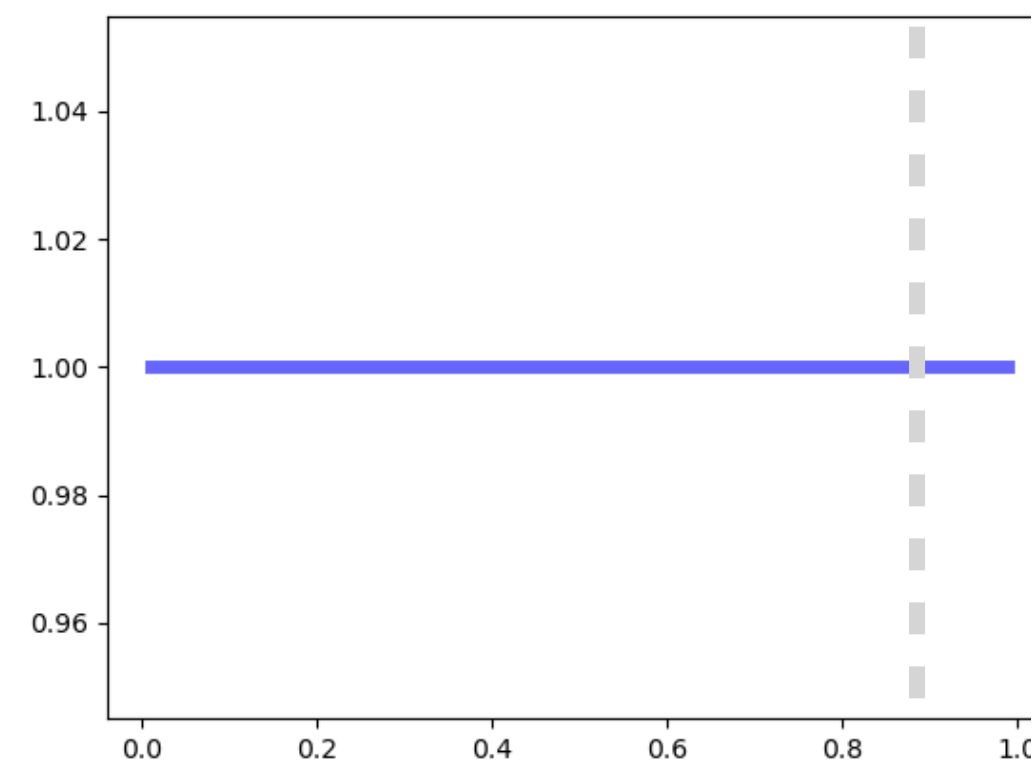
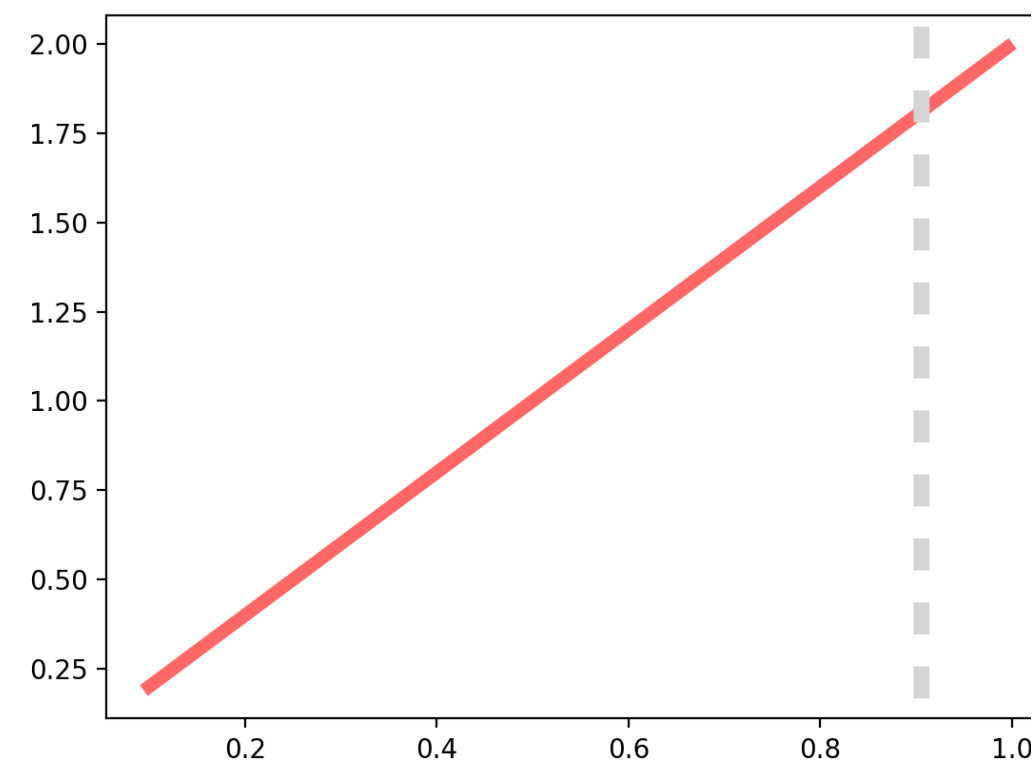




1

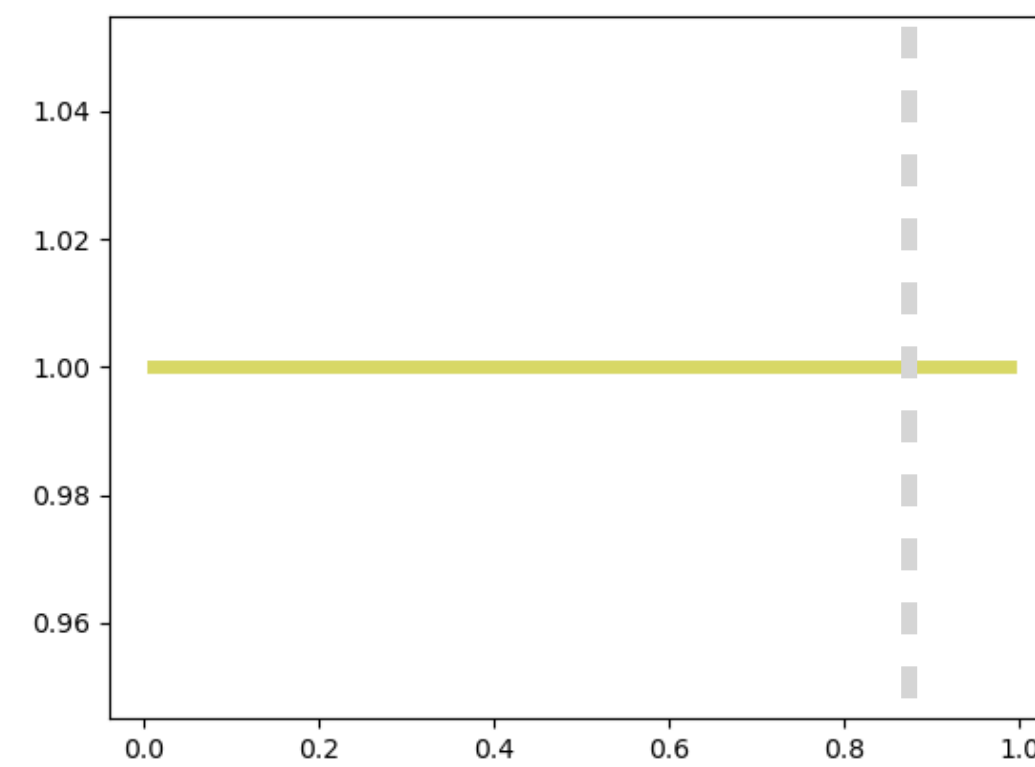
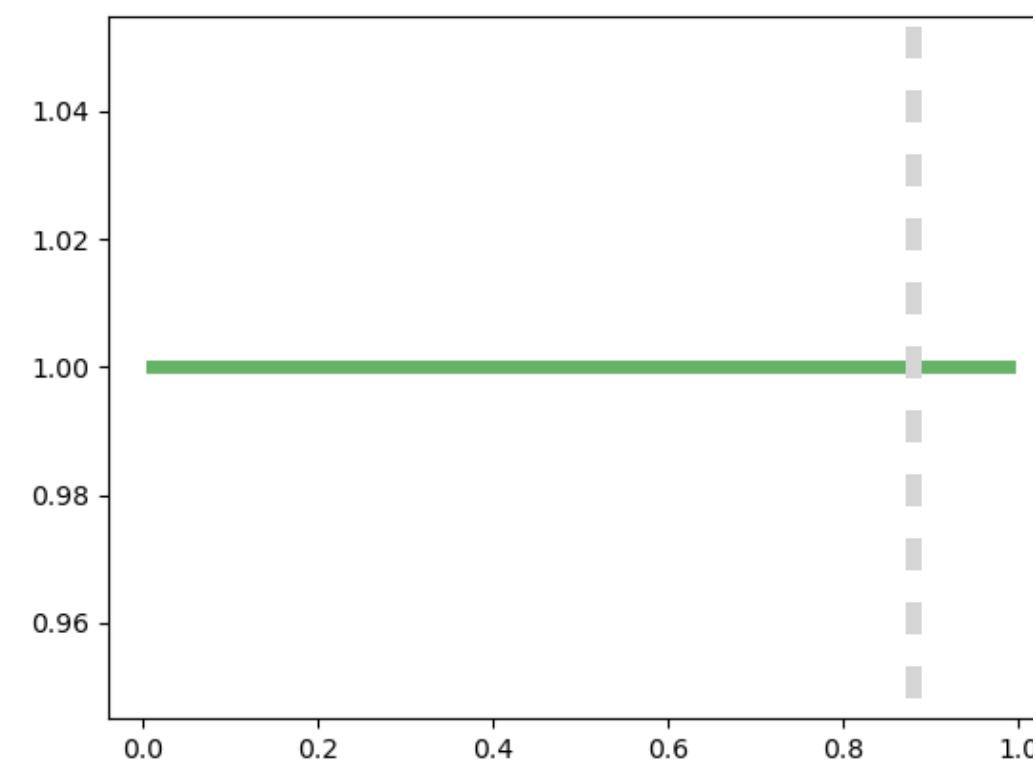
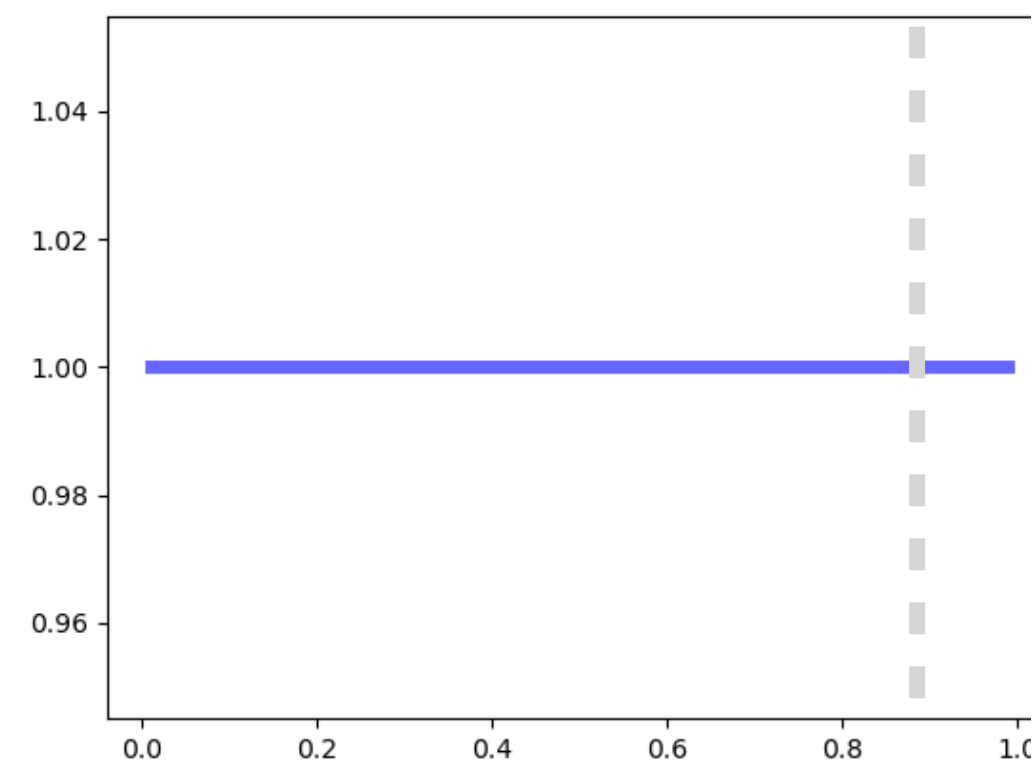
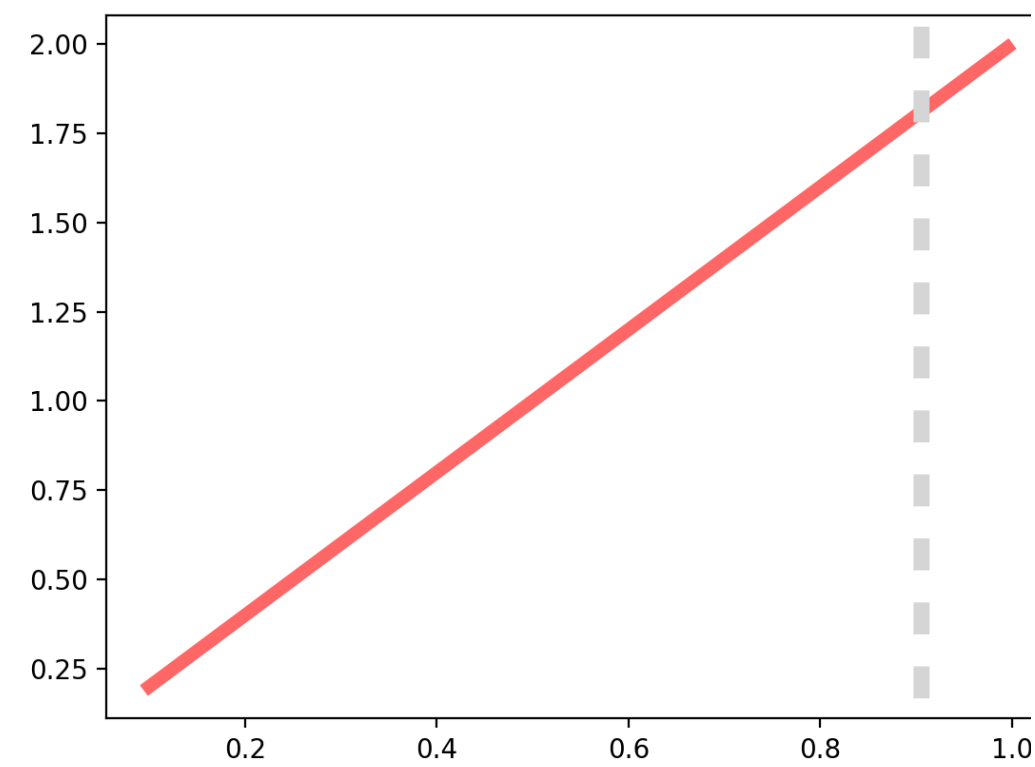


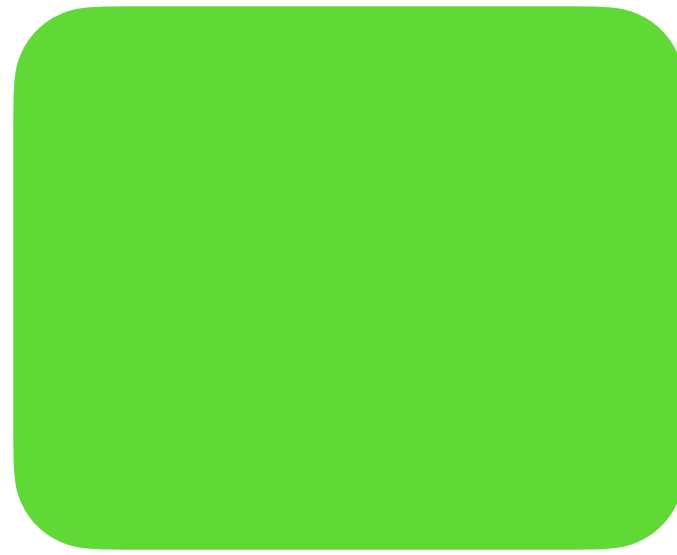
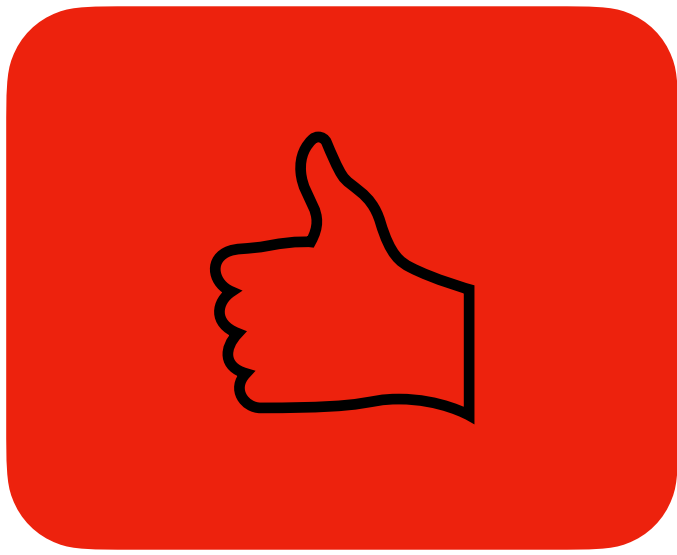
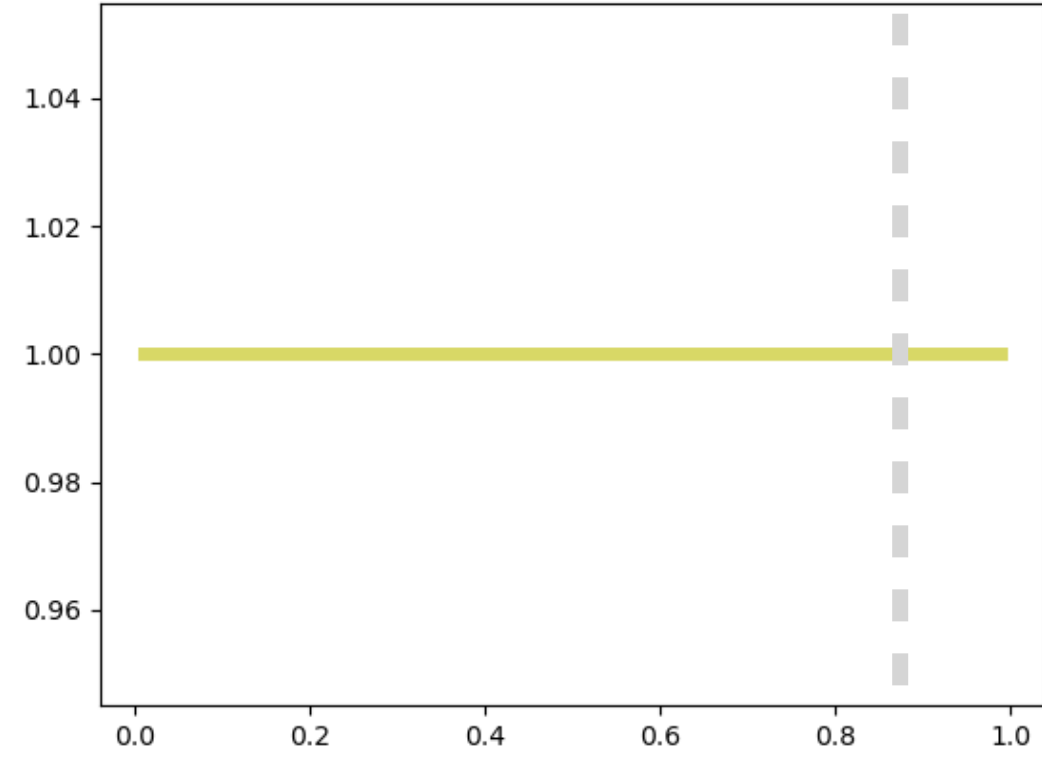
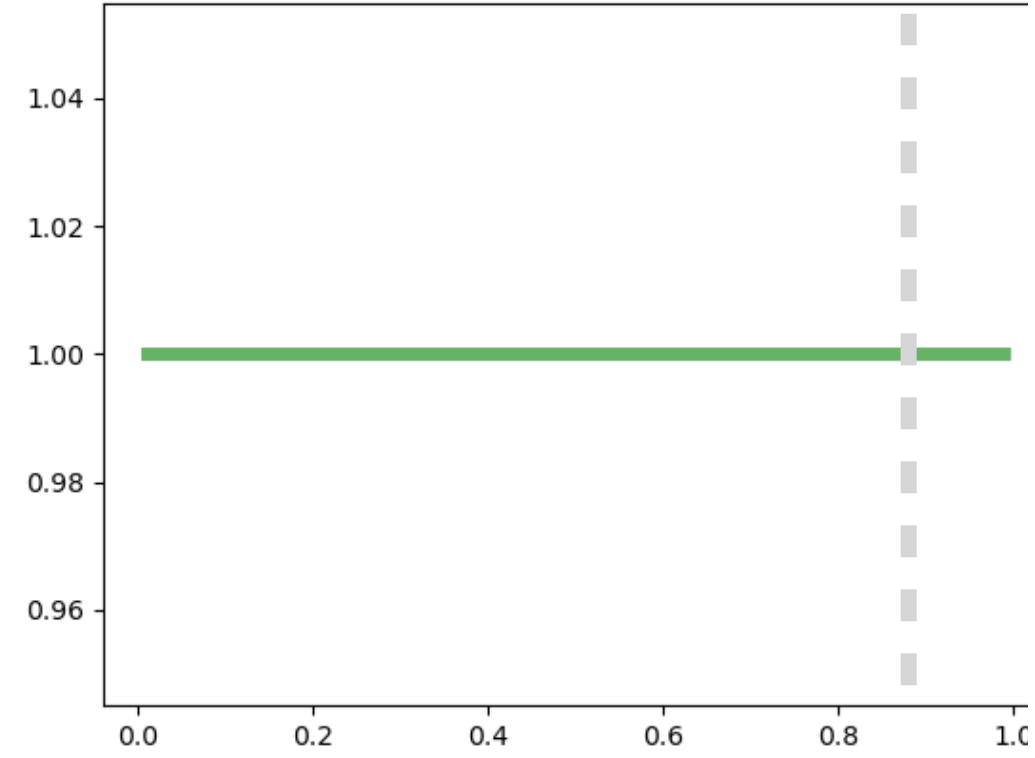
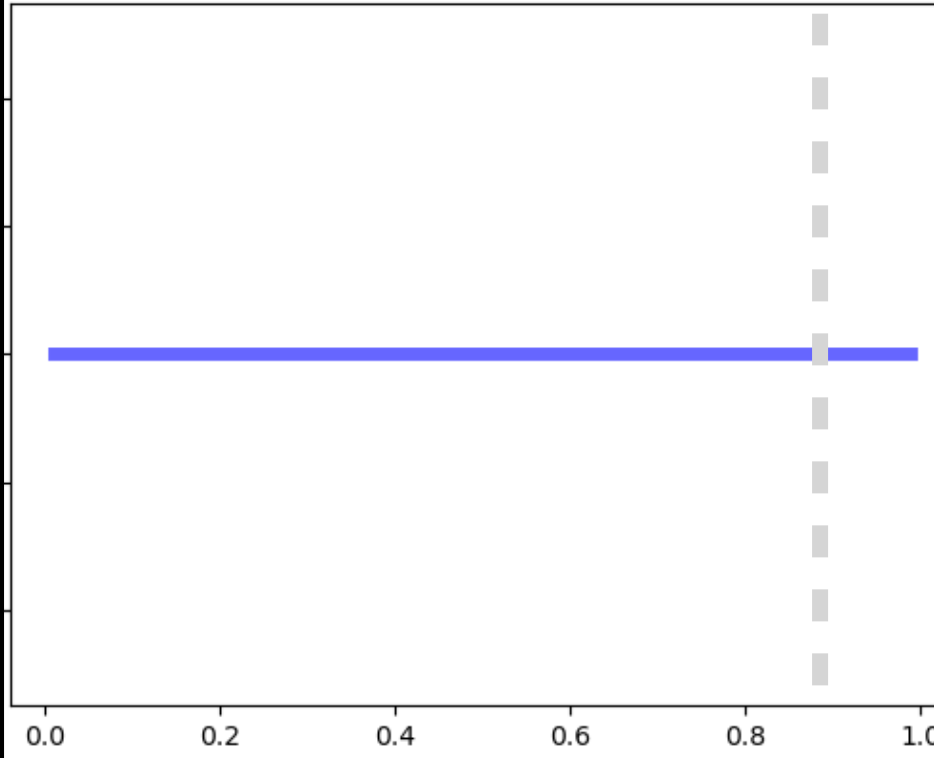
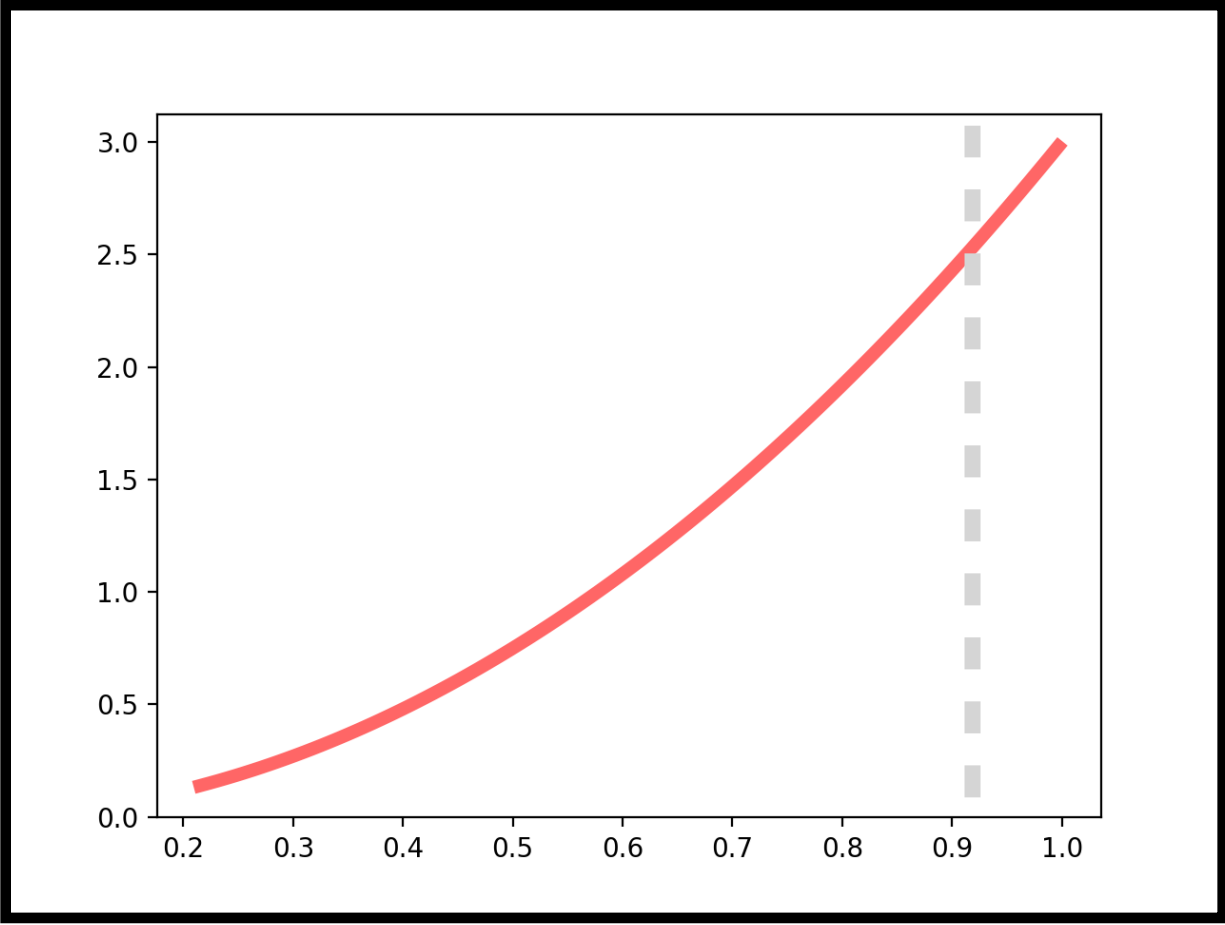


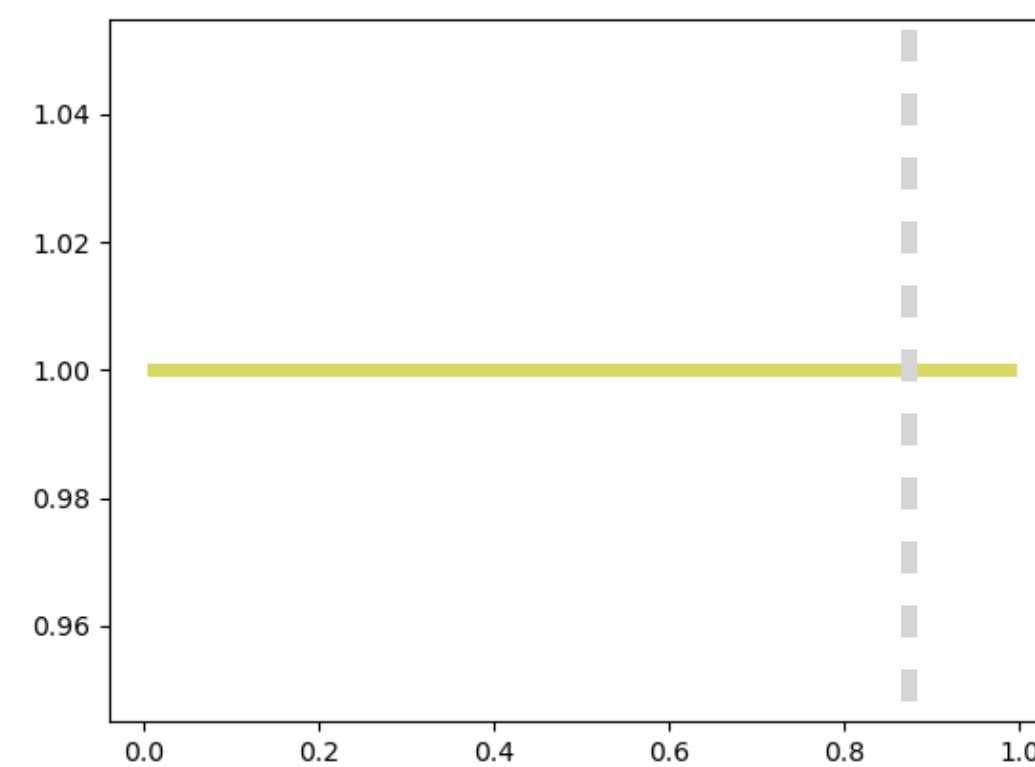
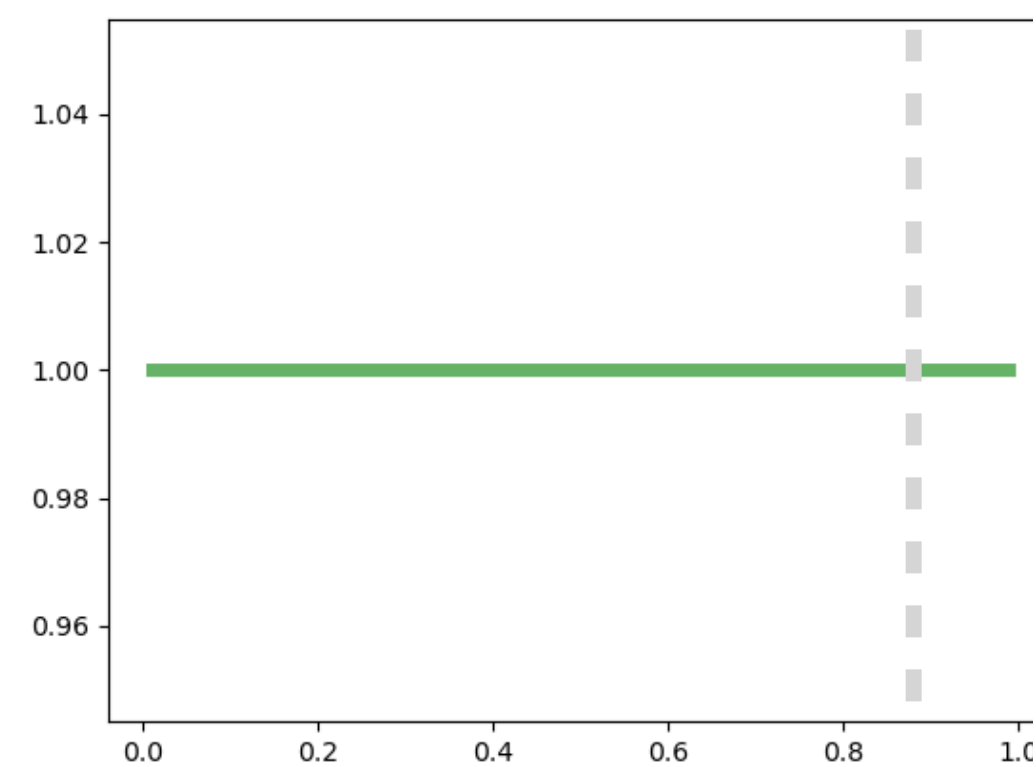
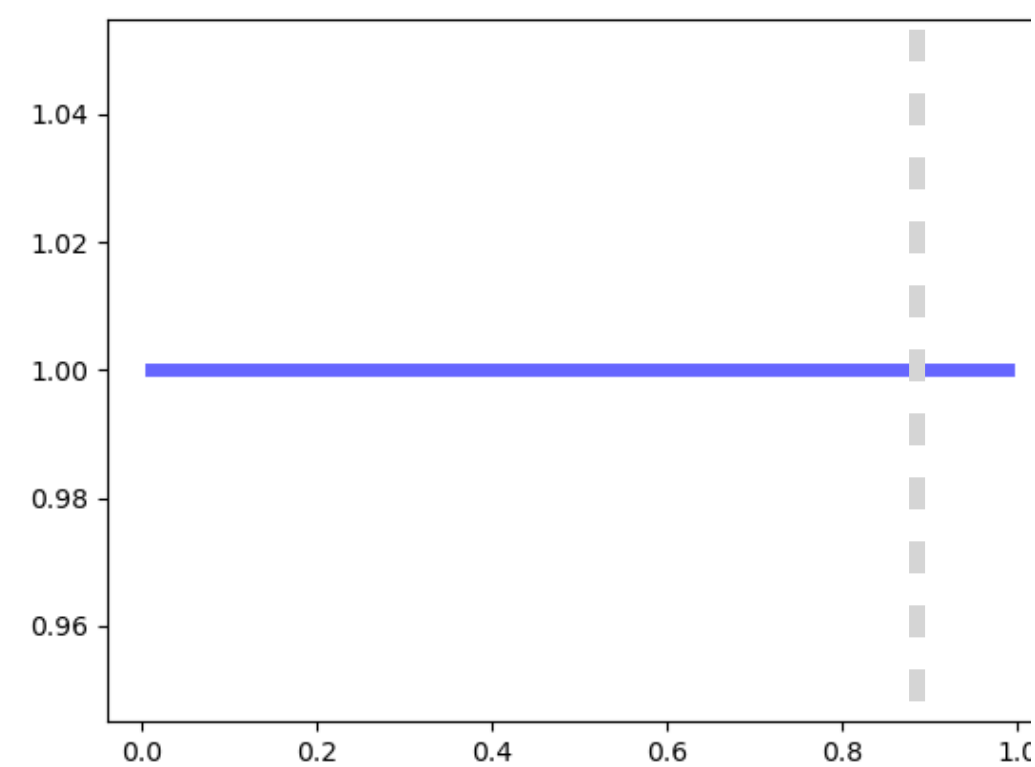
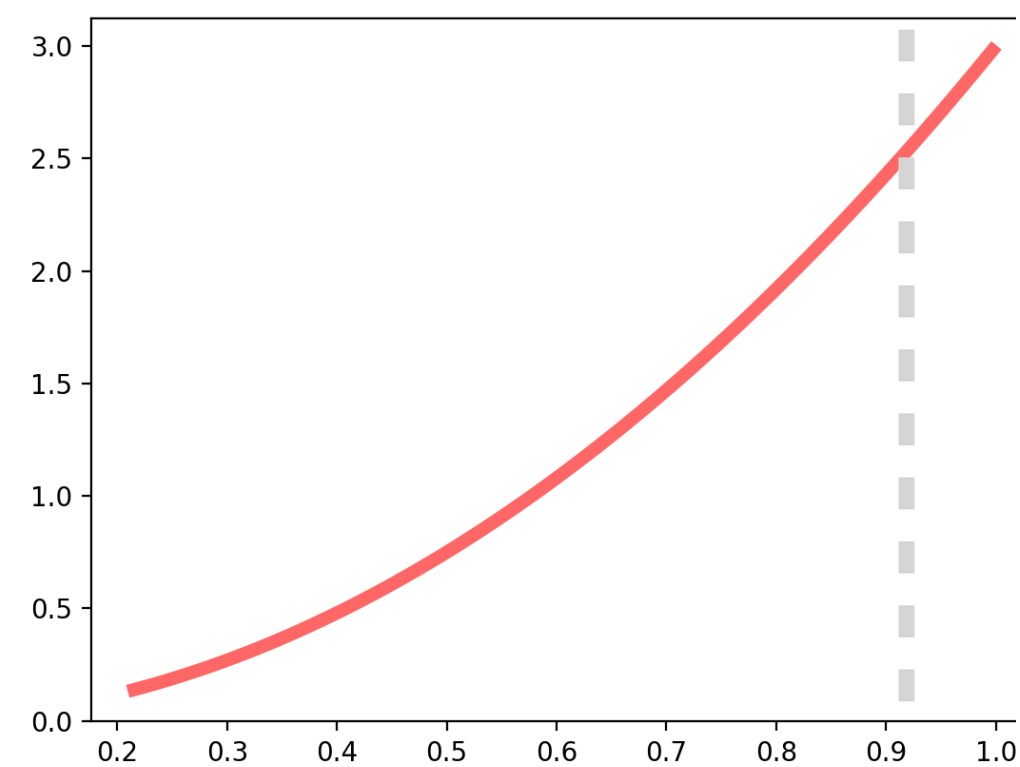


2

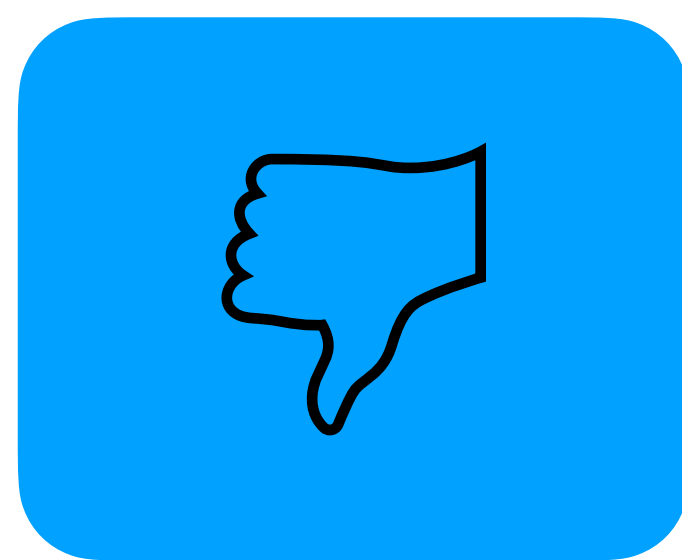
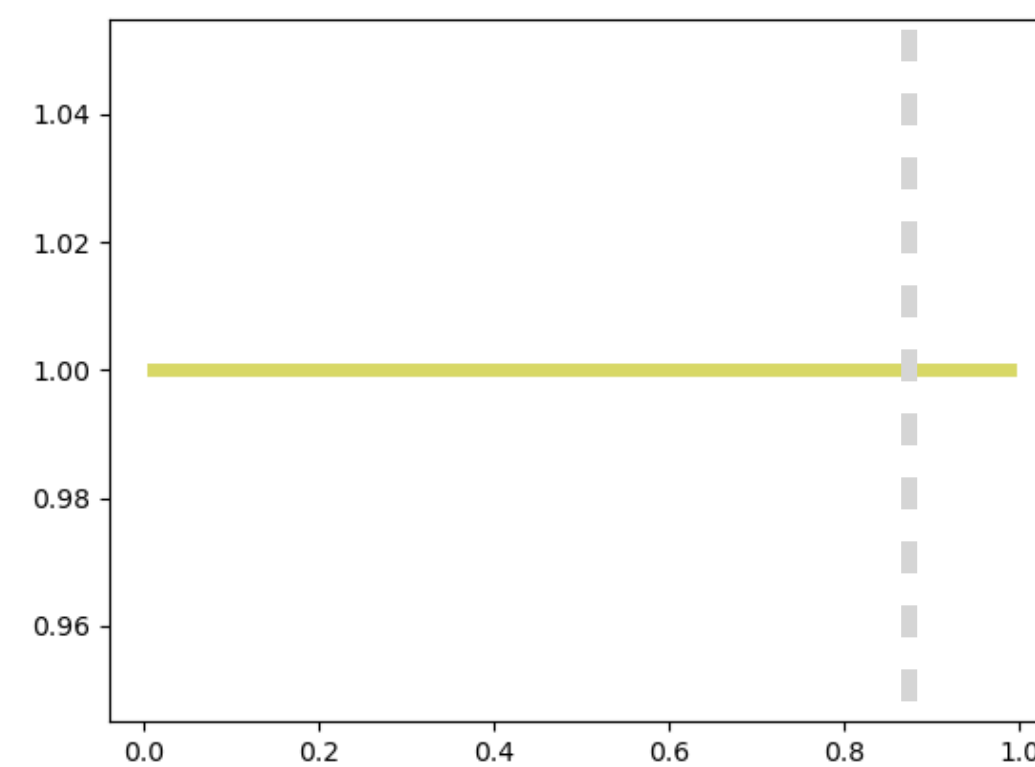
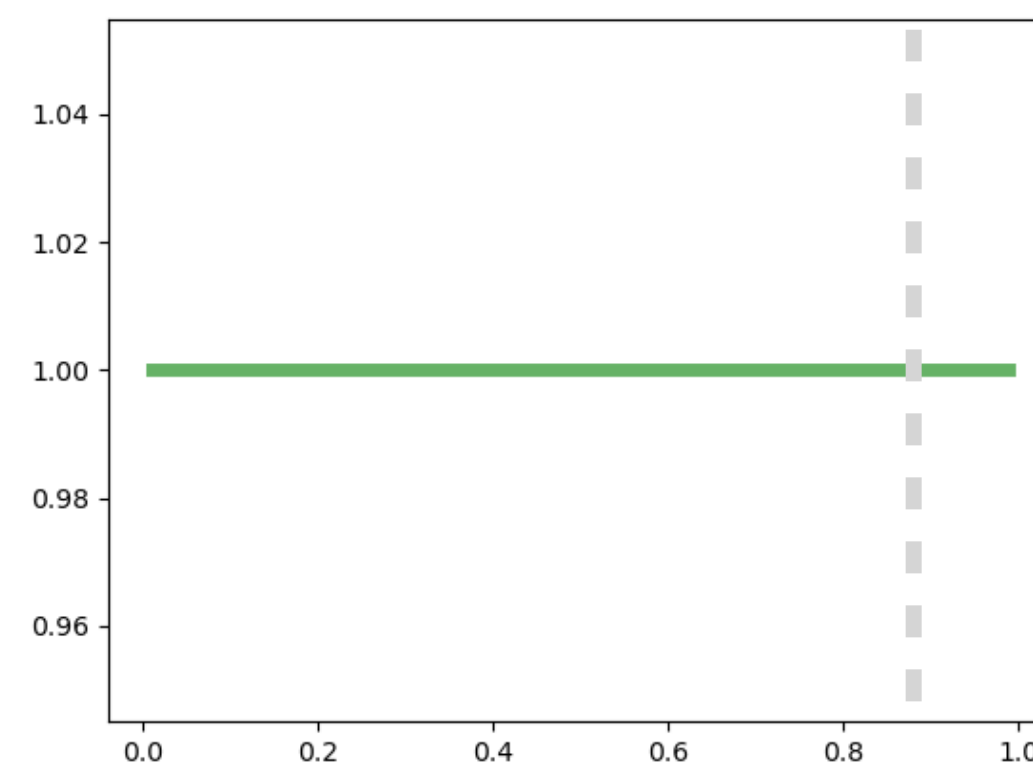
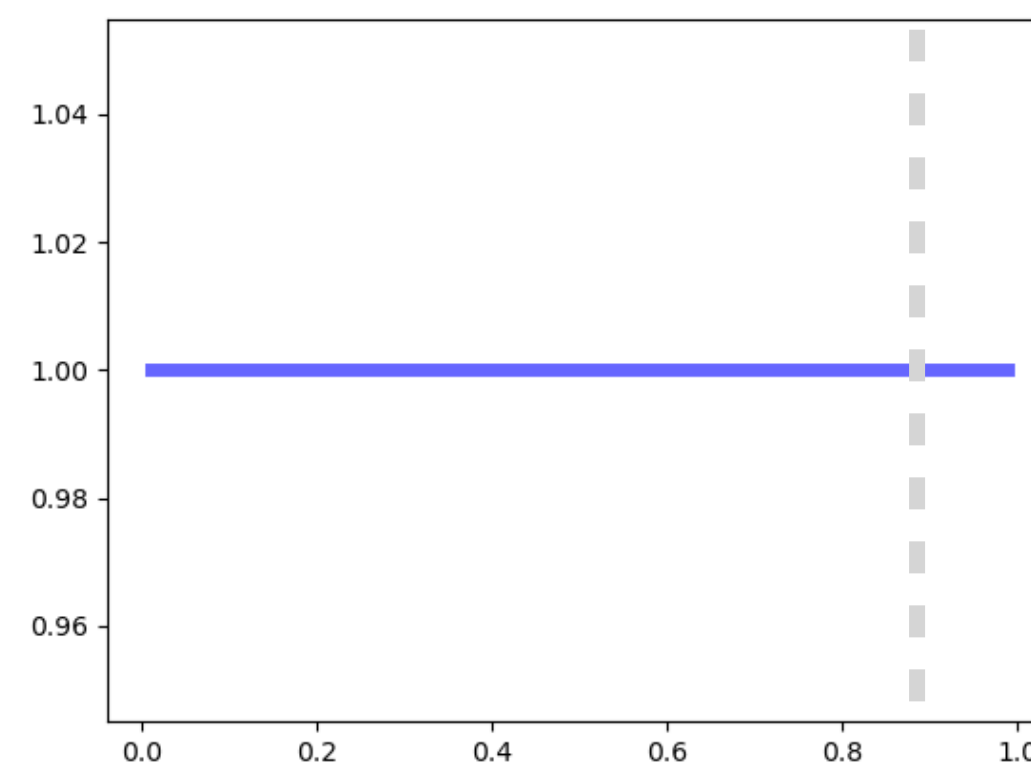
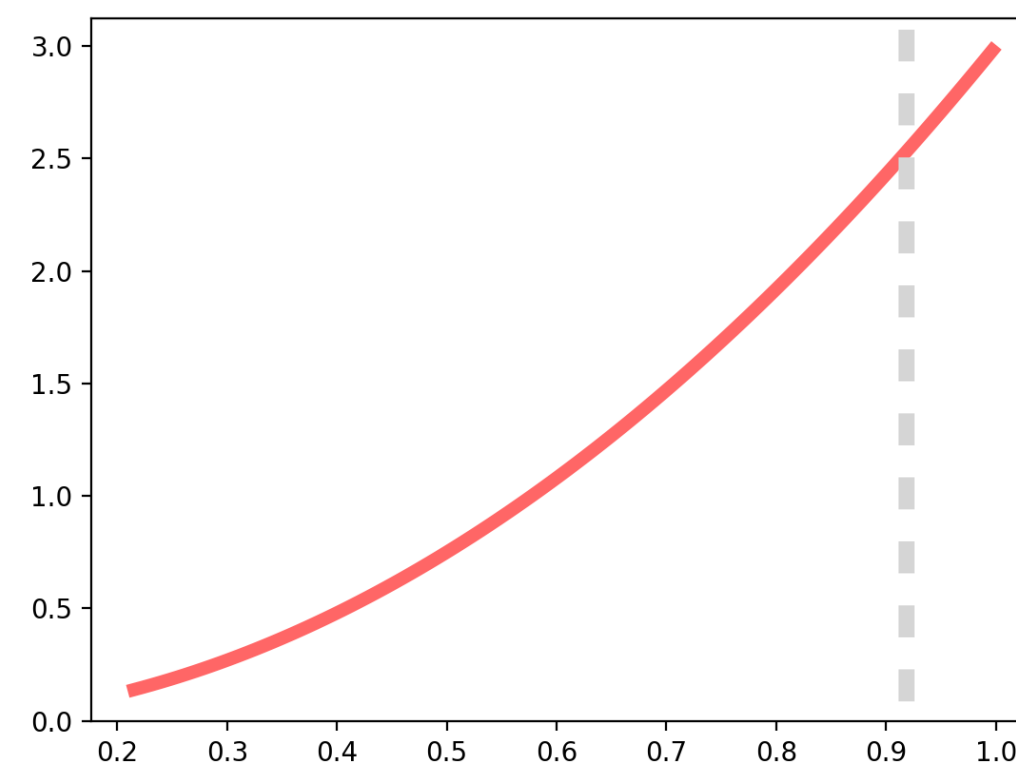


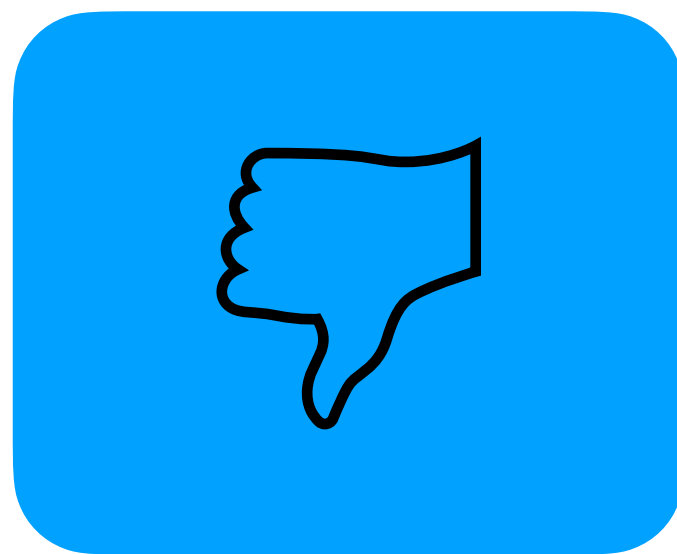
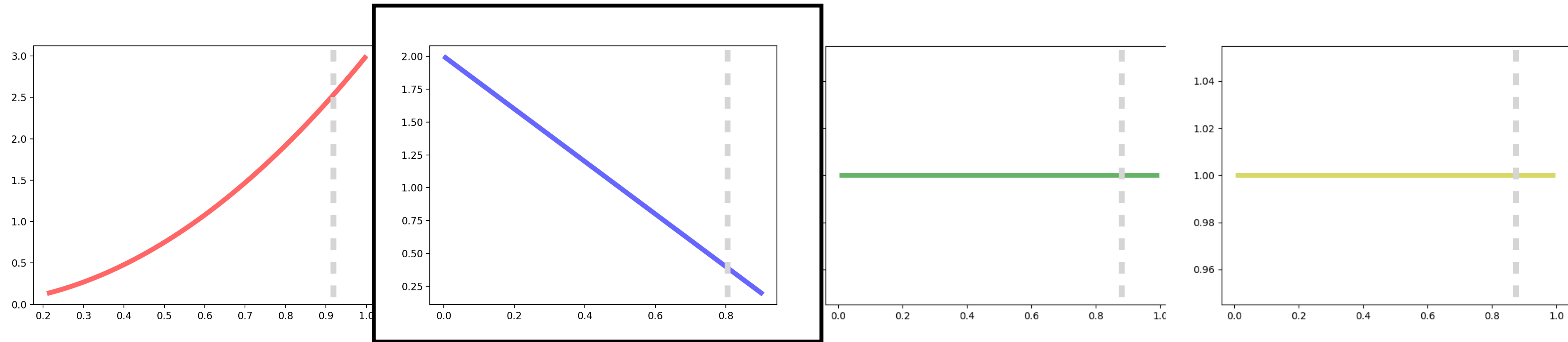


