Aversion to option loss in a restless bandit task

Danielle Navarro, Peter Tran &
Nicole Baz

UNSW SYDNEY

11 June 2018, ALG meeting, Katoomba

In real life, choice options have limited availability

Images via pixabay

In real life, choice options have limited availability

Products don't stay on sale forever

Images via pixabay

In real life, choice options have limited availability

Products don't stay on sale forever

Houses go off the market

Images via pixabay

In real life, choice options have limited availability

Products don't stay on sale forever

Houses go off the market

Possible romantic partners move on

Images via pixabay

Keeping options viable requires active maintenance

Images via pixabay

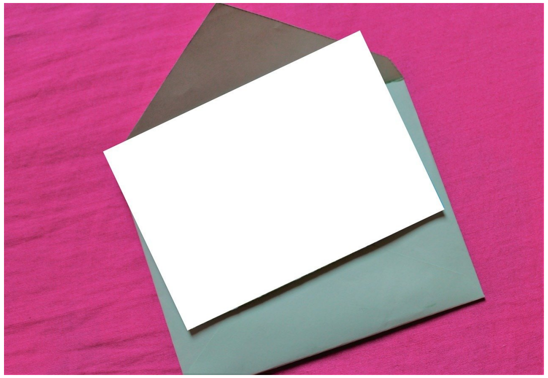
Keeping options viable requires active maintenance