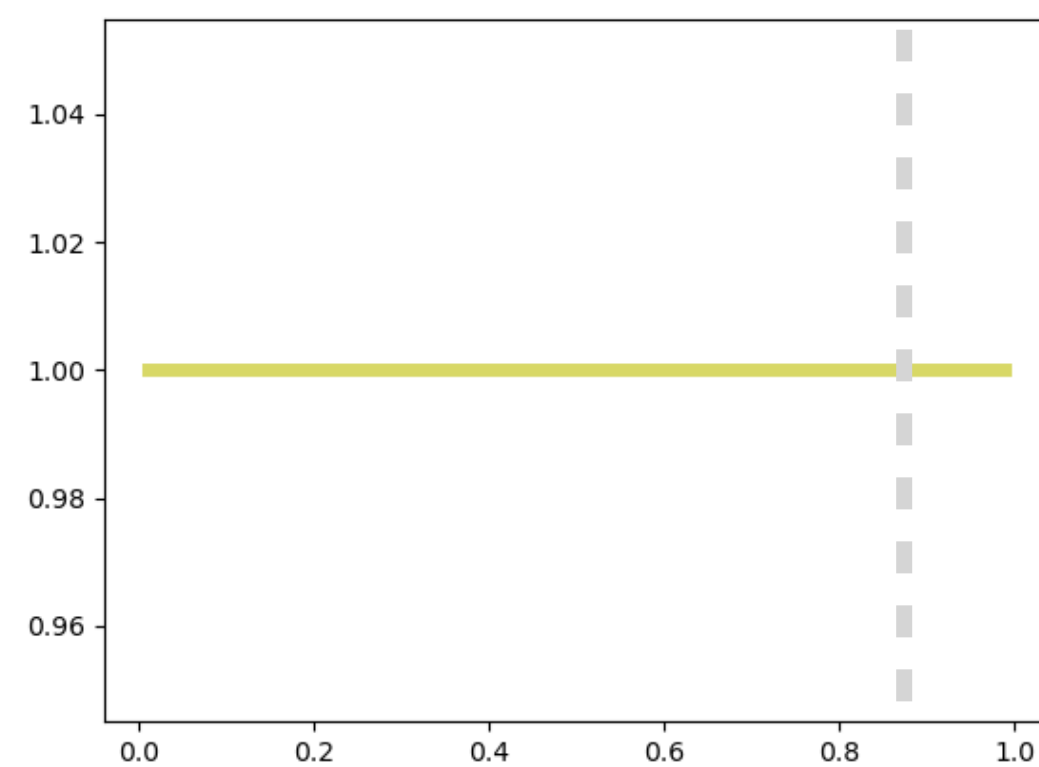
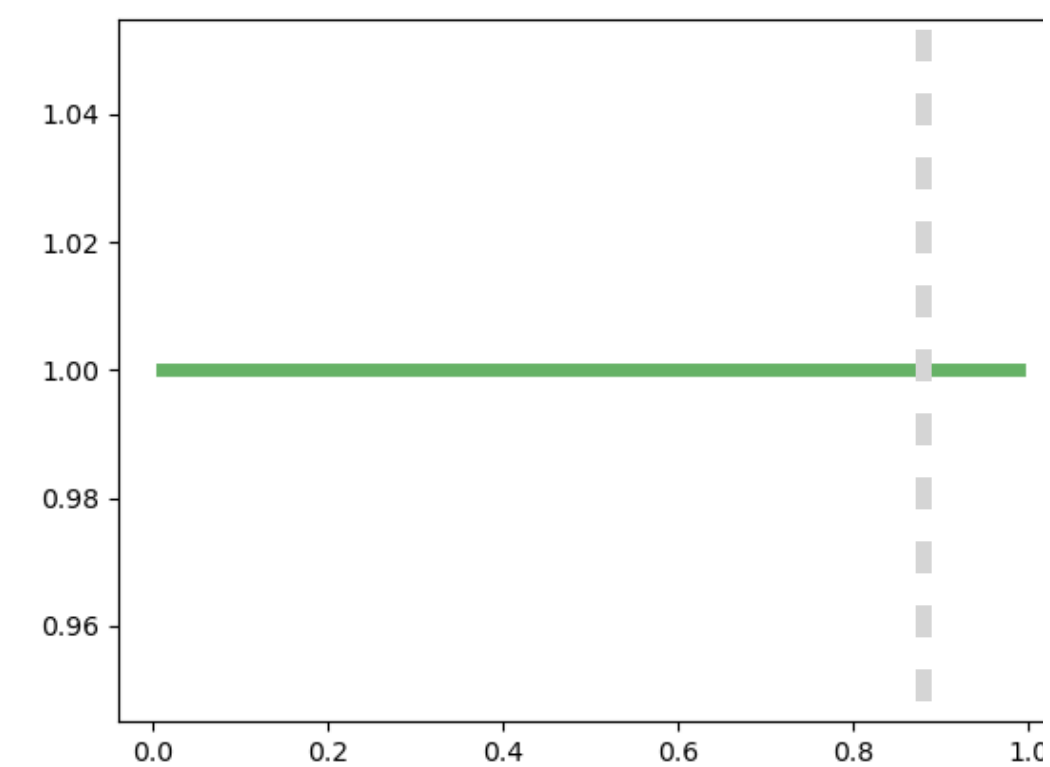
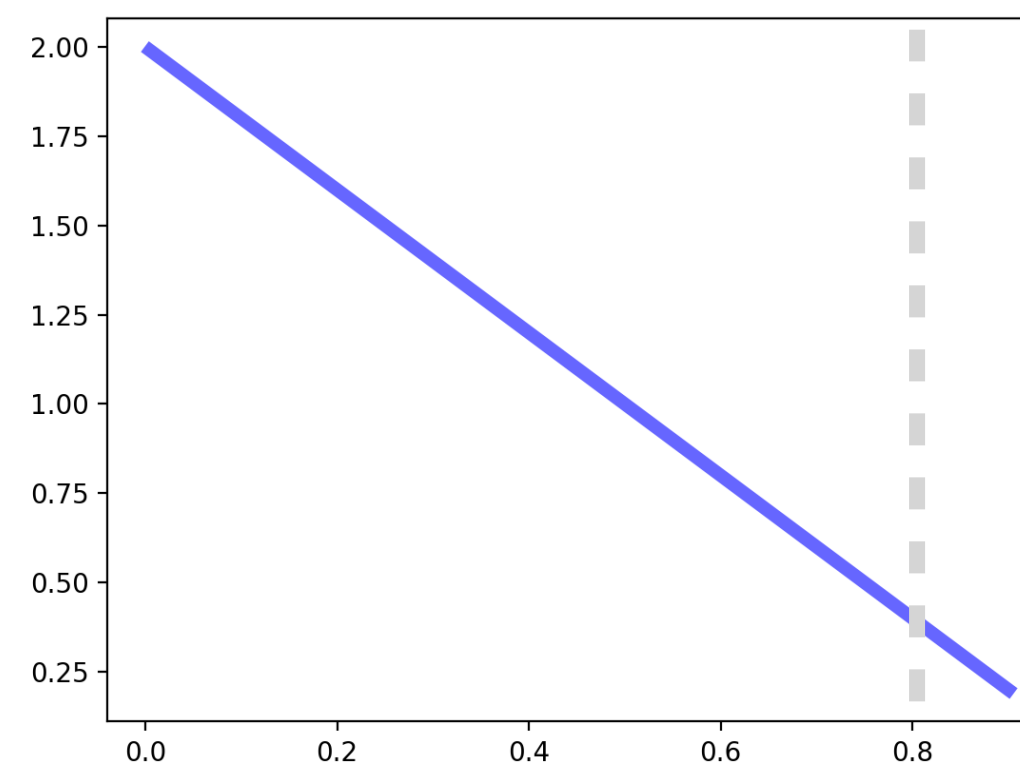
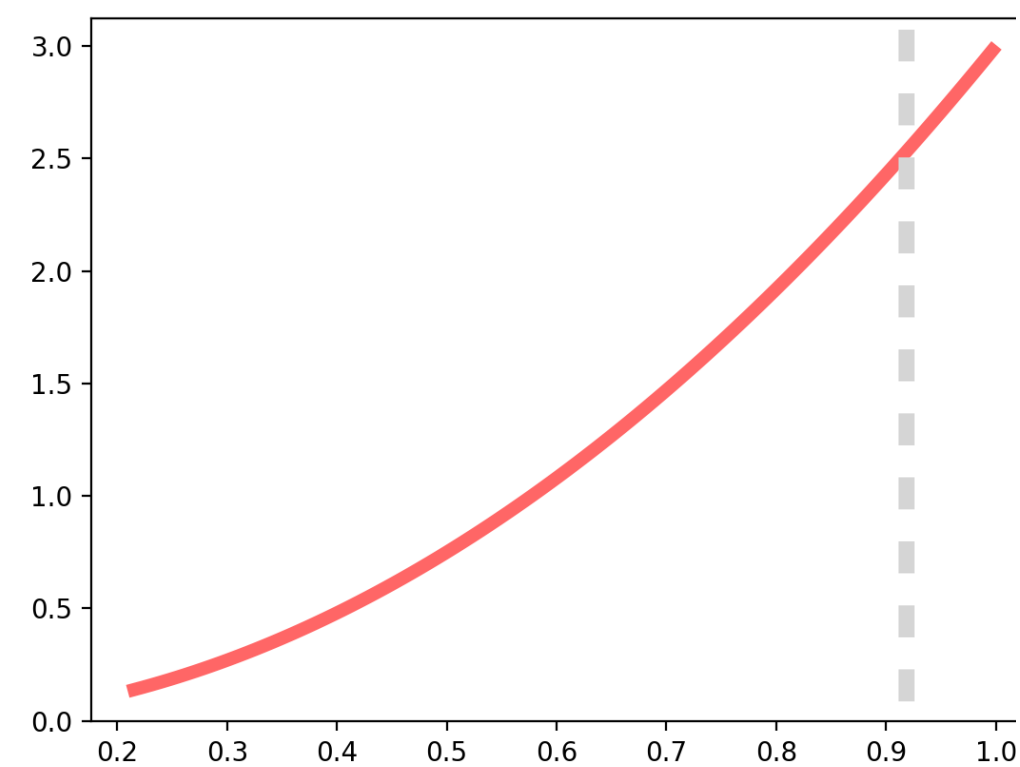


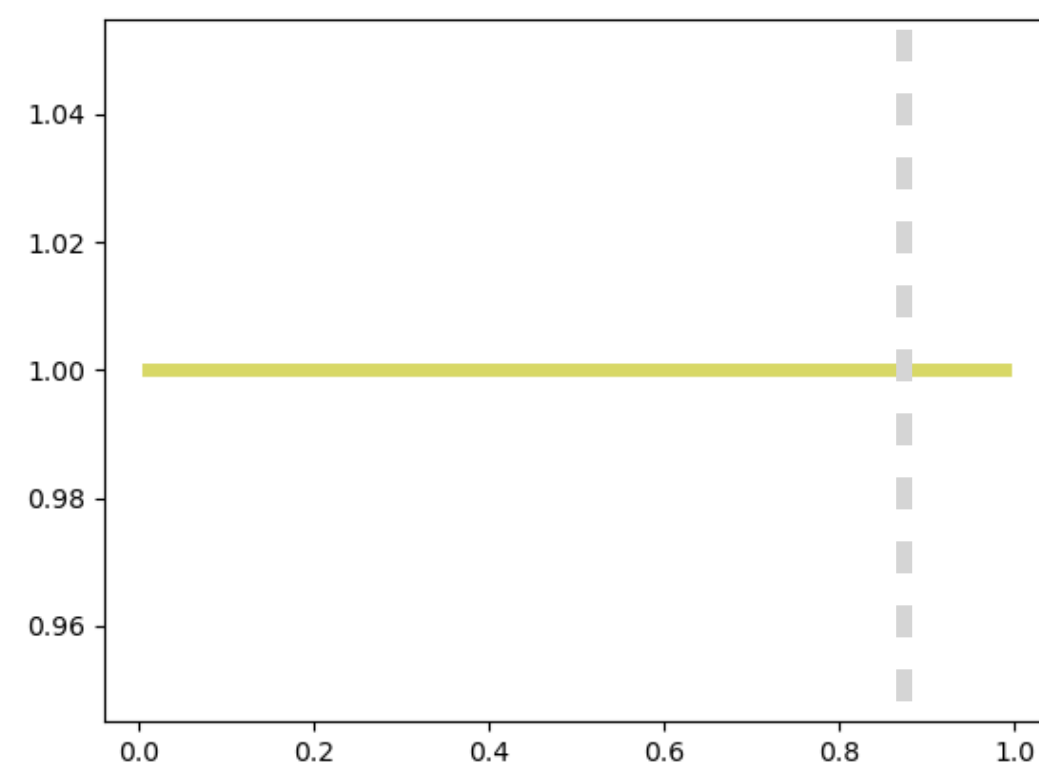
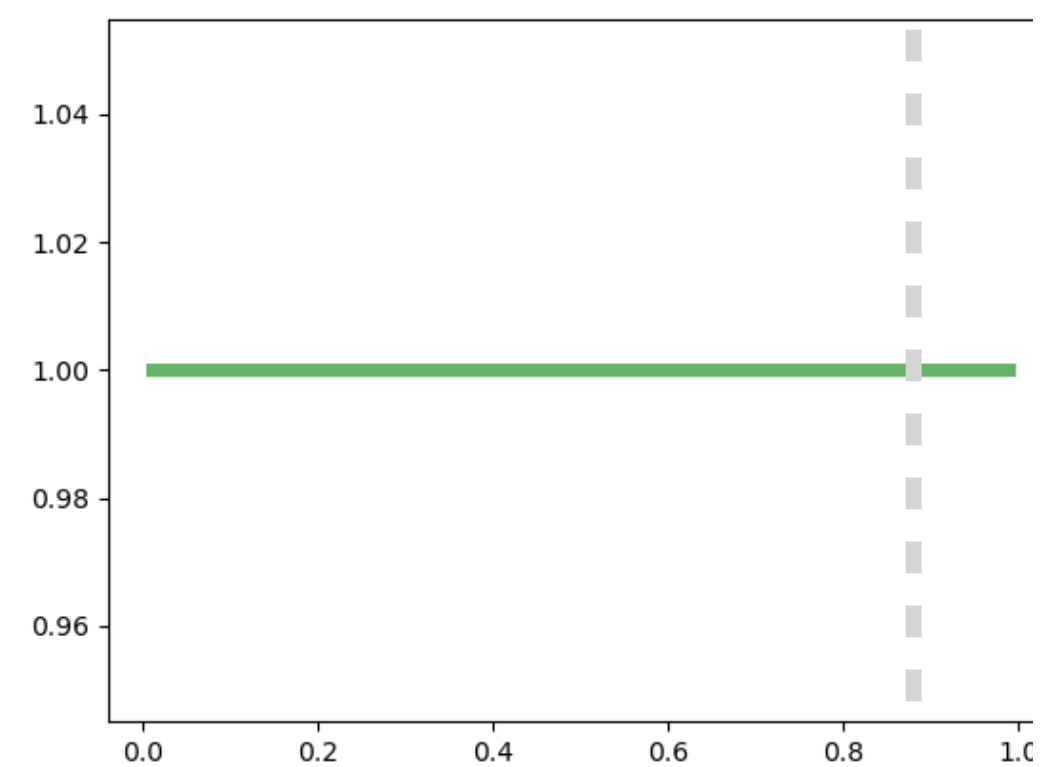
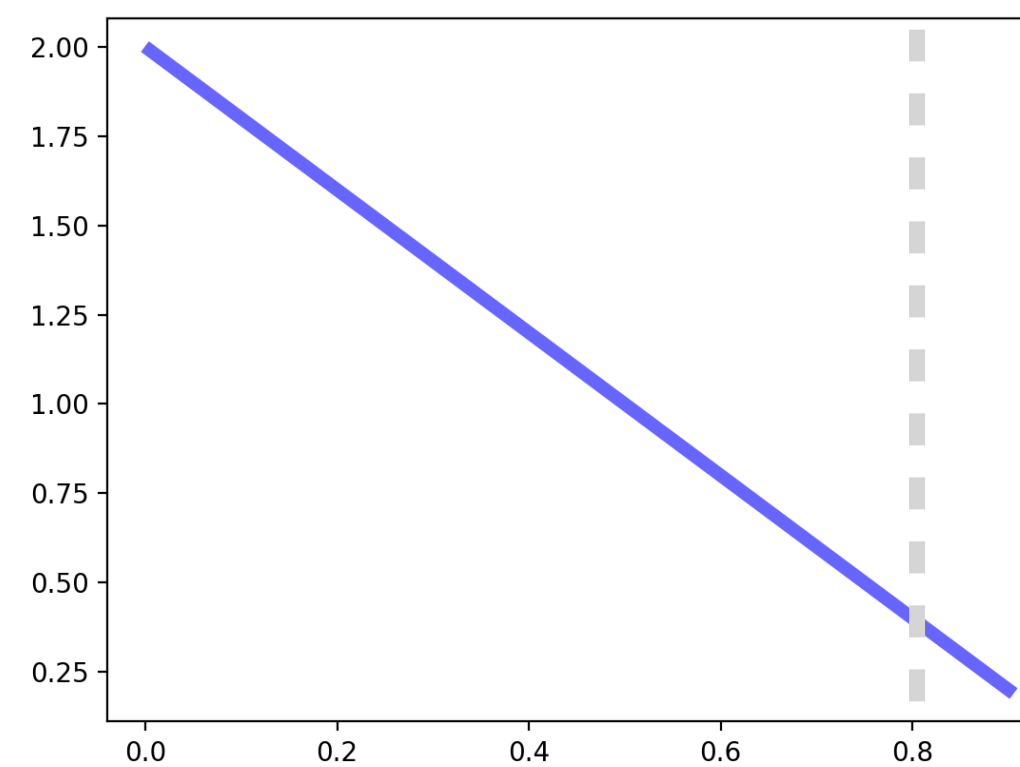
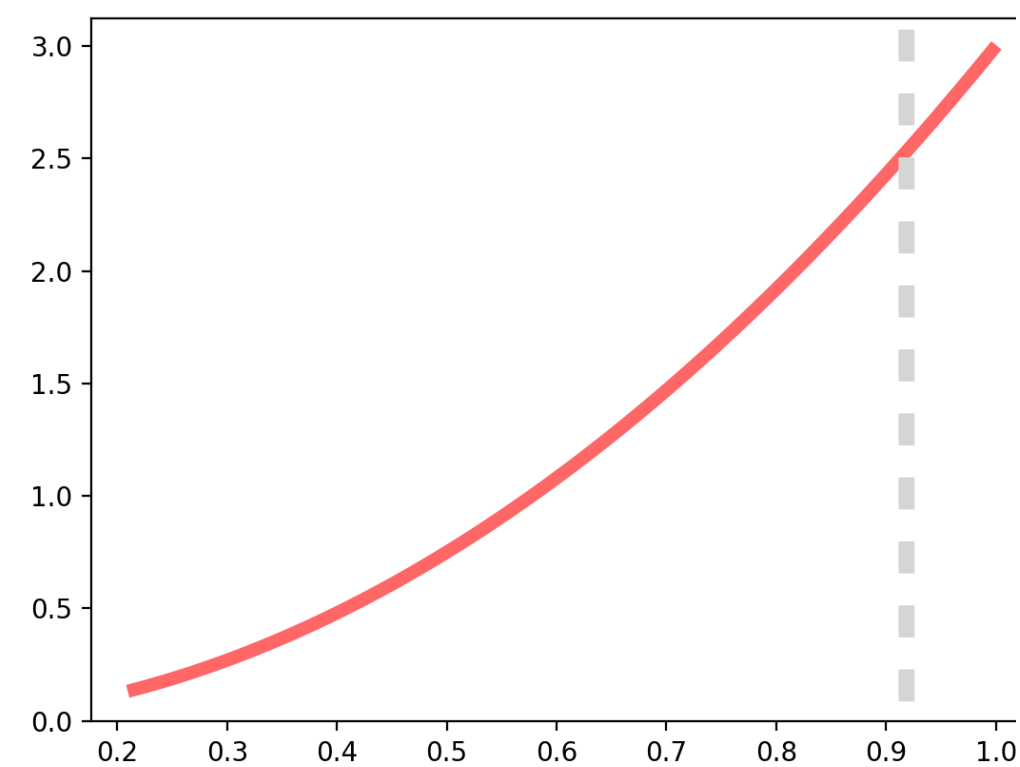


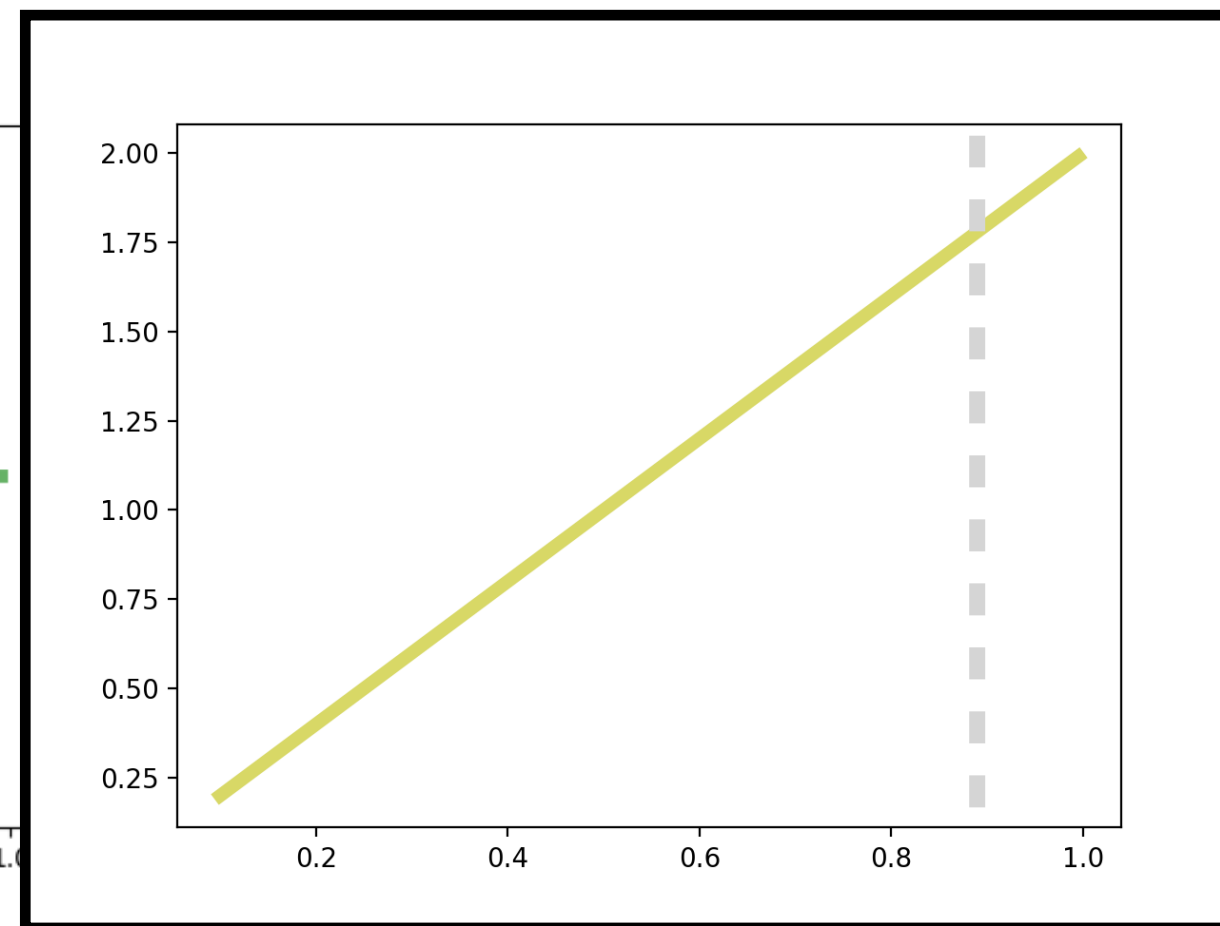
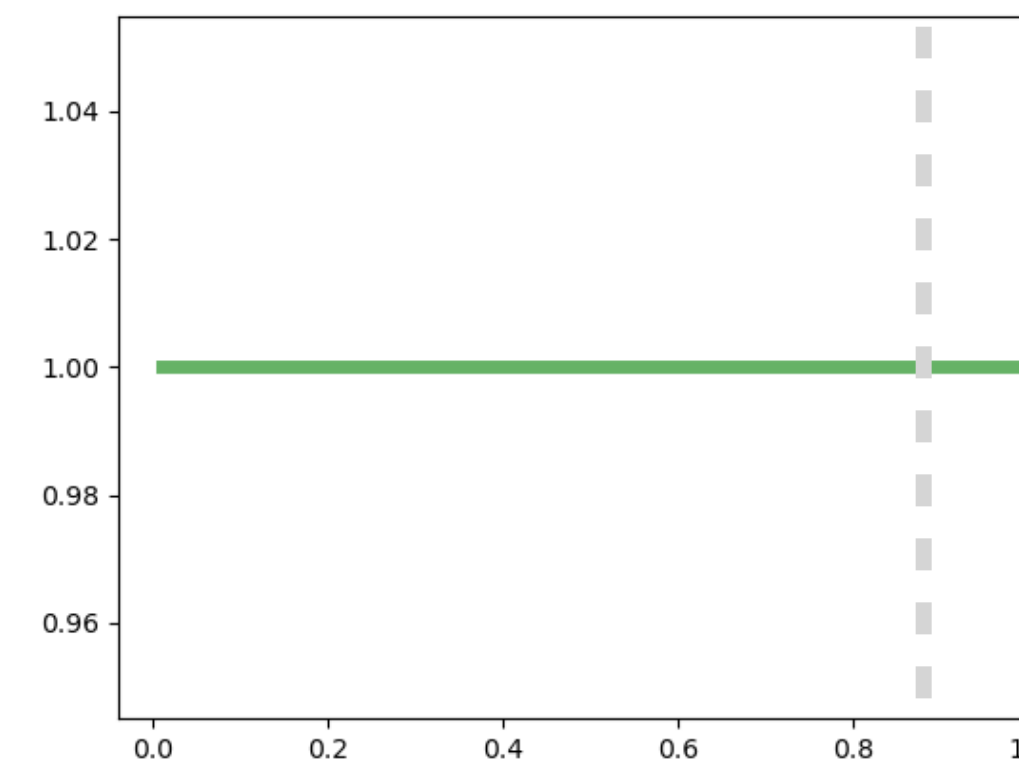
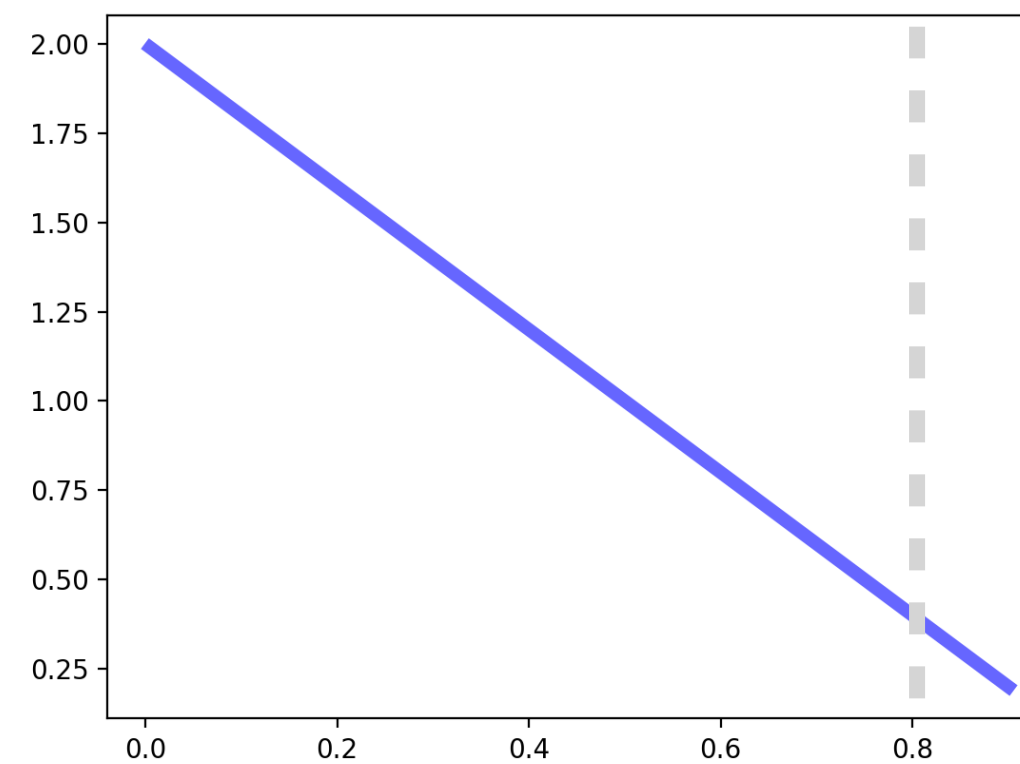
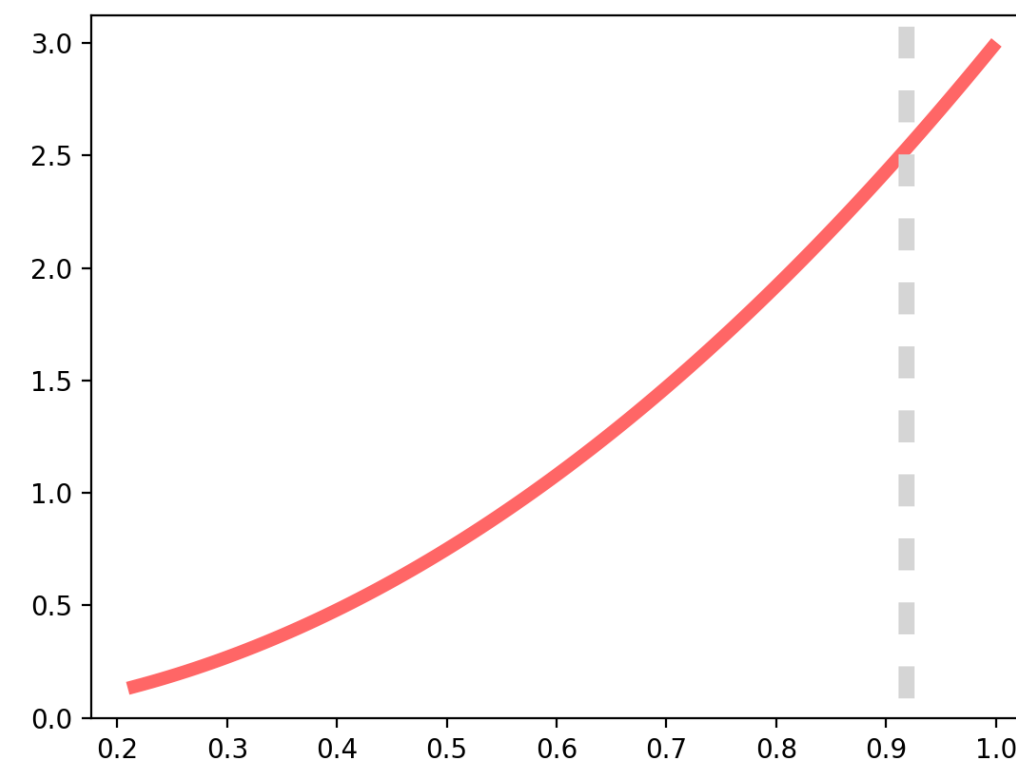
1

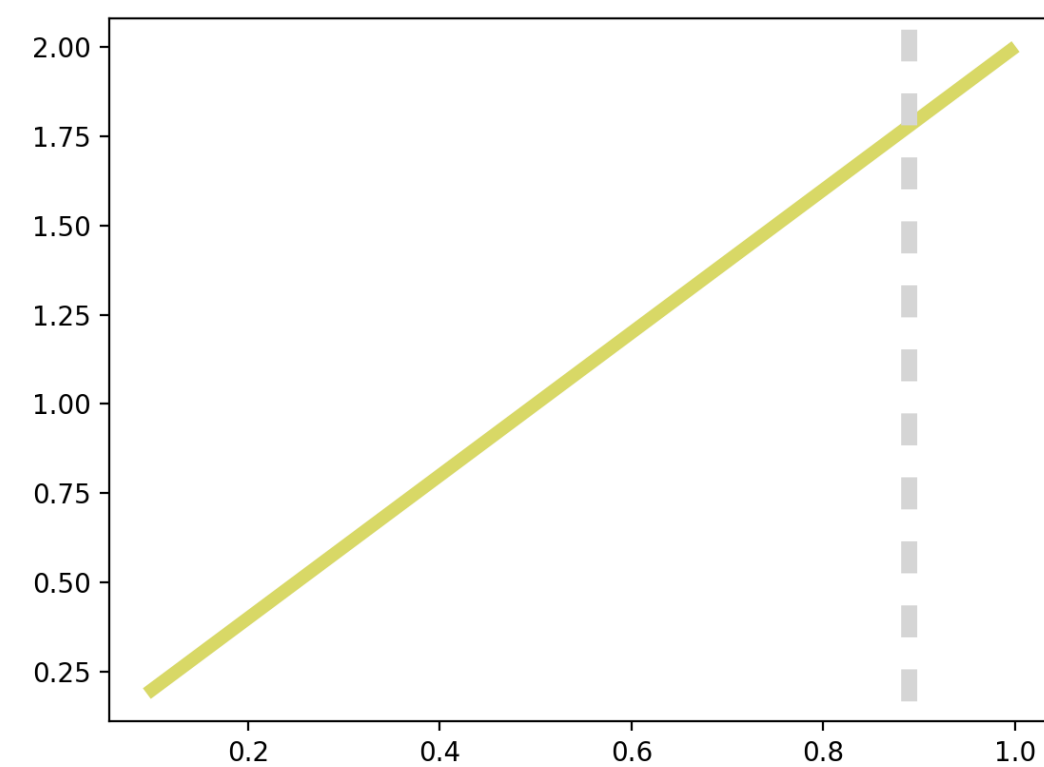
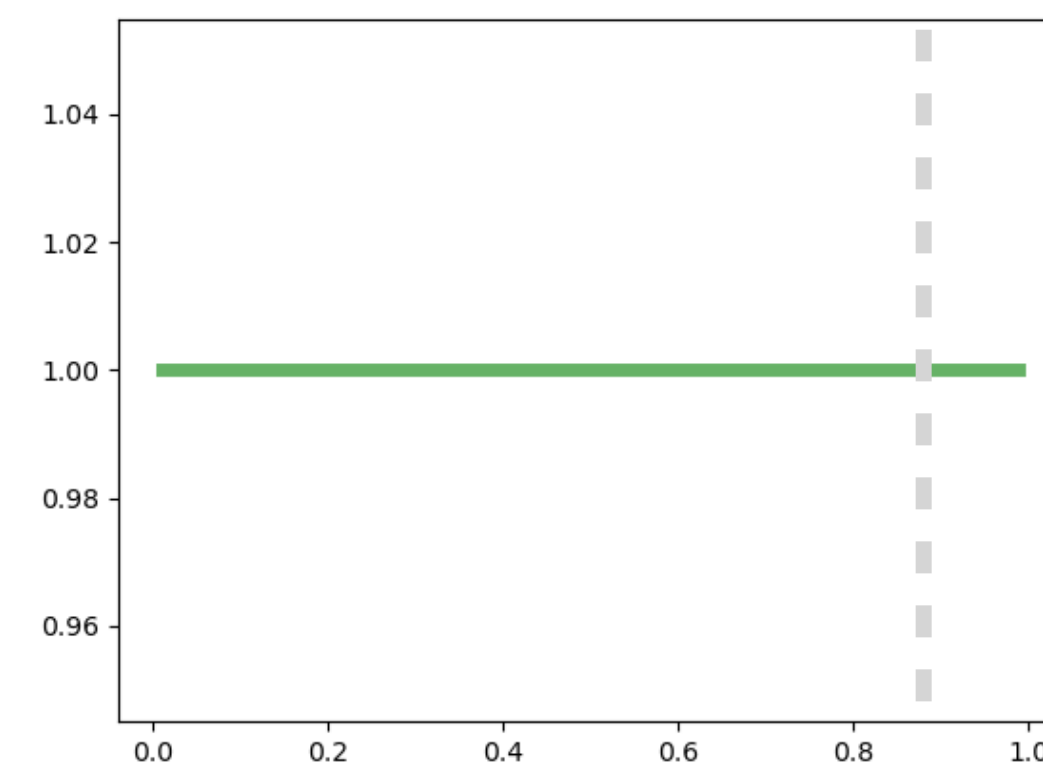
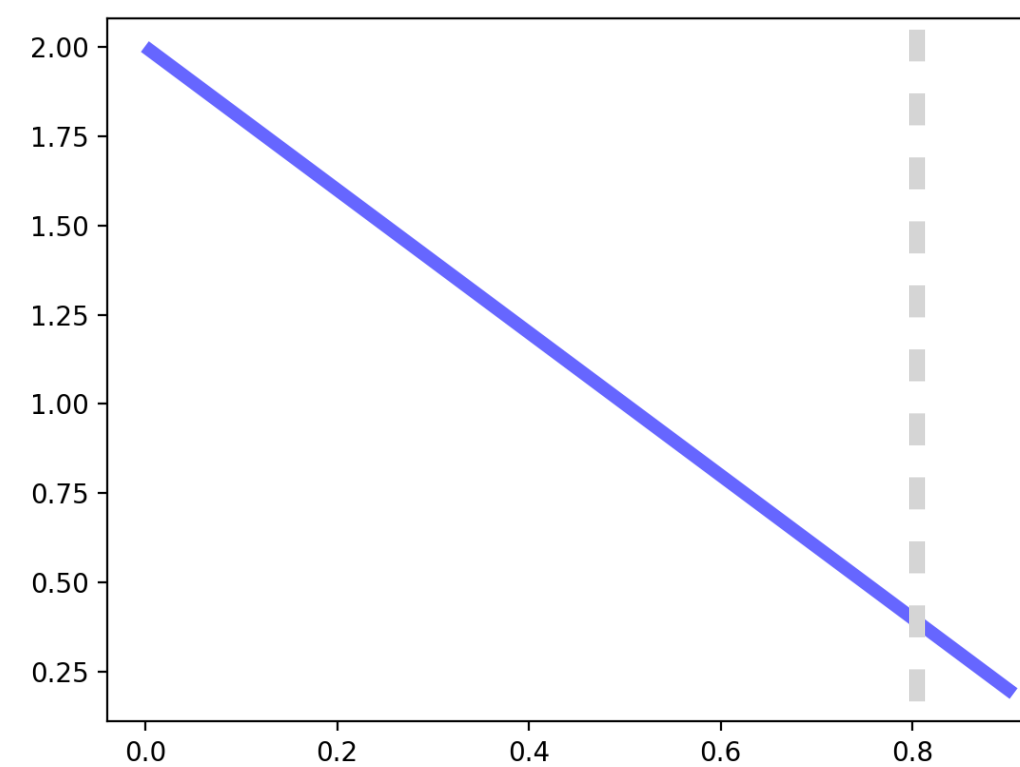
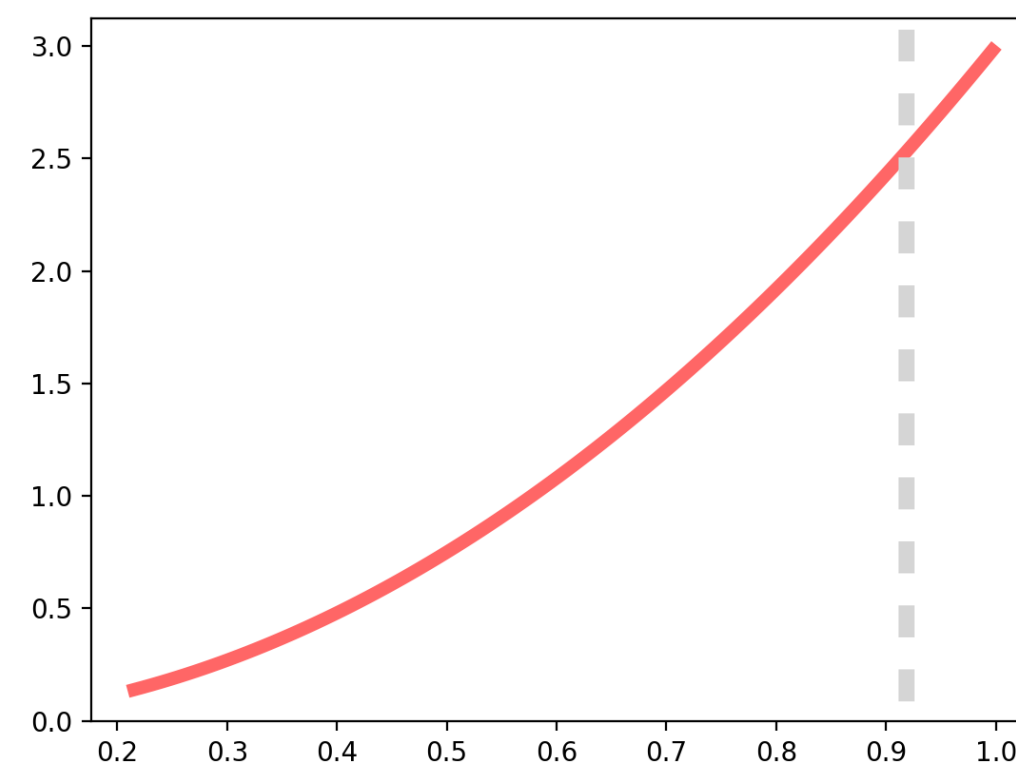




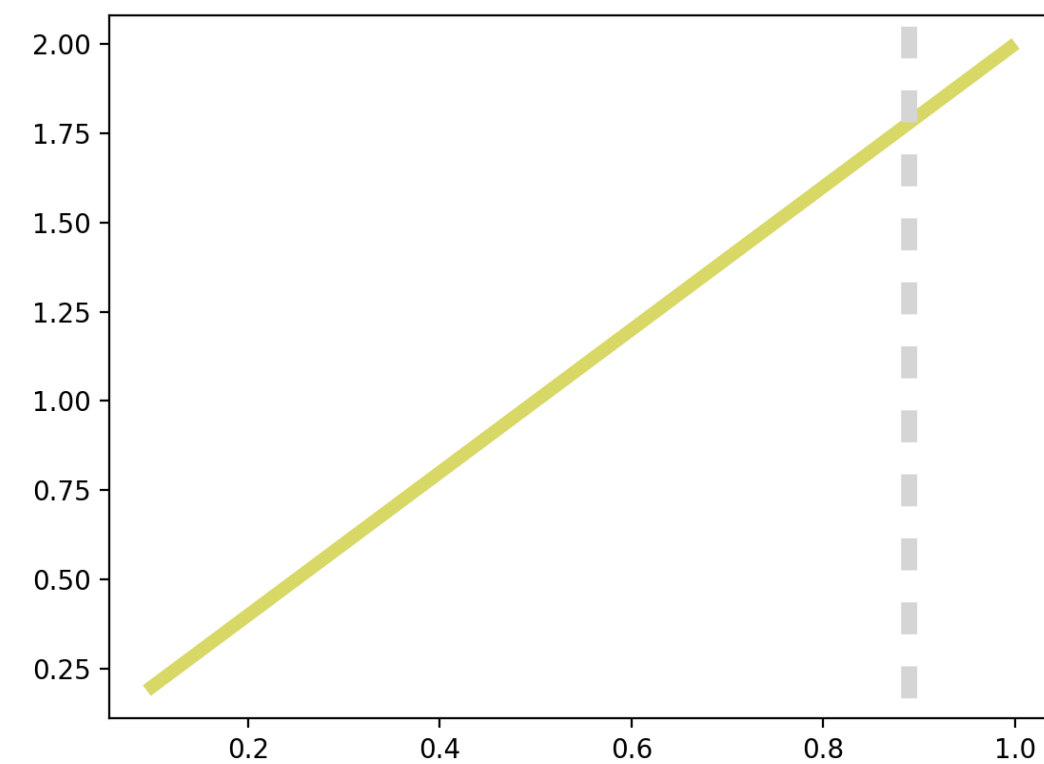
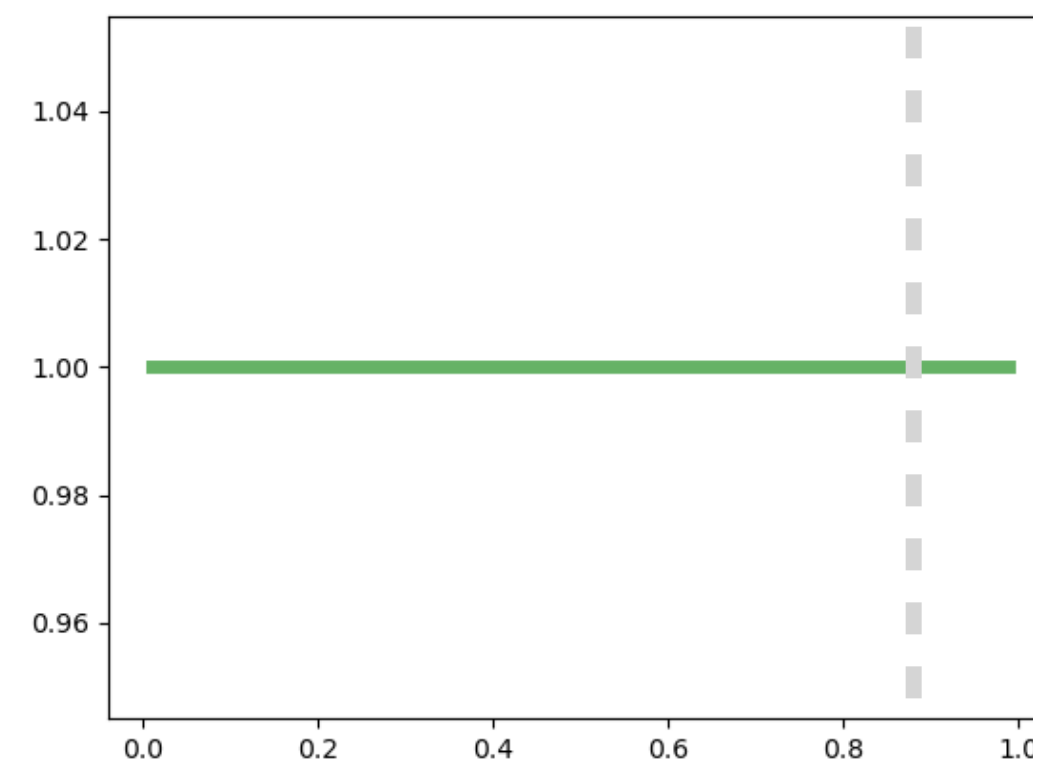
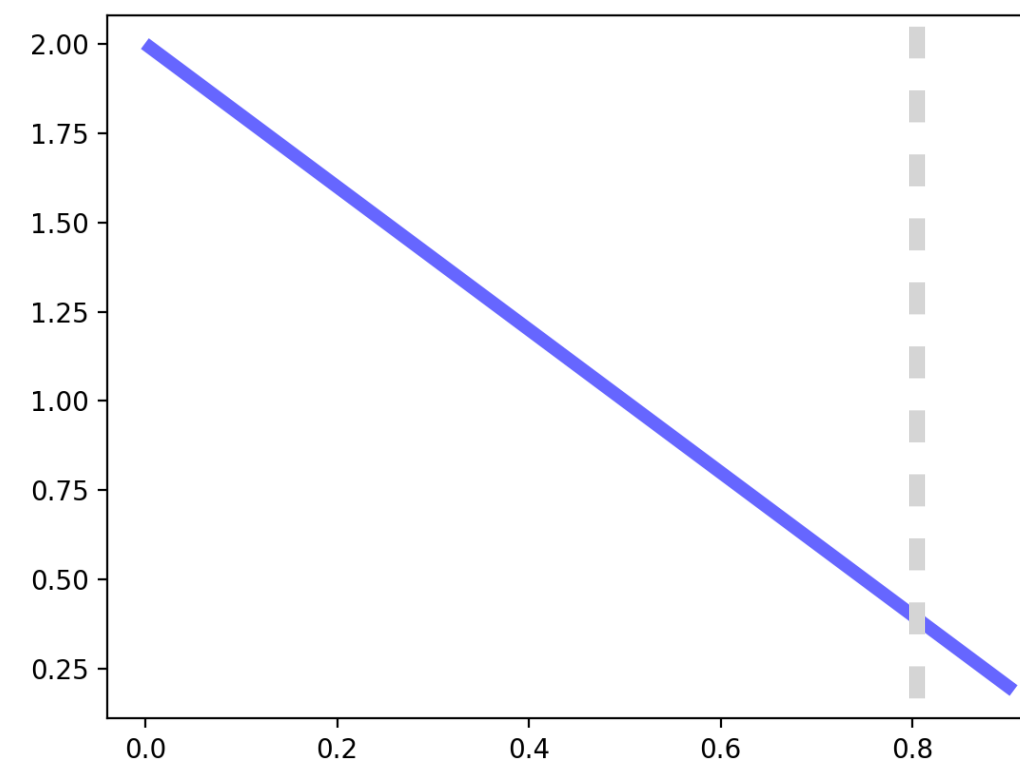
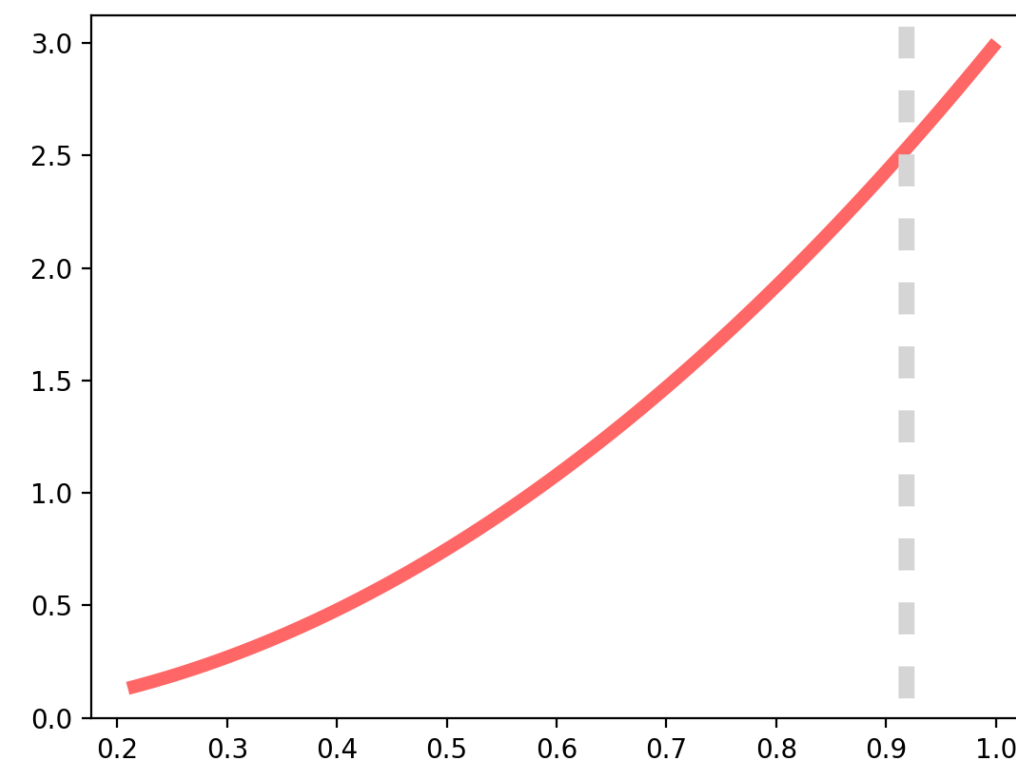


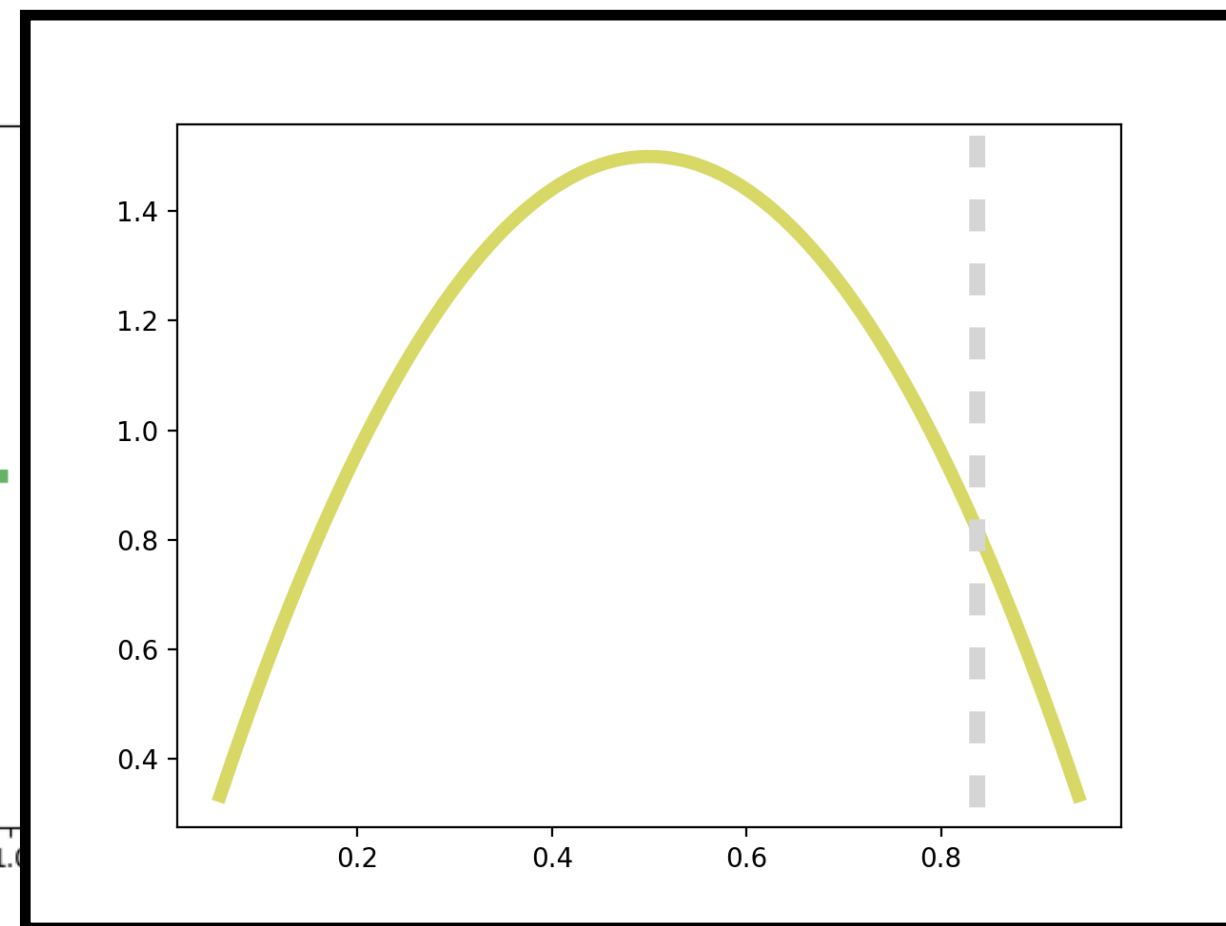
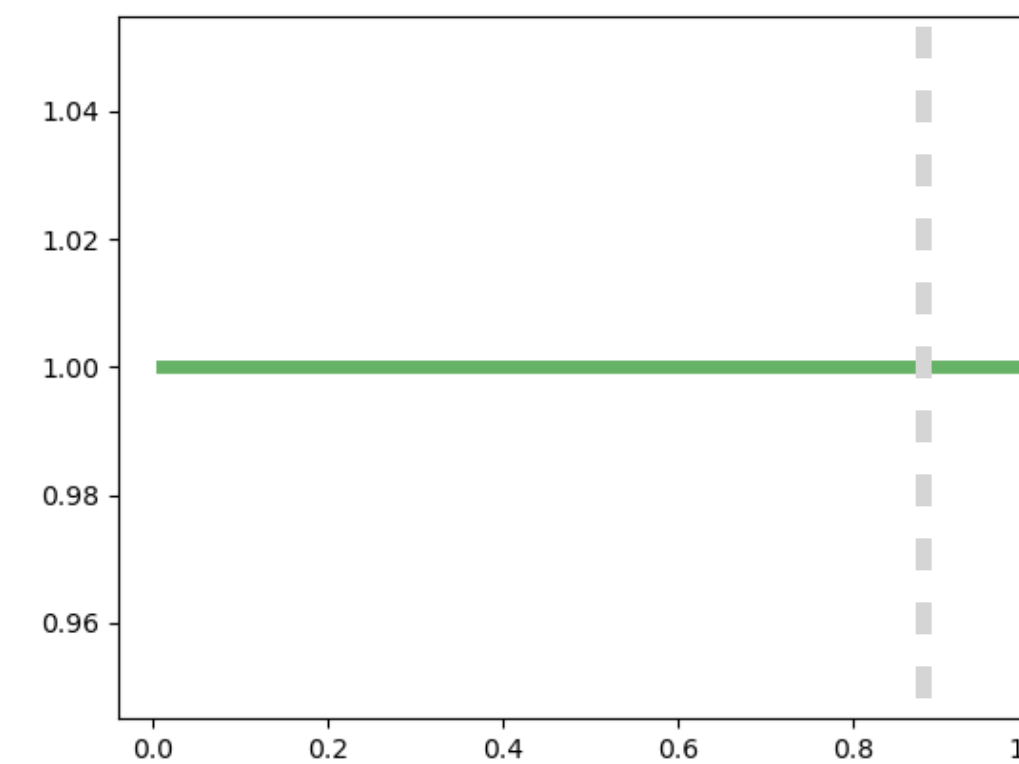
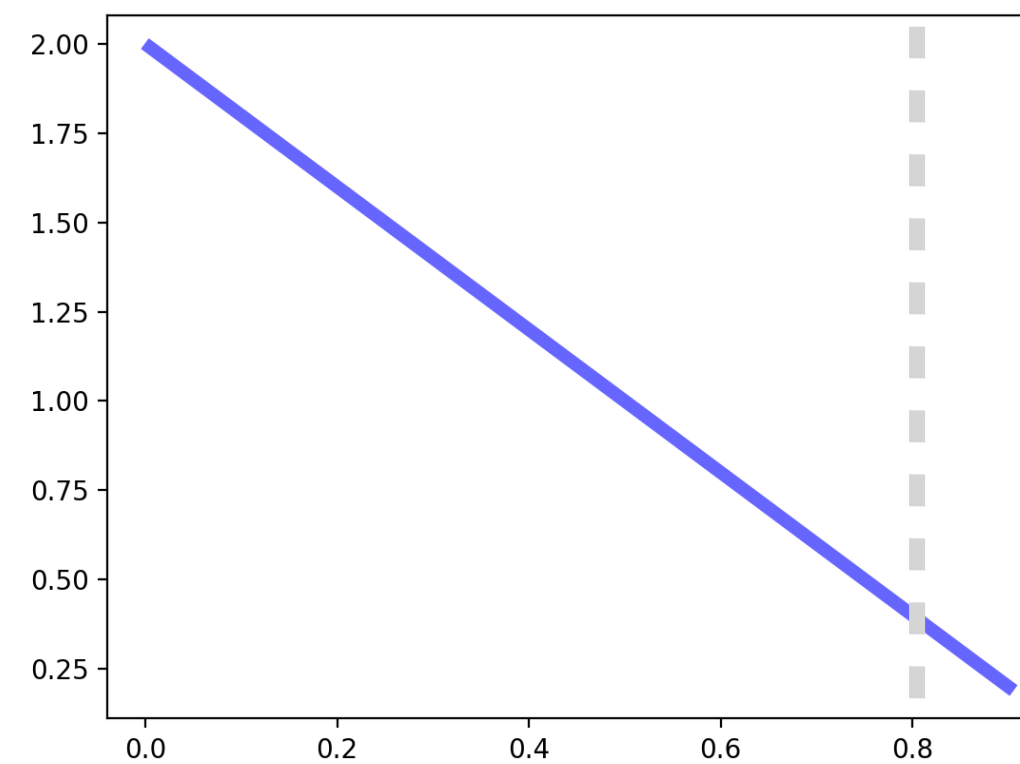
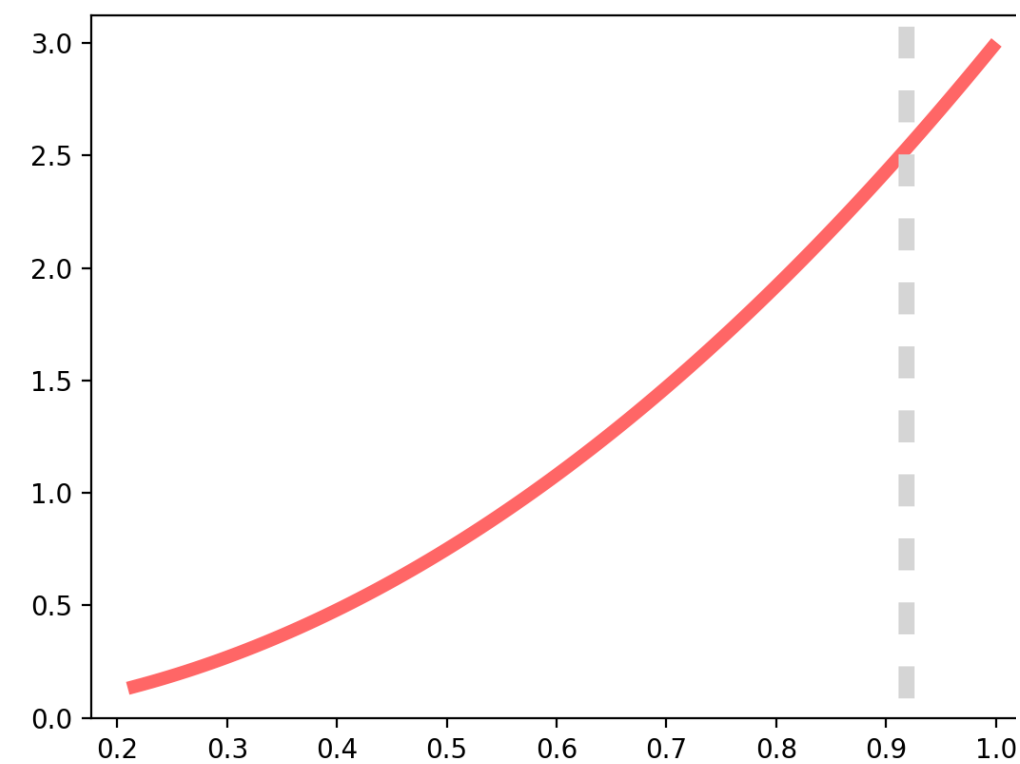


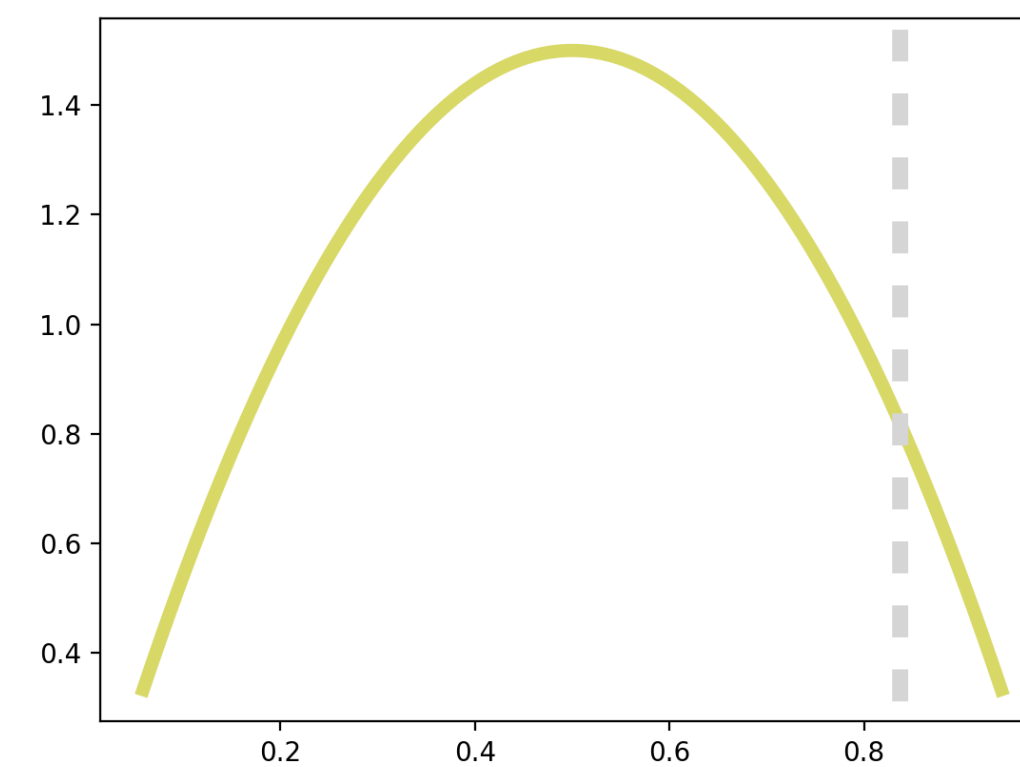
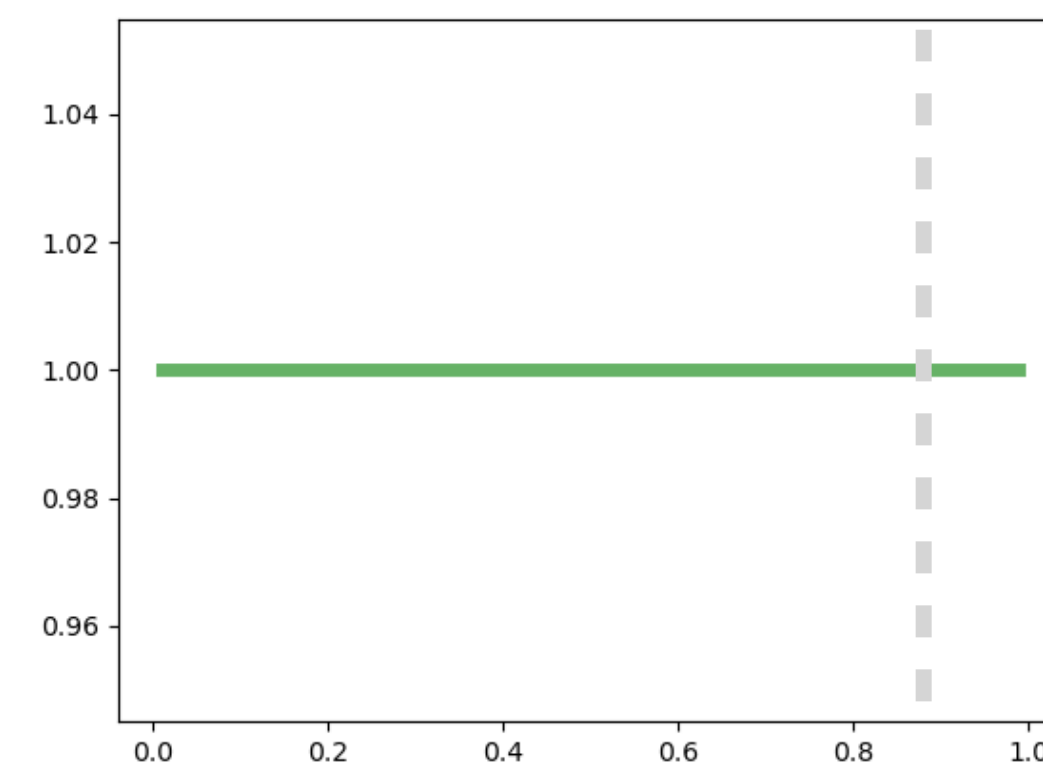
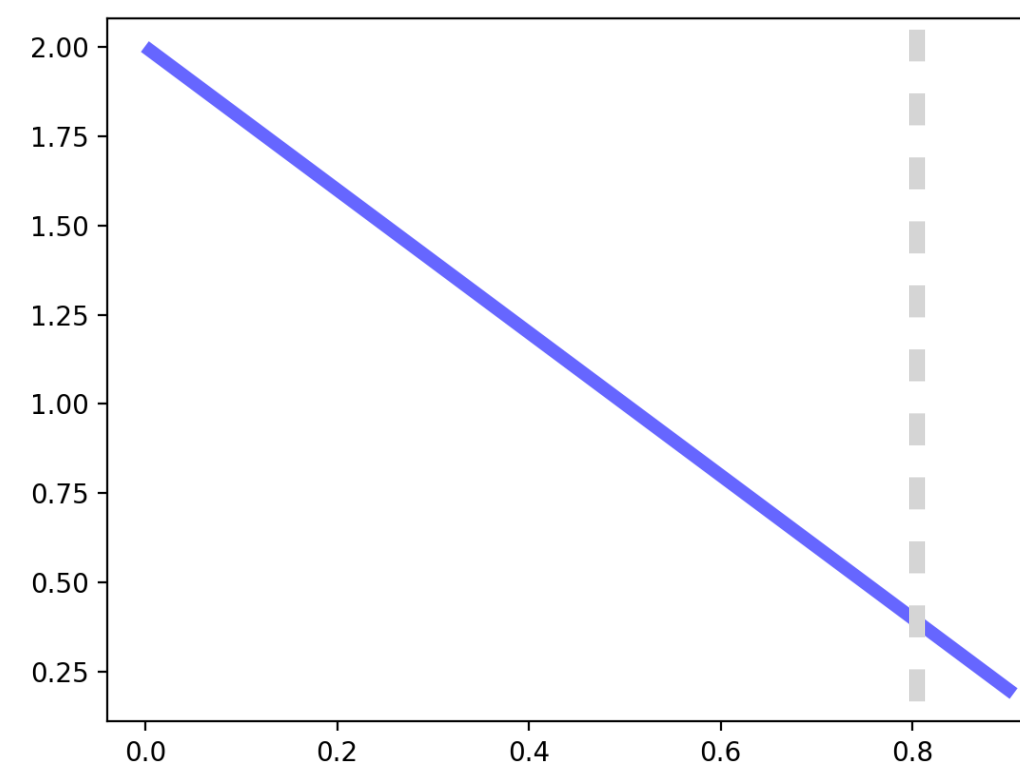
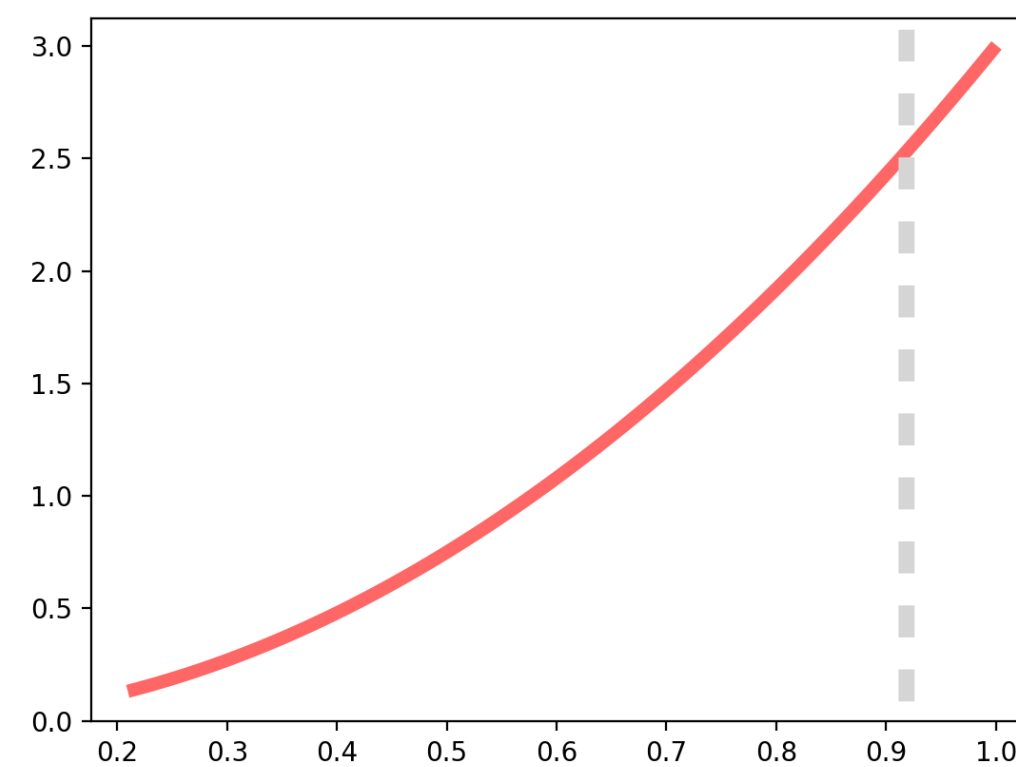




2

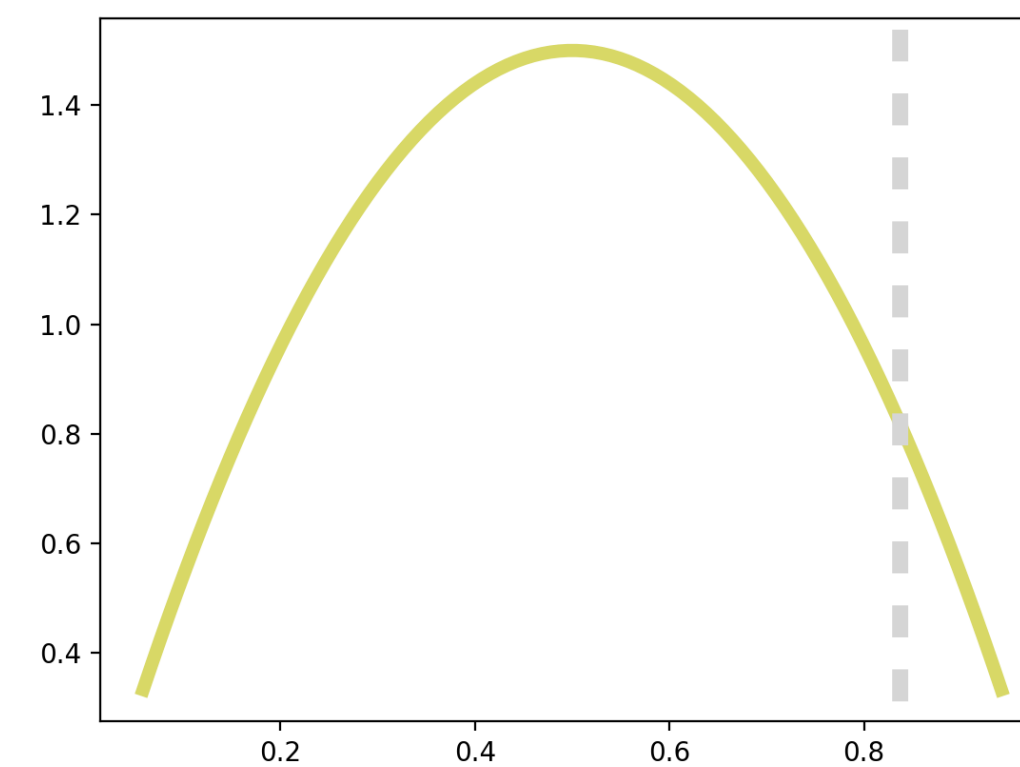
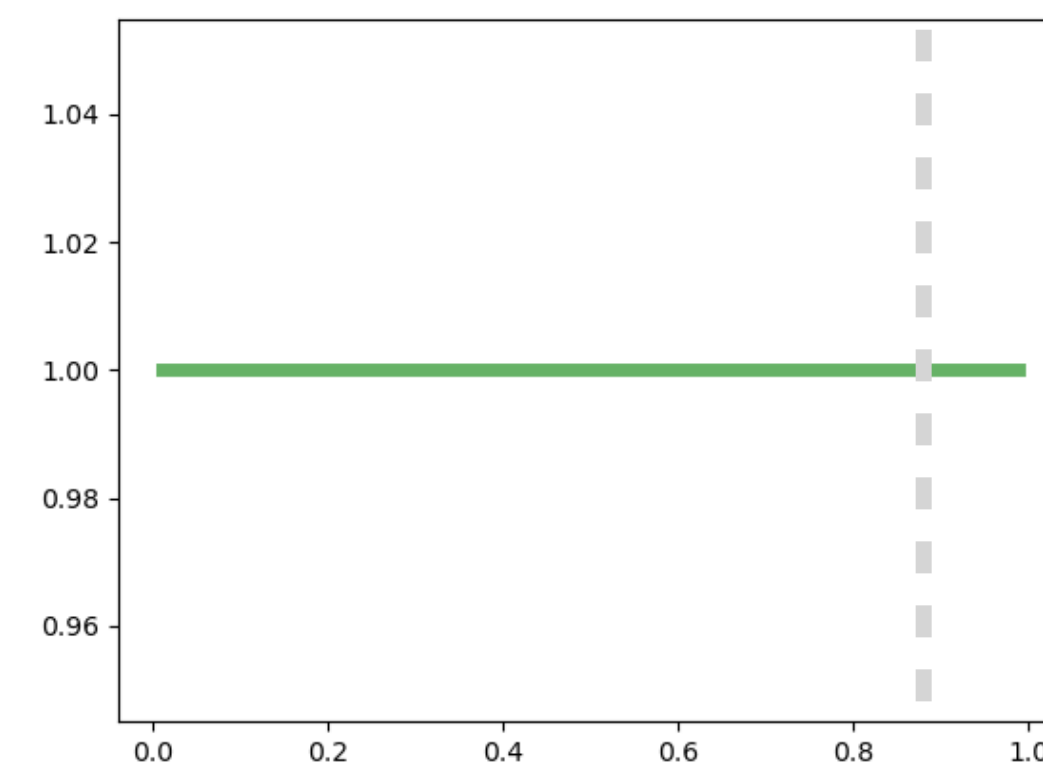
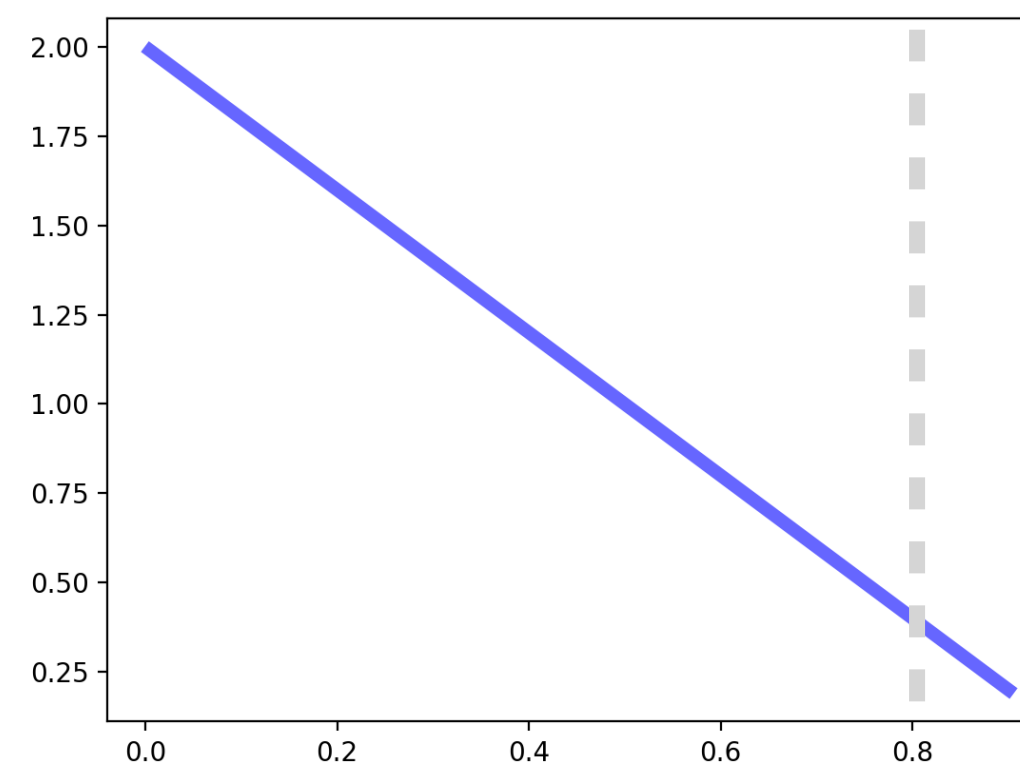
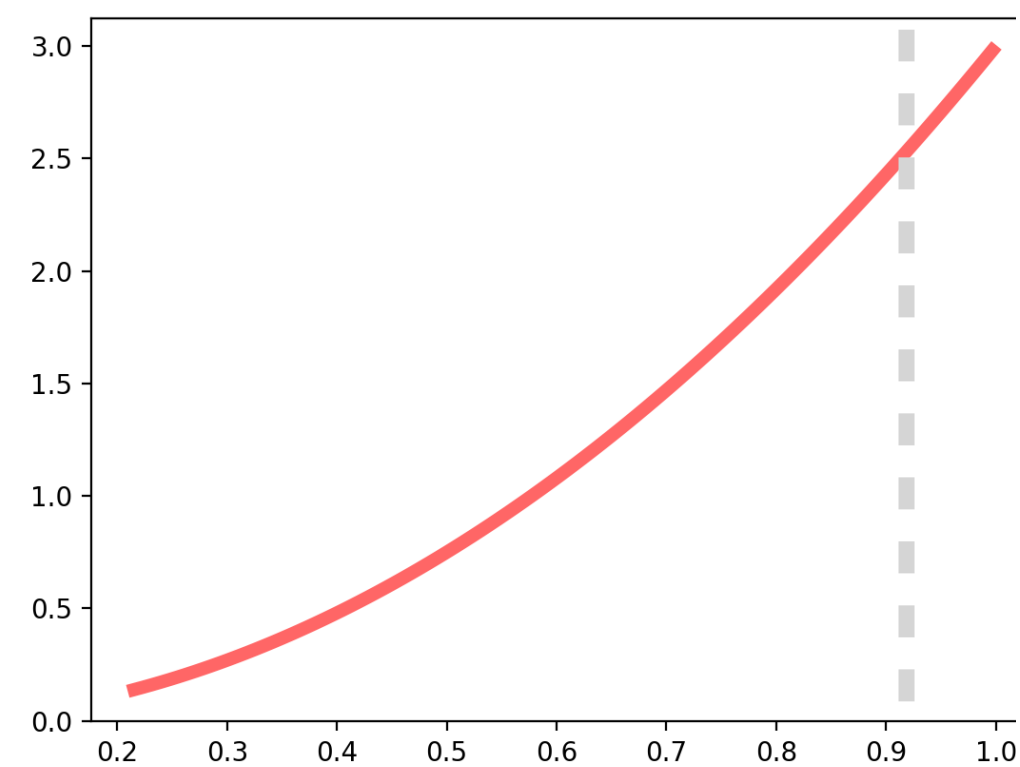


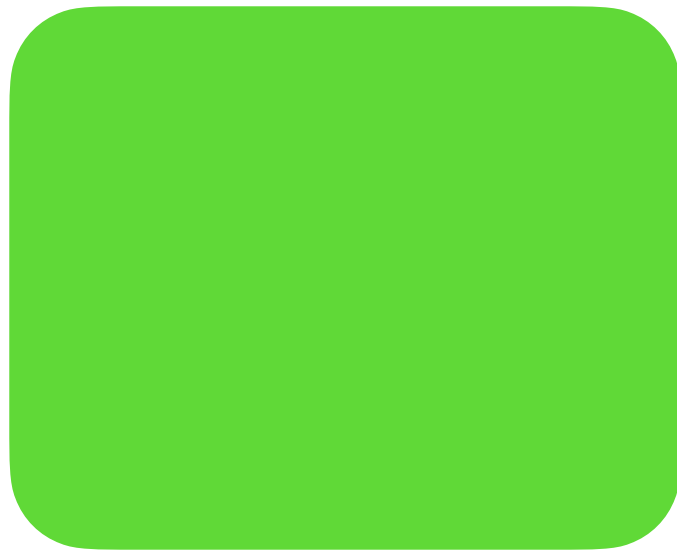
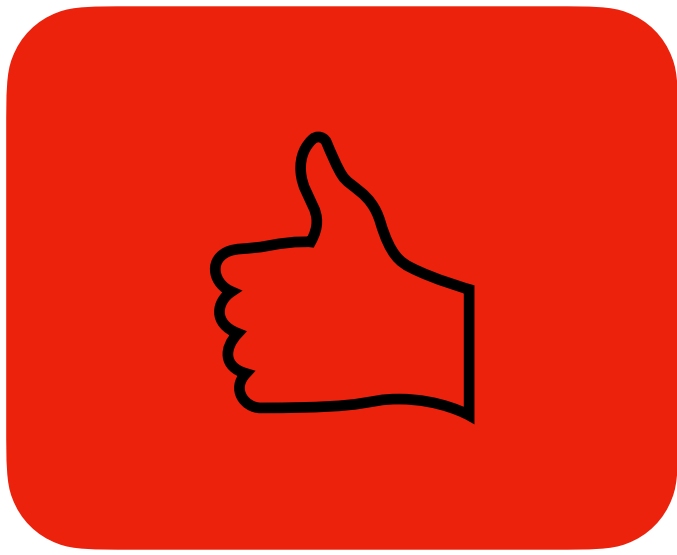
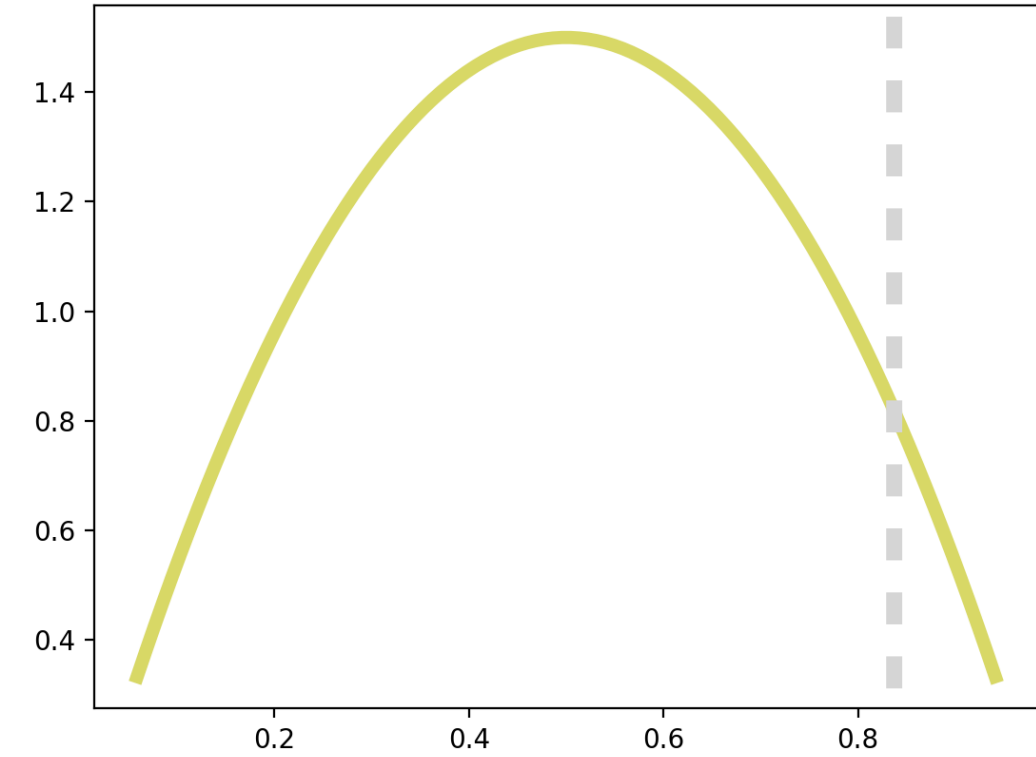
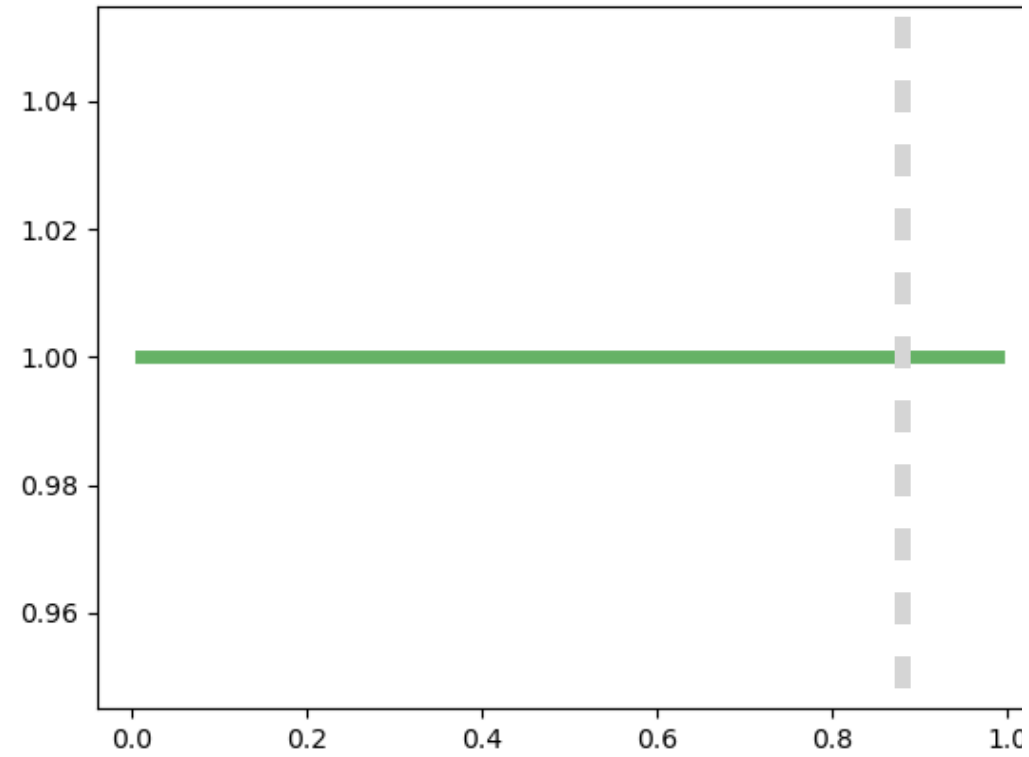
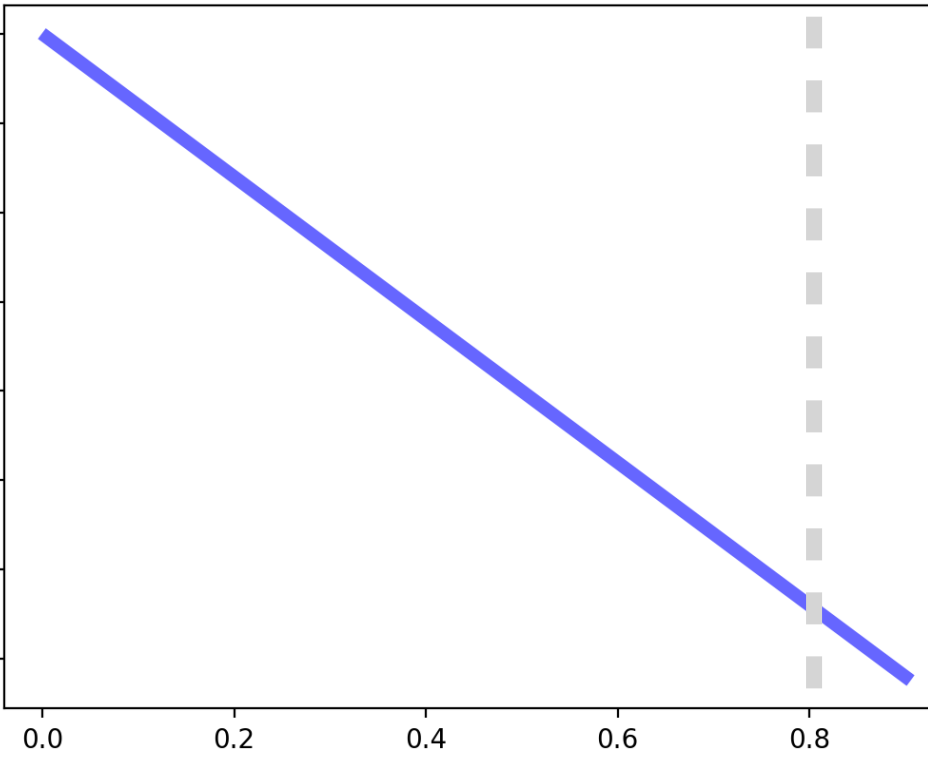
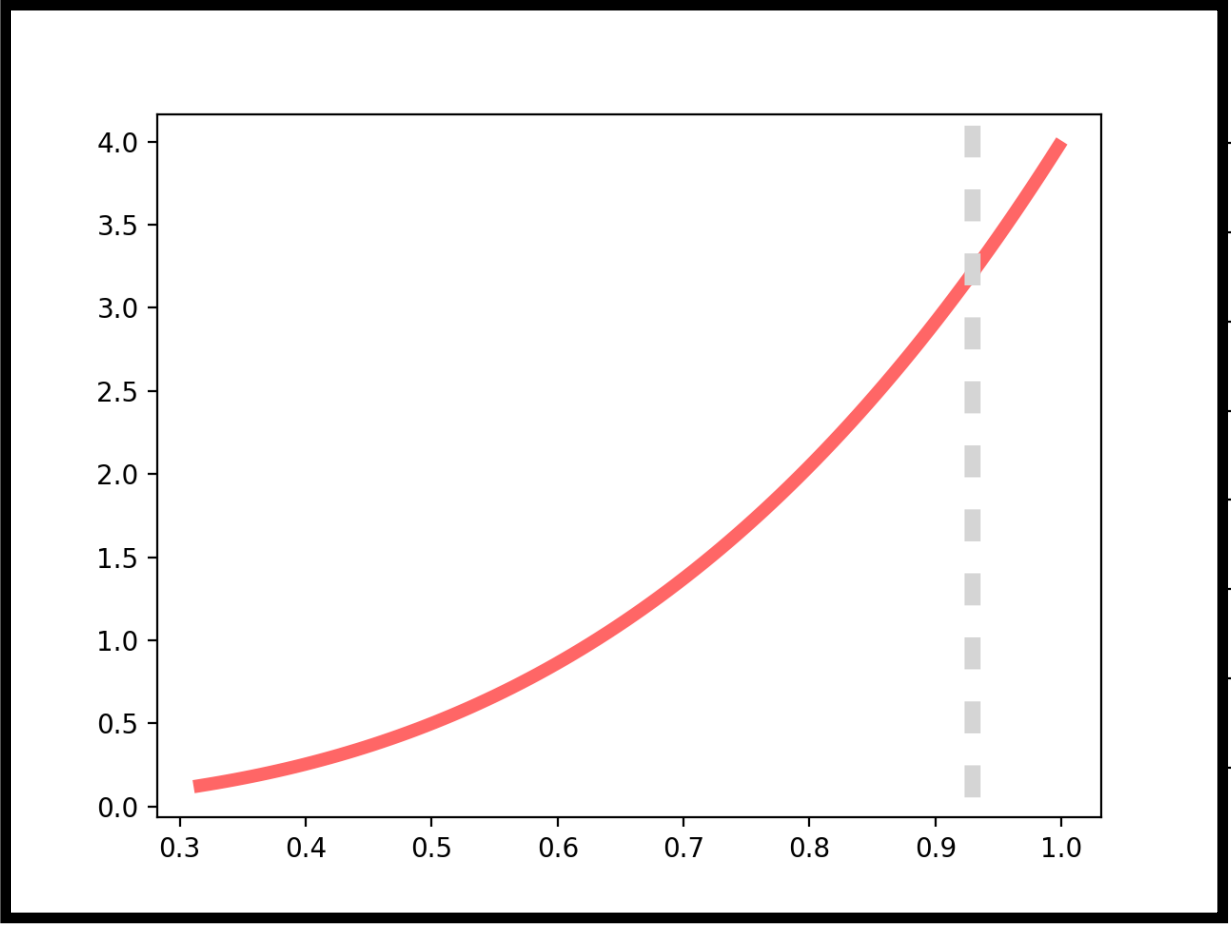


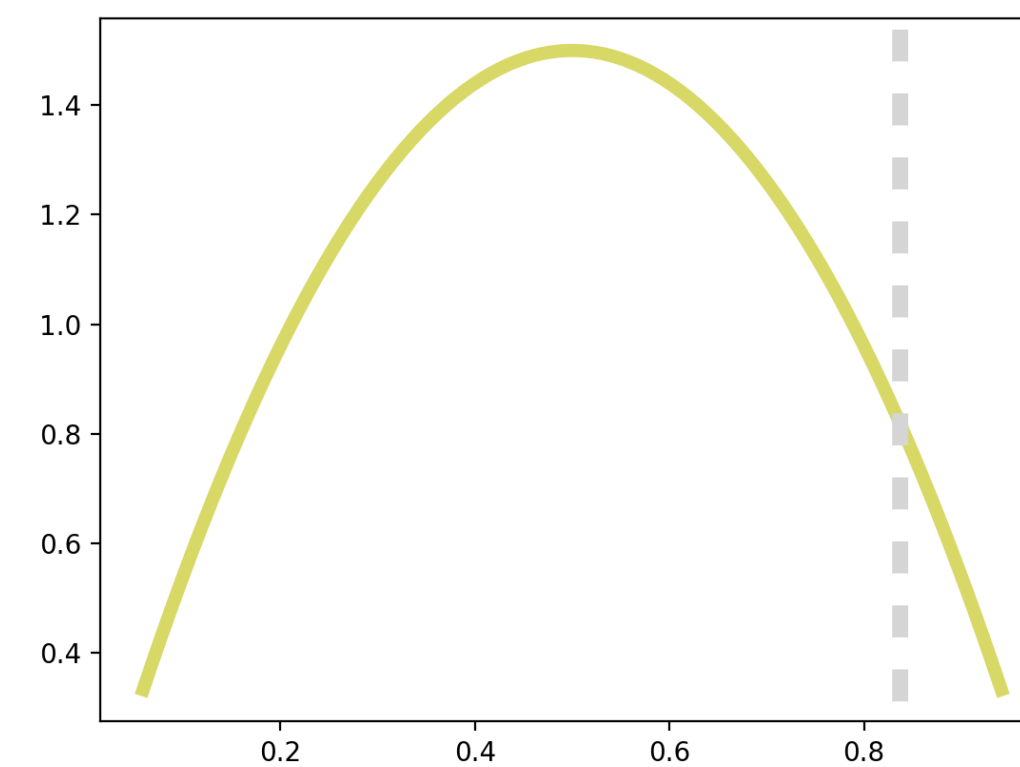
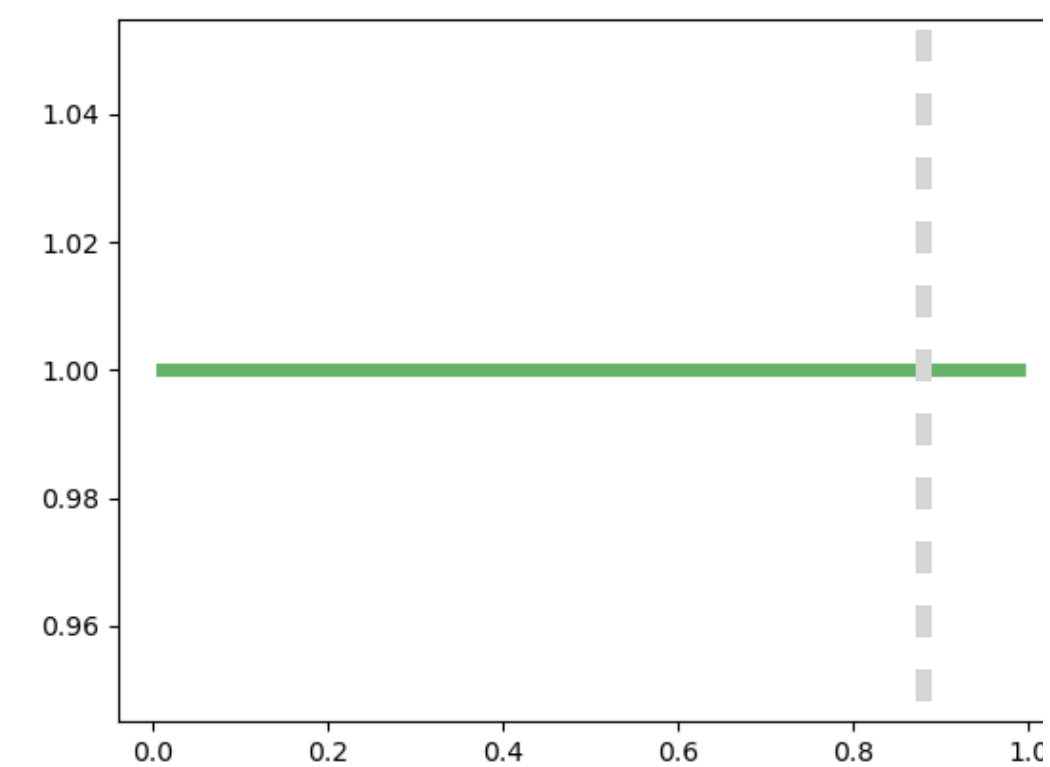
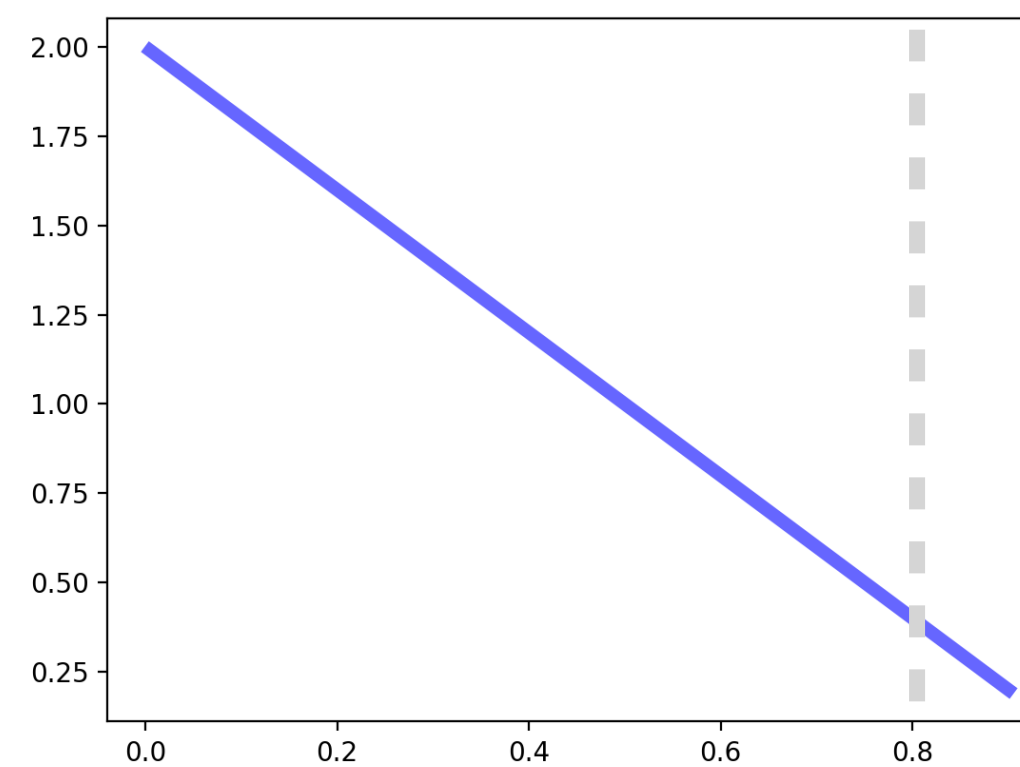
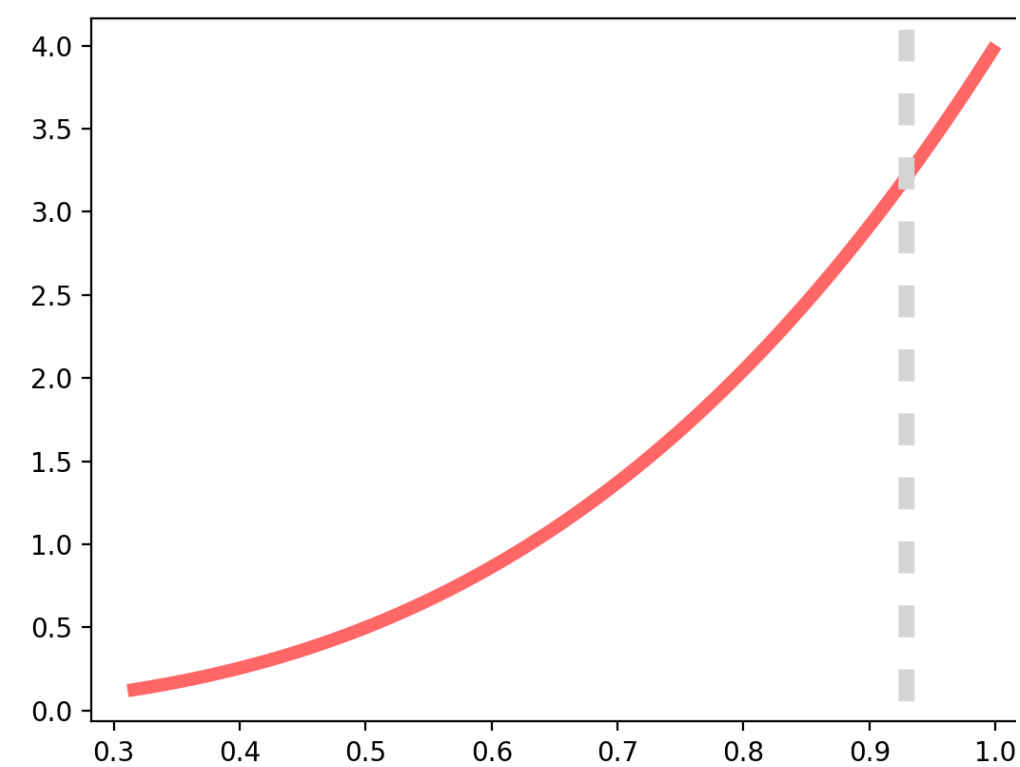