Study now for a career later

Images via pixabay

Keeping options viable requires active maintenance

Study now for a career later

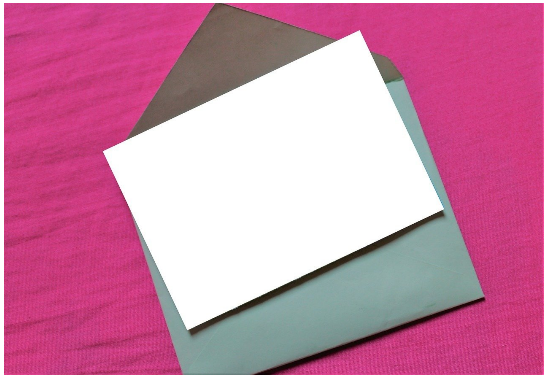


RSVP now to attend party later

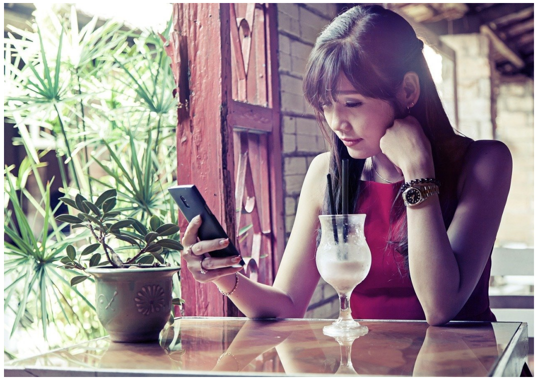
Images via pixabay

Keeping options viable requires active maintenance

Study now for a career later



RSVP now to attend party later



Show up to the first date to get invited on a second

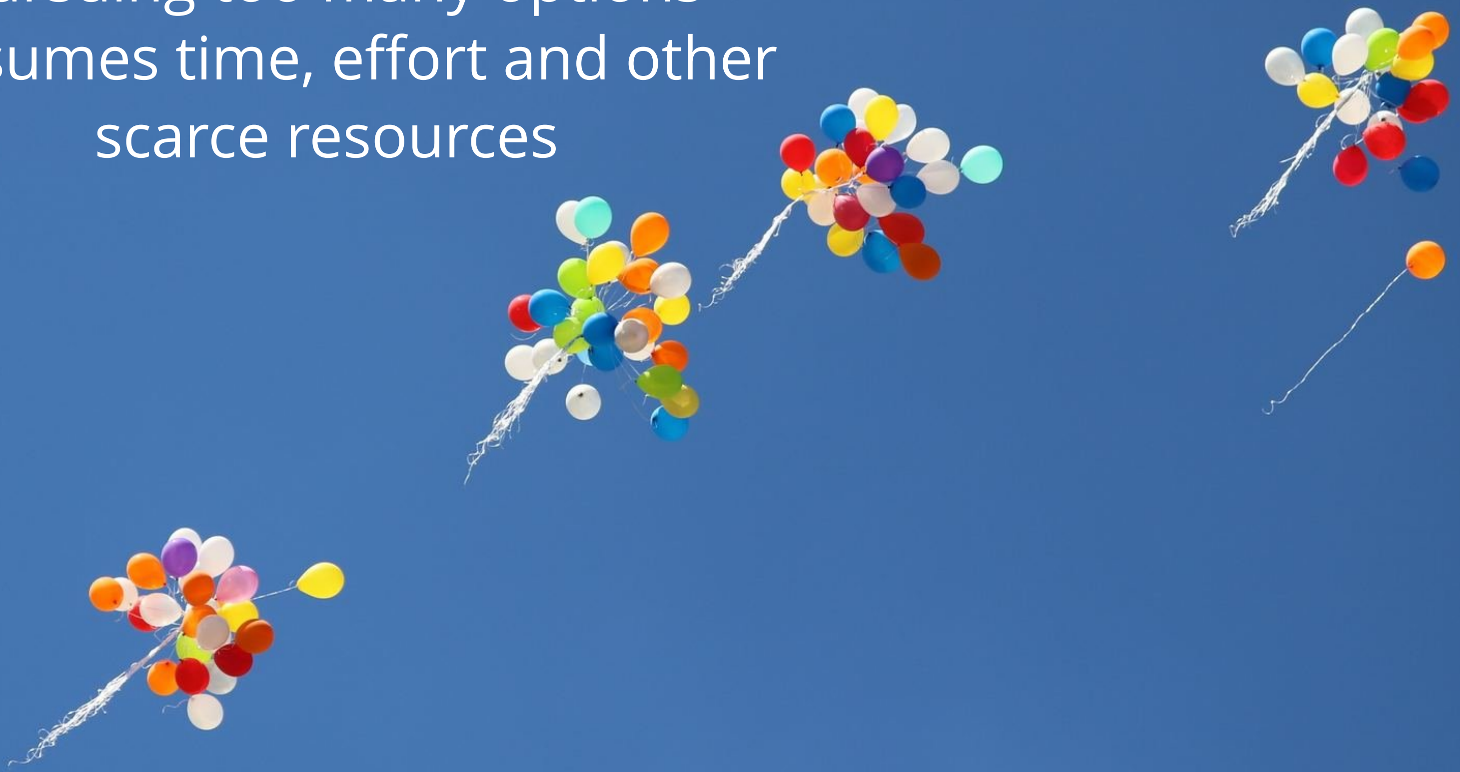
Images via pixabay

How many to pursue?