3

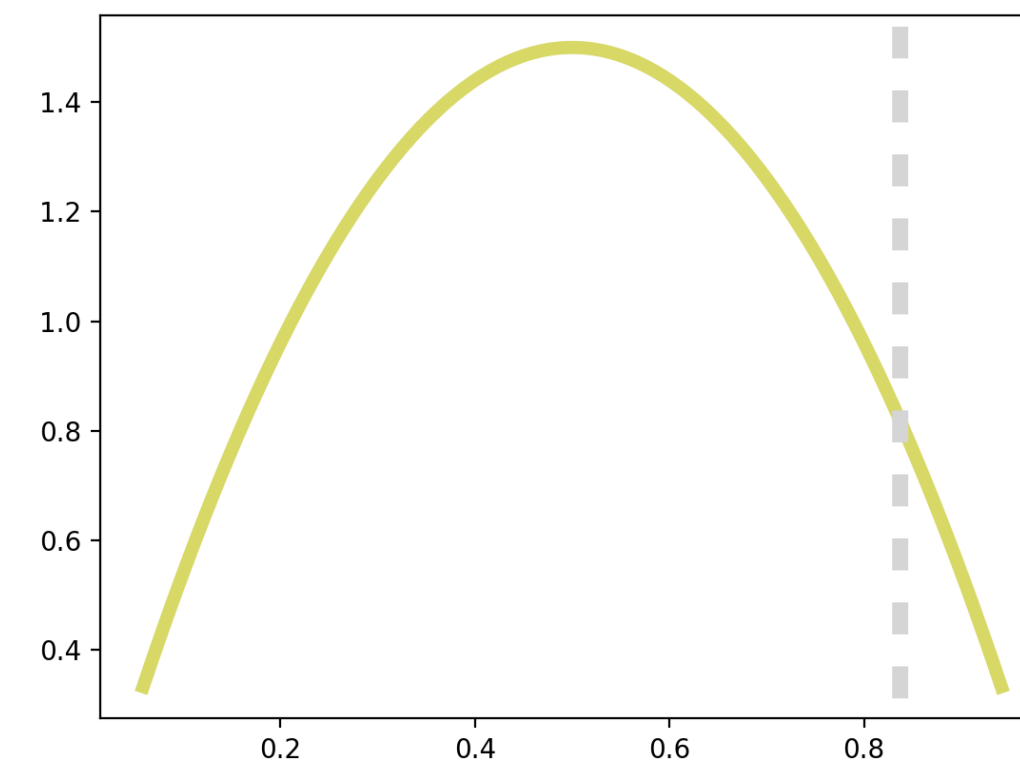
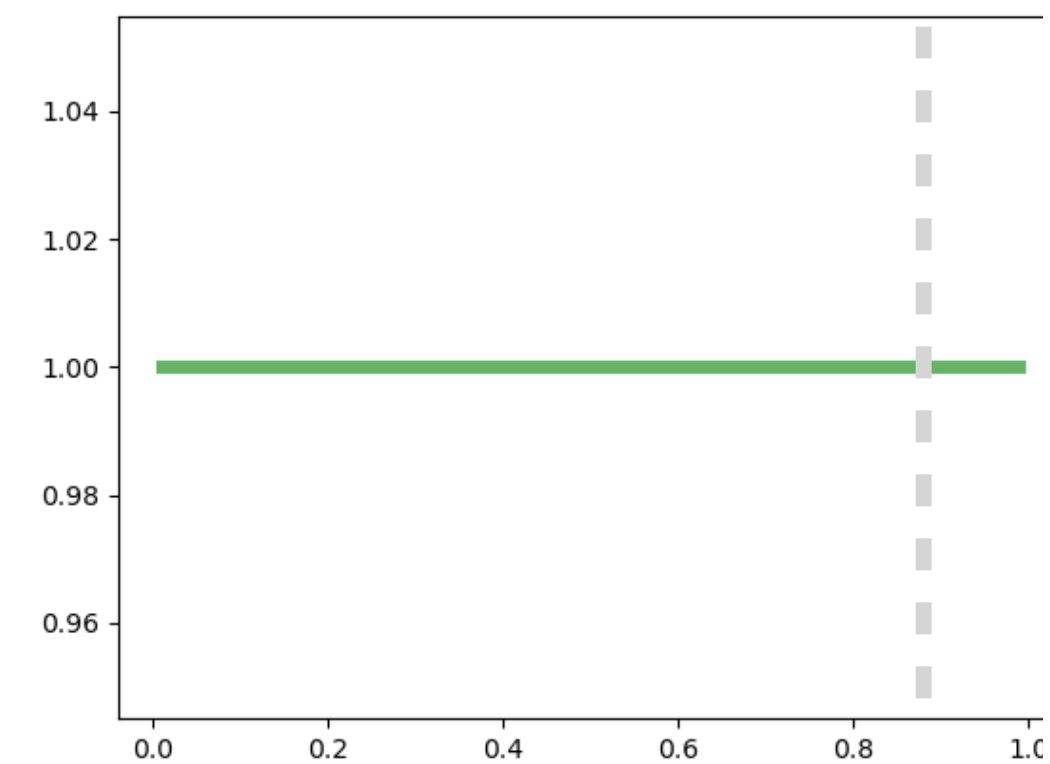
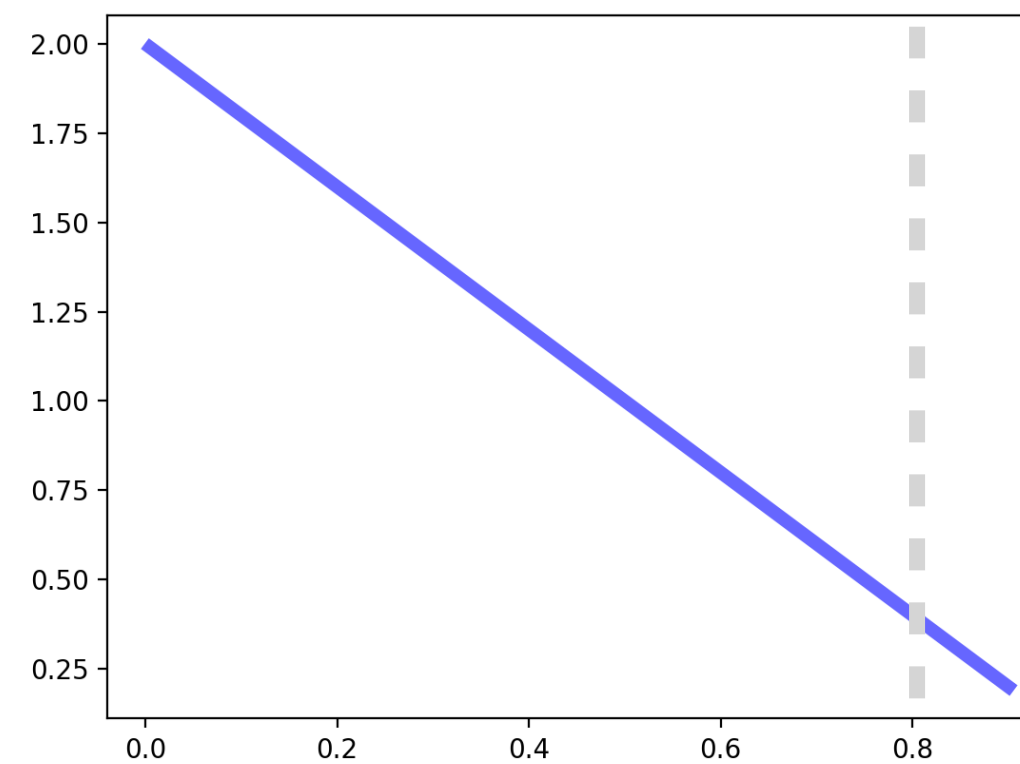
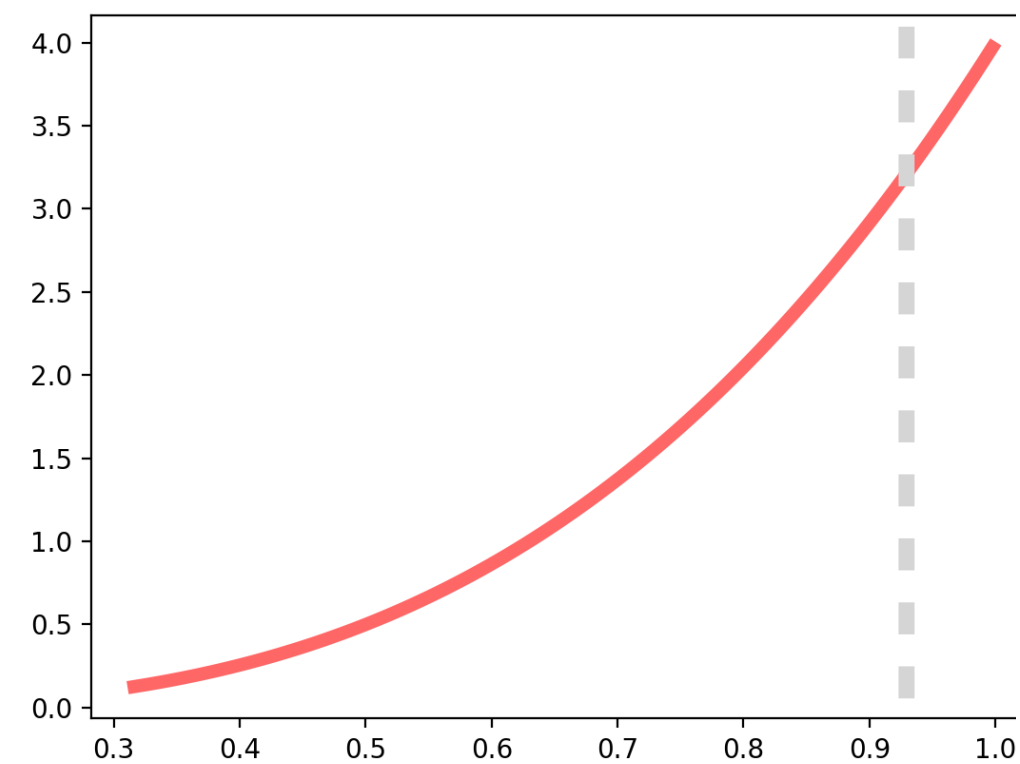


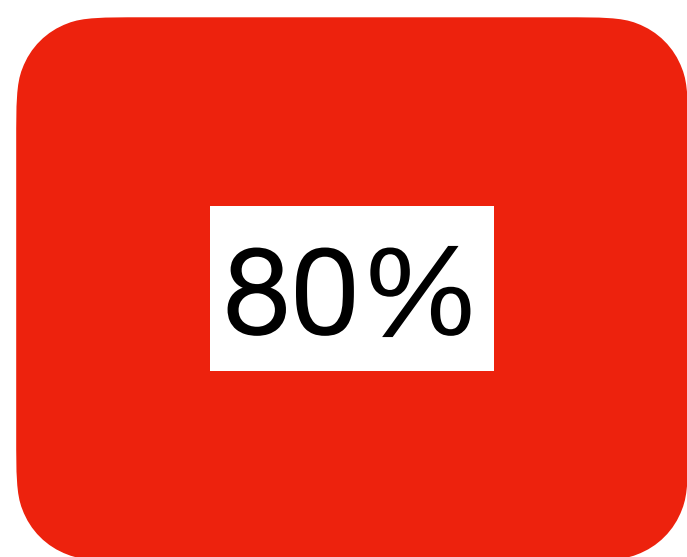
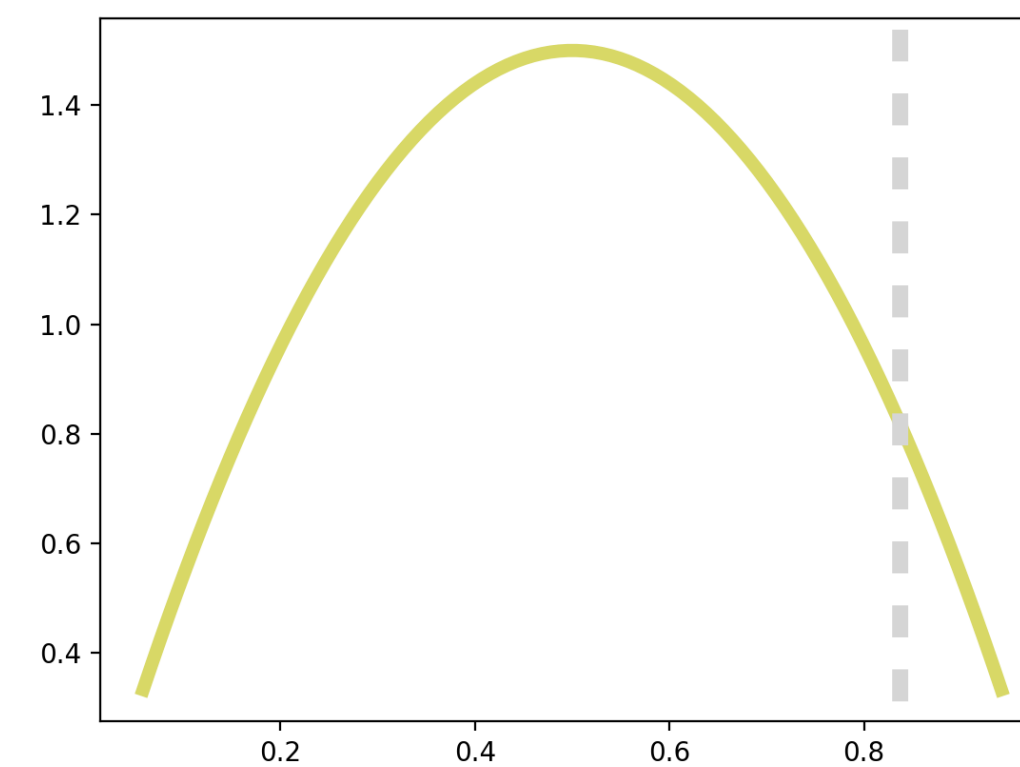
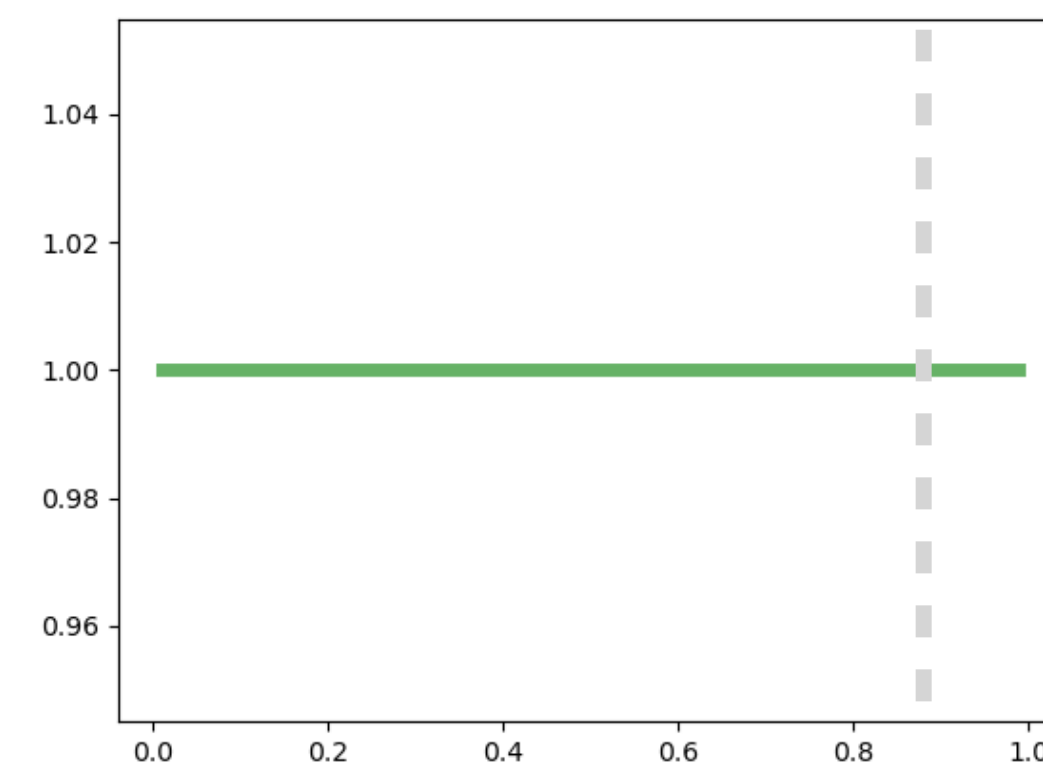
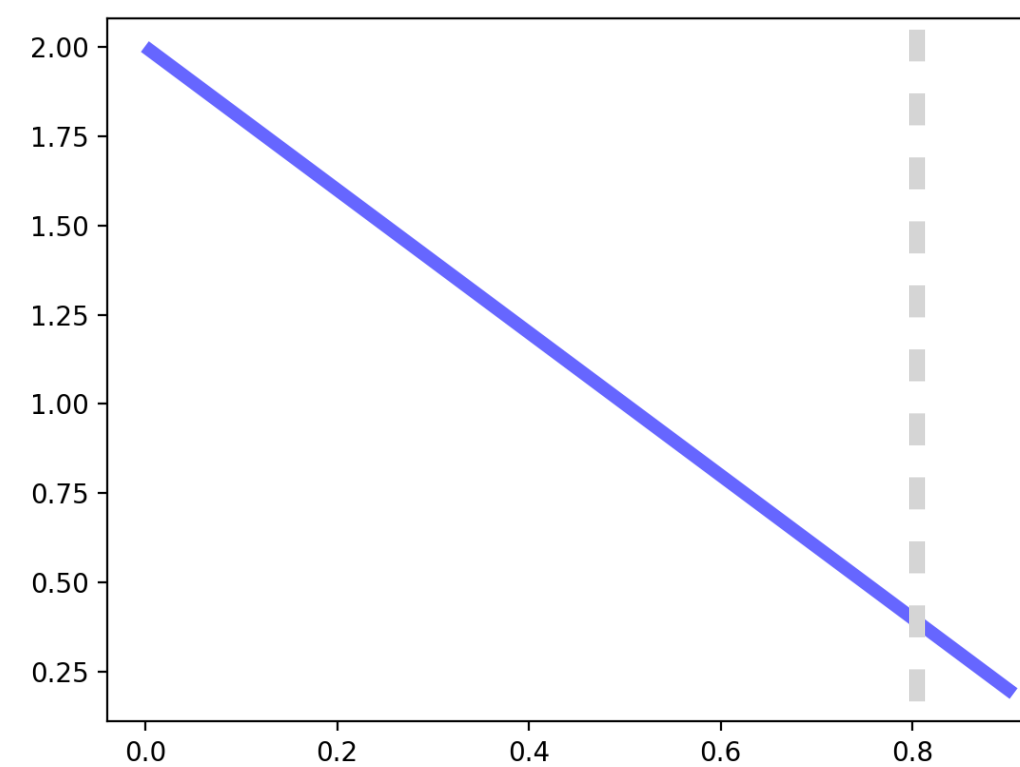
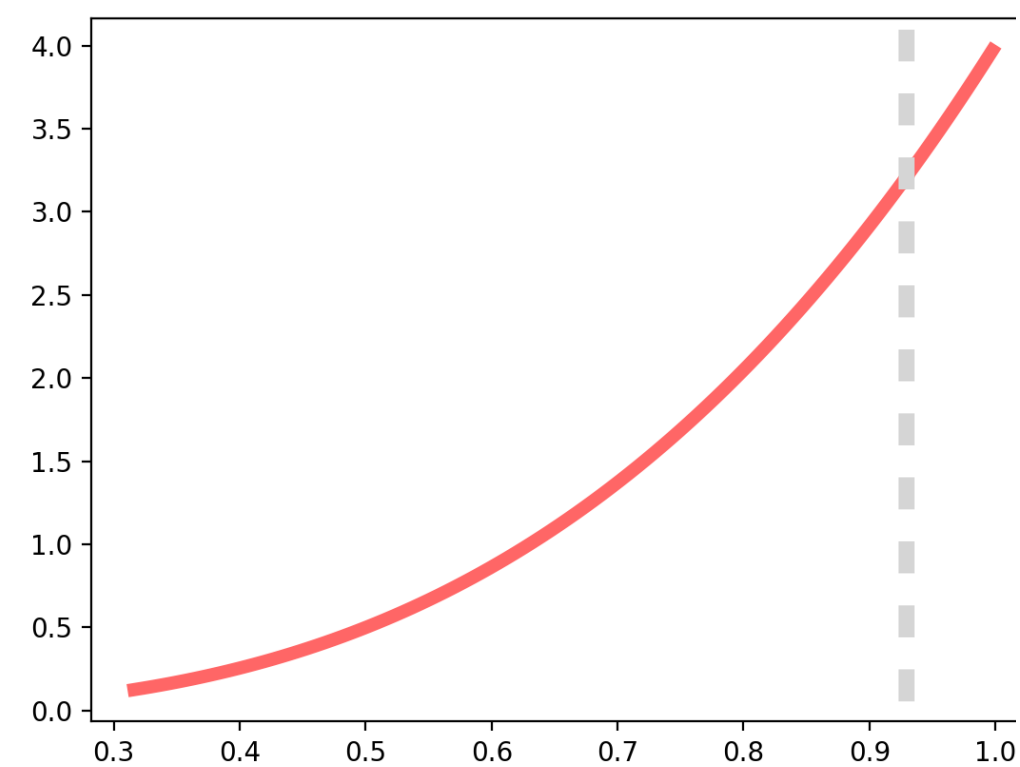




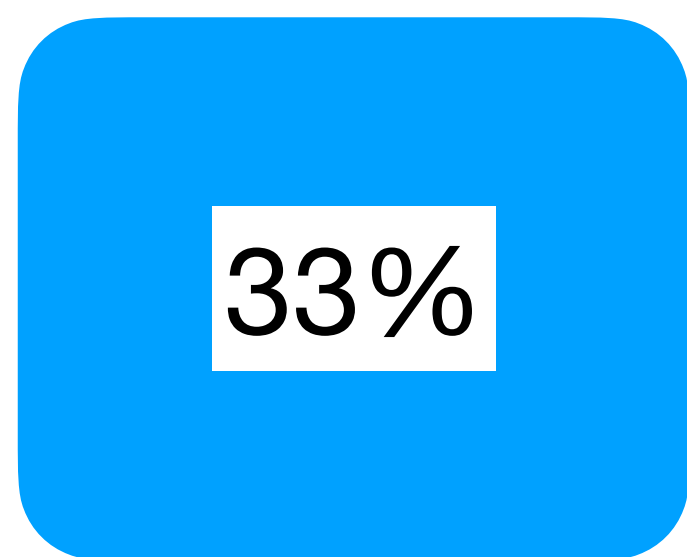


4





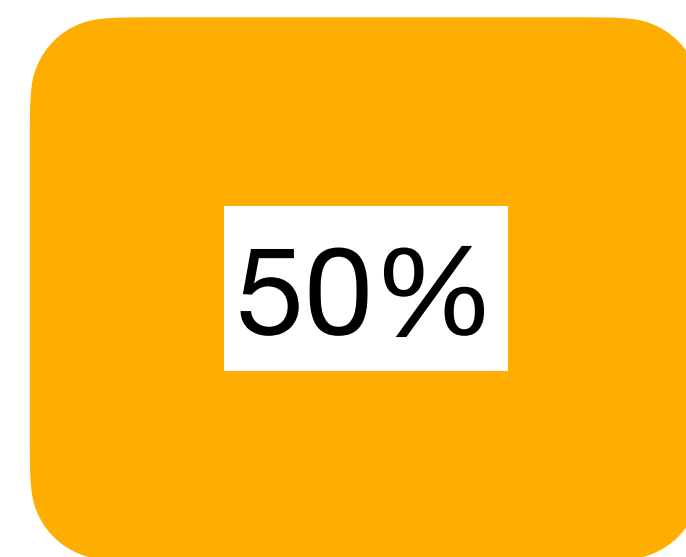
4



1



1



2

Interim summary

Rational exploration leads to inaccurate impressions

Interim summary

Rational exploration leads to inaccurate impressions

- **The goal is rational/minimal: Maximize the chance of being helped**
 - Not motivated to demean groups

Interim summary

Rational exploration leads to inaccurate impressions

- **The goal is rational/minimal: Maximize the chance of being helped**
 - Not motivated to demean groups
- **The process is rational/minimal: Belief updating**
 - Not unable to absorb new information

Interim summary

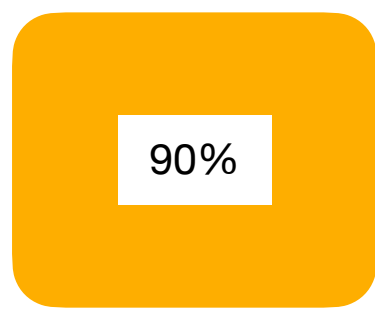
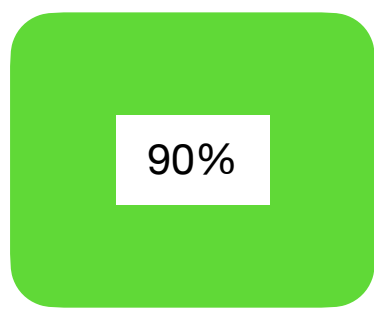
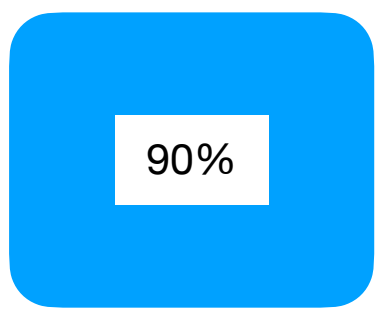
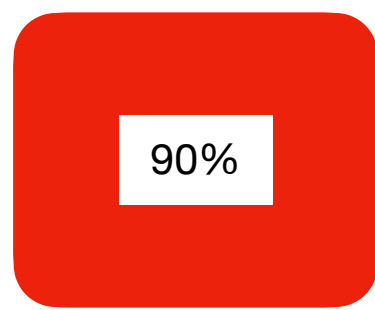
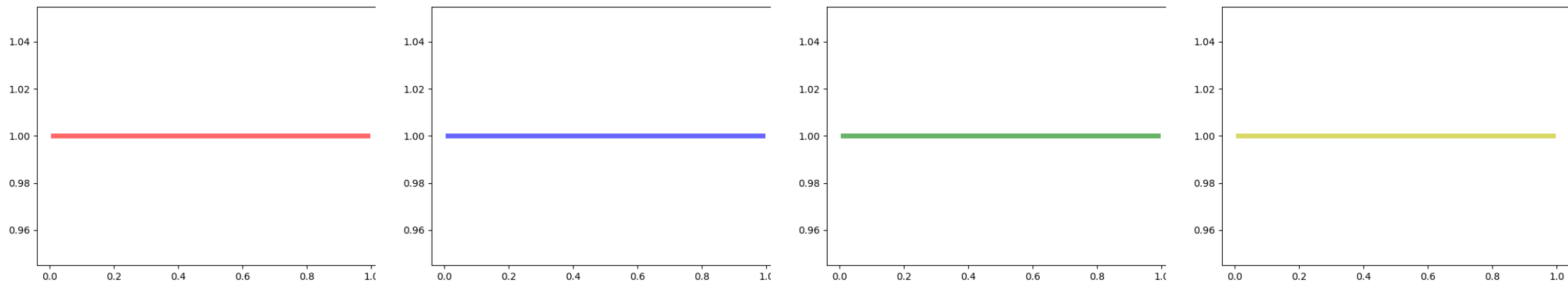
Rational exploration leads to inaccurate impressions

- **The goal is rational/minimal: Maximize the chance of being helped**
 - Not motivated to demean groups
- **The process is rational/minimal: Belief updating**
 - Not unable to absorb new information
- **The outcome is inaccurate: Find an ostensibly best group**
 - Although unintended, inaccurate impressions about the under-explored groups

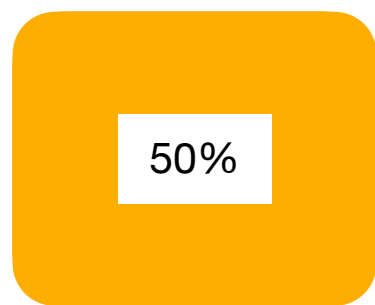
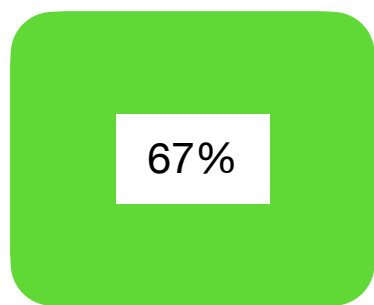
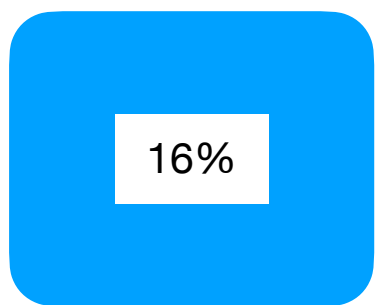
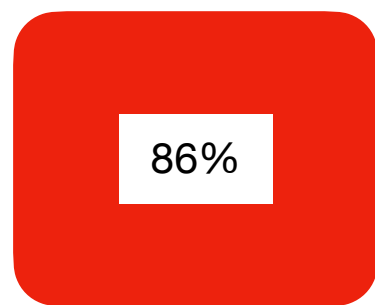
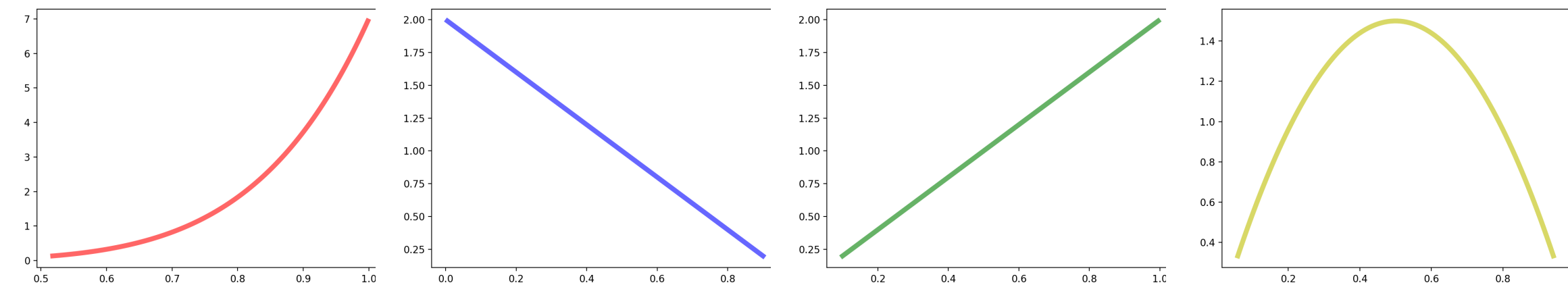
Our hypothesis

Globally inaccurate stereotypes can result from locally rational exploration

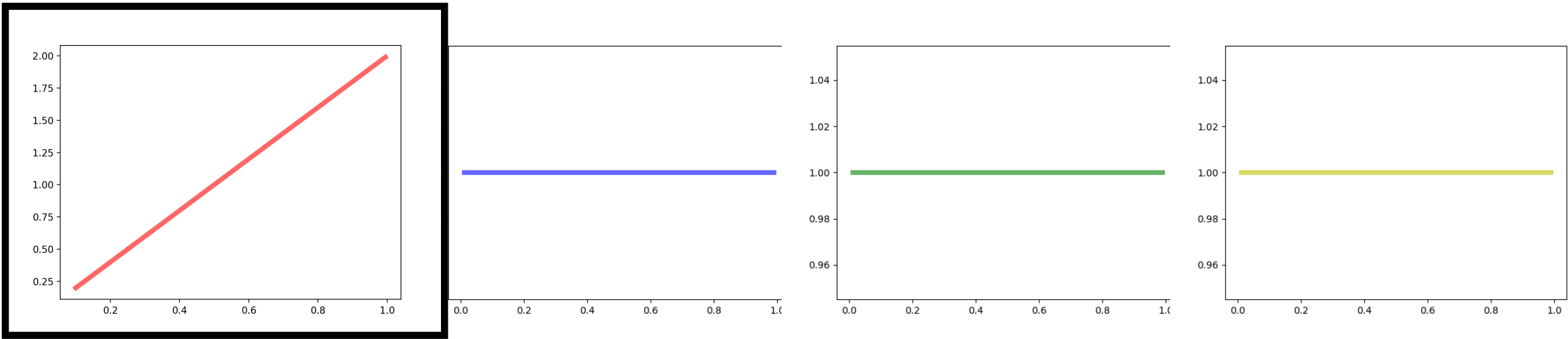
- An intuitive example
- **A formal model and simulations**
- Two human experiments
- Implications



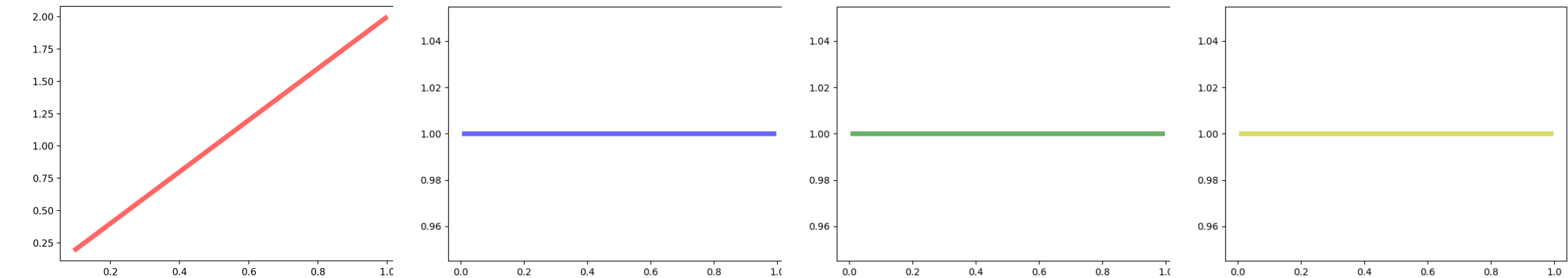
Multi-Armed Bandit Problem



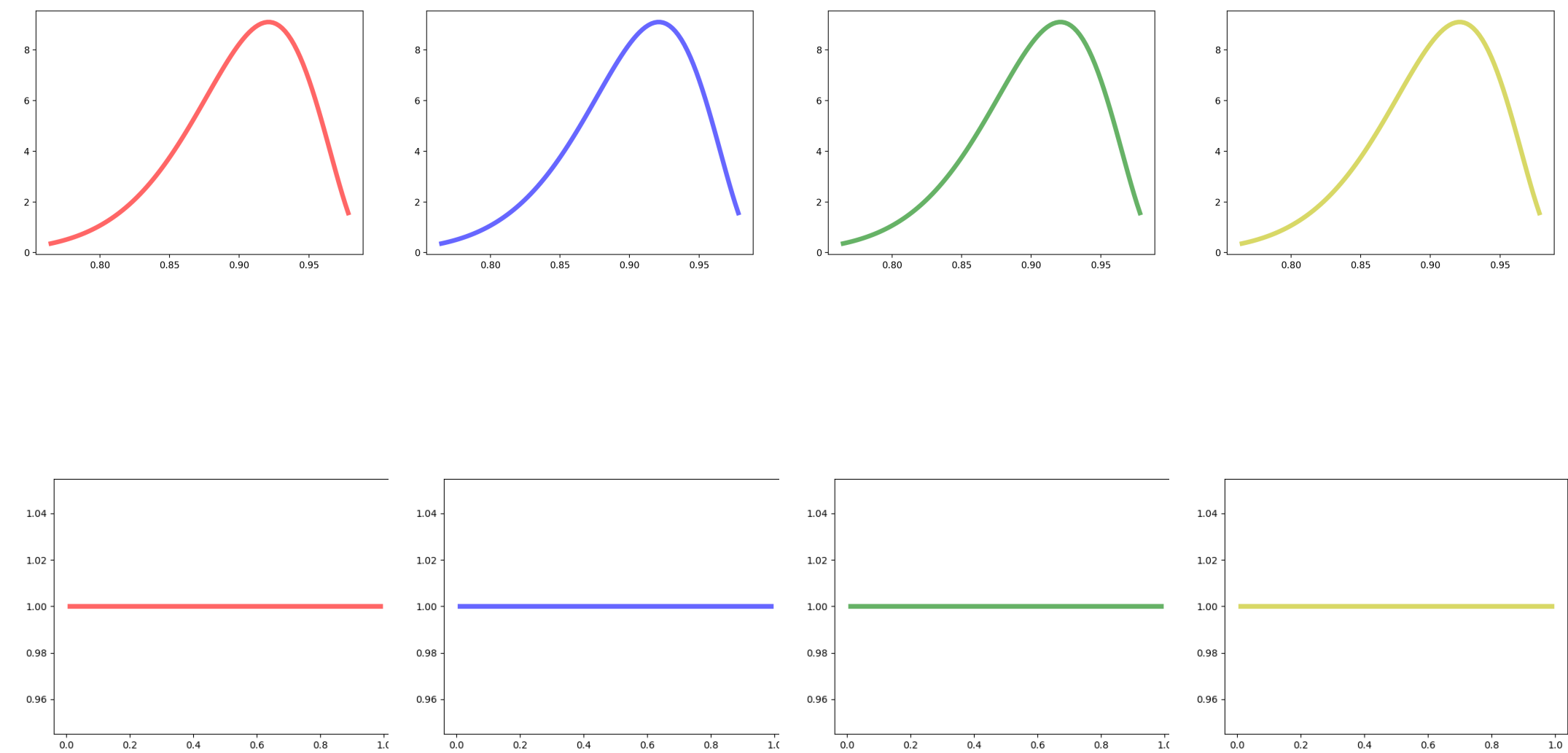
Explore v. Exploit dilemma in Reinforcement Learning (Sutton & Barto, 2018)



Thompson Sampling



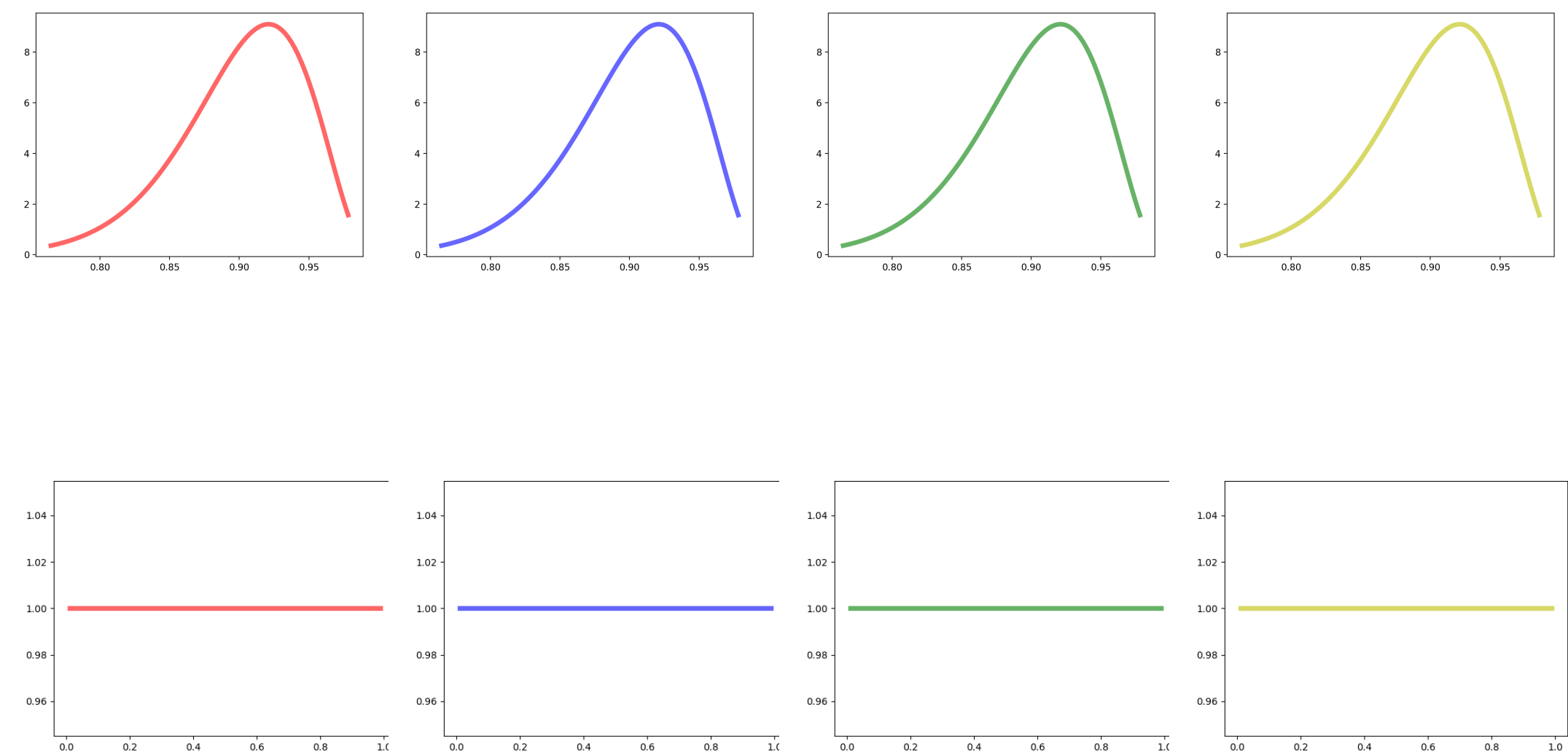
Our hypothesis



Environment structure: Identical high reward

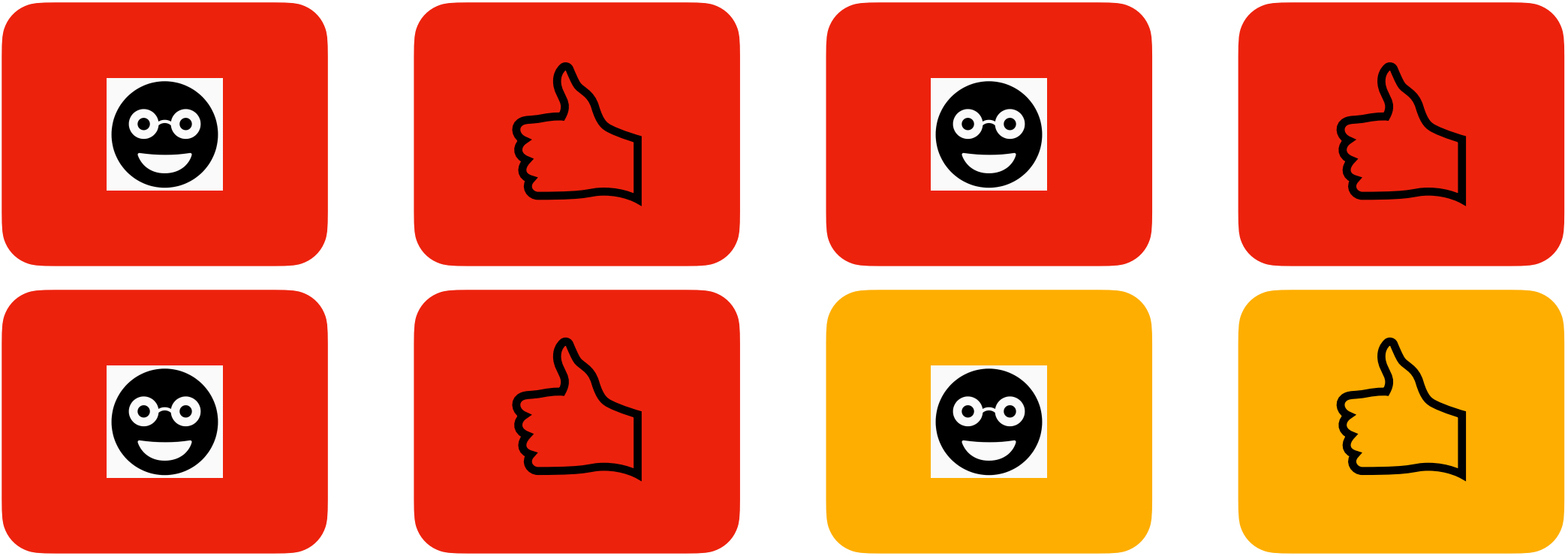
Initial structure: No prior bias

Our hypothesis



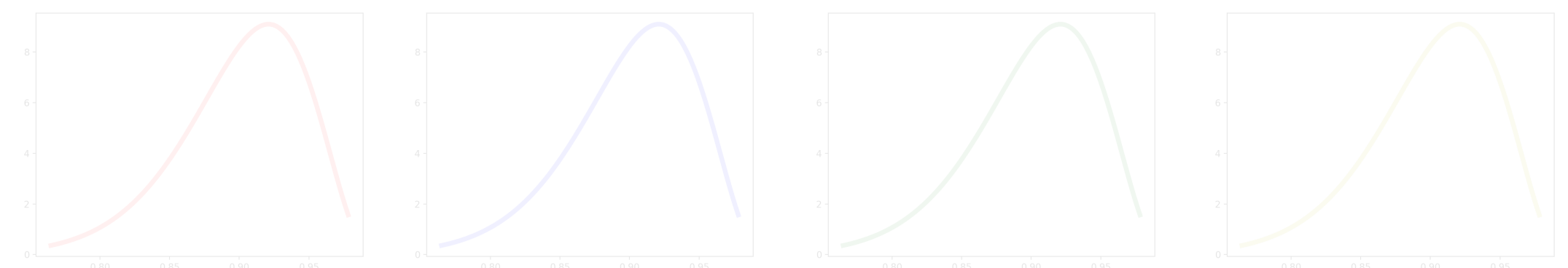
Environment structure: Identical high reward

Initial structure: No prior bias



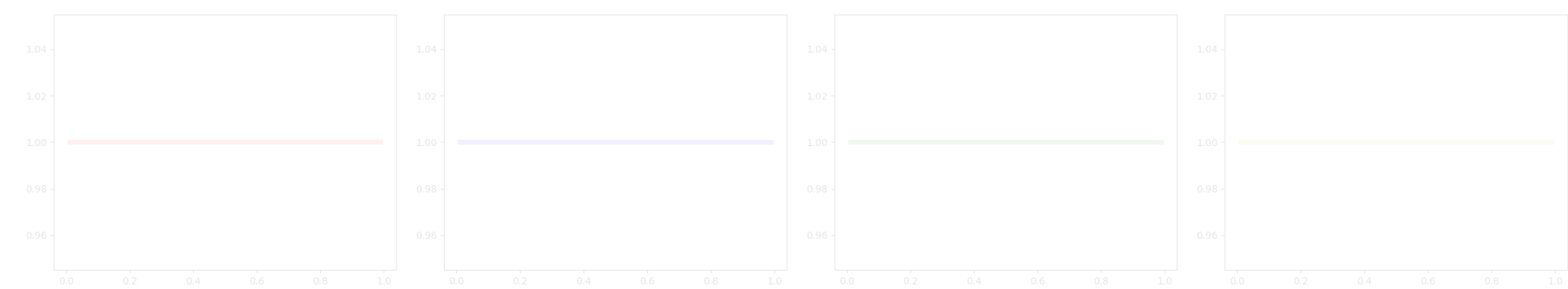
Sampling strategy: Thompson v. Random

Our hypothesis



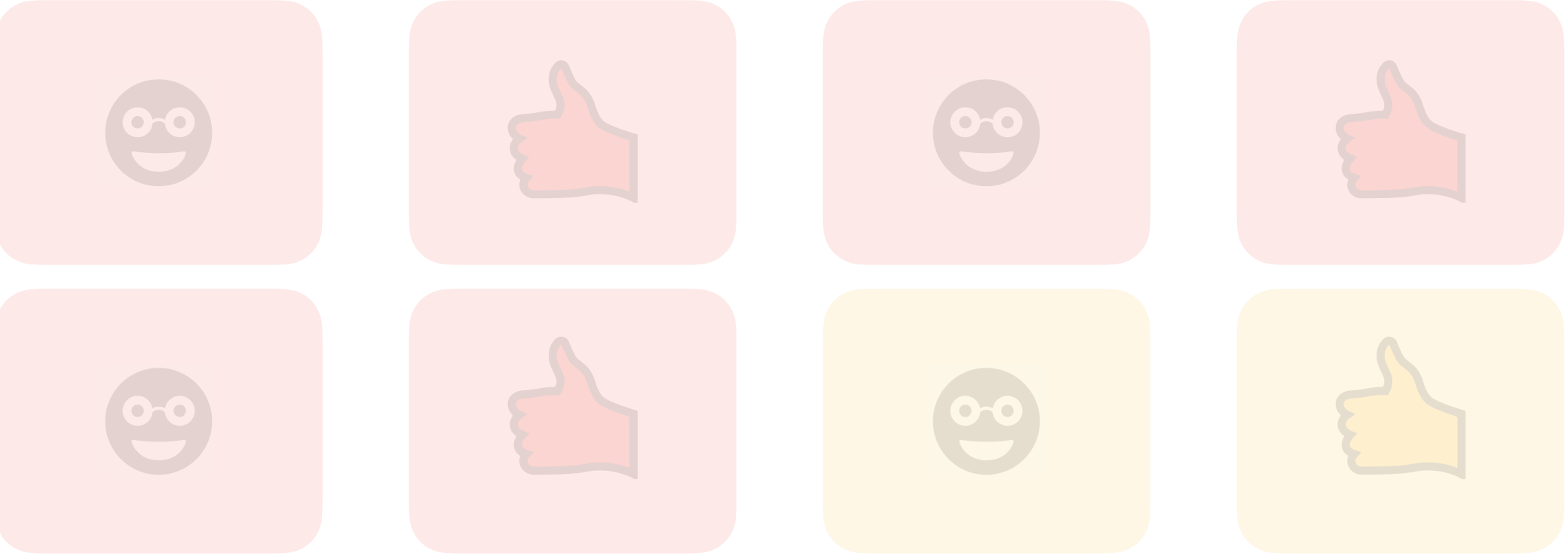
Environment structure: Identical reward

Partner choices: How many times did the model interact with each group?



Initial structure: No prior bias

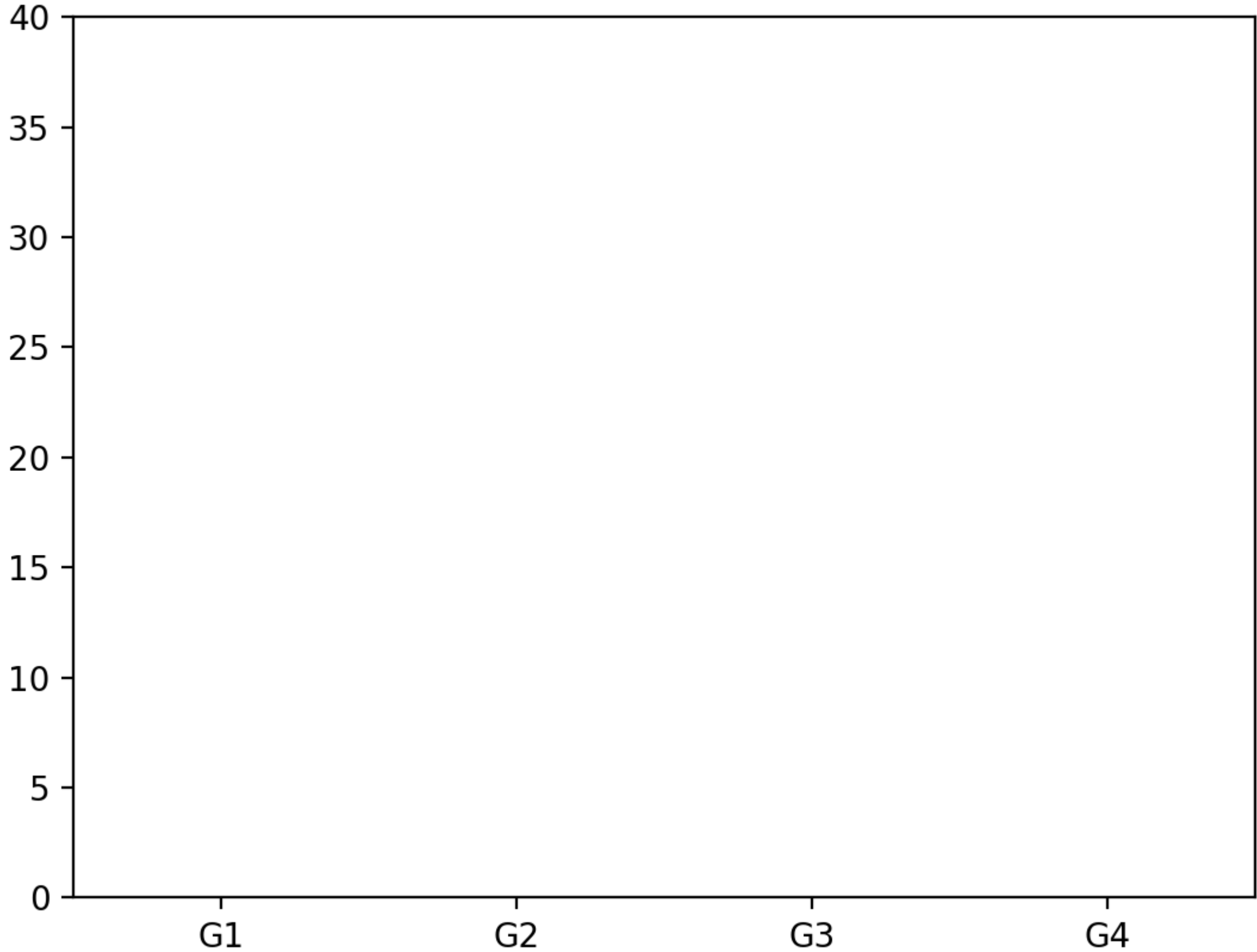
Reward estimations: What is the estimated reward for each group?



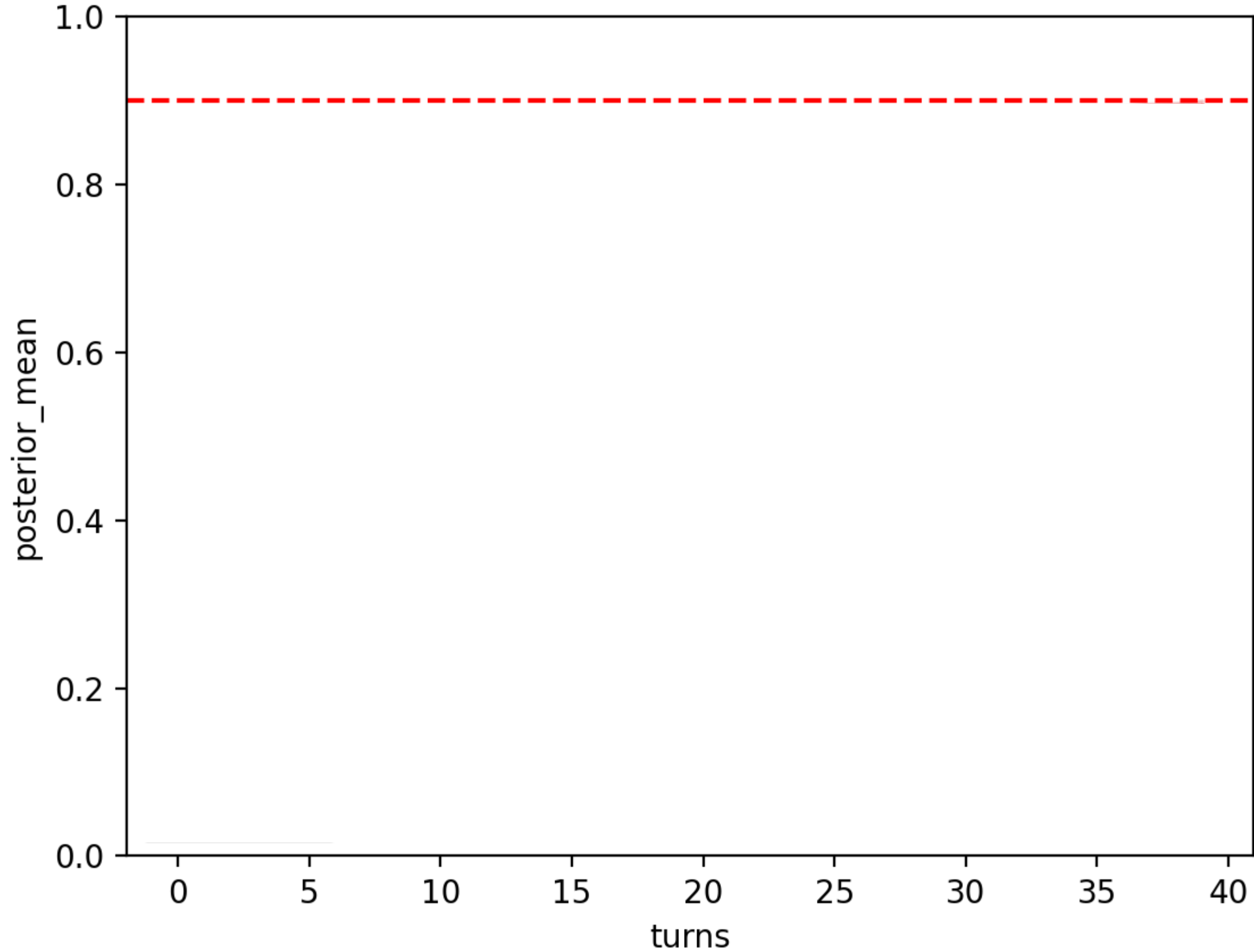
Sampling strategy: Thompson v. Random

Identical reward ($\mu = 0.9$), No prior, Random sampling

Partner choices

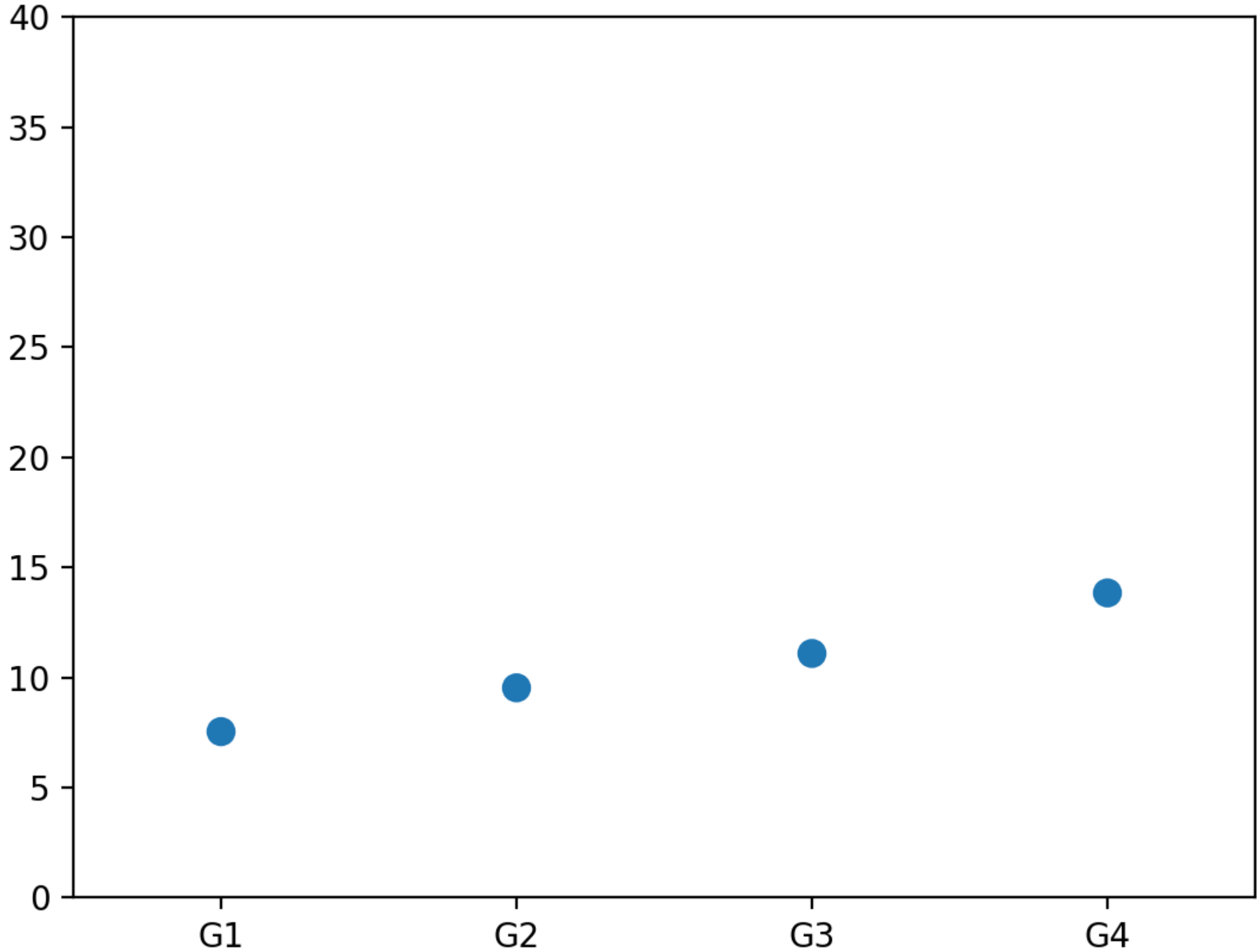


Reward estimations

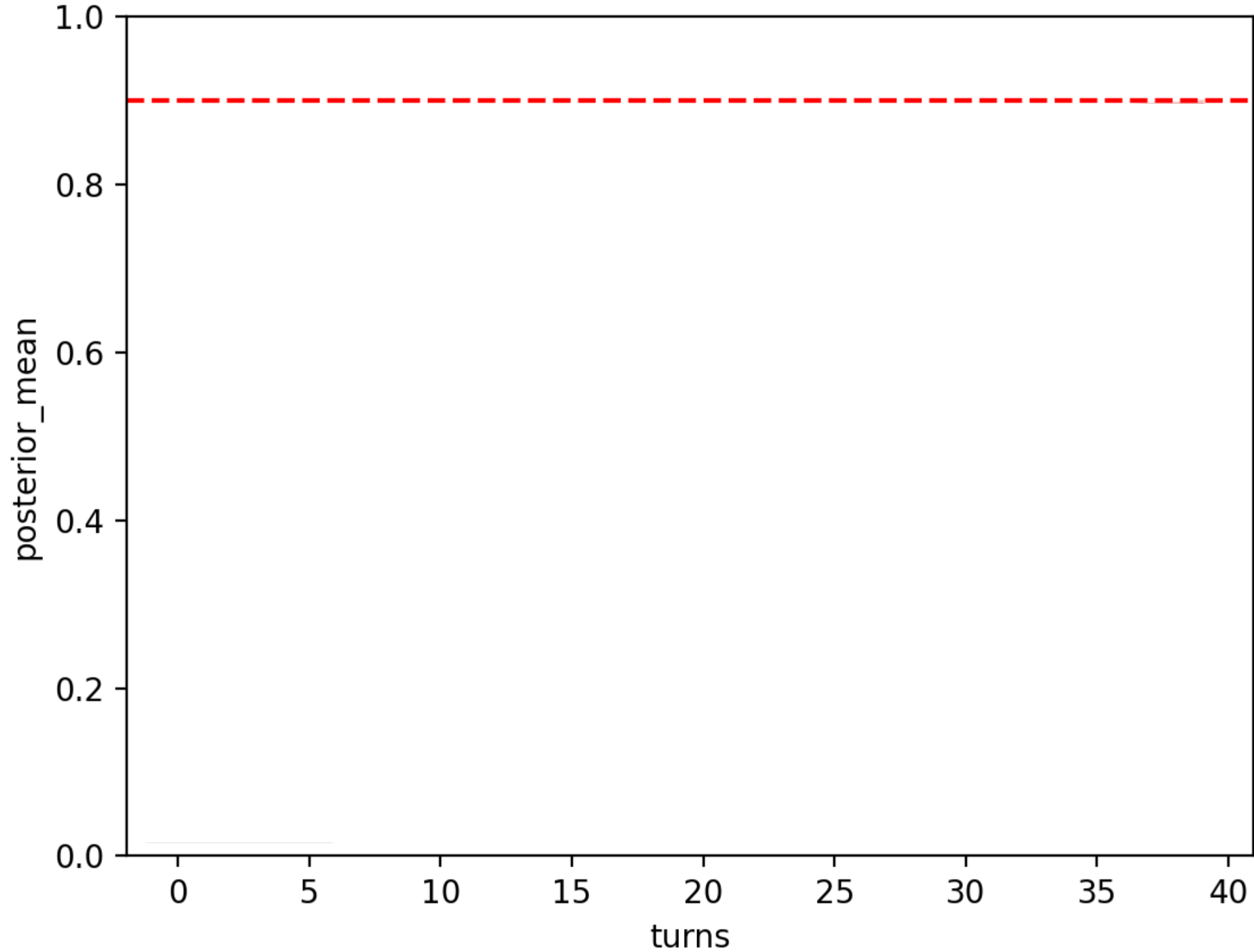


Identical reward ($\mu = 0.9$), No prior, Random sampling

Partner choices

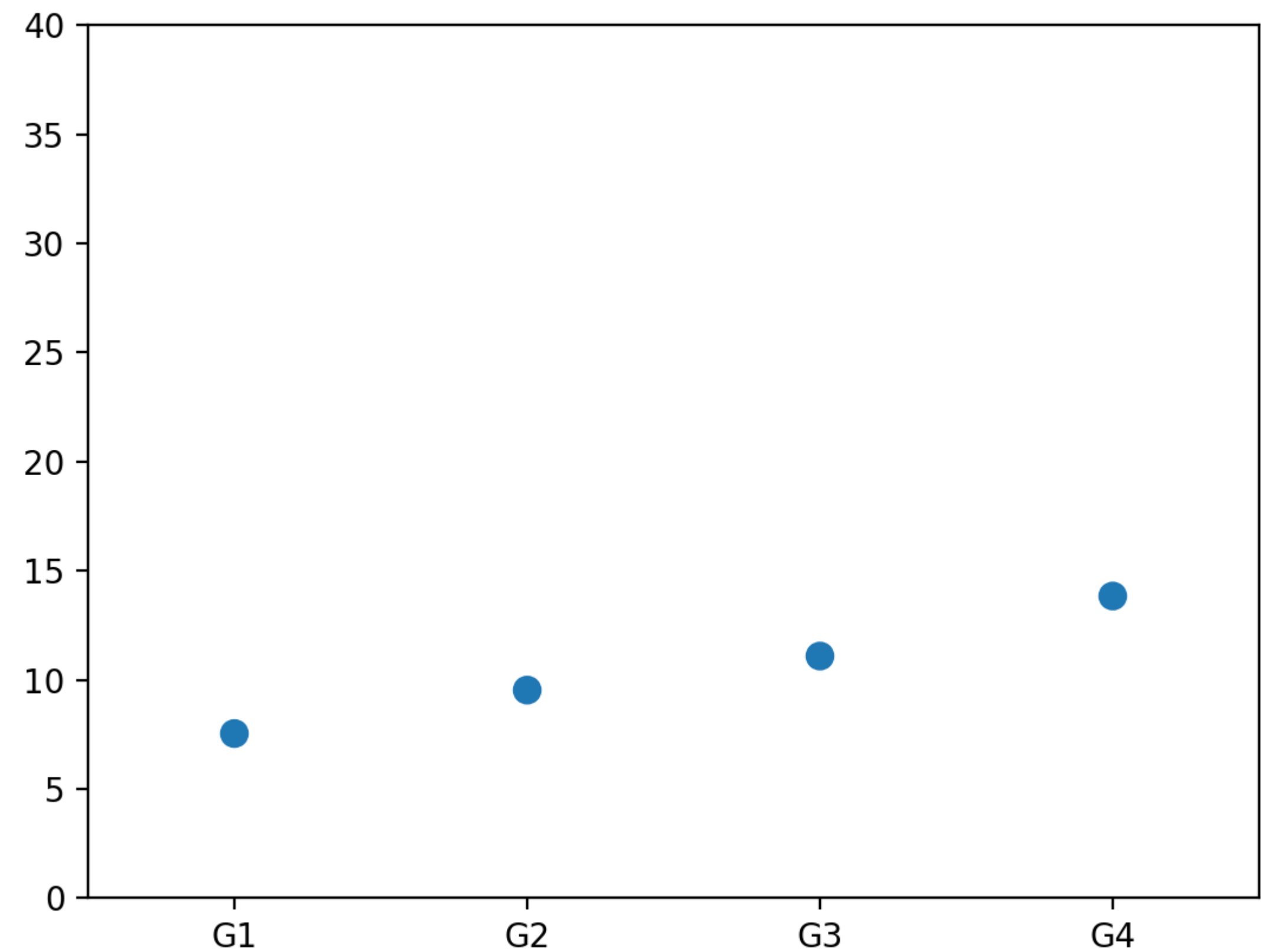


Reward estimations

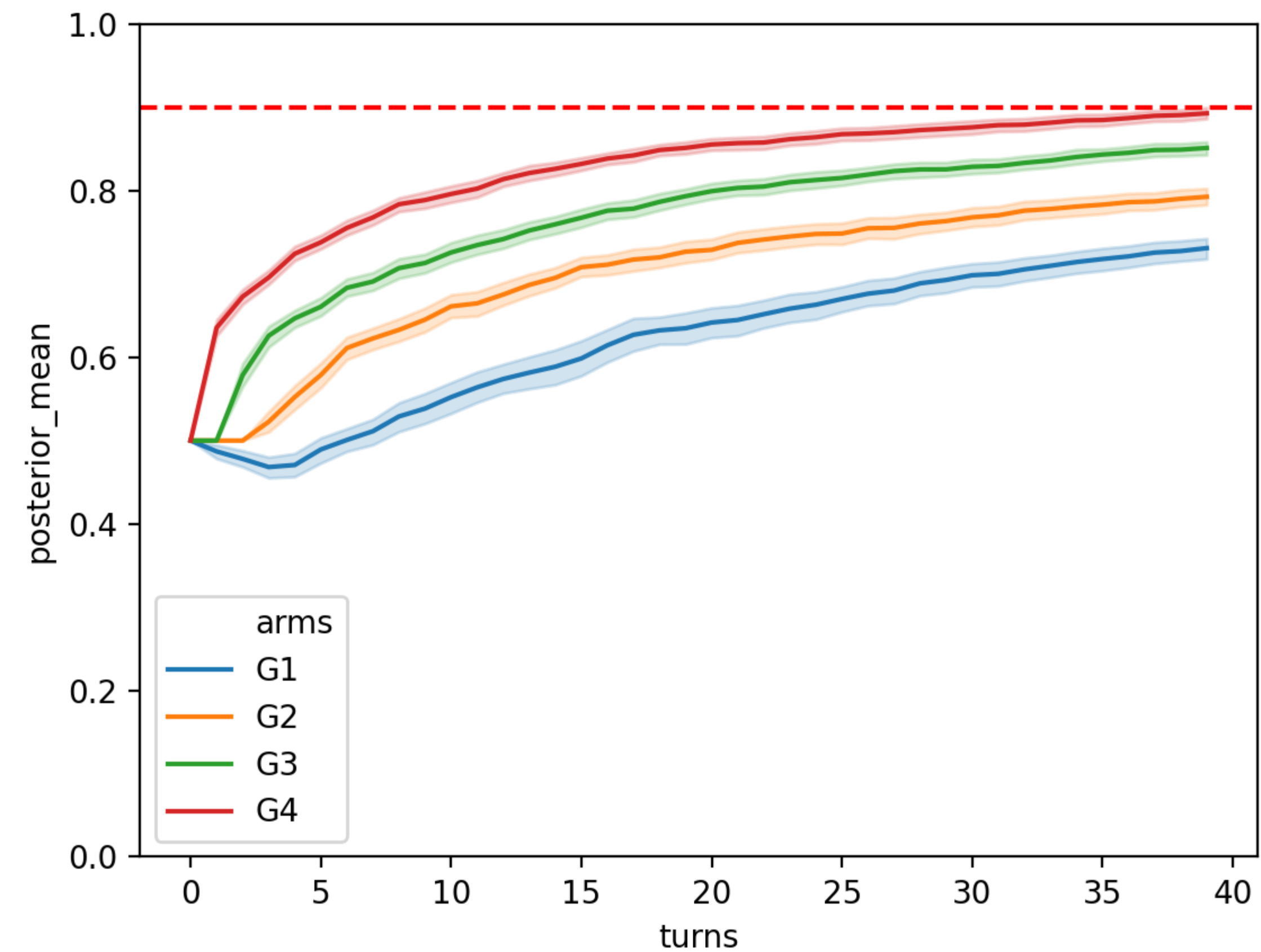


Identical reward ($\mu = 0.9$), No prior, Random sampling

Partner choices



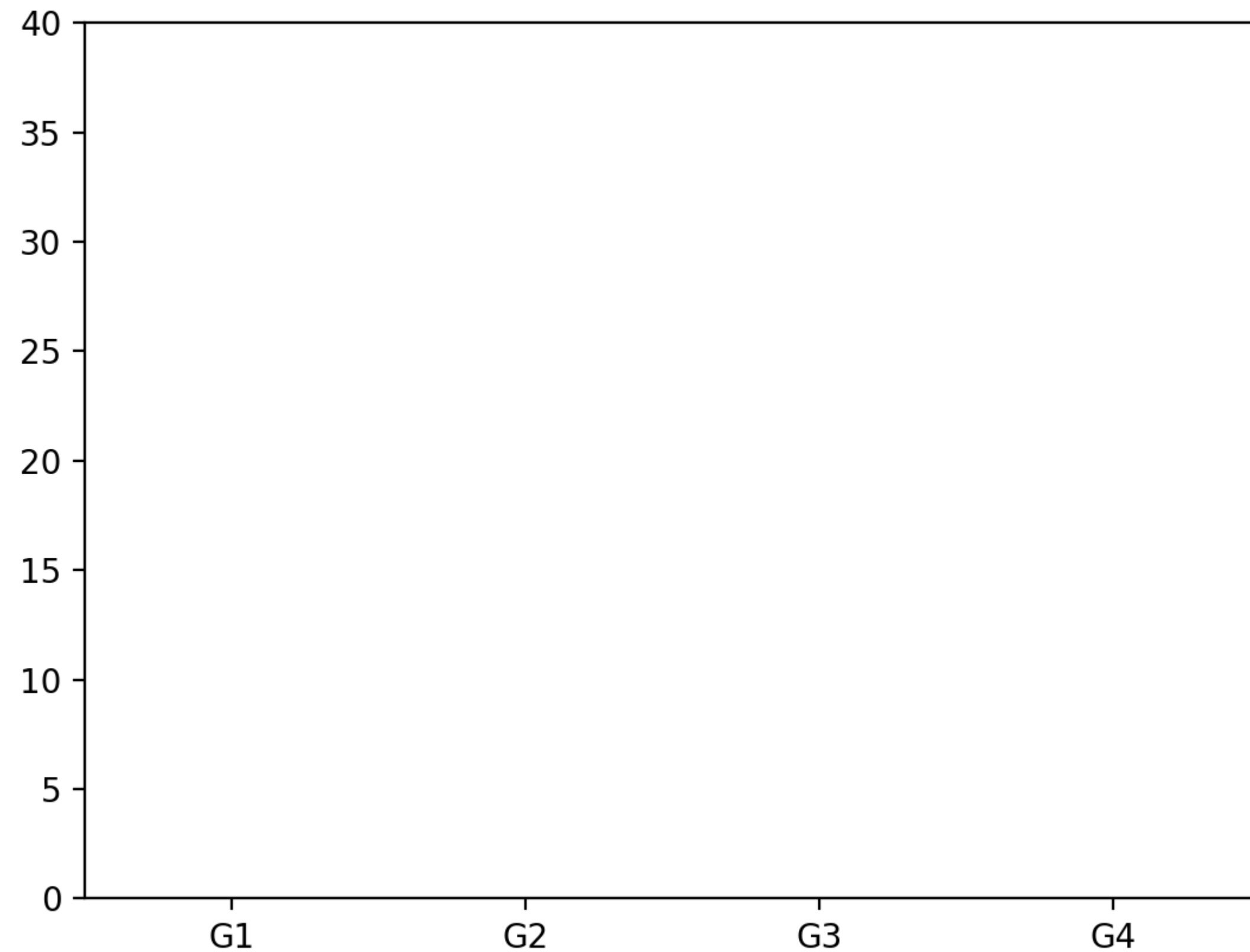
Reward estimations



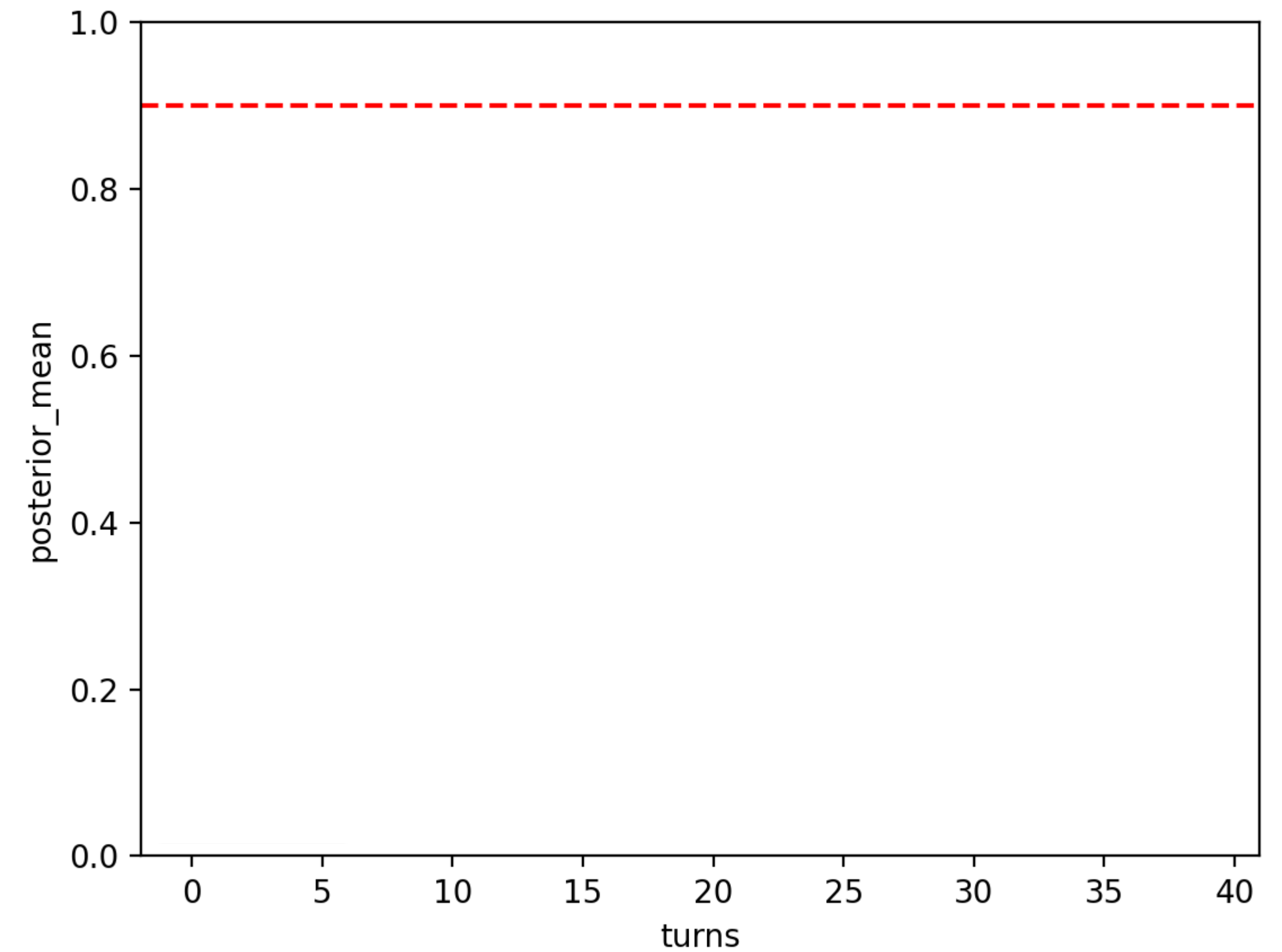
Our hypothesis

Identical reward ($\mu = 0.9$), No prior, Thompson sampling

Partner choices



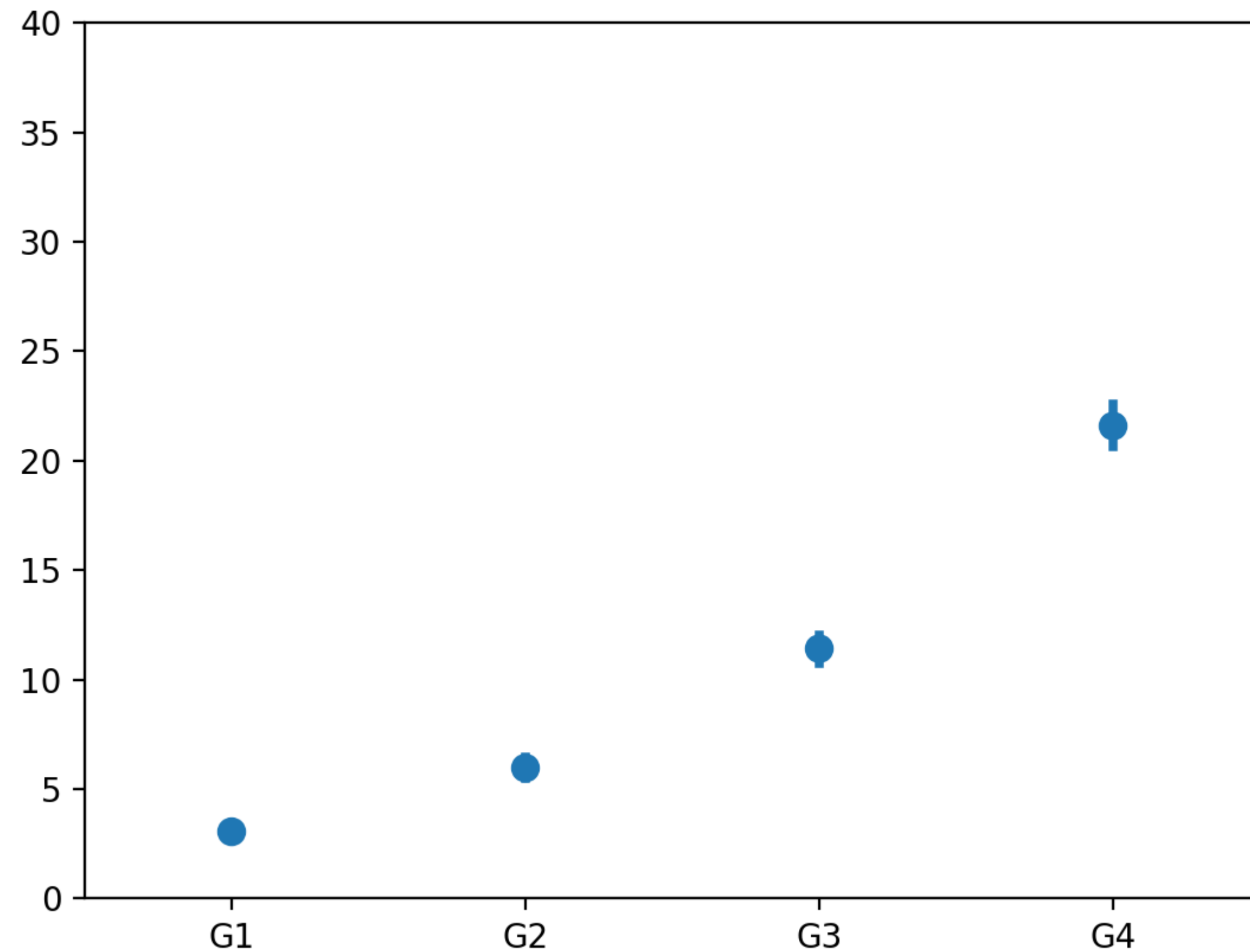
Reward estimations



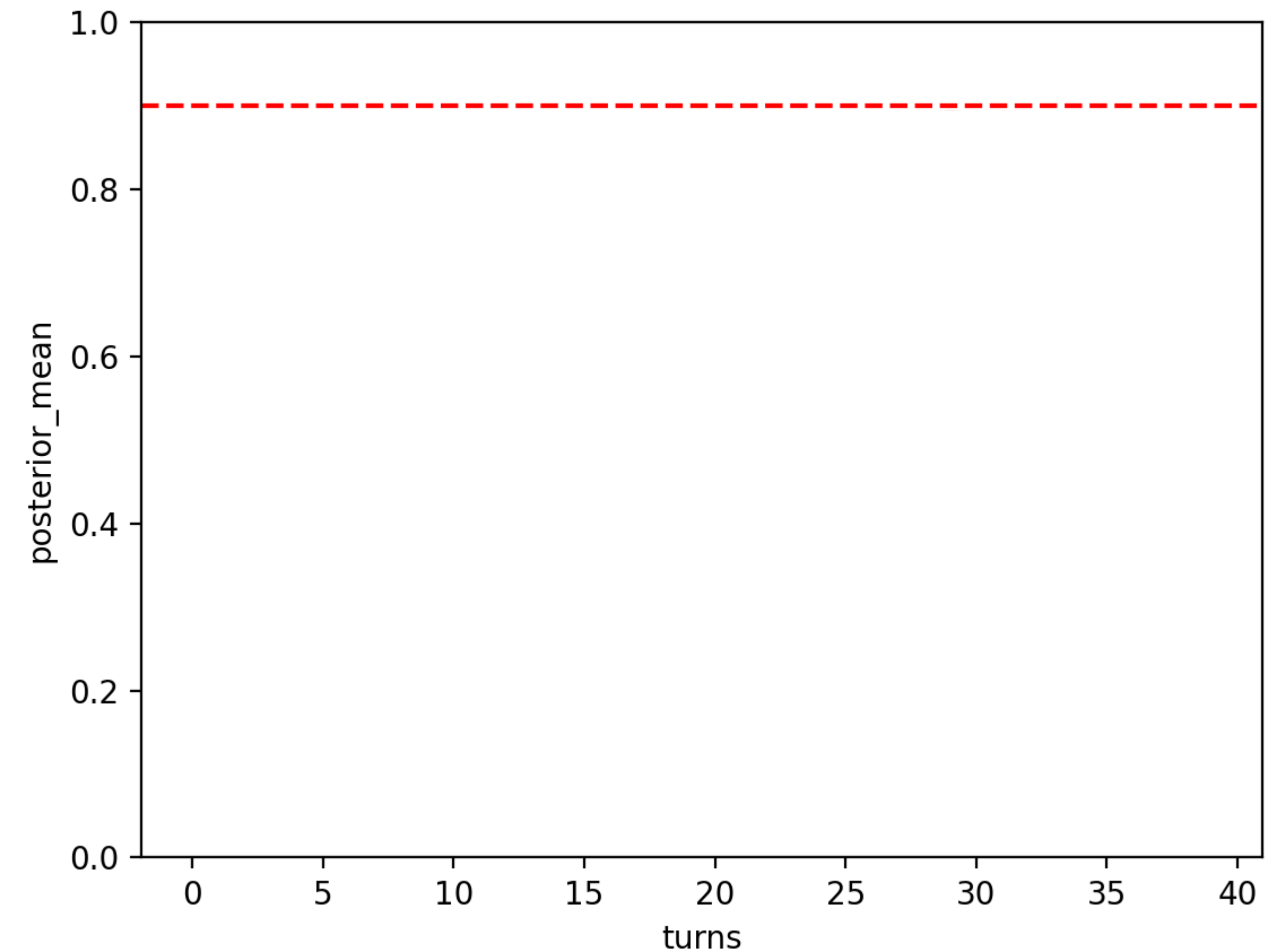
Our hypothesis

Identical reward ($\mu = 0.9$), No prior, Thompson sampling

Partner choices



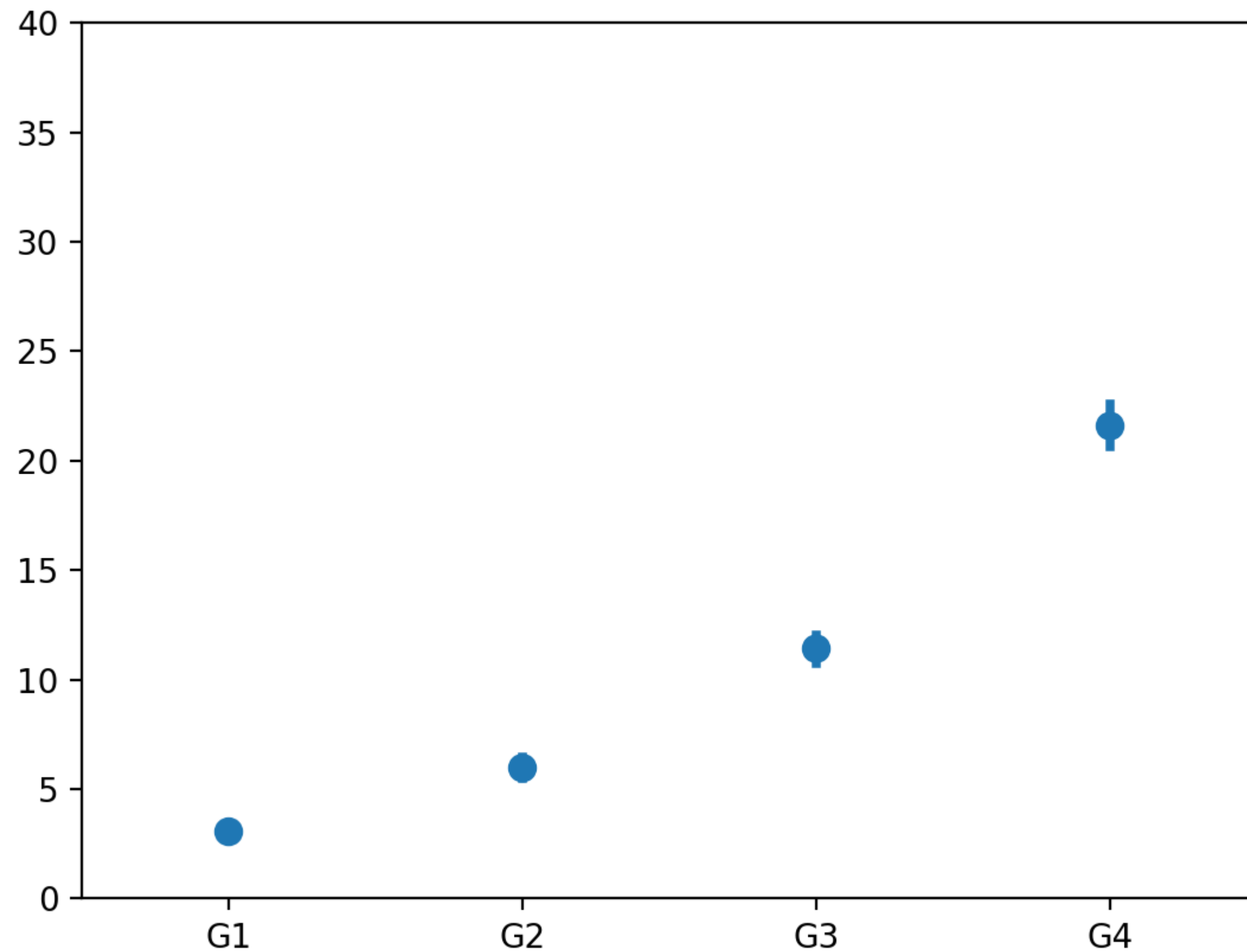
Reward estimations



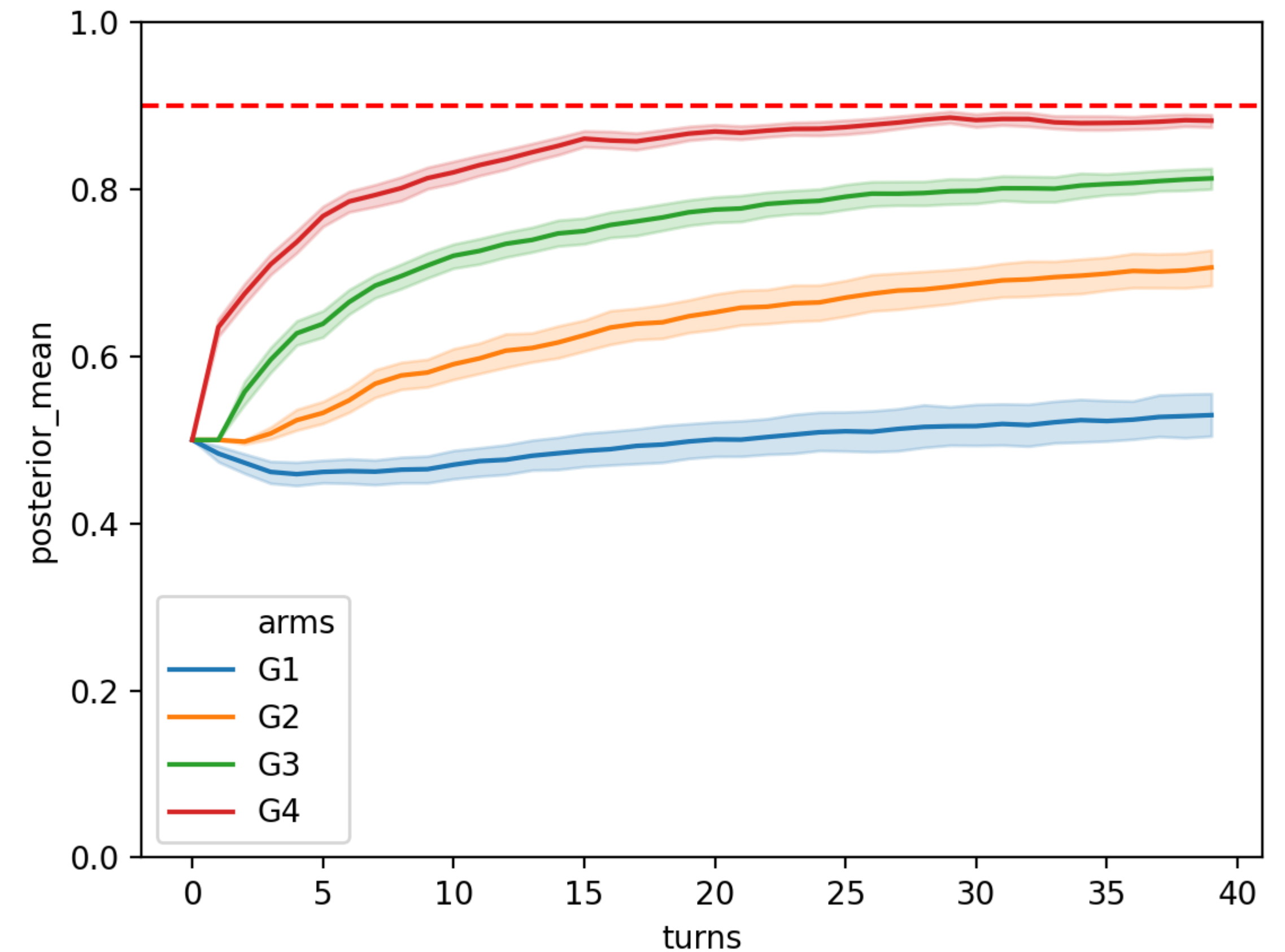
Our hypothesis

Identical reward ($\mu = 0.9$), No prior, Thompson sampling

Partner choices



Reward estimations



Interim summary

Selective interactions and biased impressions emerged when

Interim summary

Selective interactions and biased impressions emerged when

- The environment gives identical and high rewards;
- Prior bias is unnecessary;
- The agent uses adaptive sampling strategies

Interim summary

Selective interactions and biased impressions emerged when

- The environment gives identical and high rewards;
- Prior bias is unnecessary;
- The agent uses adaptive sampling strategies

How about human participants?

Our hypothesis

Globally inaccurate stereotypes can result from locally rational exploration

- An intuitive example
- A formal model and simulations
- **Two human experiments**
- Implications

Explore Toma City

A city with 100,000 residents



Ready to play?

Let's meet some new Tufas, Aimas, Rekus, and Wekis!

Tufa



Aima



Reku



Weki



Explore Toma City

Independent variables:

- **Underlying rewards:**
 - Unbeknown to participants, identical ($\mu = 0.9$) v. different average rewards
- **Prior stereotypes:**
 - No description v. Rekus are warm and competent
- **Sampling strategy:**
 - You choose v. You meet [random]

Explore Toma City

Dependent variables:

- **Partner choices:**
 - You choose a [group] vs. You meet a [group]
- **Reward estimations:**
 - For each group, how many times out of 100 do you think working with a person from that group would result you in earning 1 point?

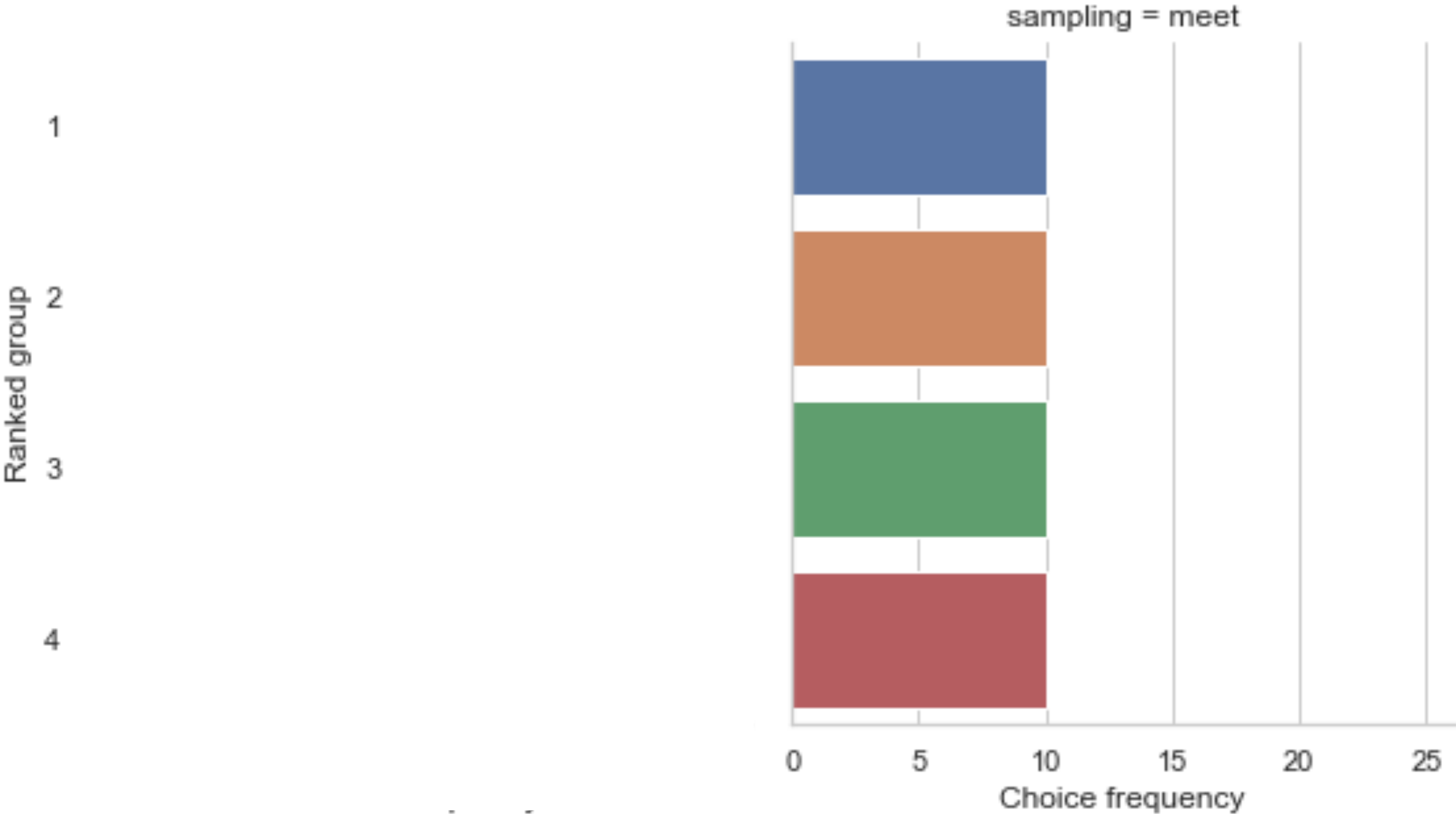
Explore Toma City

400 online participants in Study 1 (N = 2000 in Study 2)

Partner choices

Reward estimations

reward = identical & bias = no

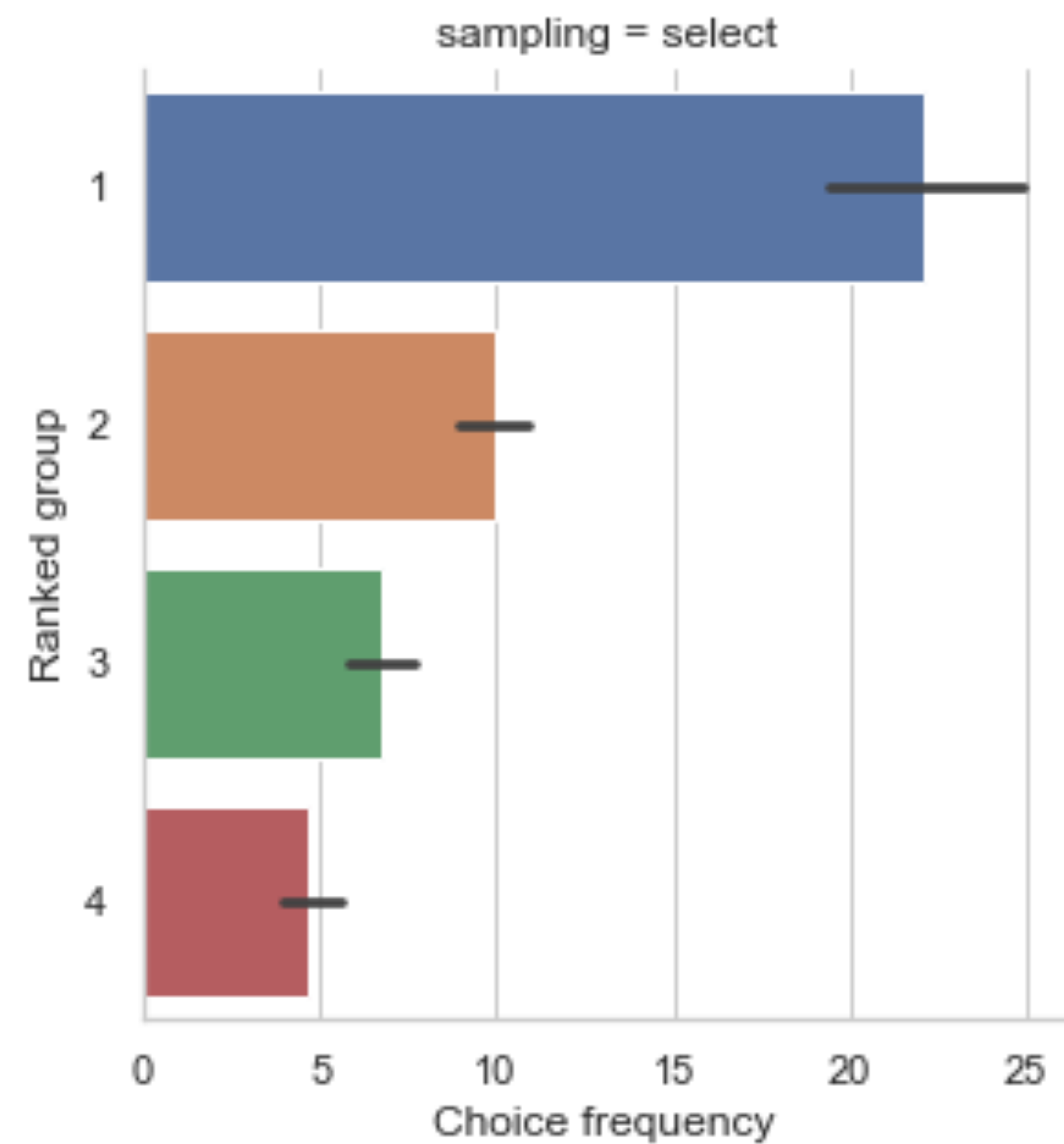


Explore Toma City

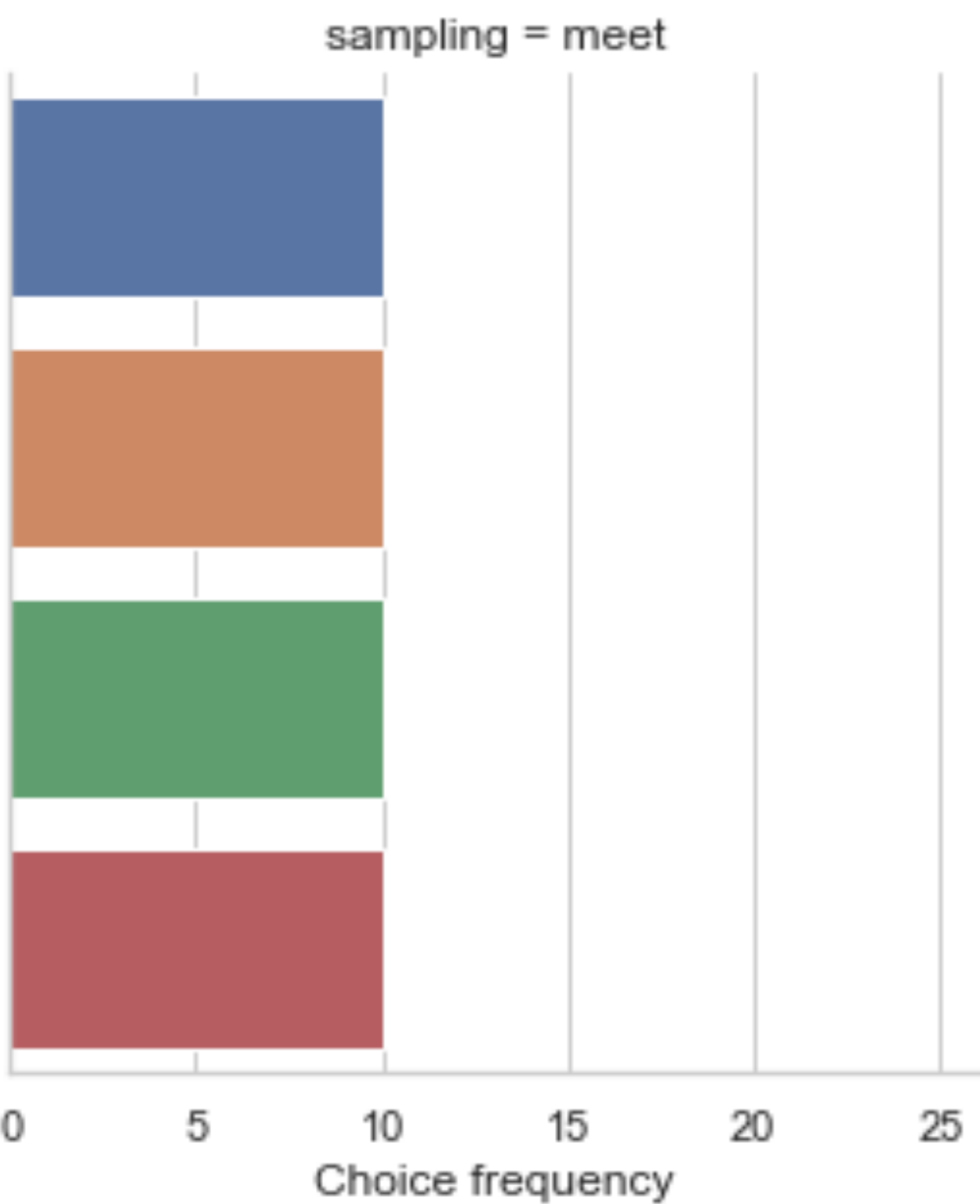
400 online participants in Study 1 (N = 2000 in Study 2)

Partner choices

reward = identical & bias = no



Reward estimations

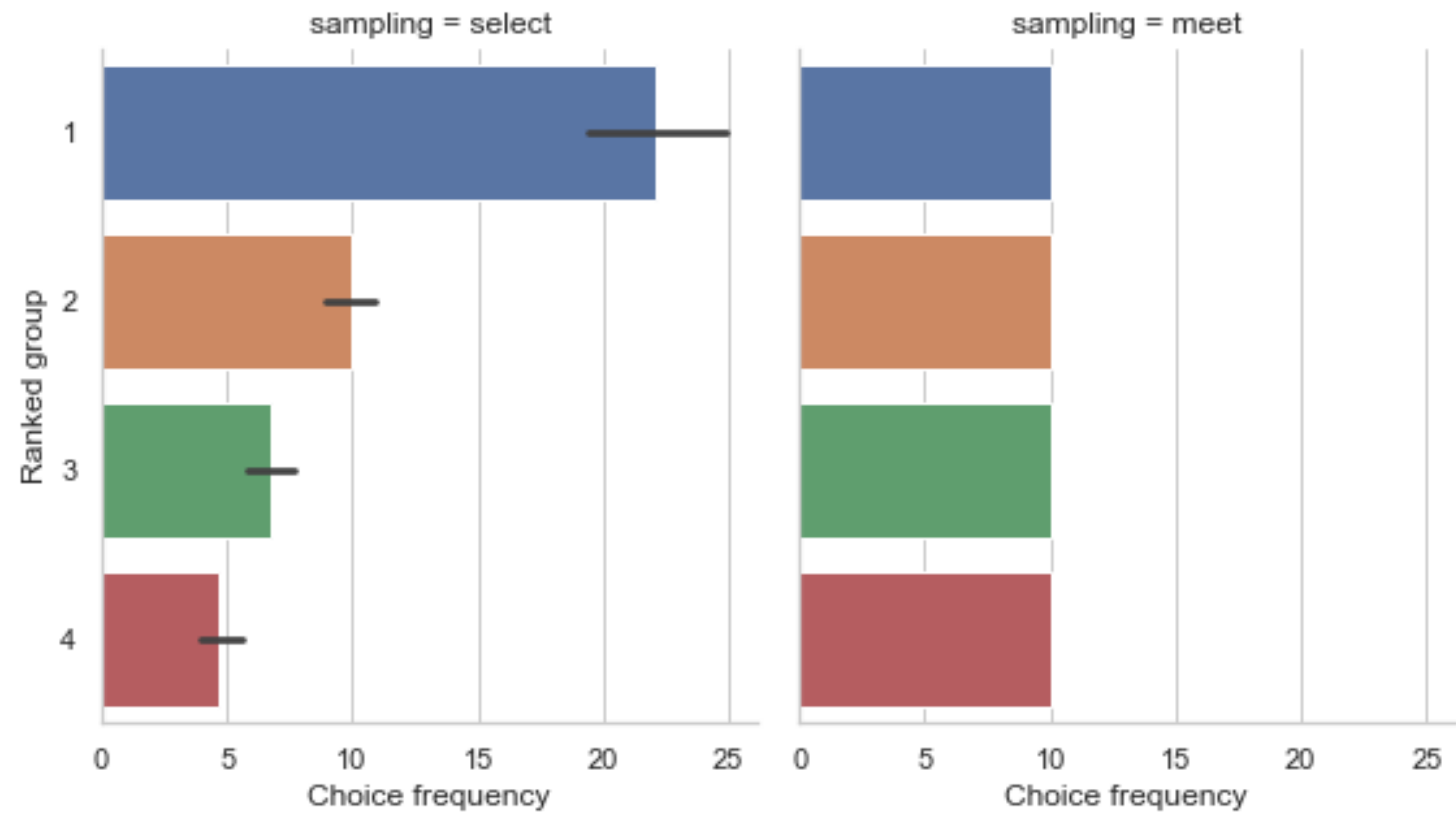


Explore Toma City

400 online participants in Study 1 (N = 2000 in Study 2)