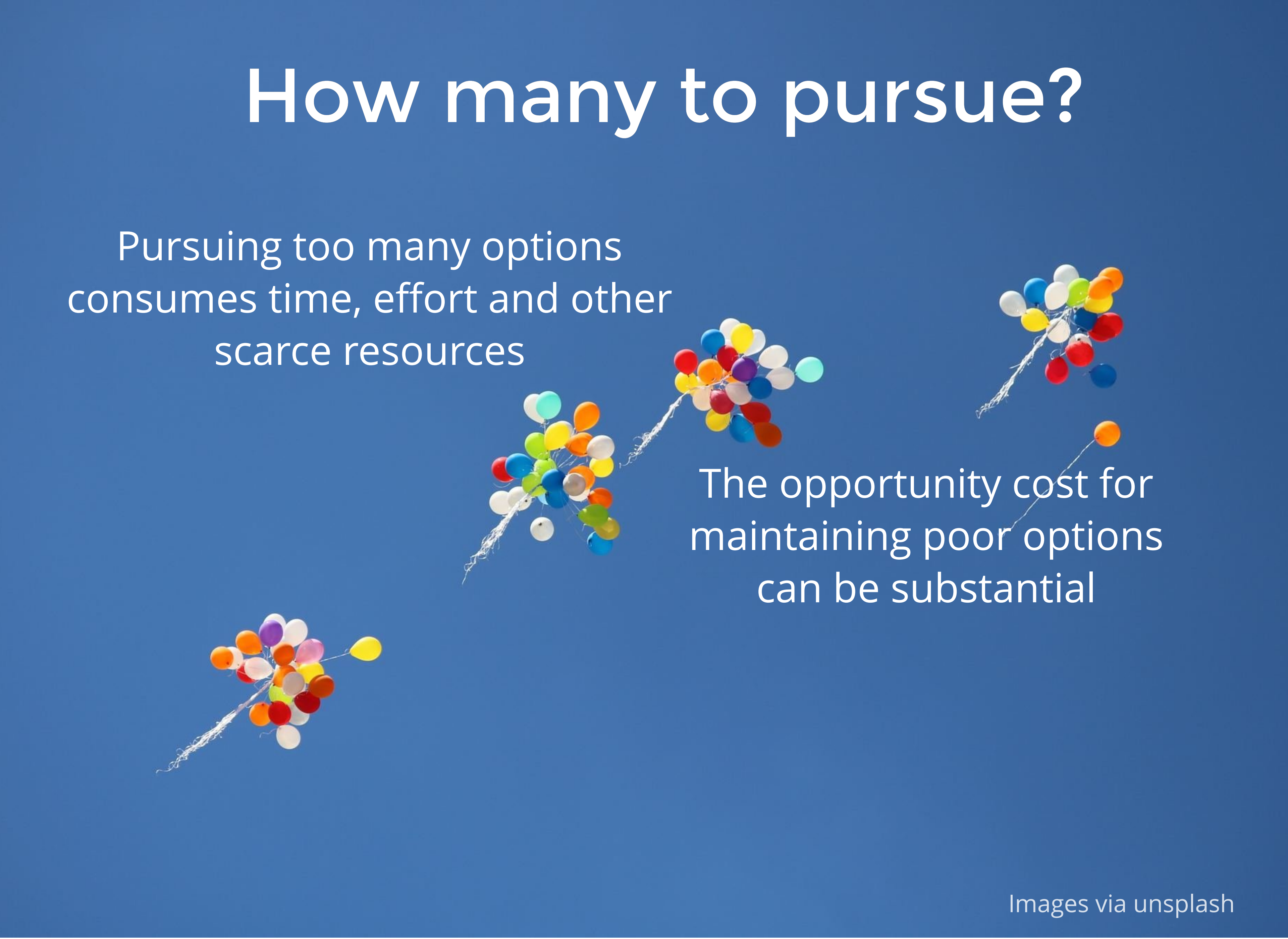
Images via unsplash

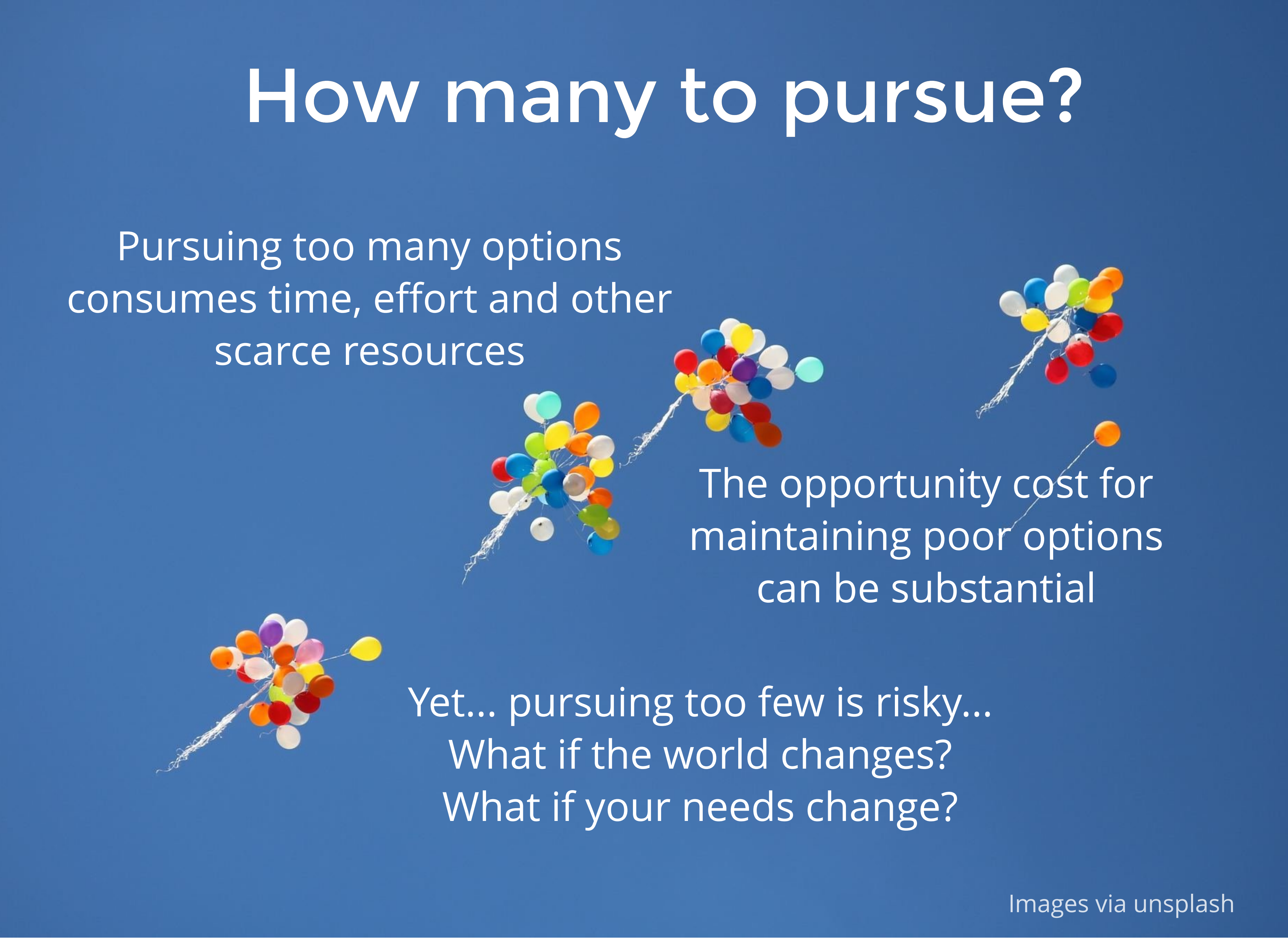
How many to pursue?