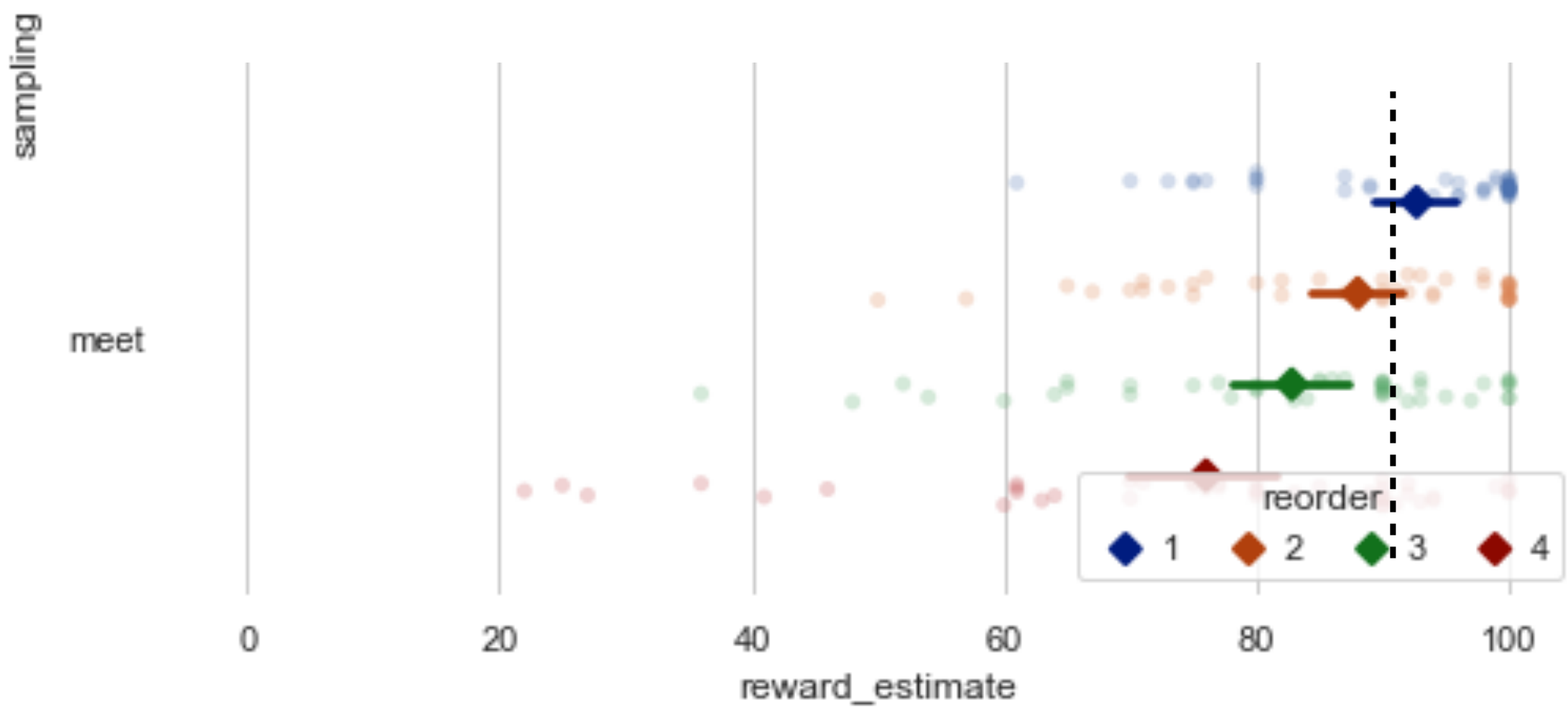
Partner choices

reward = identical & bias = no



Reward estimations

reward = identical & bias = no

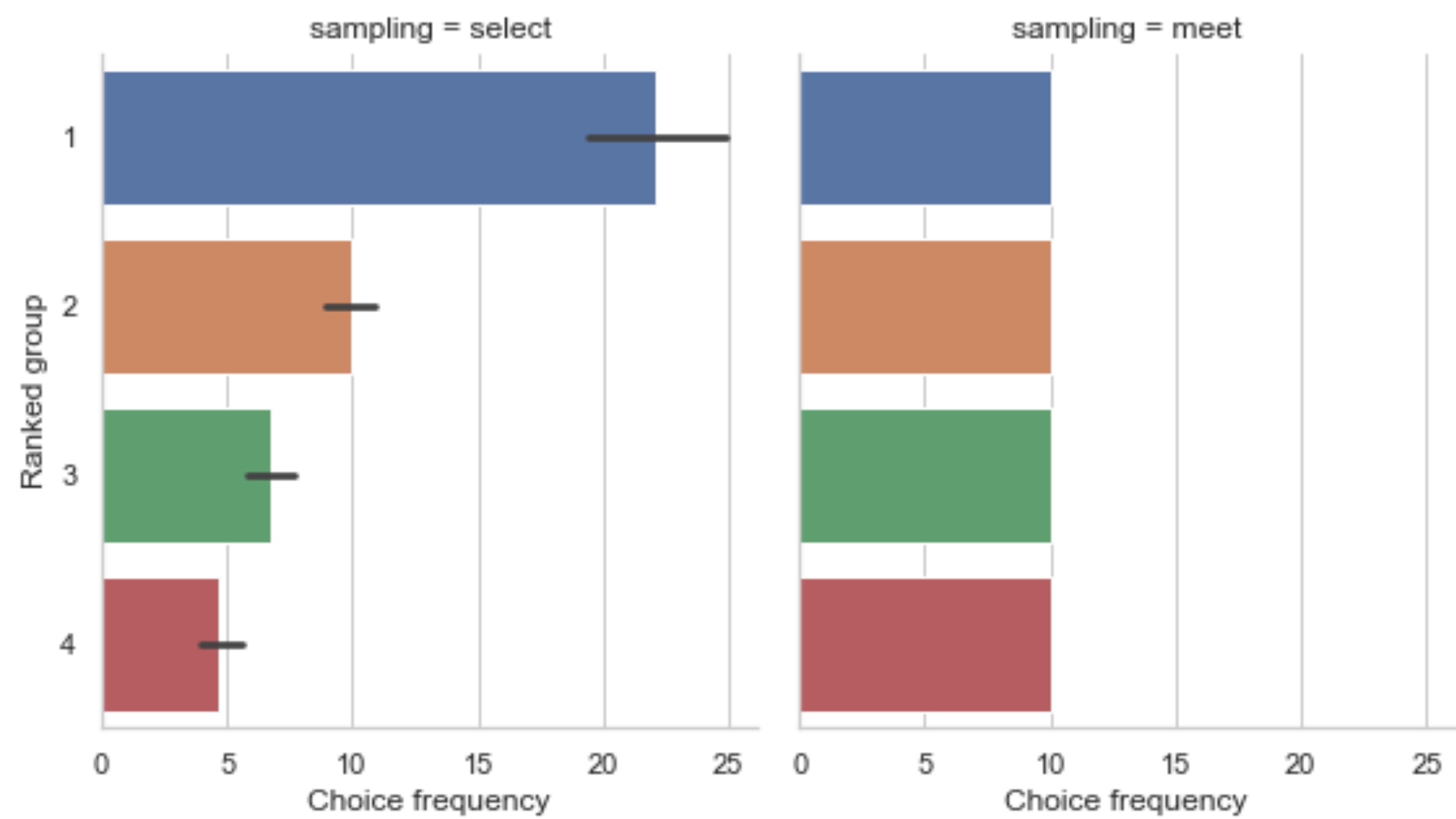


Explore Toma City

400 online participants in Study 1 (N = 2000 in Study 2)

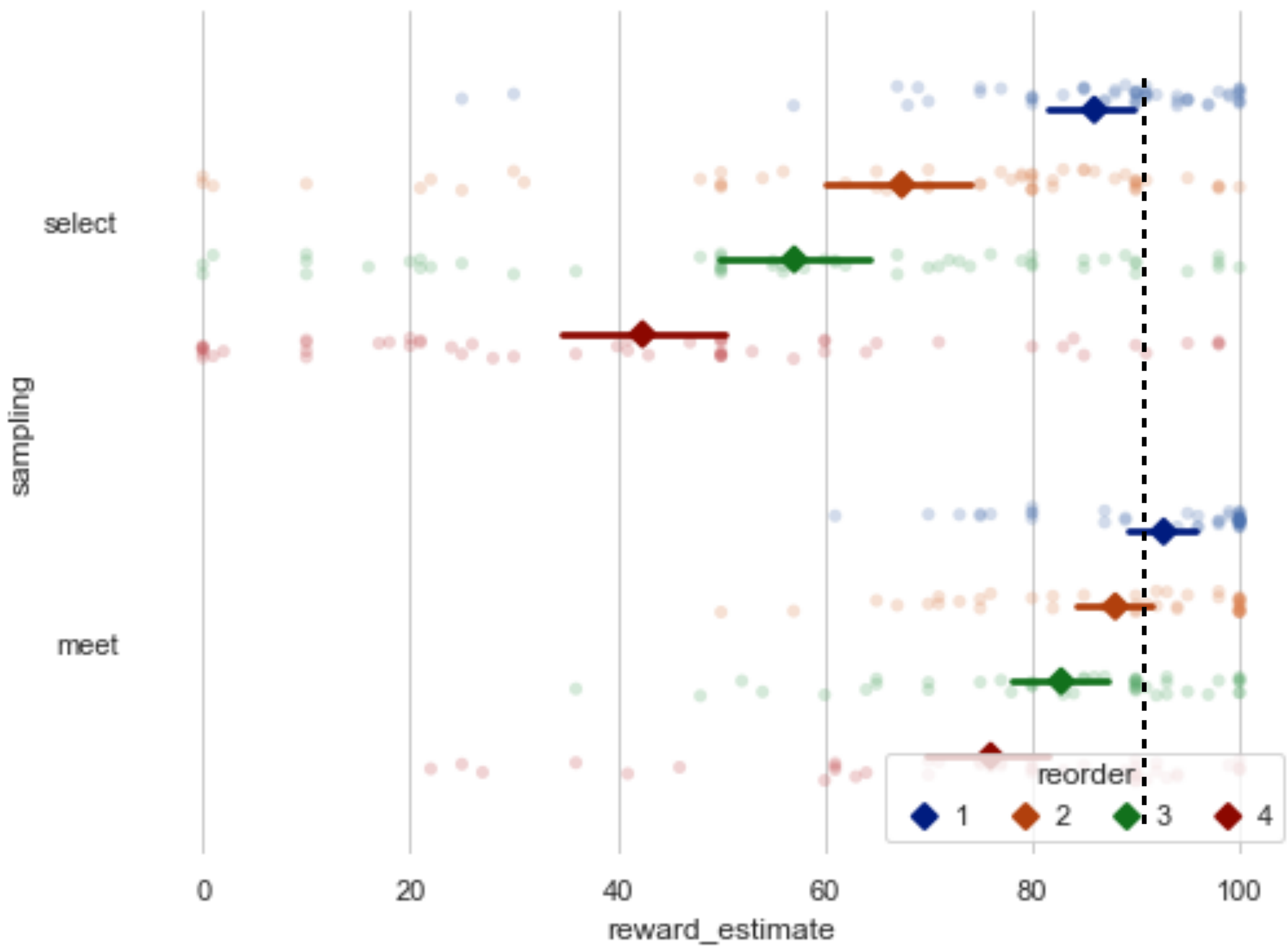
Partner choices

reward = identical & bias = no



Reward estimations

reward = identical & bias = no



Interim summary

How do humans behave in an environment with identical and high rewards & no prior bias

- **Replicated model predictions.**

Interim summary

How do humans behave in an environment with identical and high rewards & no prior bias

- **Replicated model predictions.**

Selective interactions and biased impressions?

- **Yes.**

Interim summary

How do humans behave in an environment with identical and high rewards & no prior bias

- **Replicated model predictions.**

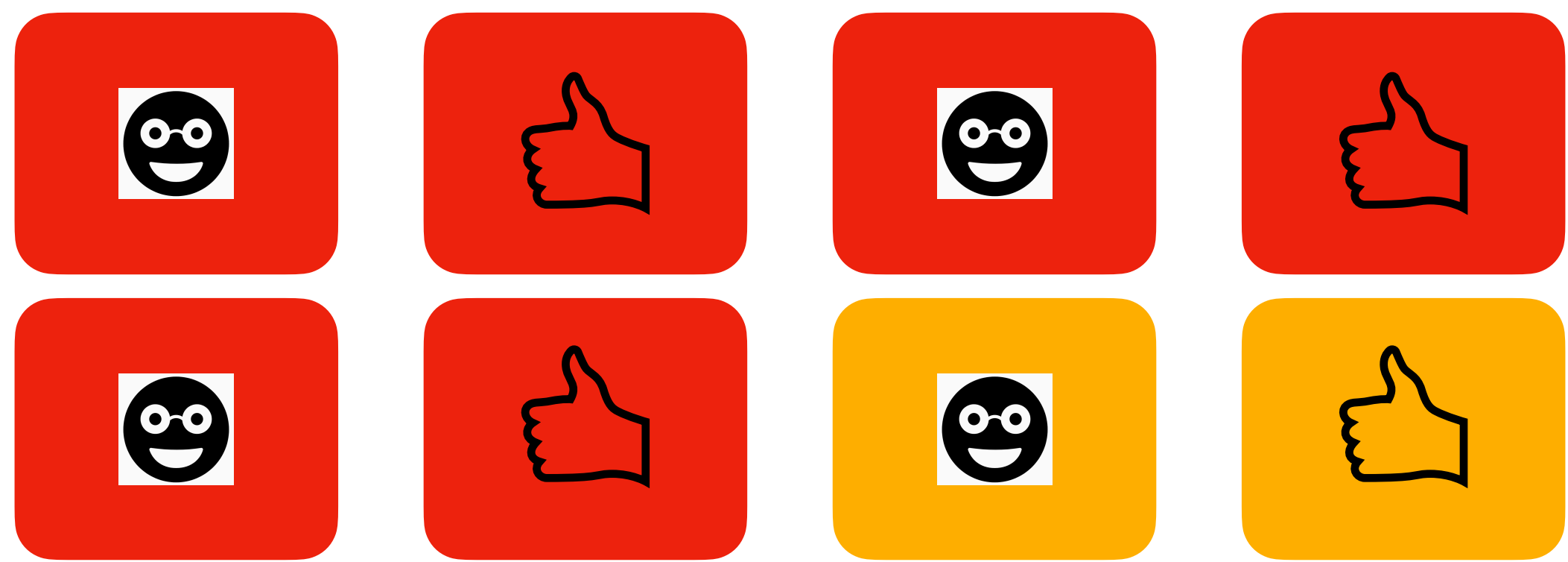
Selective interactions and biased impressions?

- **Yes.**

Thompson sampling v. Random sampling

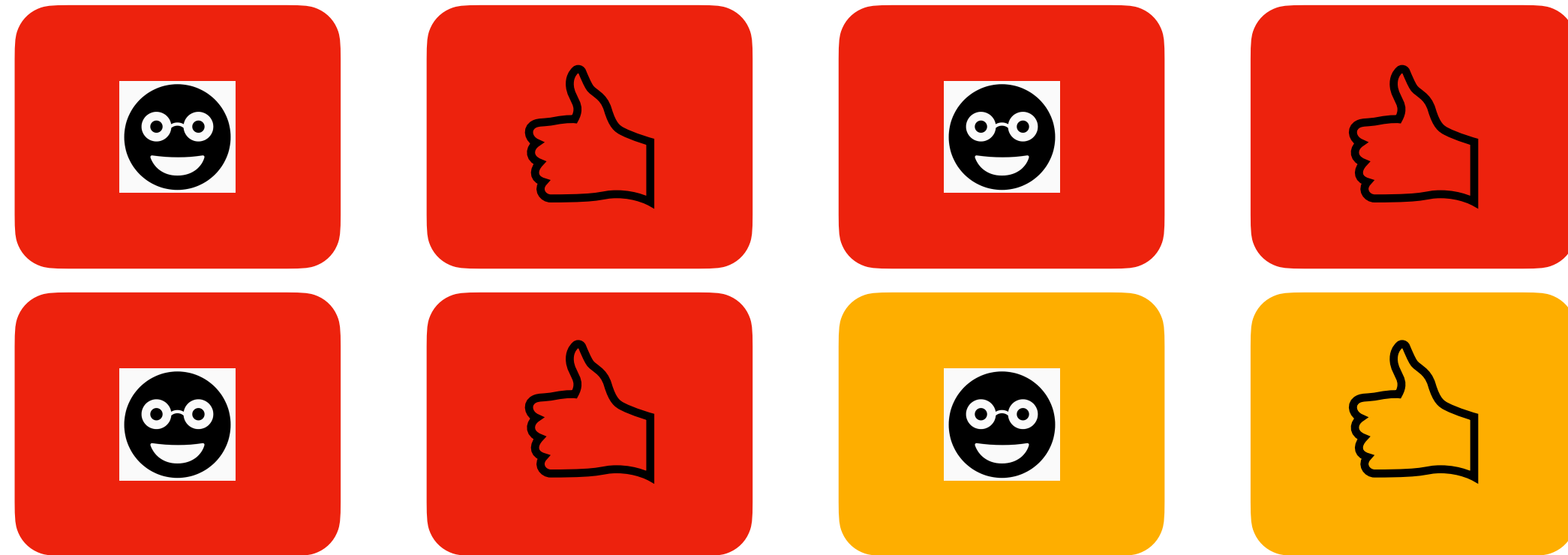
- **Consistent with rational strategy thus biased.**

Confounds and Mechanisms



You choose v. You meet [random draw]

Confounds and Mechanisms

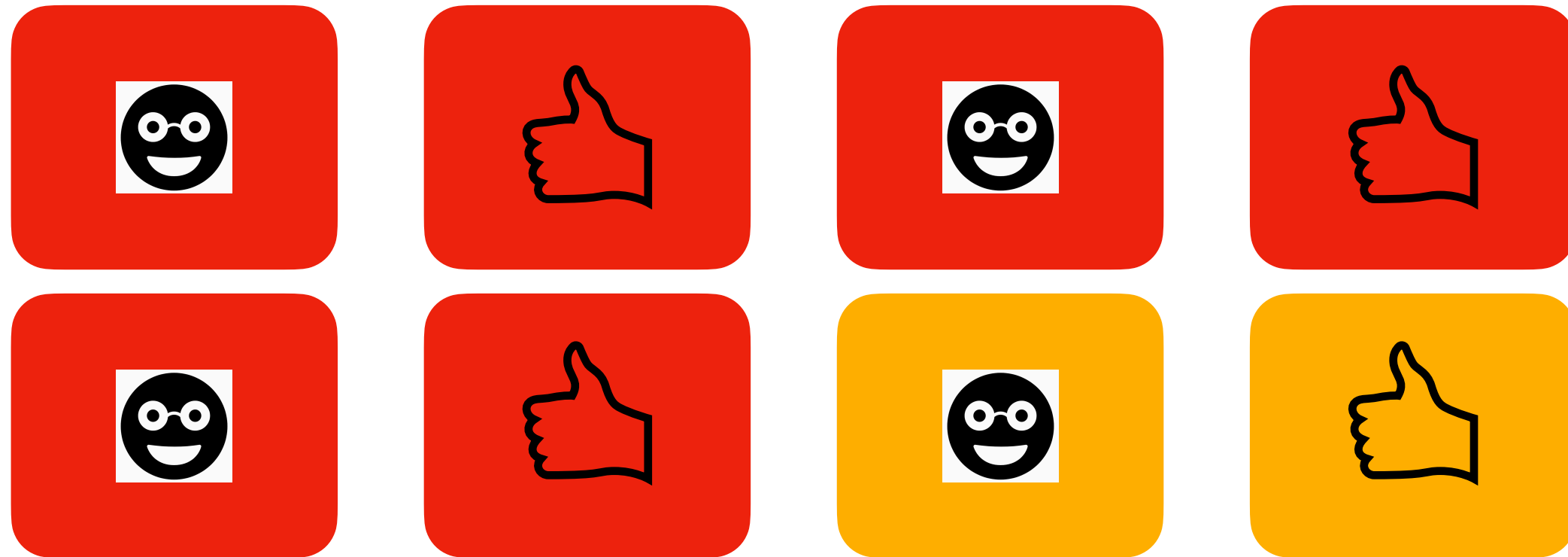


You choose v. You meet [random draw]

1. Selective sample or sense of control?

2. Order of rewards or mere presence of choices?

Confounds and Mechanisms



You choose v. You meet [random draw]

1. Selective sample or sense of control?
 - Yoke both choices and rewards
2. Order of rewards or mere presence of choices?
 - Yoke choices but not rewards

Our hypothesis

Globally inaccurate stereotypes can result from locally rational exploration

- An intuitive example
- A formal model and simulations
- Two human experiments
- **Implications**

Conclusion

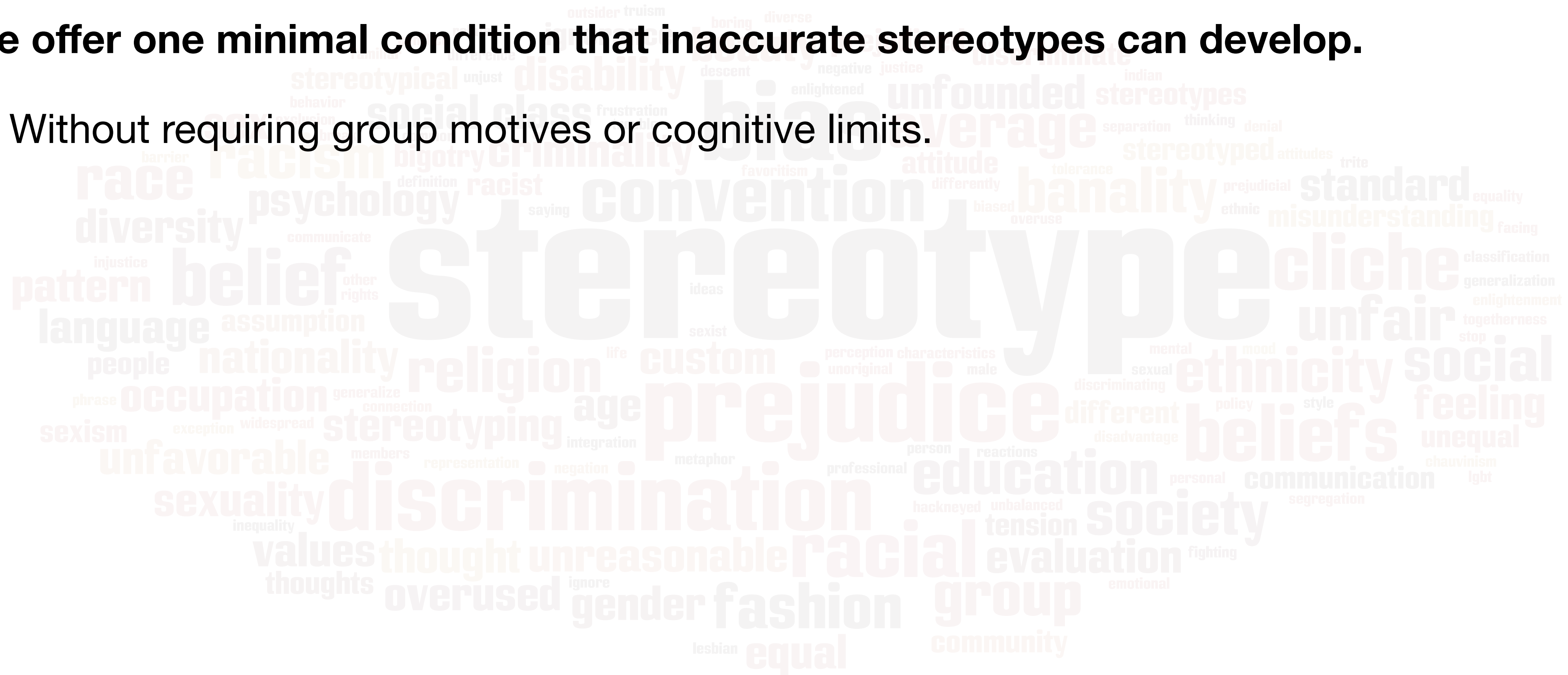
WHY do people develop inaccurate stereotypes in the first place?



Conclusion

WHY do people develop inaccurate stereotypes in the first place?

- **We offer one minimal condition that inaccurate stereotypes can develop.**
 - Without requiring group motives or cognitive limits.



Conclusion

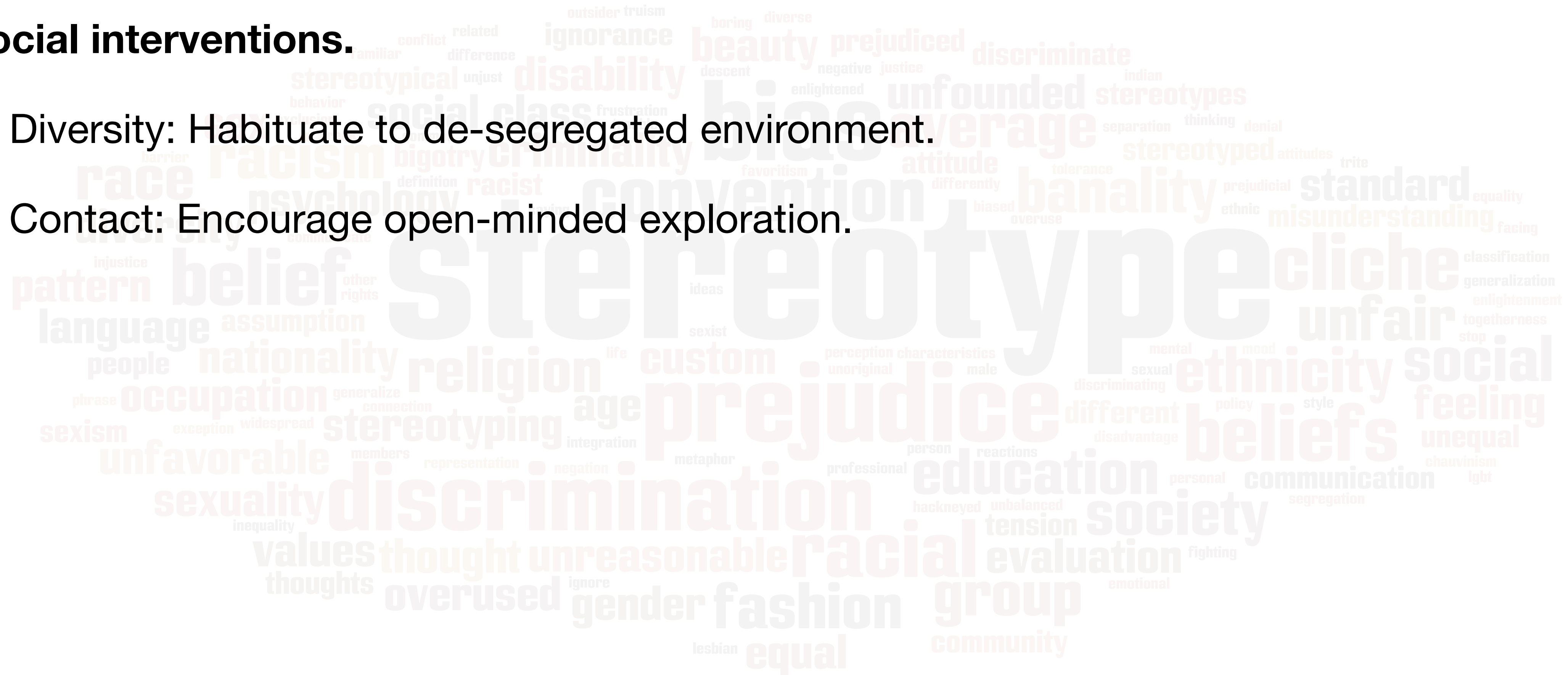
WHY do people develop inaccurate stereotypes in the first place?

- We offer one minimal condition that inaccurate stereotypes can develop.
 - Without requiring group motives or cognitive limits.
- **A functional minimal-process explanation.**
 - The mind adapts to a particular social environment.
 - Rational local choices still lead to inaccurate overall impressions.

Implication

Inaccurate stereotypes can result from rational exploration

- **Social interventions.**
 - Diversity: Habituate to de-segregated environment.
 - Contact: Encourage open-minded exploration.



Implication

Inaccurate stereotypes can result from rational exploration

- **Social interventions.**

- Diversity: Habituate to de-segregated environment.
- Contact: Encourage open-minded exploration.

- **Algorithmic interventions.**

- ML fairness: The origin of unequal base rates.
- Filter bubble: Recommendation systems.

Thanks to my advisors for collaboration and labs for feedback

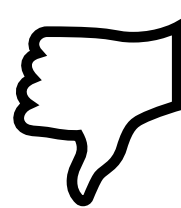
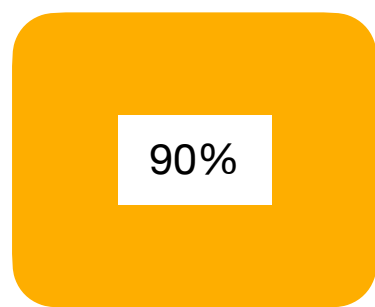
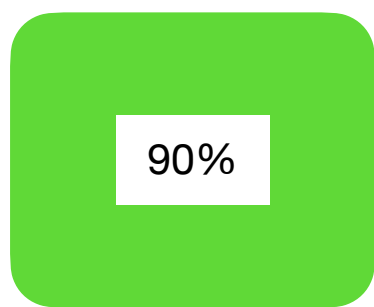
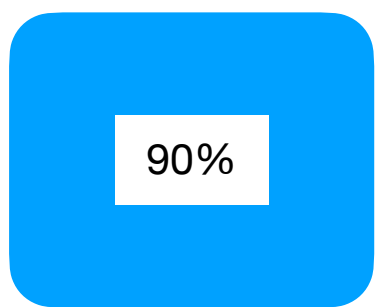
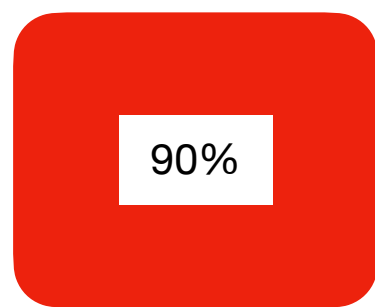
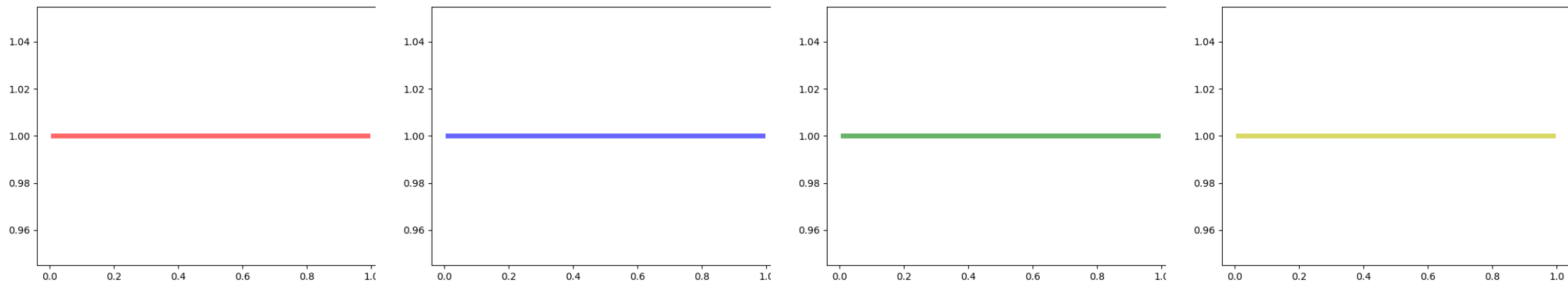


Thank you!

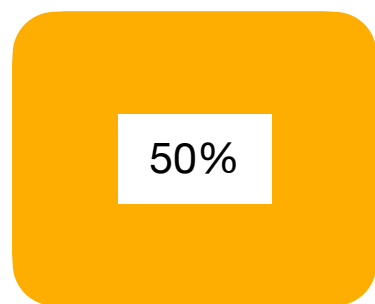
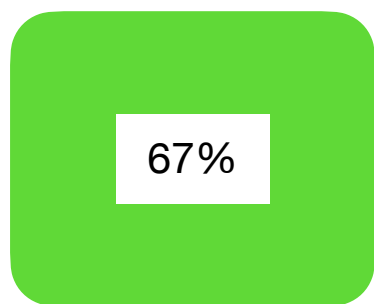
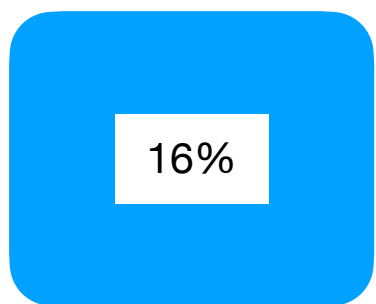
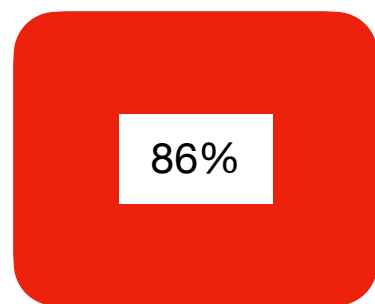
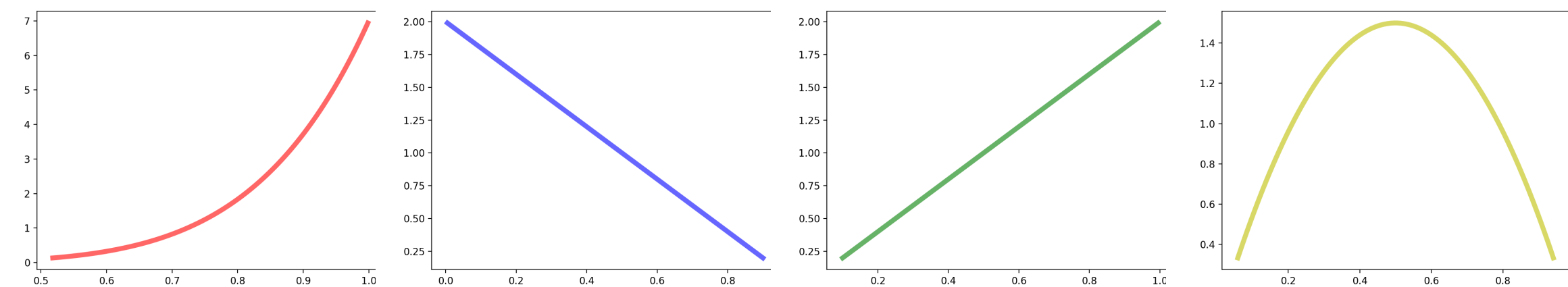
Our hypothesis

Inaccurate stereotypes can result from rational explorations

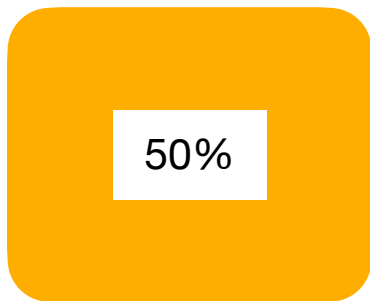
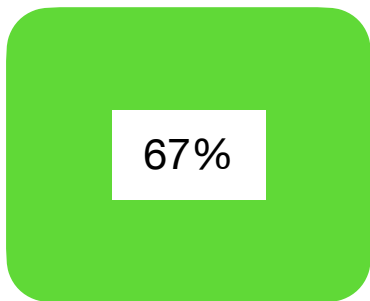
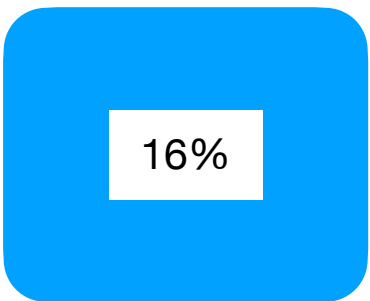
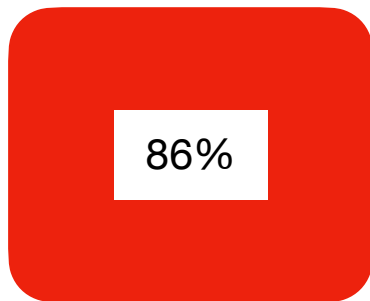
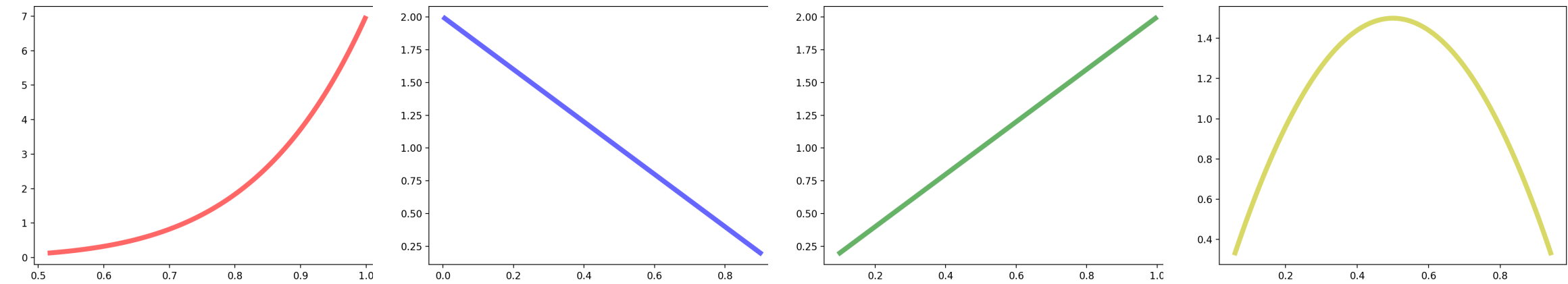
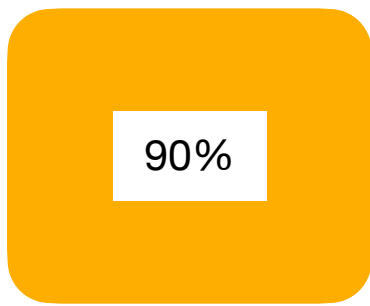
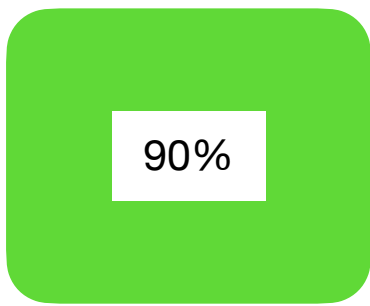
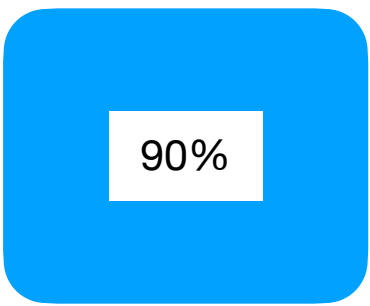
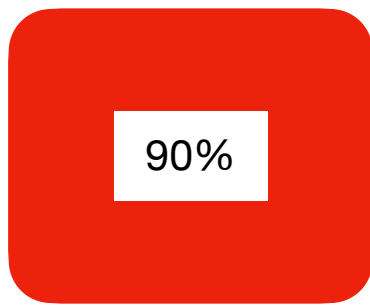
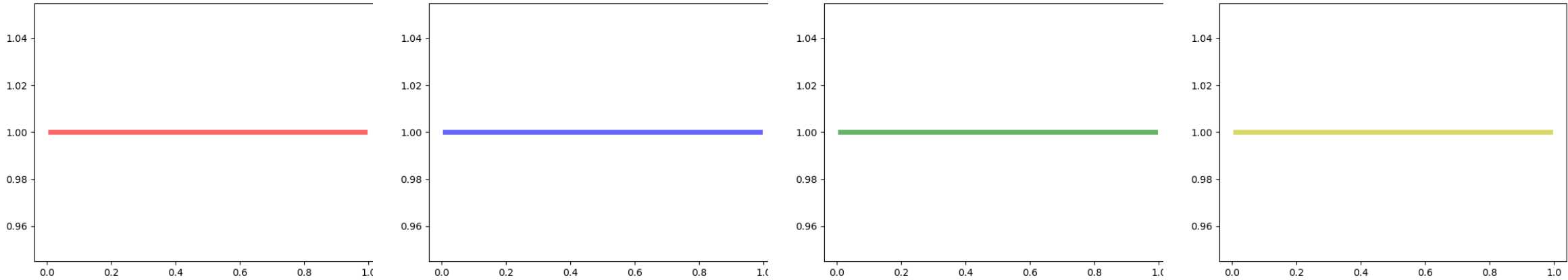
- An intuitive example
- **A formal model and simulations**
- Two human experiments
- Implications



Multi-Armed Bandit Problem



Explore v. Exploit dilemma in Reinforcement Learning (Sutton & Barto, 2018)

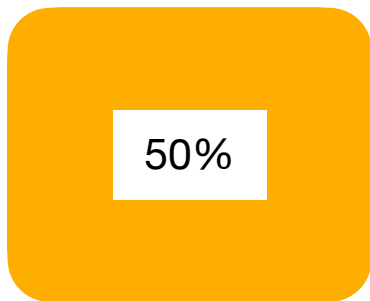
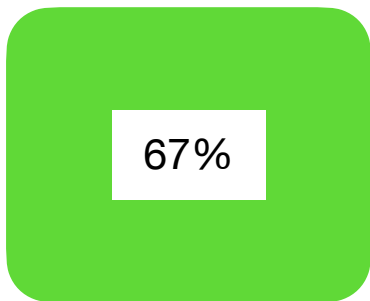
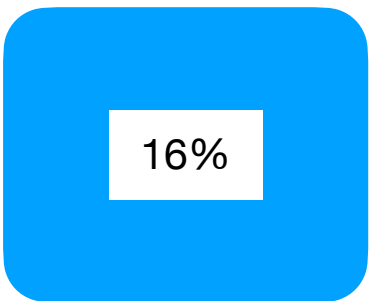
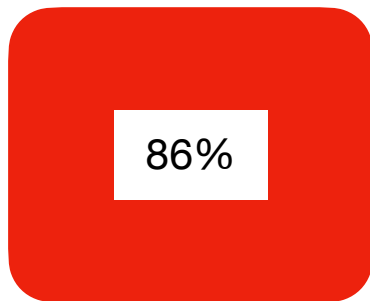
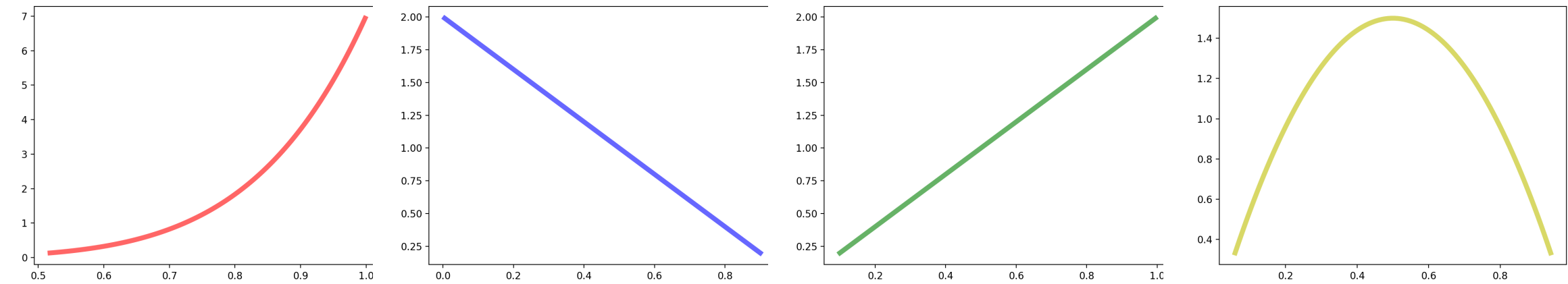
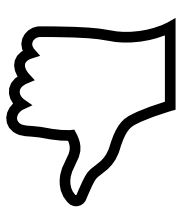
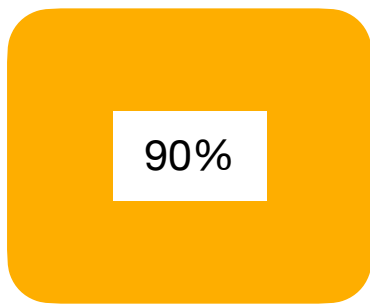
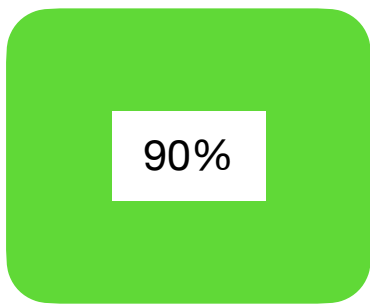
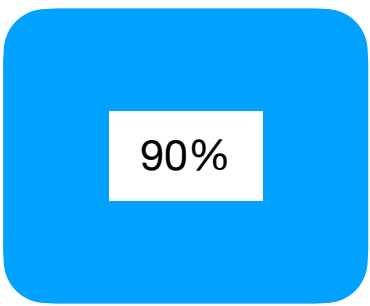
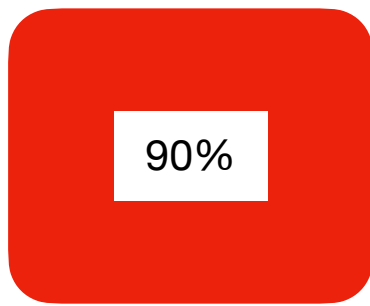
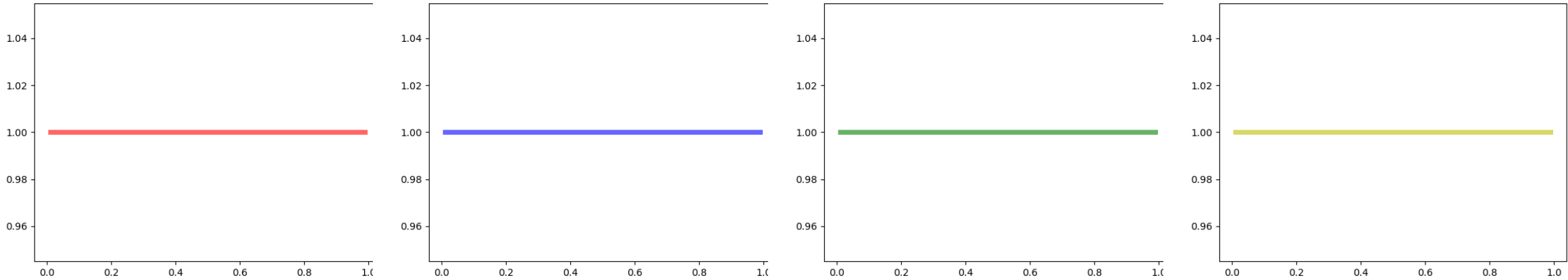


$$\theta_k$$

$$r_{t(k)} \sim \text{Bern}(\theta_k)$$

$$\sum_{t=1}^T r_{t(k)}$$

$$R = E\left[\sum_{t=1}^T Q_{t(k^*)} - \sum_{t=1}^T r_{t(k)}\right]$$

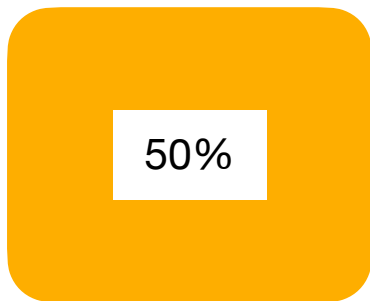
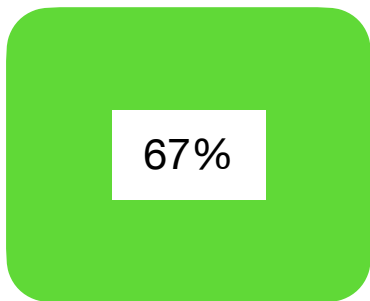
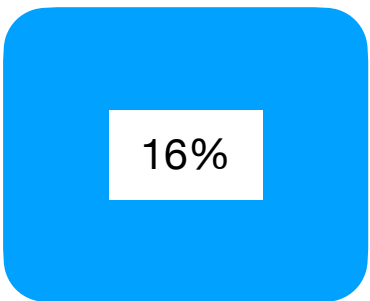
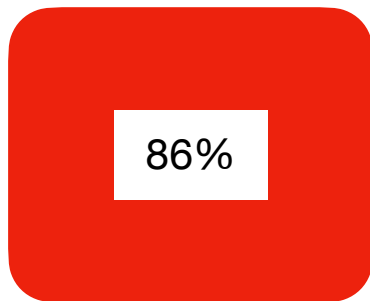
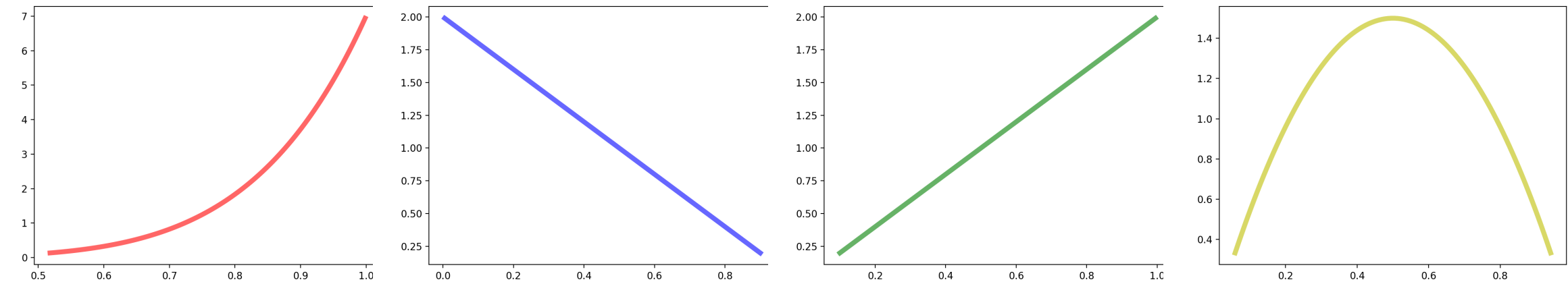
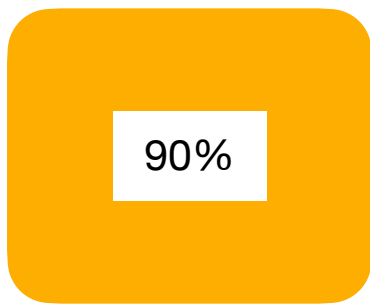
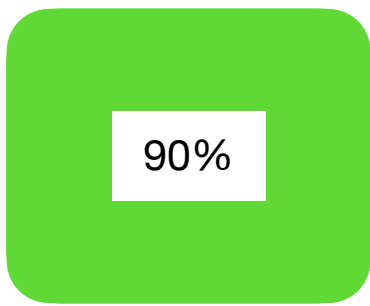
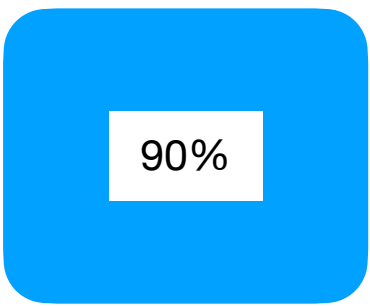
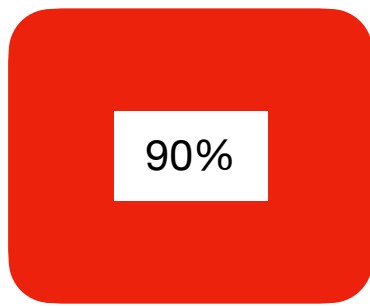
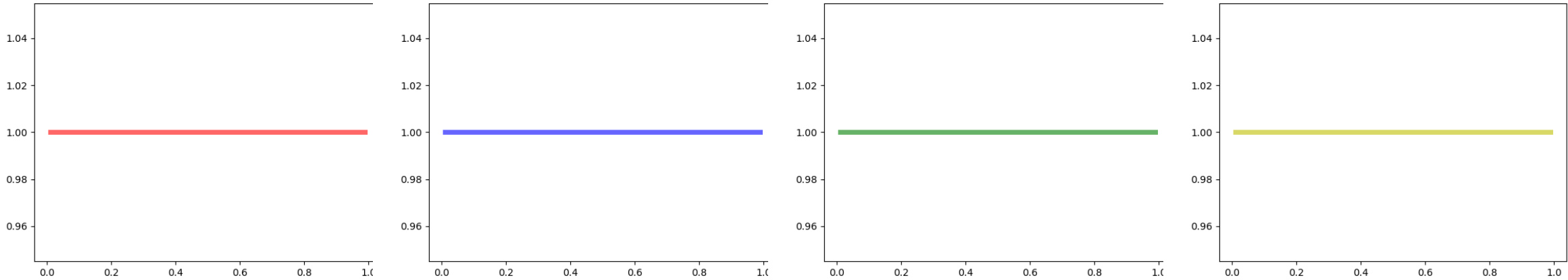


$$\theta_k$$

$$r_{t(k)} \sim \text{Bern}(\theta_k)$$

$$\sum_{t=1}^T r_{t(k)}$$

$$R = E\left[\sum_{t=1}^T Q_{t(k^*)} - \sum_{t=1}^T r_{t(k)}\right]$$

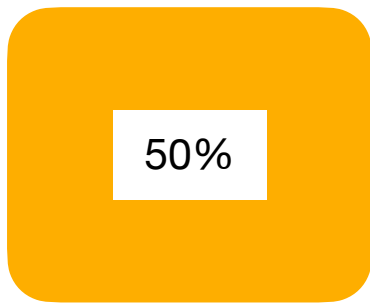
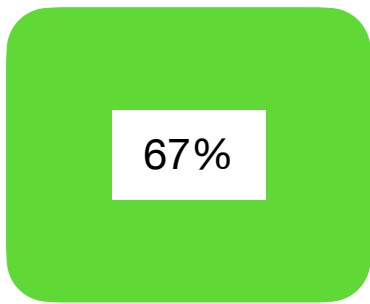
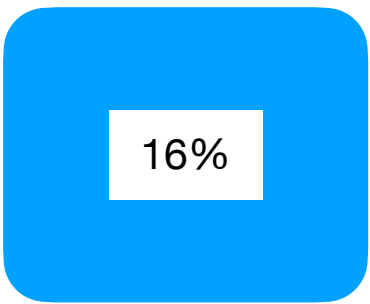
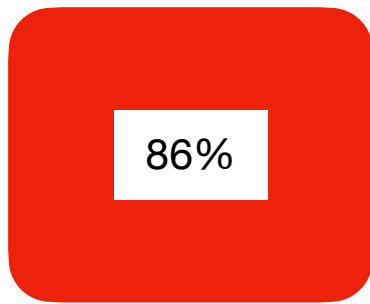
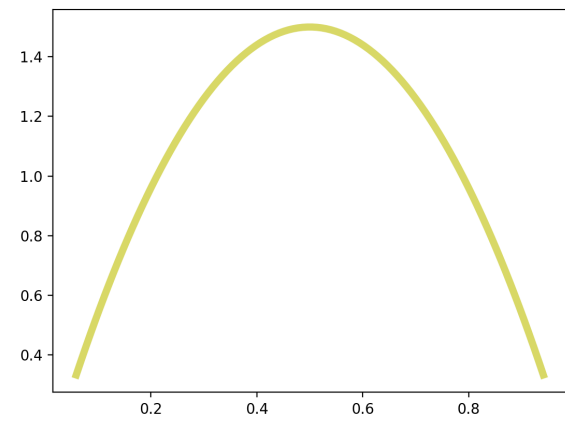
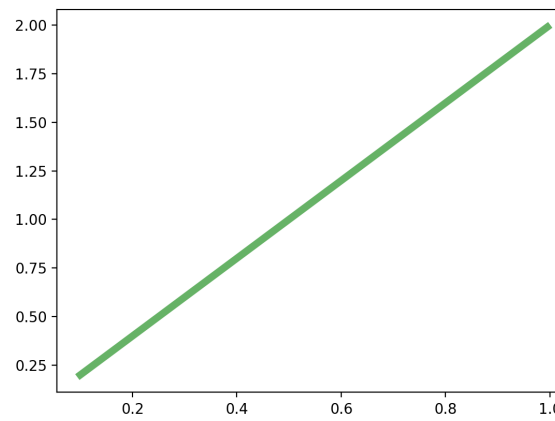
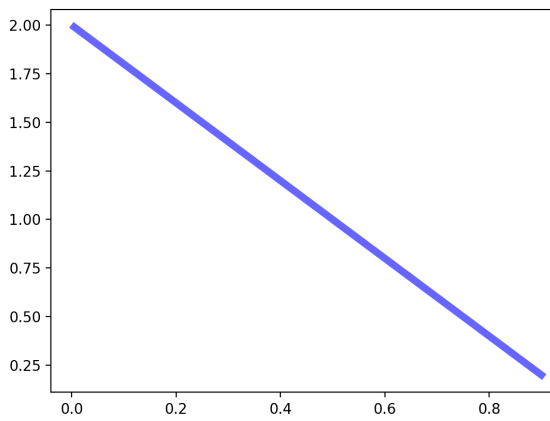
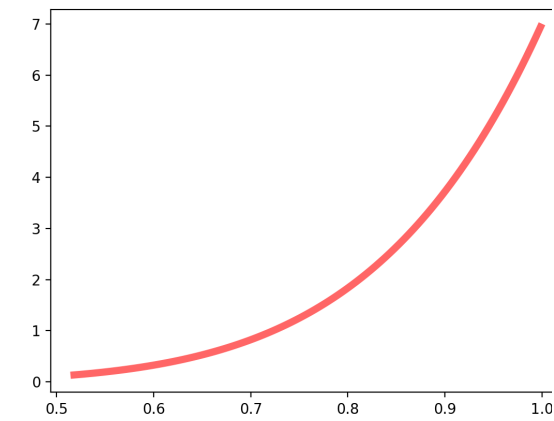
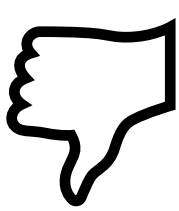
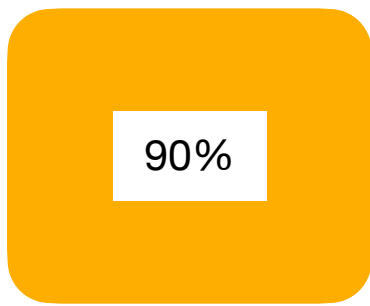
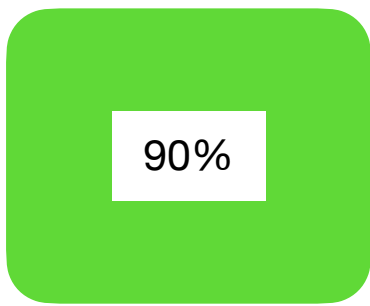
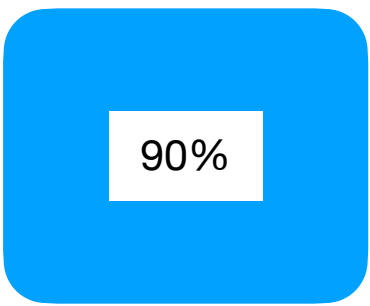
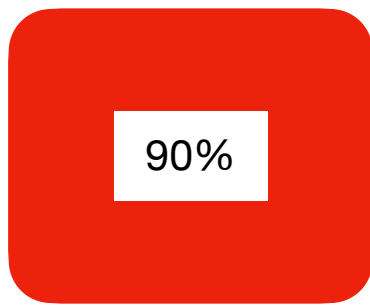
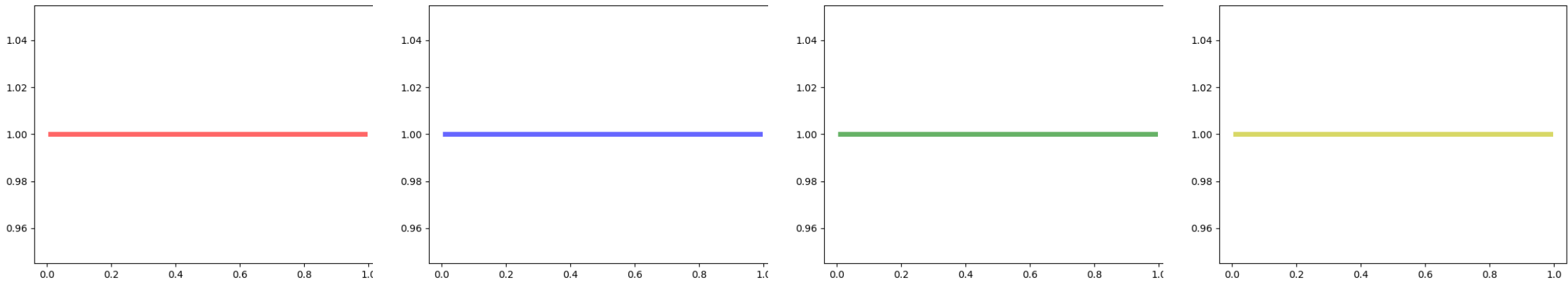


$$\theta_k$$

$$r_{t(k)} \sim \text{Bern}(\theta_k)$$

$$\sum_{t=1}^T r_{t(k)}$$

$$R = E\left[\sum_{t=1}^T Q_{t(k^*)} - \sum_{t=1}^T r_{t(k)}\right]$$

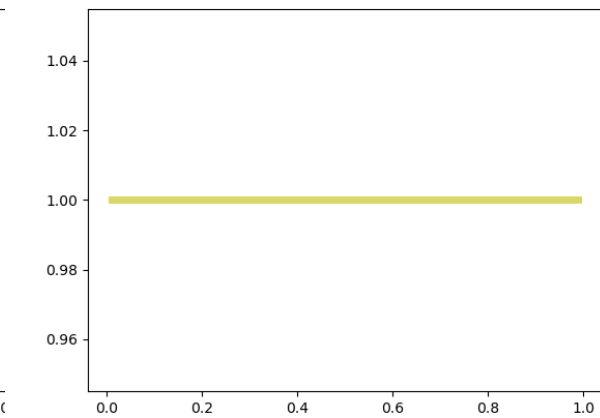
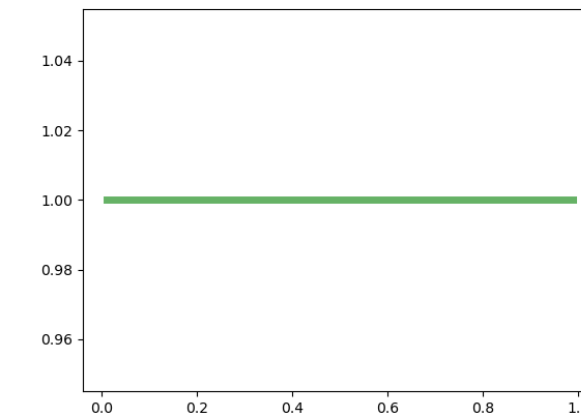
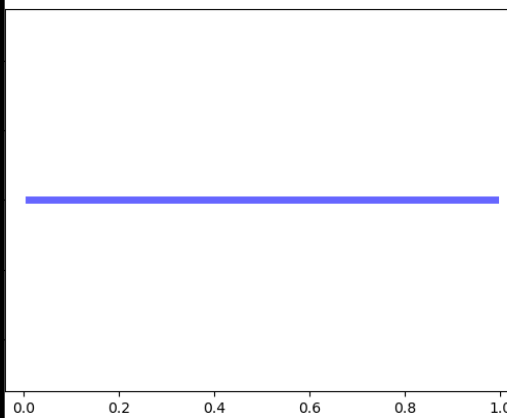
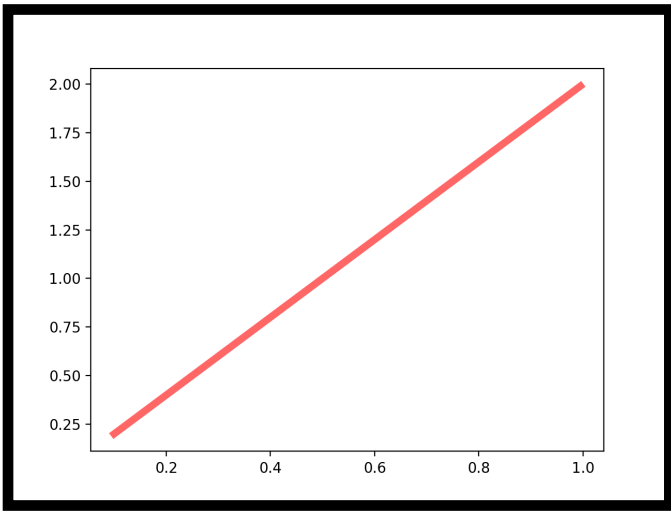


$$\theta_k$$

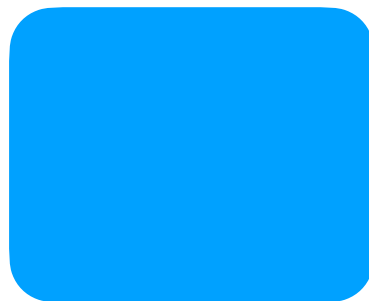
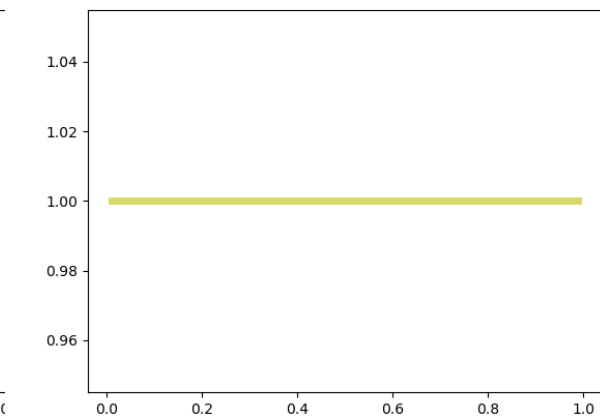
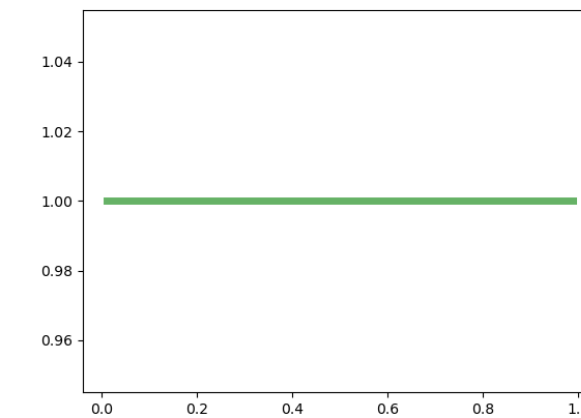
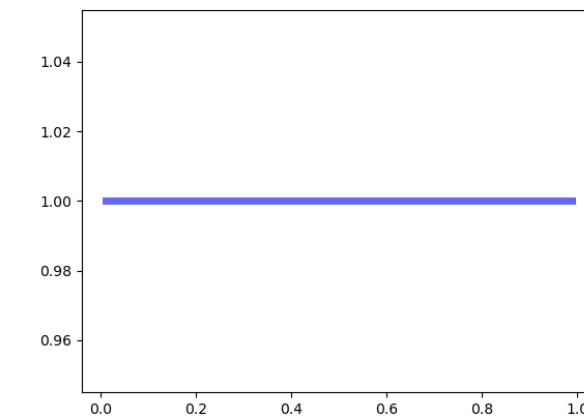
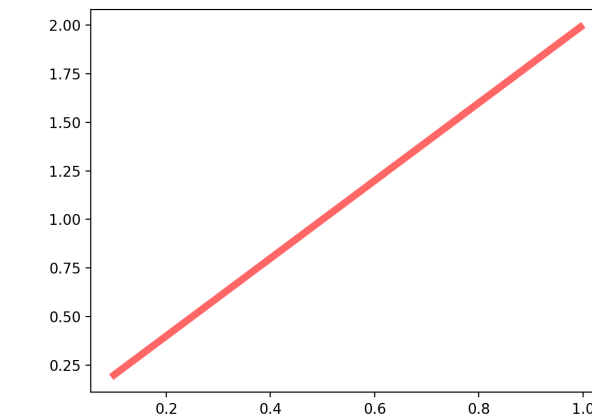
$$r_{t(k)} \sim \text{Bern}(\theta_k)$$

$$\sum_{t=1}^T r_{t(k)}$$

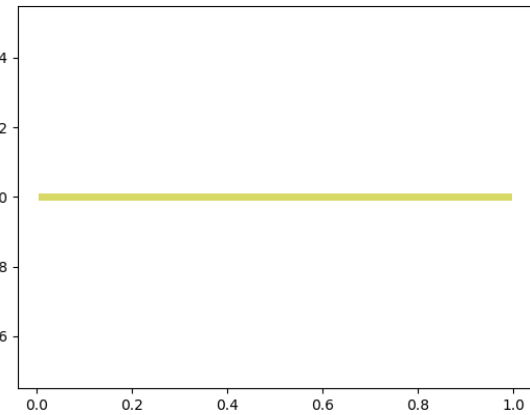
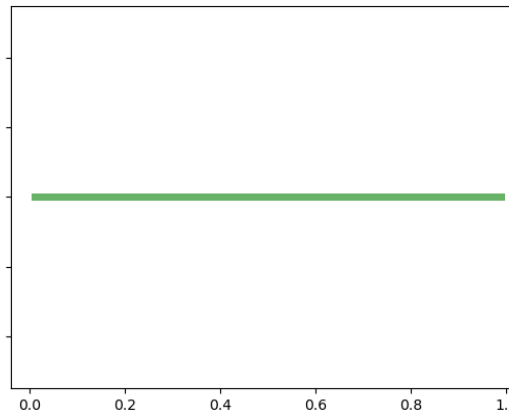
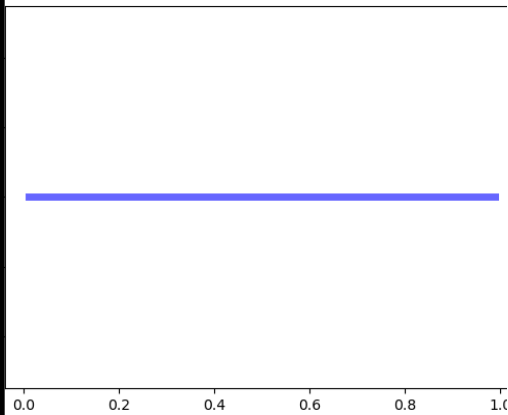
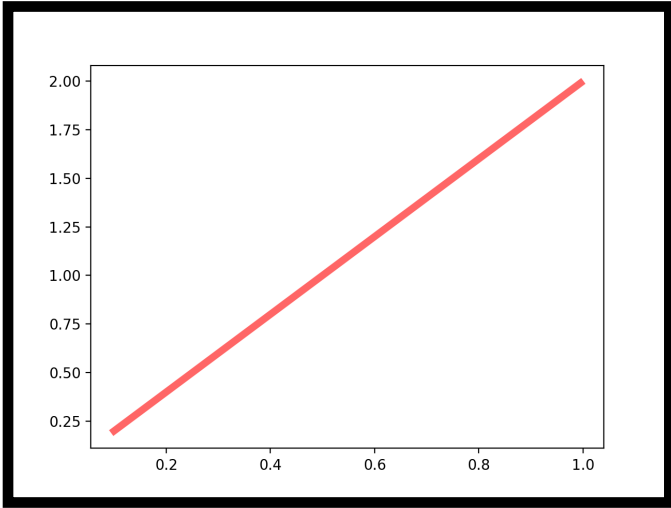
$$R = E\left[\sum_{t=1}^T Q_{t(k^*)} - \sum_{t=1}^T r_{t(k)}\right]$$



Thompson Sampling

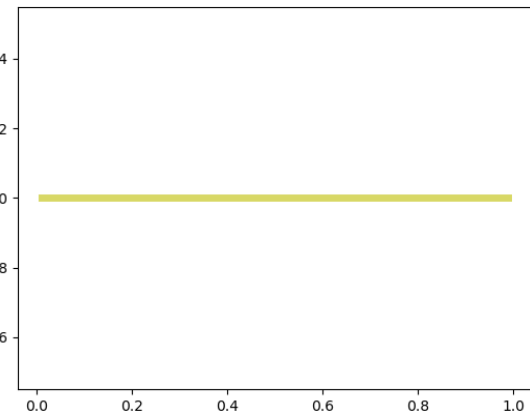
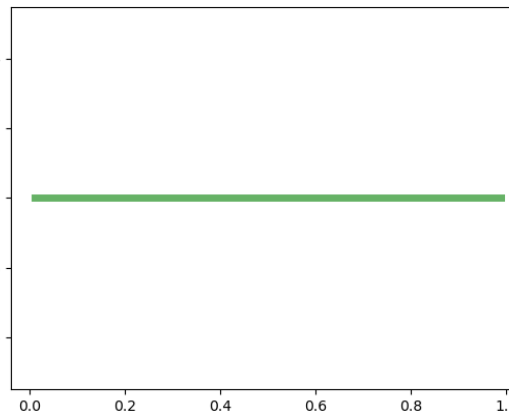
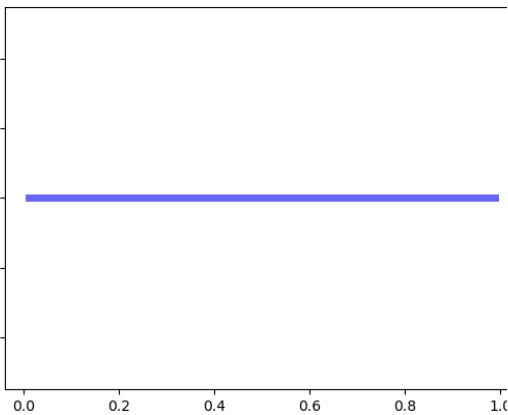
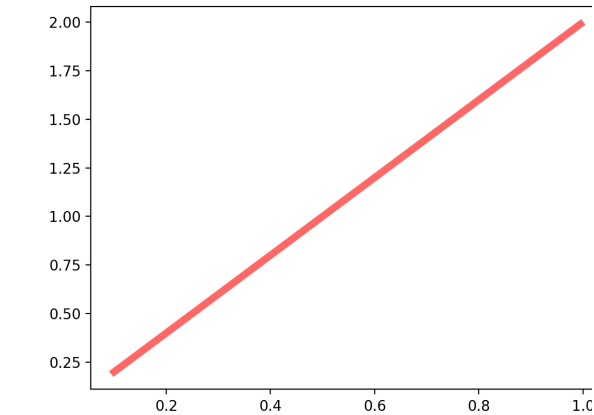


Thompson sampling (Thompson, 1933; Agrawal & Goyal, 2012)



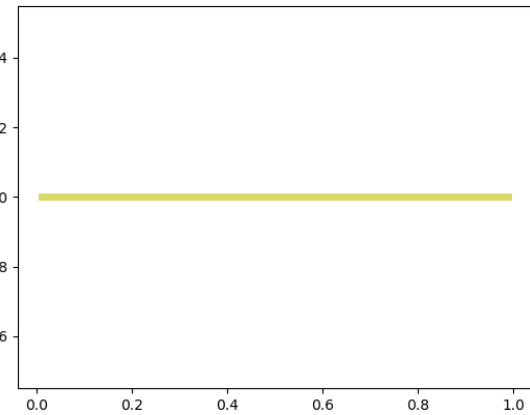
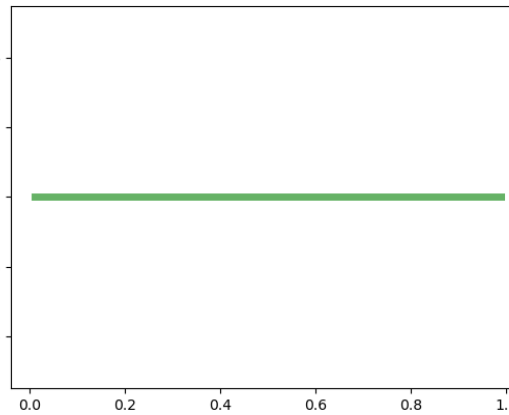
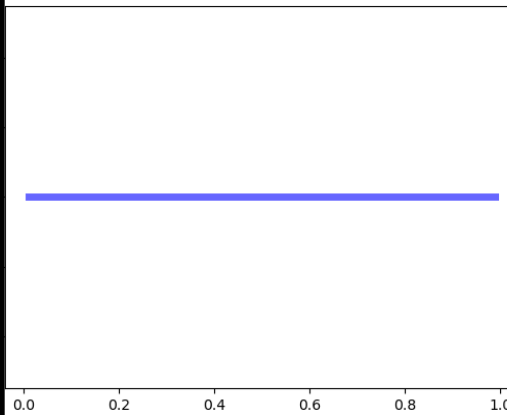
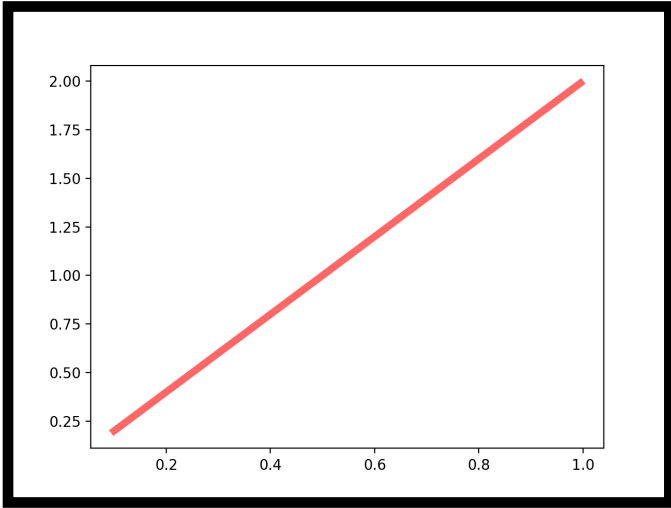
$$\theta_k \sim \text{Beta}(\alpha, \beta)$$

$$E = \frac{\alpha}{\alpha + \beta}$$



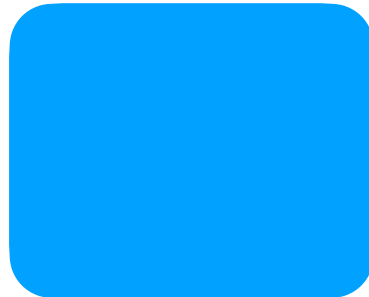
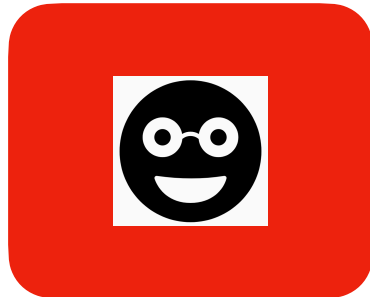
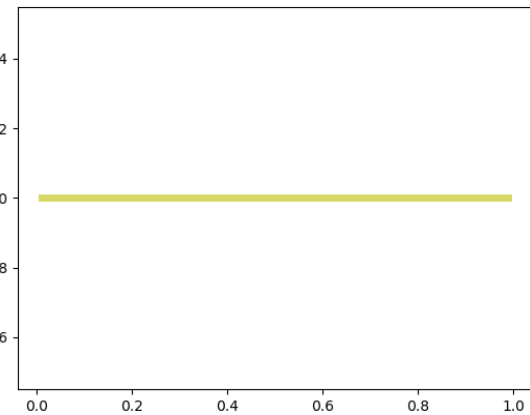
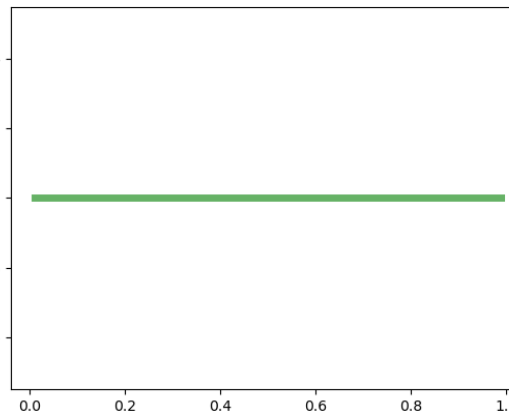
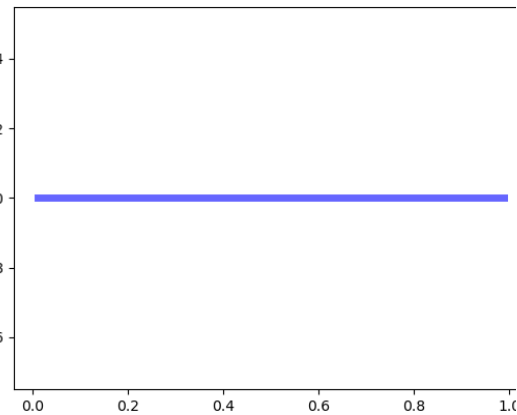
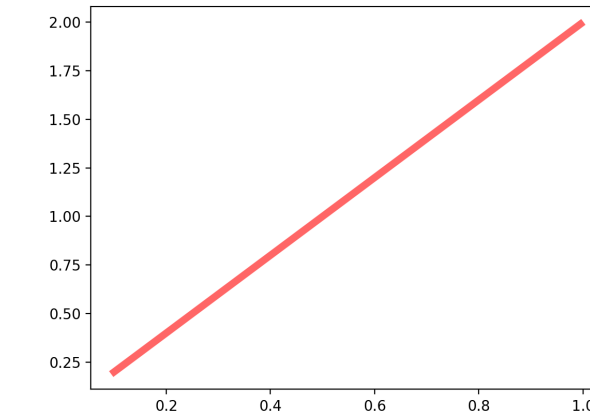
$$S_{k(t)}, F_{k(t)}$$

$$E = \frac{\alpha + S_{k(t)}}{\alpha + S_{k(t)} + \beta + F_{k(t)}}$$



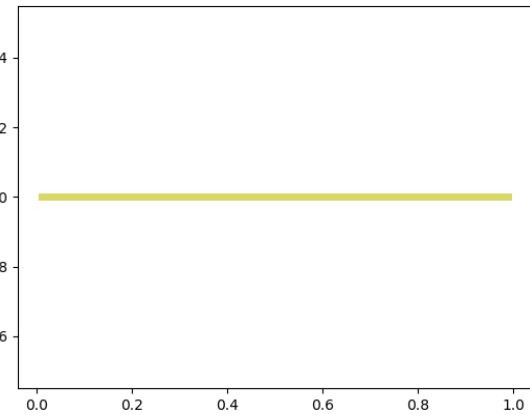
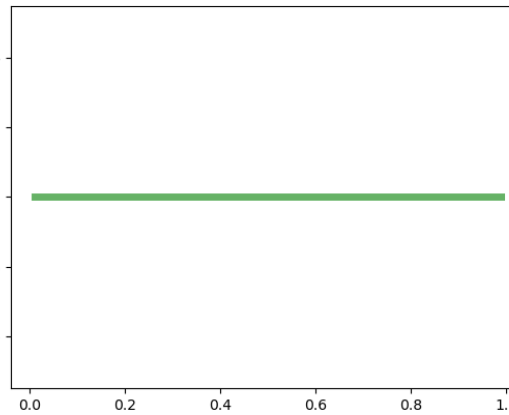
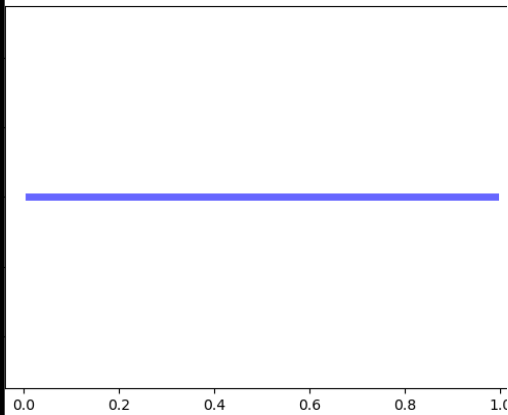
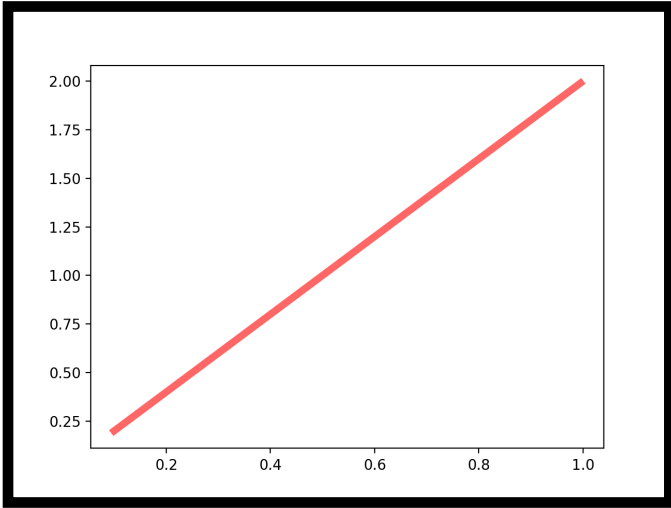
$$\theta_k \sim \text{Beta}(\alpha, \beta)$$

$$E = \frac{\alpha}{\alpha + \beta}$$



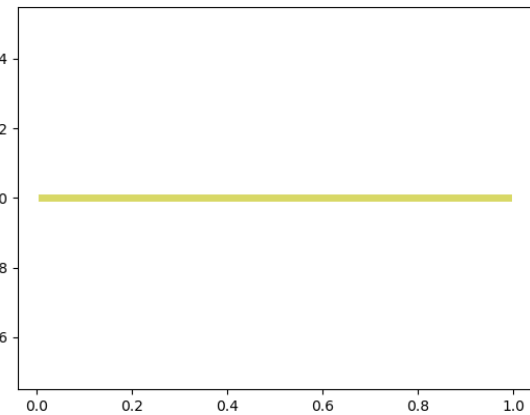
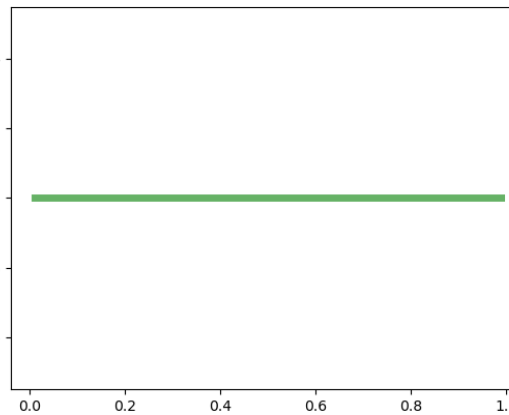
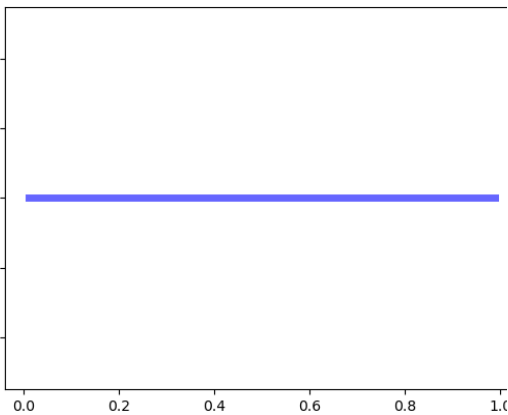
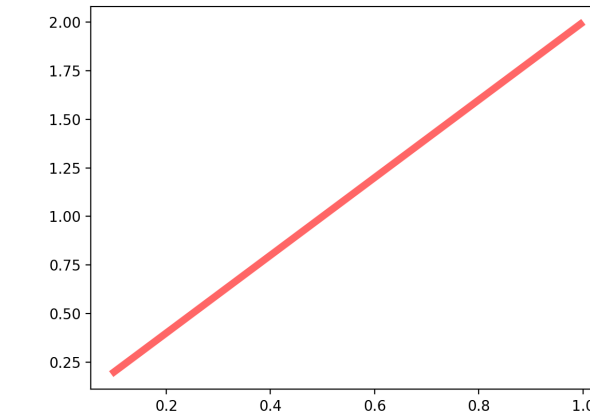
$$S_{k(t)}, F_{k(t)}$$

$$E = \frac{\alpha + S_{k(t)}}{\alpha + S_{k(t)} + \beta + F_{k(t)}}$$



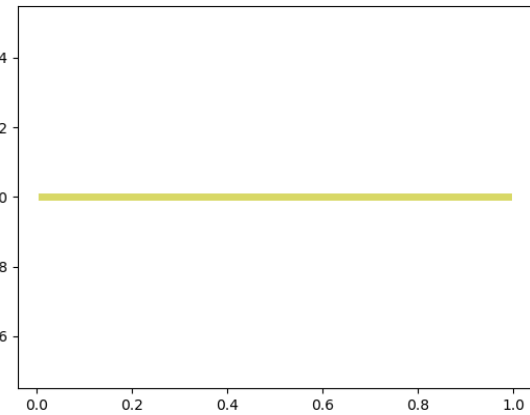
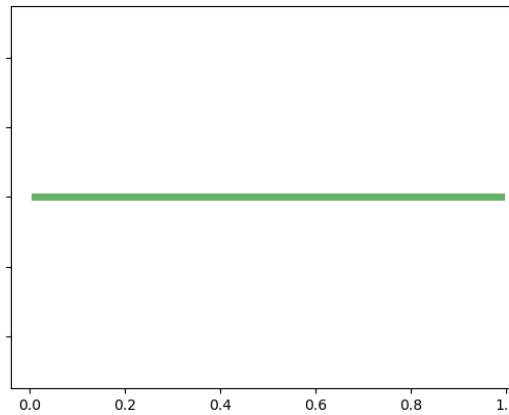
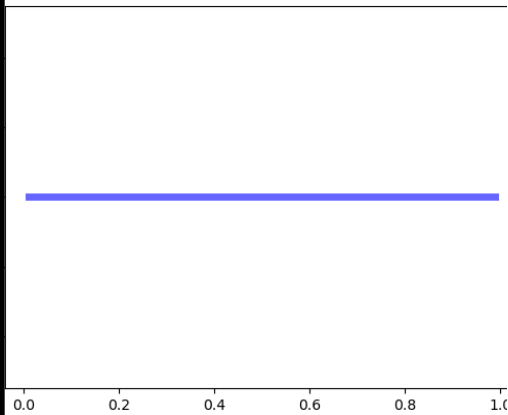
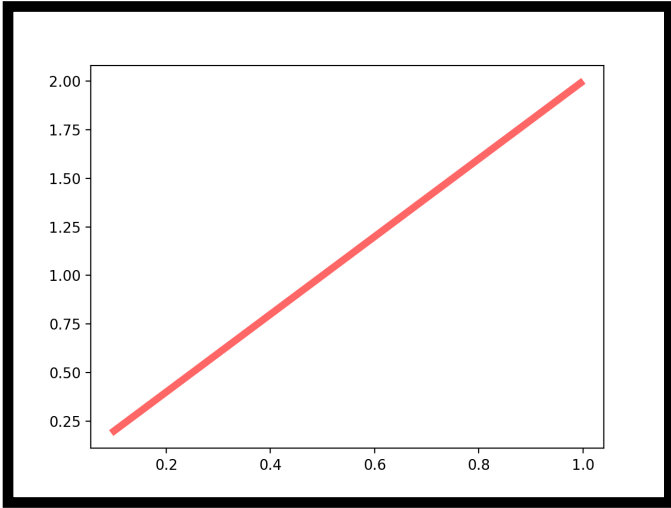
$$\theta_k \sim \text{Beta}(\alpha, \beta)$$

$$E = \frac{\alpha}{\alpha + \beta}$$



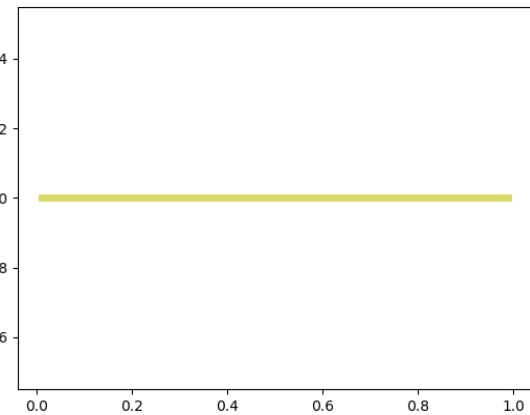
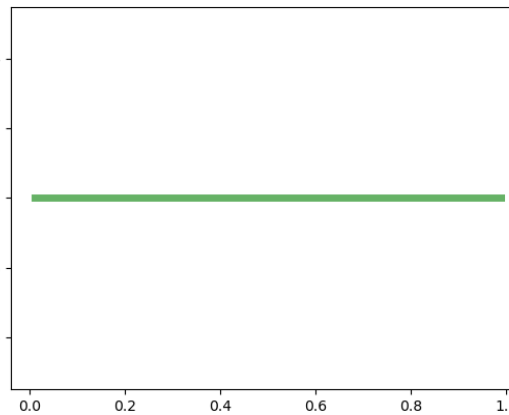
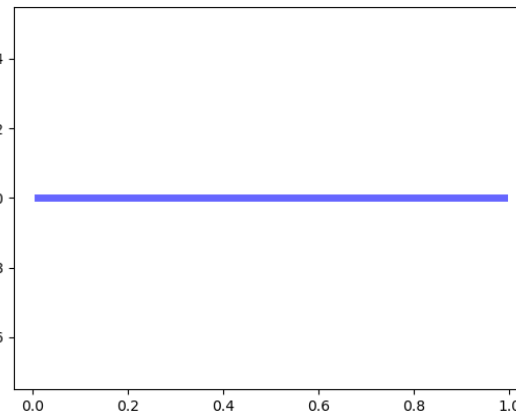
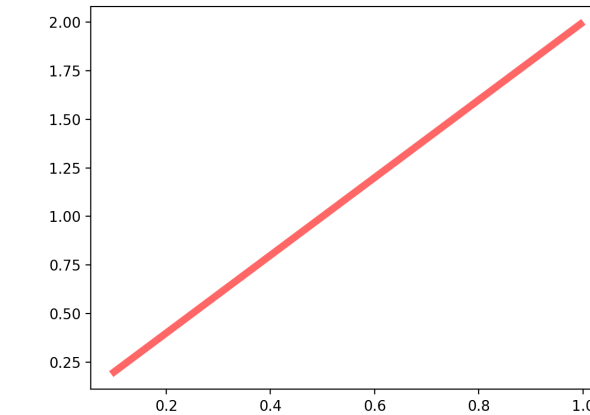
$$S_{k(t)}, F_{k(t)}$$

$$E = \frac{\alpha + S_{k(t)}}{\alpha + S_{k(t)} + \beta + F_{k(t)}}$$



$$\theta_k \sim \text{Beta}(\alpha, \beta)$$

$$E = \frac{\alpha}{\alpha + \beta}$$



$$S_{k(t)}, F_{k(t)}$$

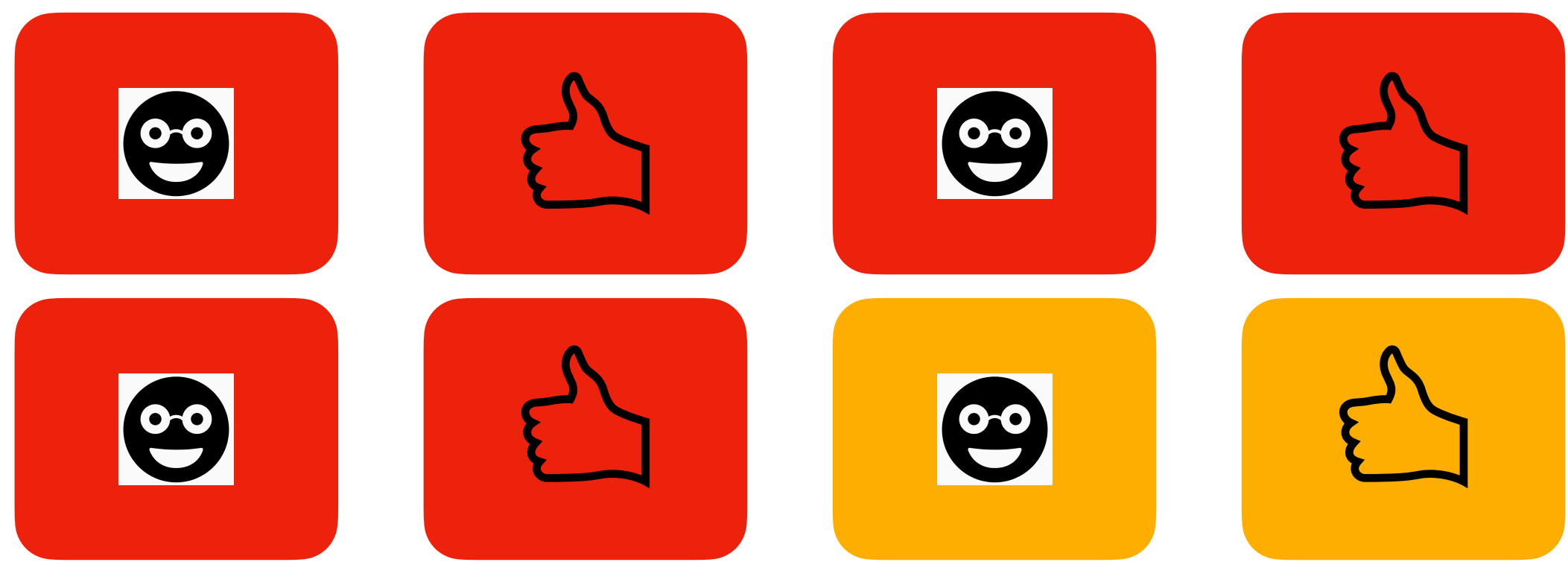
$$E = \frac{\alpha + S_{k(t)}}{\alpha + S_{k(t)} + \beta + F_{k(t)}}$$

Our hypothesis

Inaccurate stereotypes can result from rational explorations

- An intuitive example
- A formal model and simulations
- **Two human experiments**
- Implications

Confounds and Mechanisms



You choose v. You meet [random draw]

Confounds and Mechanisms

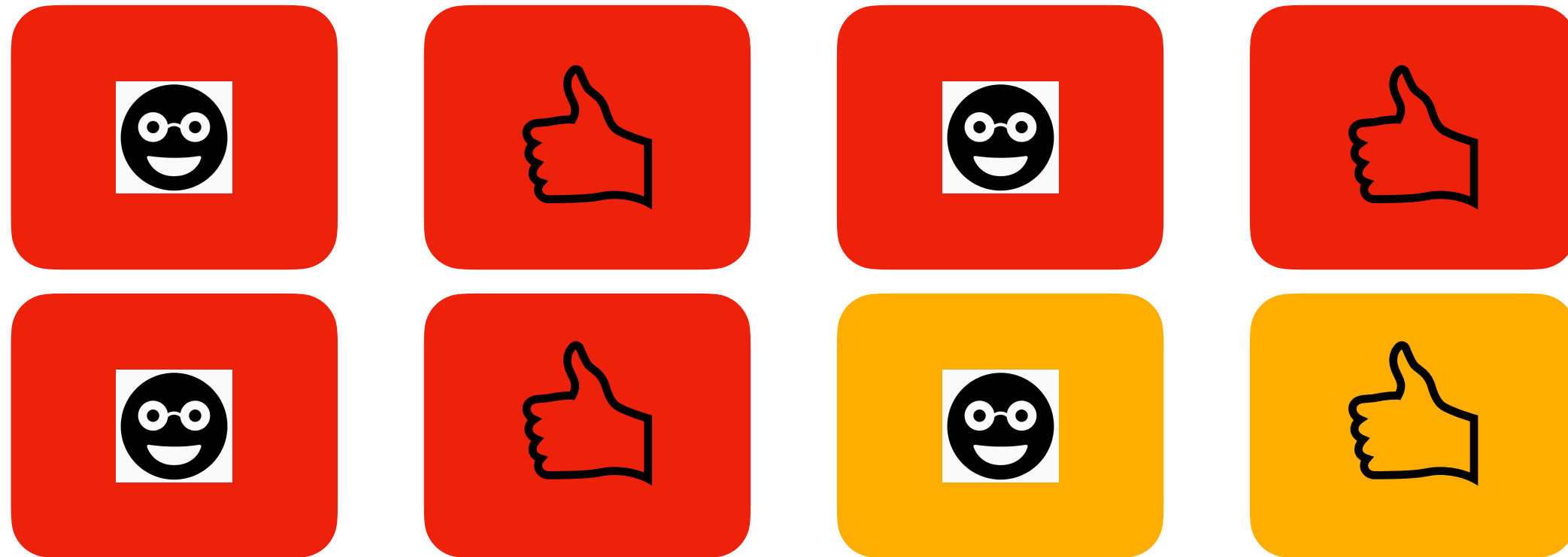


You choose v. You meet [random draw]

1. Selective sample or sense of control?

2. Order of rewards or mere presence of choices?

Confounds and Mechanisms



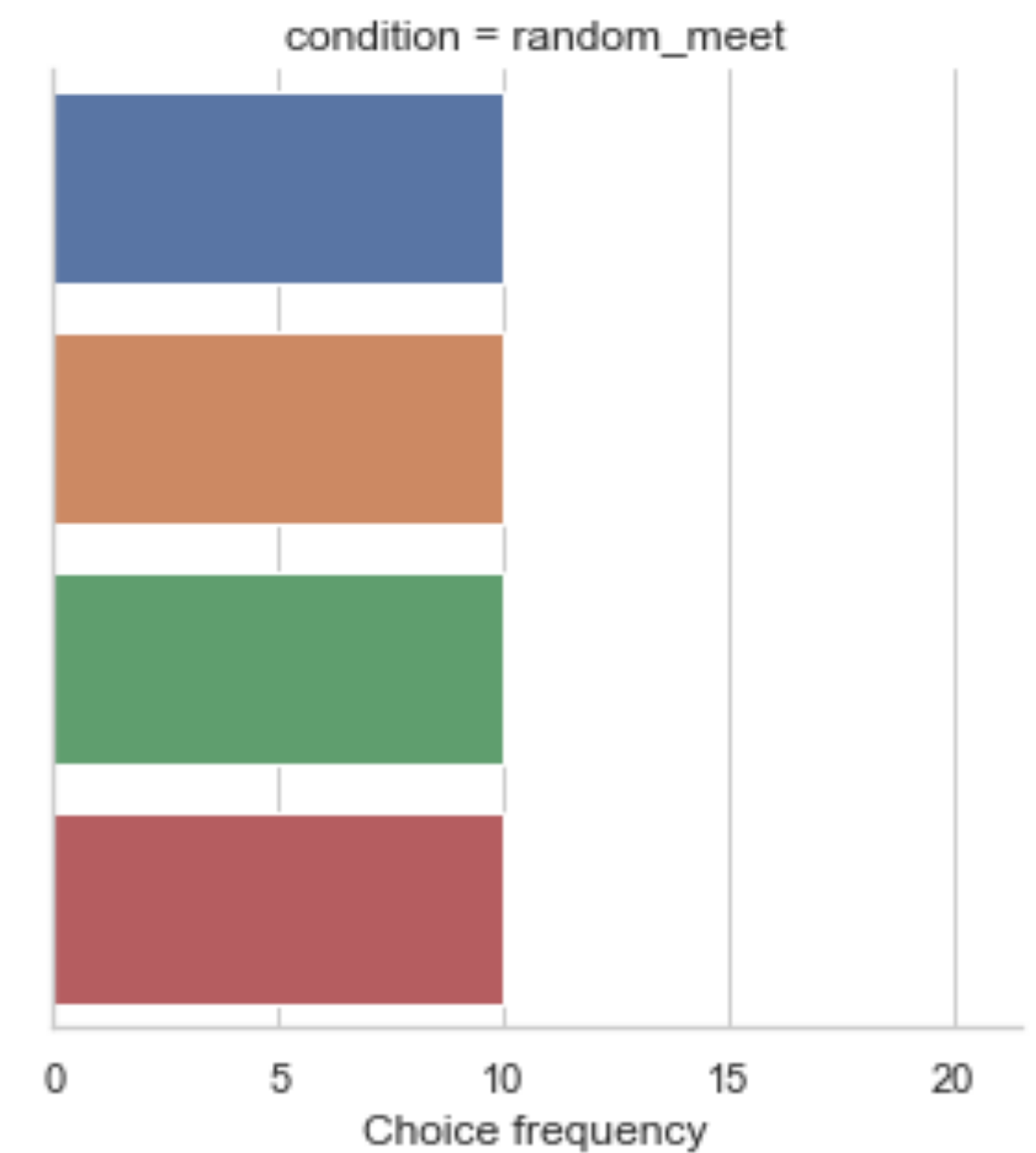
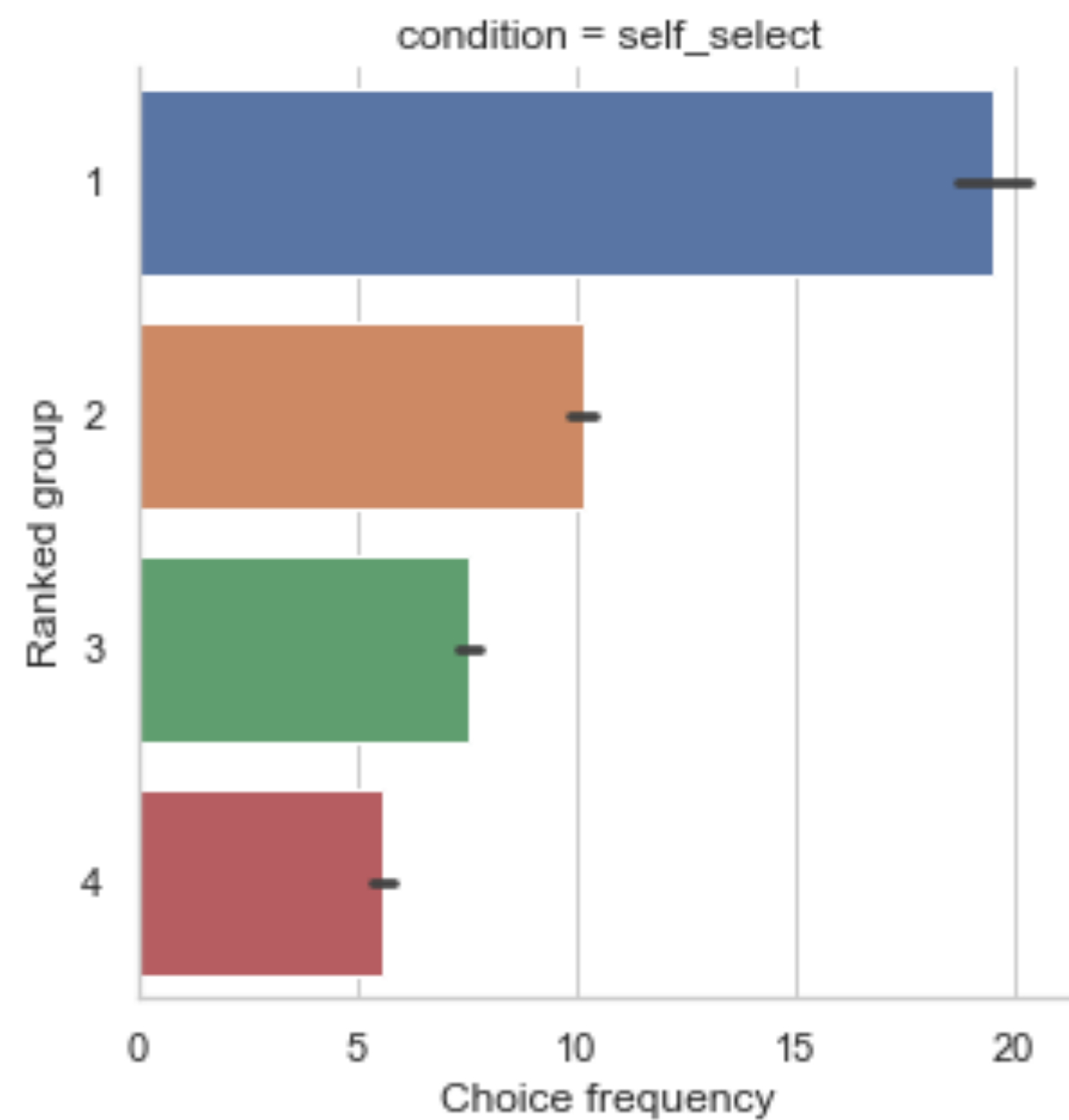
You choose v. You meet [random draw]

1. Selective sample or sense of control?
 - Yoke both choices and rewards
2. Order of rewards or mere presence of choices?
 - Yoke choices but not rewards