Pursuing too many options consumes time, effort and other scarce resources



Images via unsplash


How many to pursue?
Pursuing too many options consumes time, effort and other scarce resources
The opportunity cost for maintaining poor options can be substantial
Images via unsplash


How many to pursue?
Pursuing too many options consumes time, effort and other scarce resources
The opportunity cost for maintaining poor options can be substantial
Yet... pursuing too few is risky...
What if the world changes?
What if your needs change?
Images via unsplash

Existing literature?

- Vanishing options tasks
  - Shin & Ariely (2004)
  - Ejova et al (2009)
  - Neth et al (2014)

"doors" problems

Existing literature?

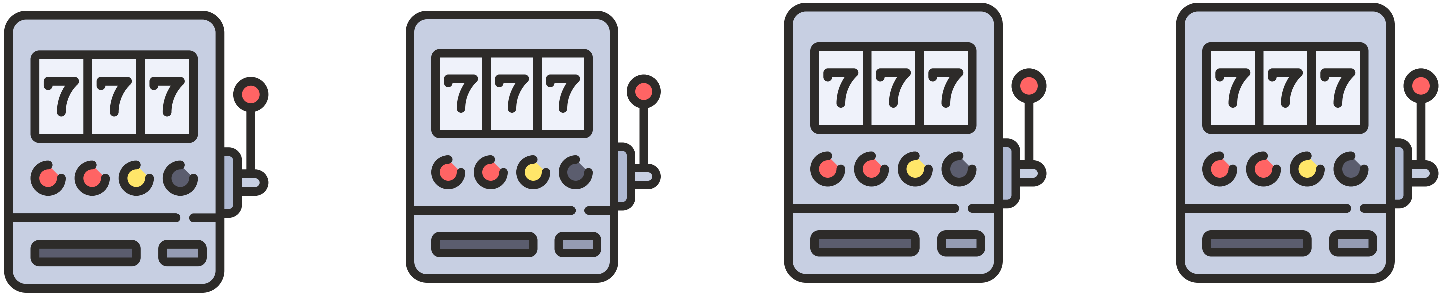
- Vanishing options tasks
  - Shin & Ariely (2004)
  - Ejova et al (2009)
  - Neth et al (2014)

"doors" problems

- Other related literature
  - Endowment effect (Kahneman & Tversky 1979)
  - RL models with prospect curves (e.g., Speekenbrink & Konstantinidis 2015)