Partner choices: replication

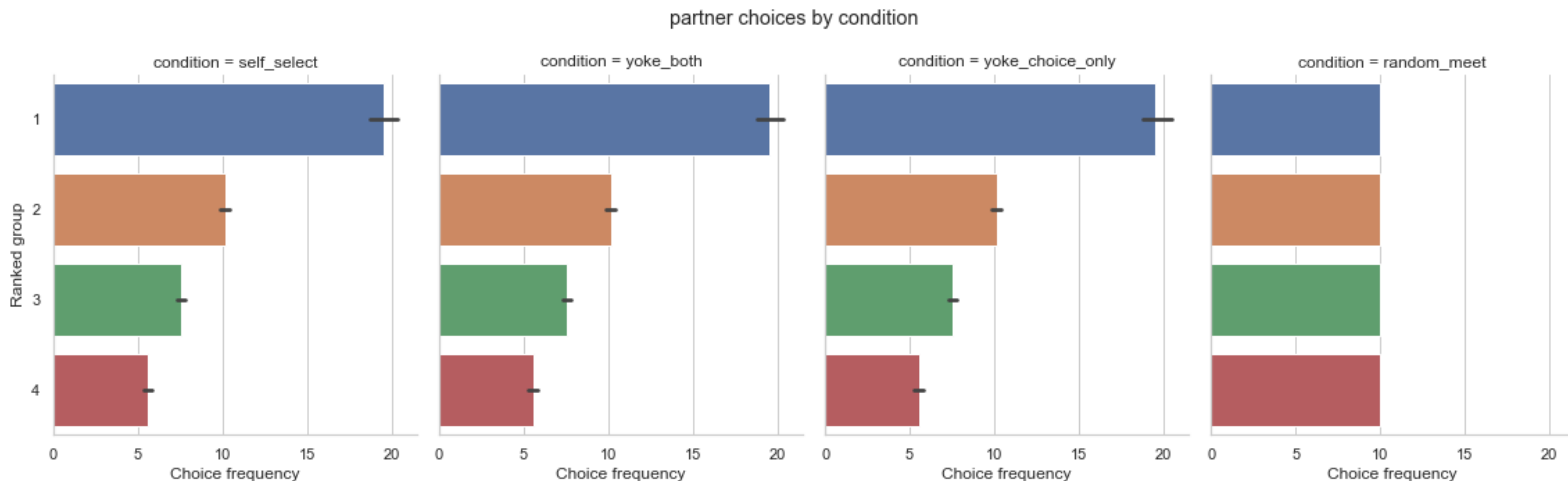
2000 online participants (500 in each 1 by 4)

partner choices by condition



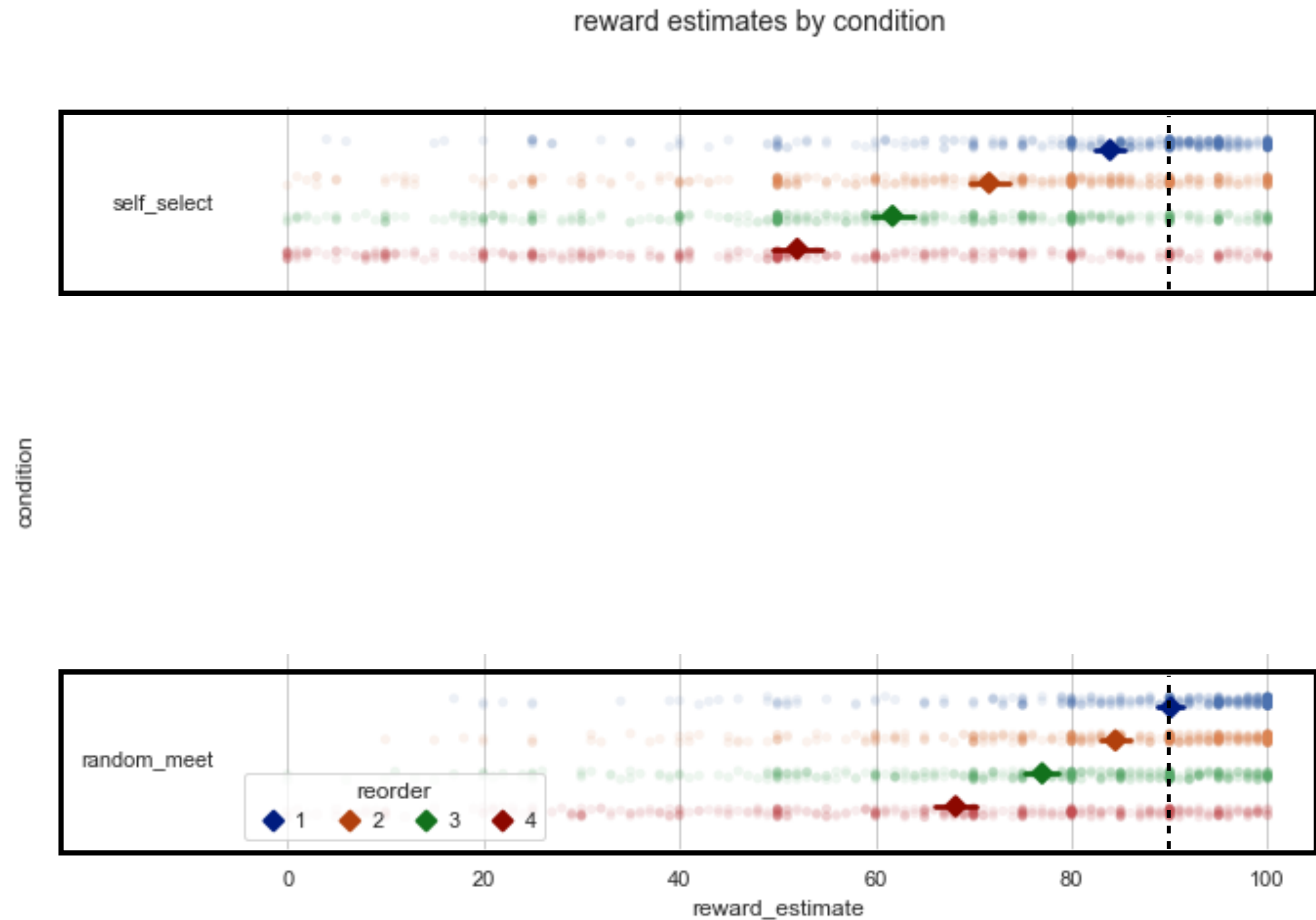
Partner choices: yoke

2000 online participants (500 in each 1 by 4)



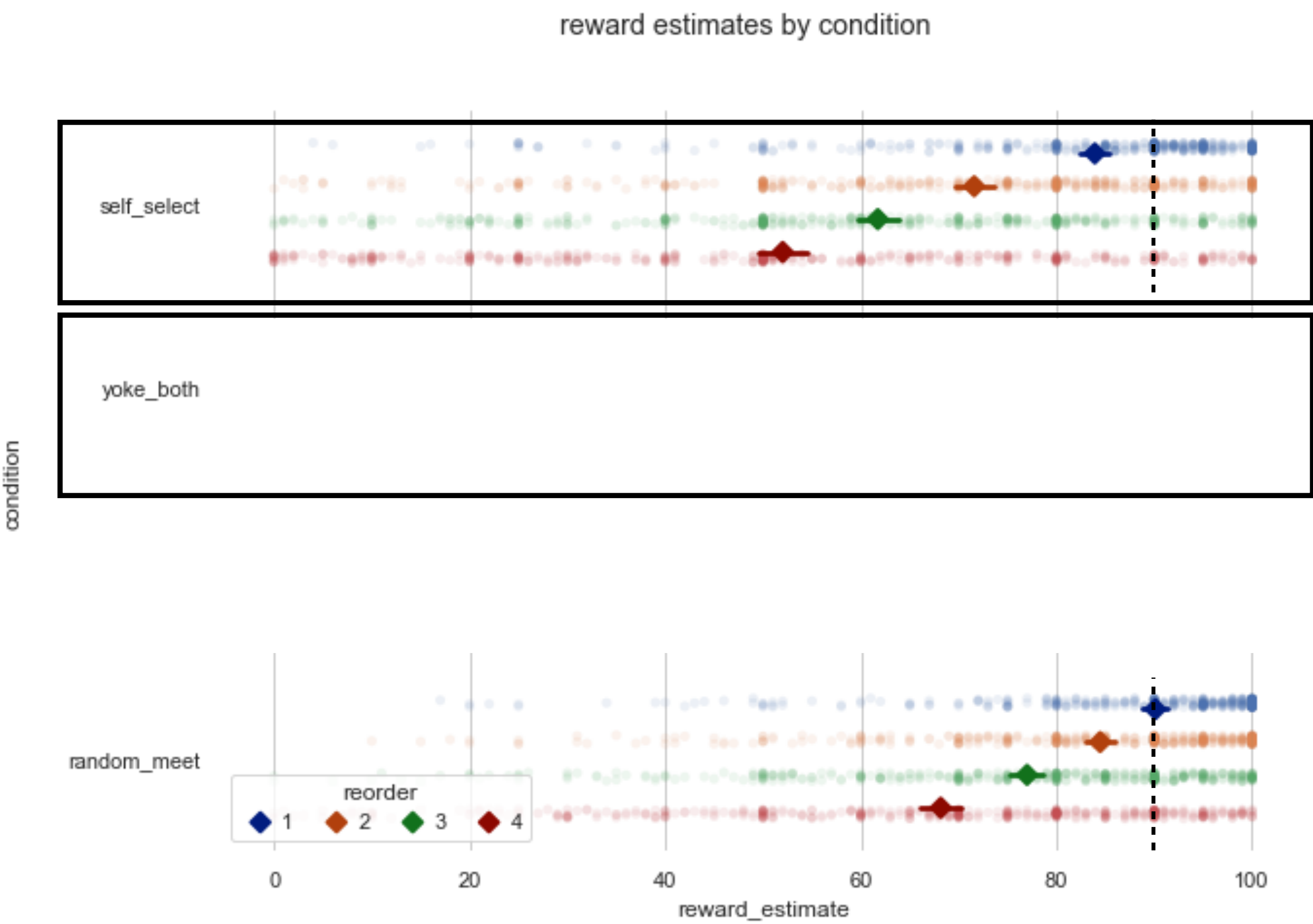
Reward estimations: replication

2000 online participants (500 in each 1 by 4)



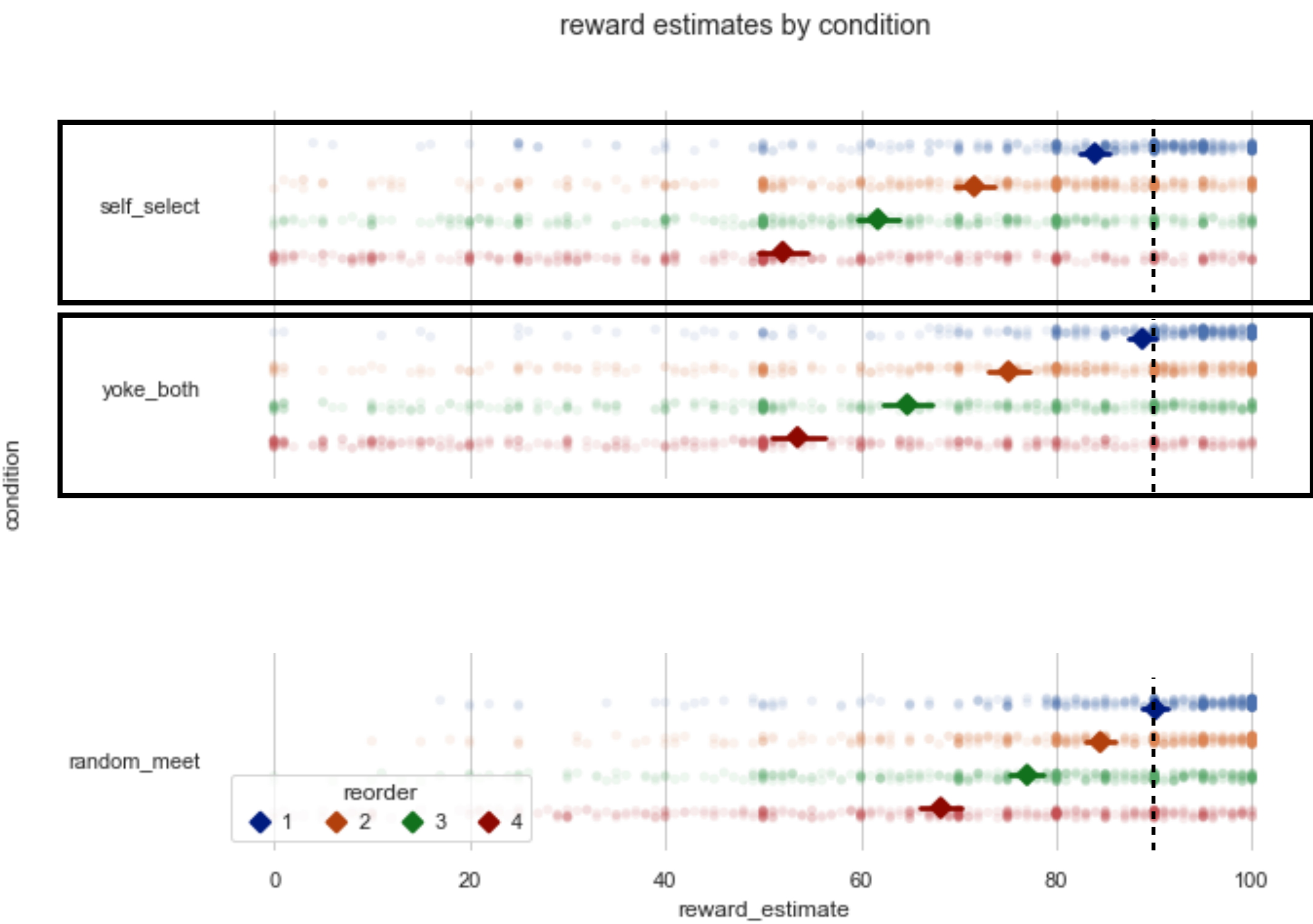
Reward estimations: selective sample or sense of control

2000 online participants (500 in each 1 by 4)



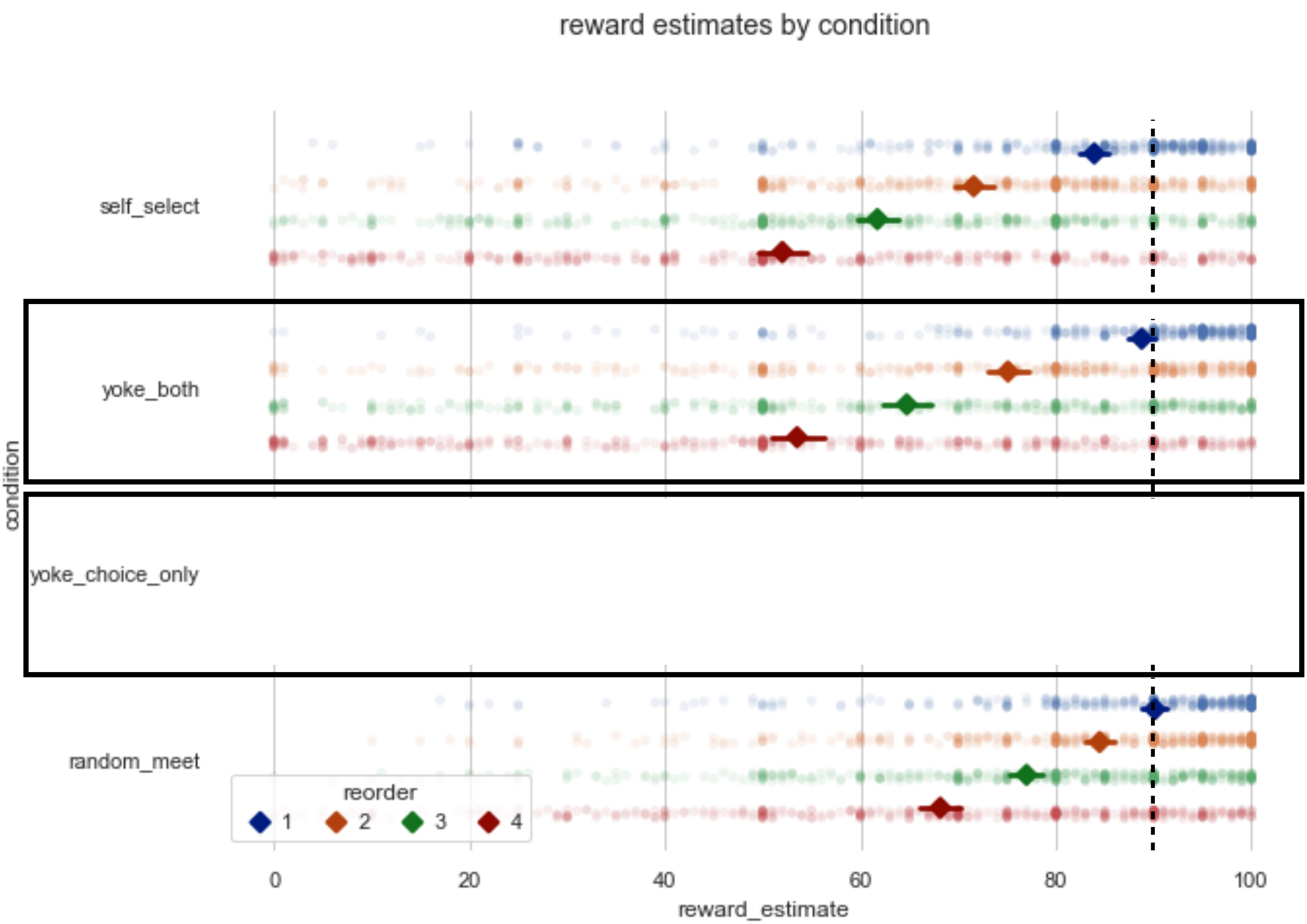
Reward estimations: selective sample or sense of control

2000 online participants (500 in each 1 by 4)



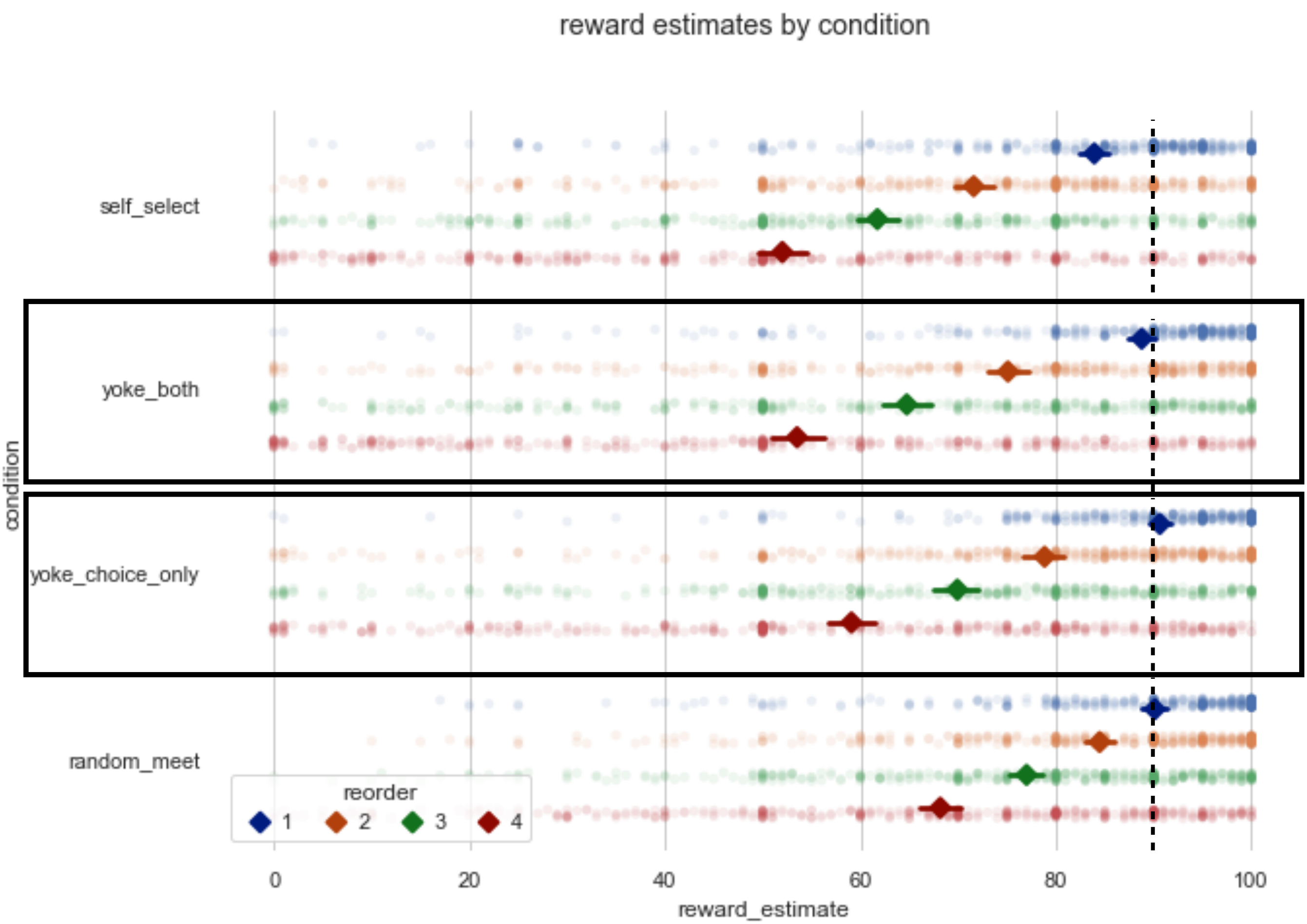
Reward estimations: order of rewards or mere presence of choices

2000 online participants (500 in each 1 by 4)



Reward estimations: order of rewards or mere presence of choices

2000 online participants (500 in each 1 by 4)



Interim summary

Replicating prior experiment?

- Yes

Interim summary

Replicating prior experiment?

- Yes

Selective samples or sense of control?

- Selective samples

Interim summary

Replicating prior experiment?

- **Yes**

Selective samples or sense of control?

- **Selective samples**

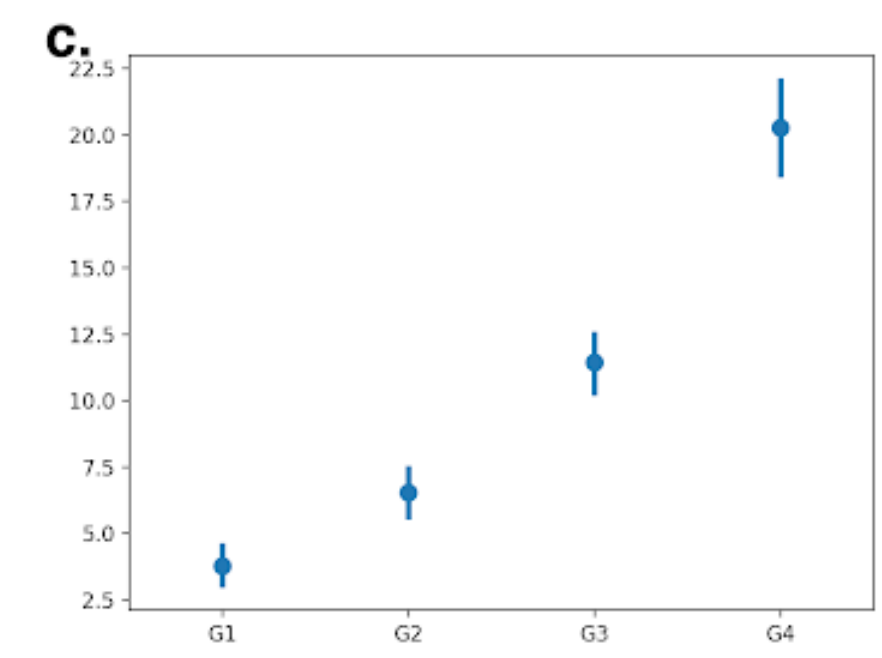
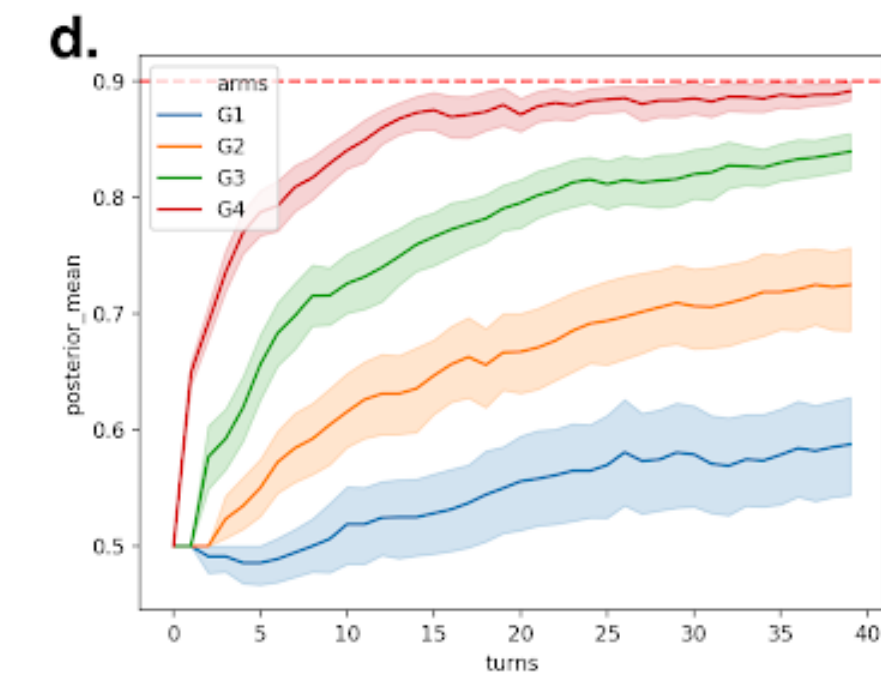
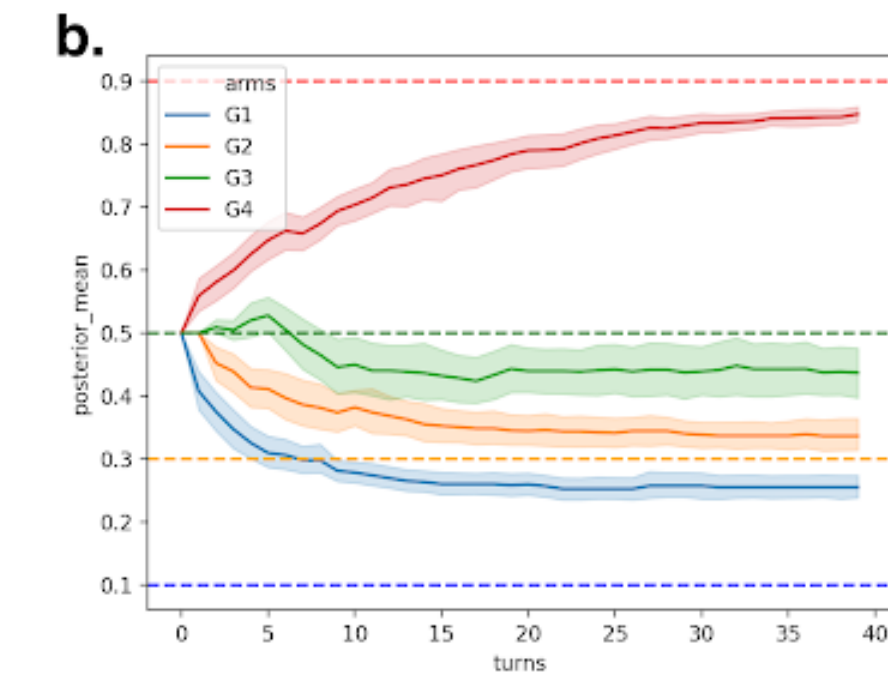
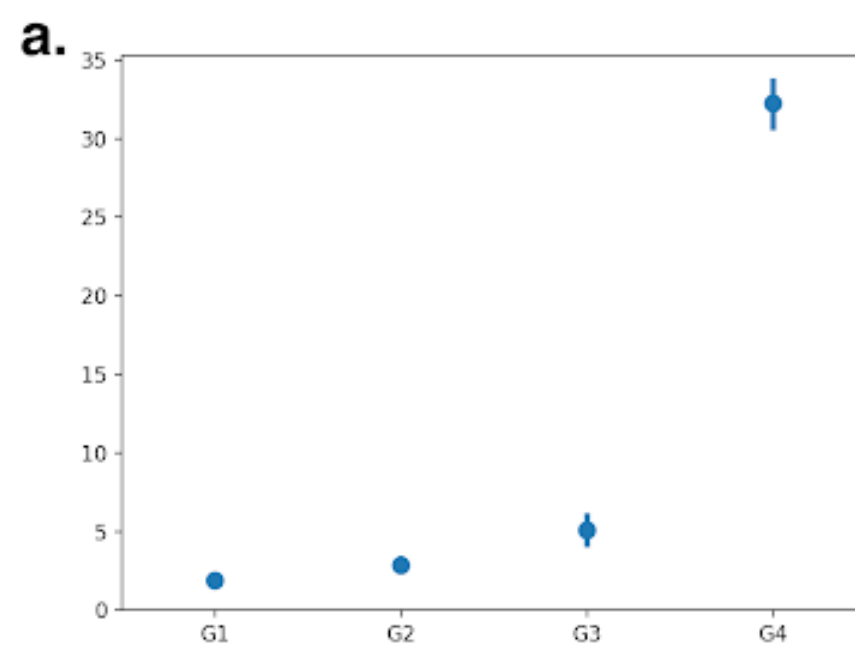
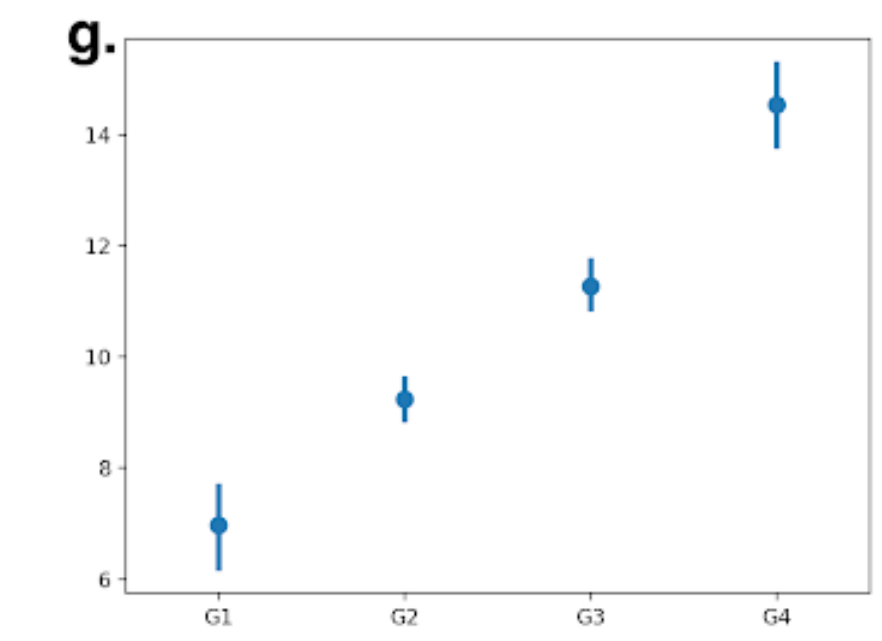
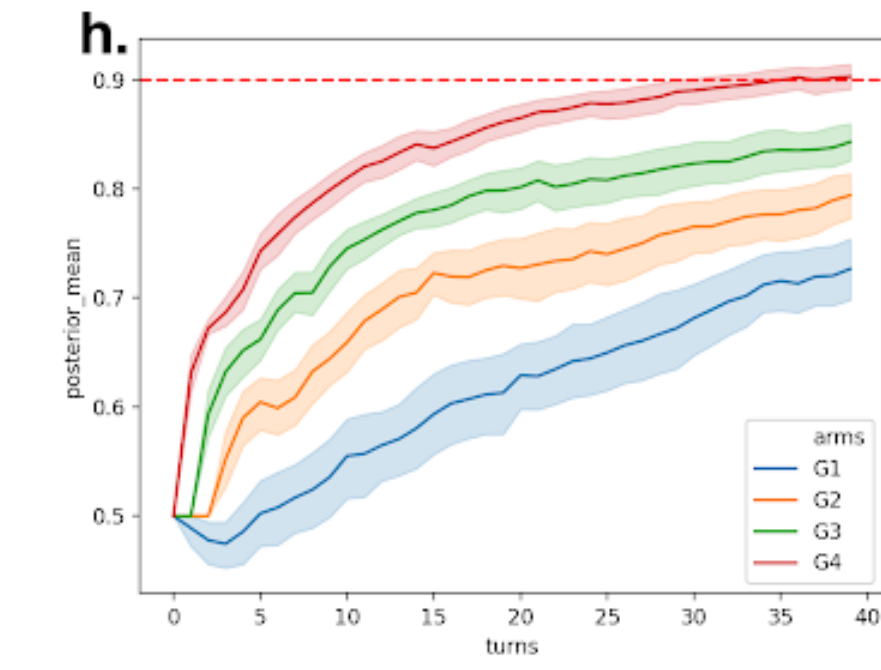
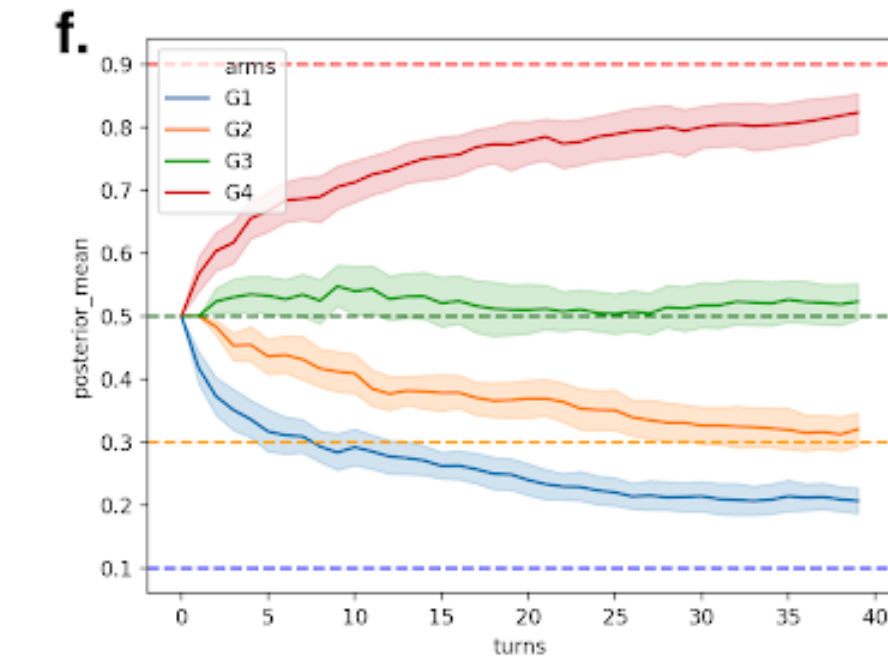
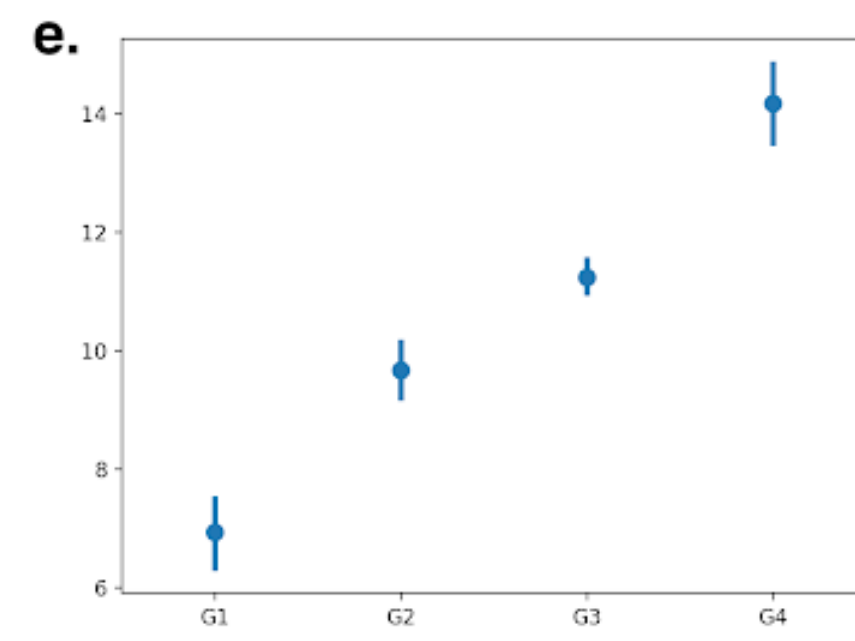
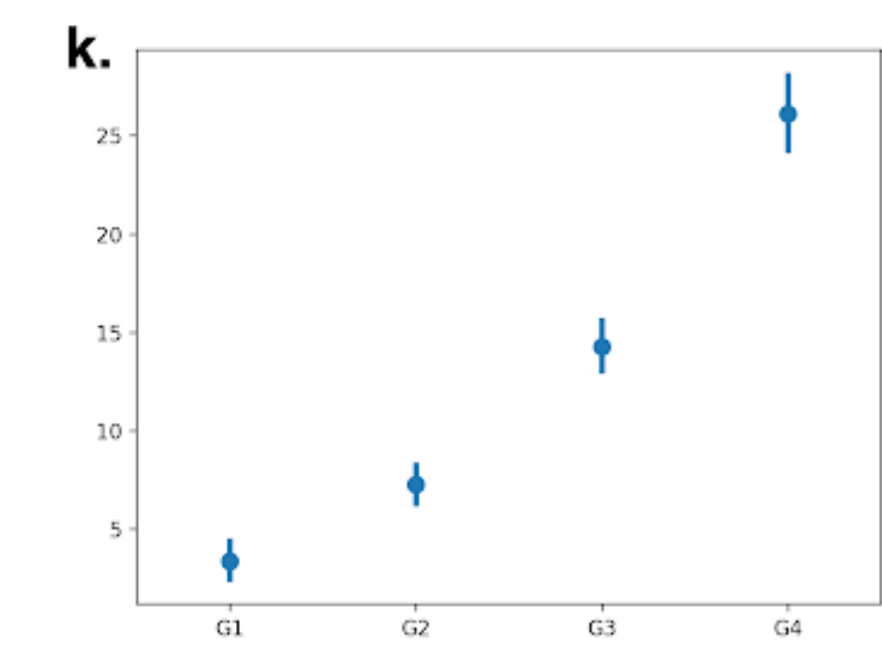
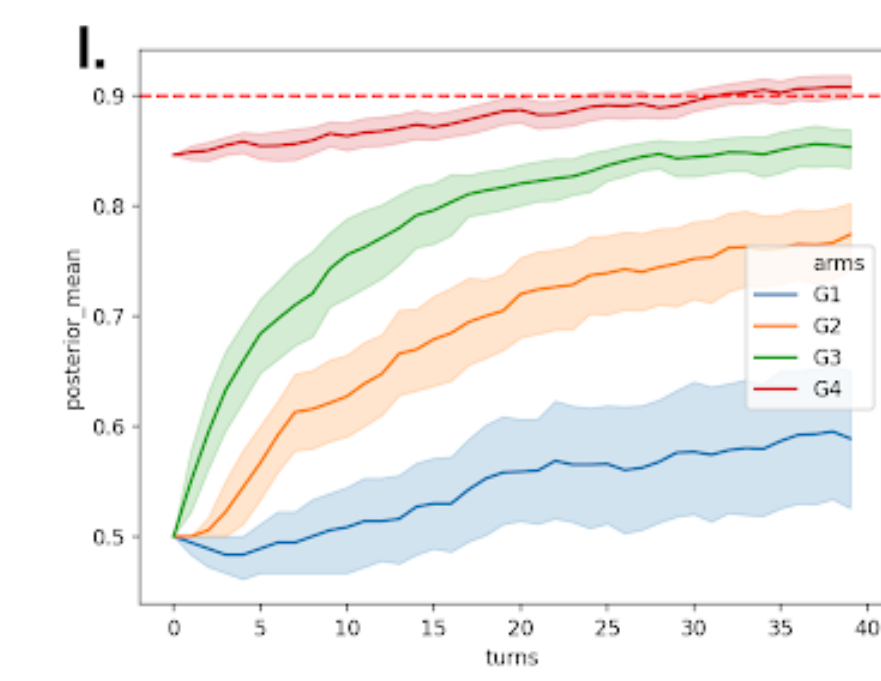
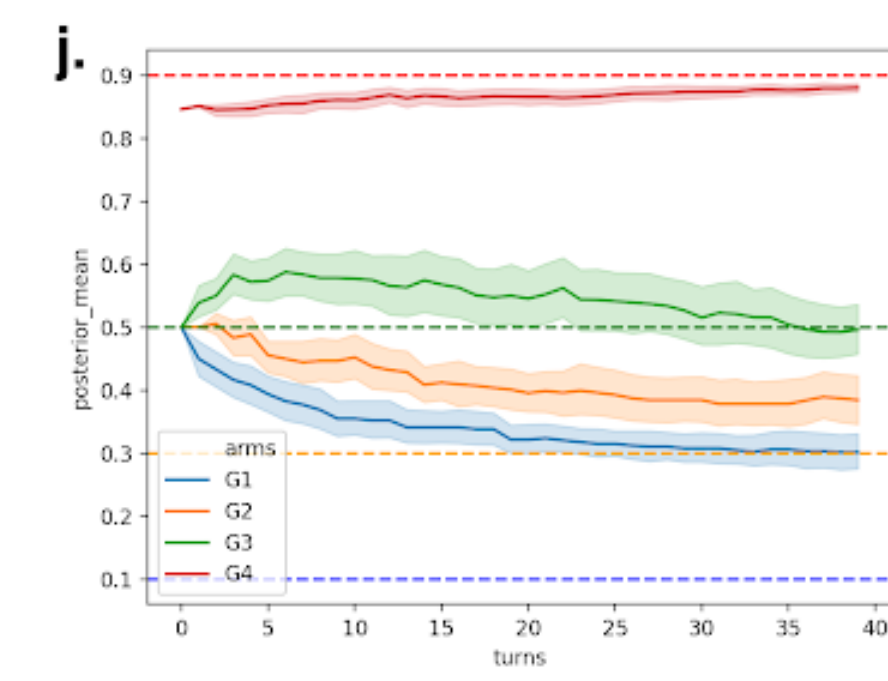
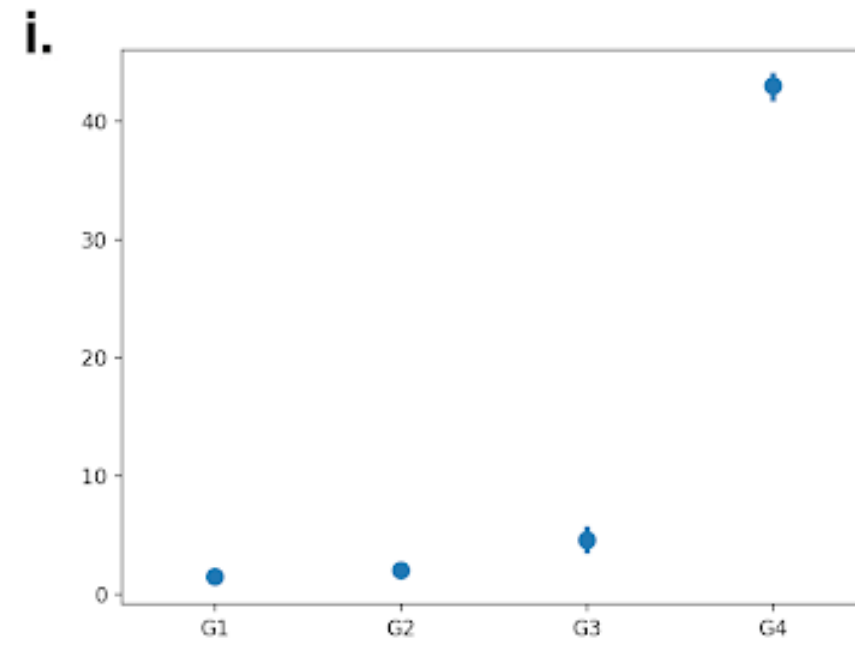
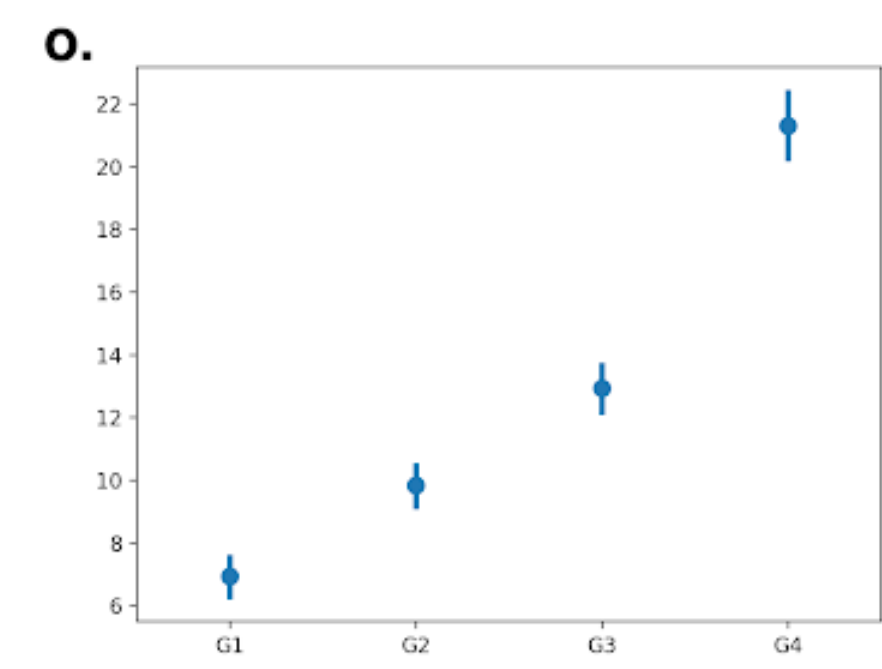
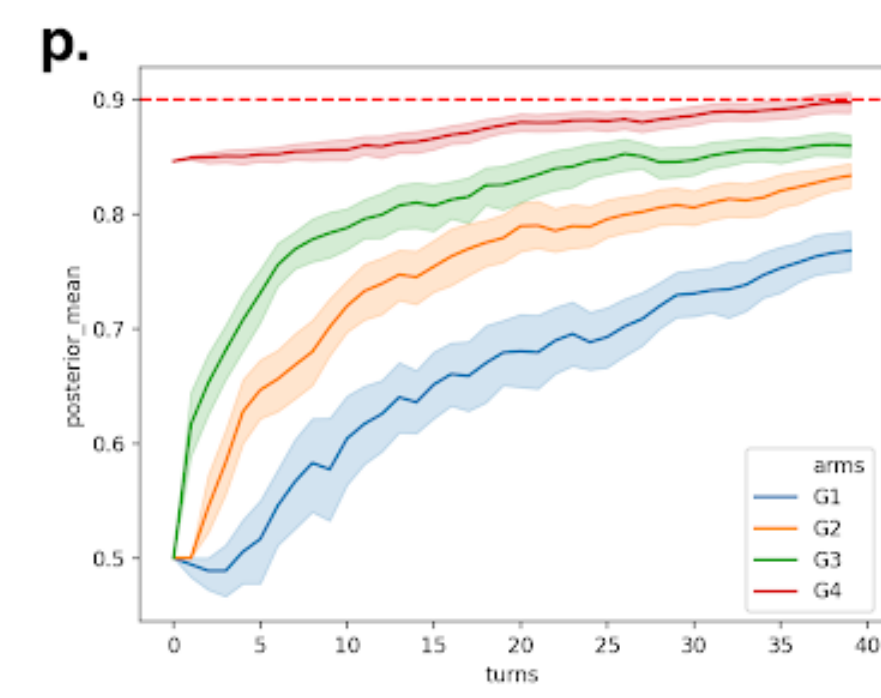
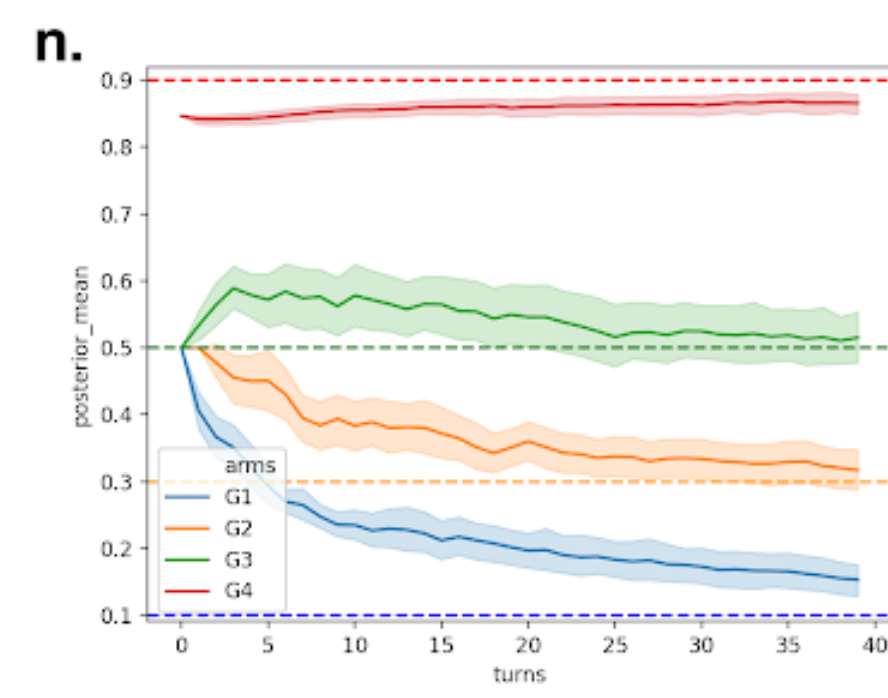
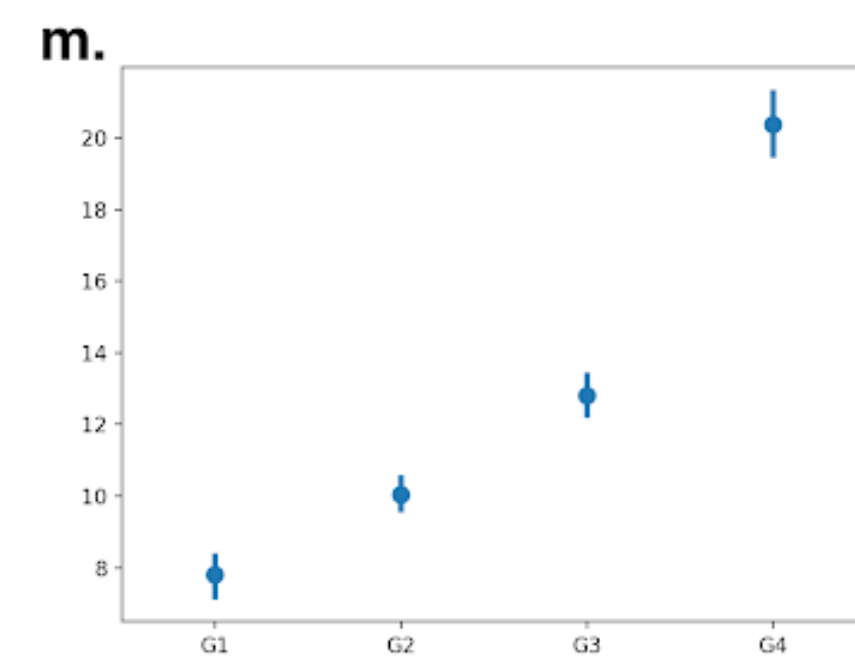
Order of rewards or mere presence of choices?

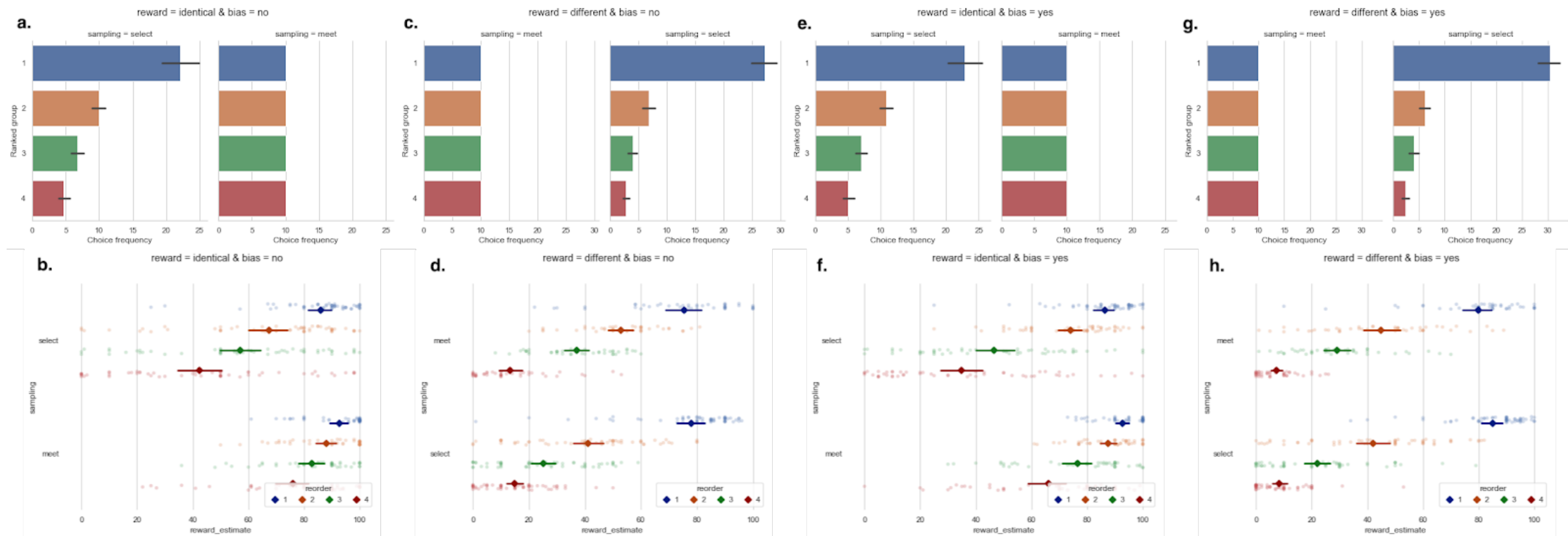
- **Order of rewards**

Appendix

Related works

- **Denrell, Experience sampling**
- **Gureckis, Learning trap**
- **Fiedler, Reward rich**





Analytic solutions

Dynamic programming