Instantiation within a bandit task

image source: wikipedia



M1 M2 M3 M4

image source: flaticon

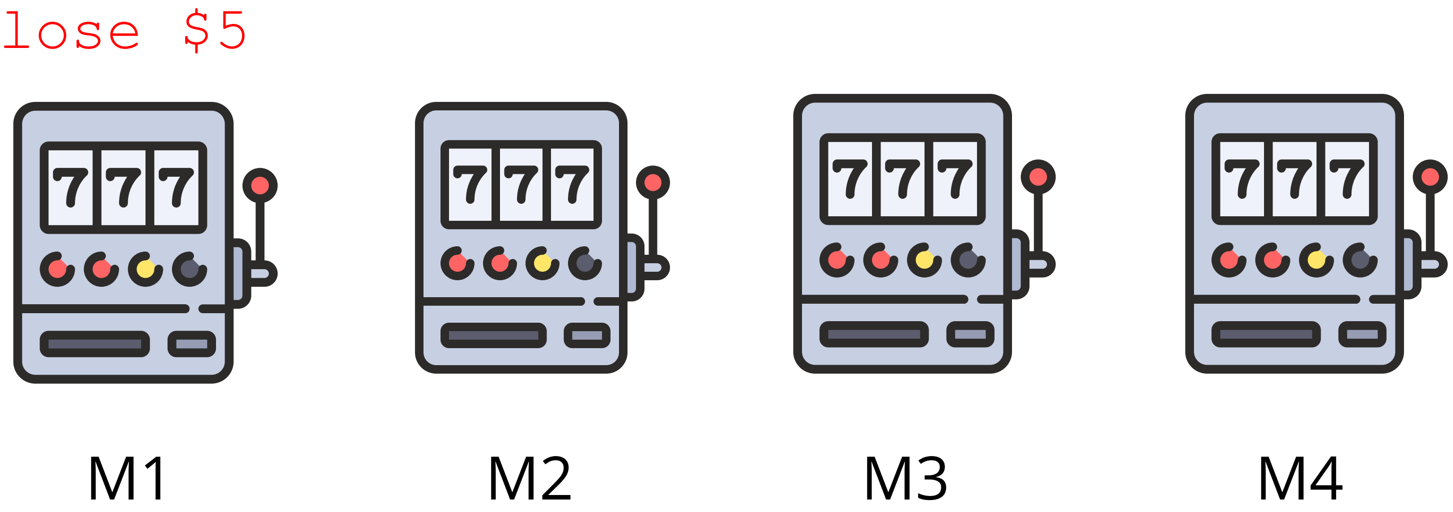


image source: flaticon

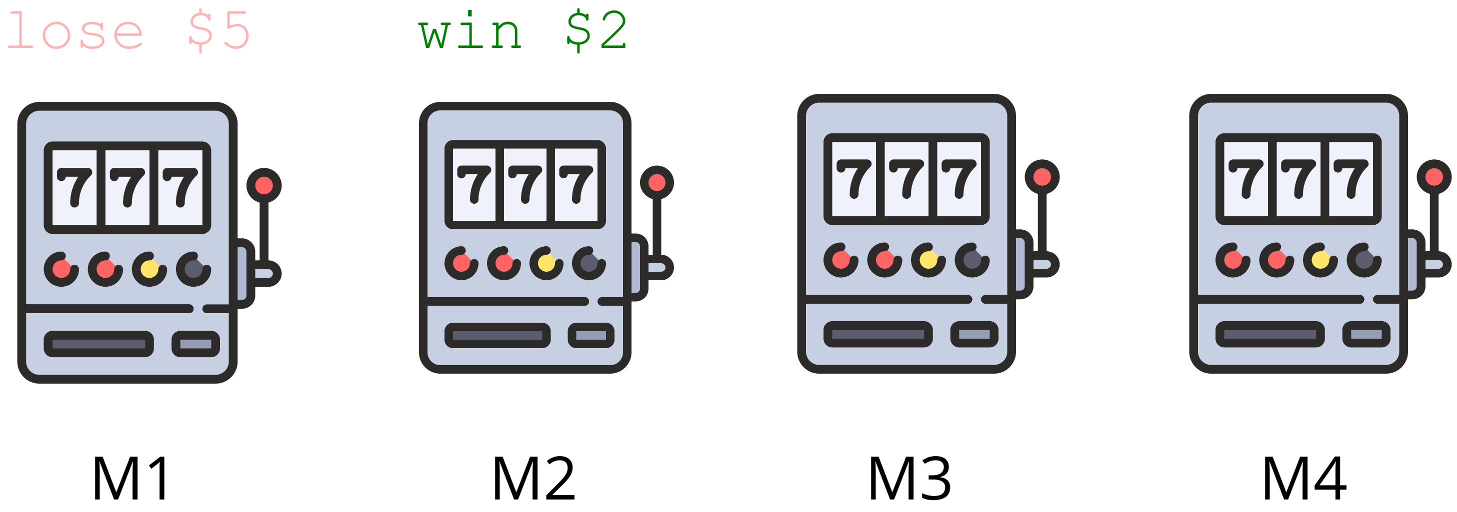


image source: flaticon

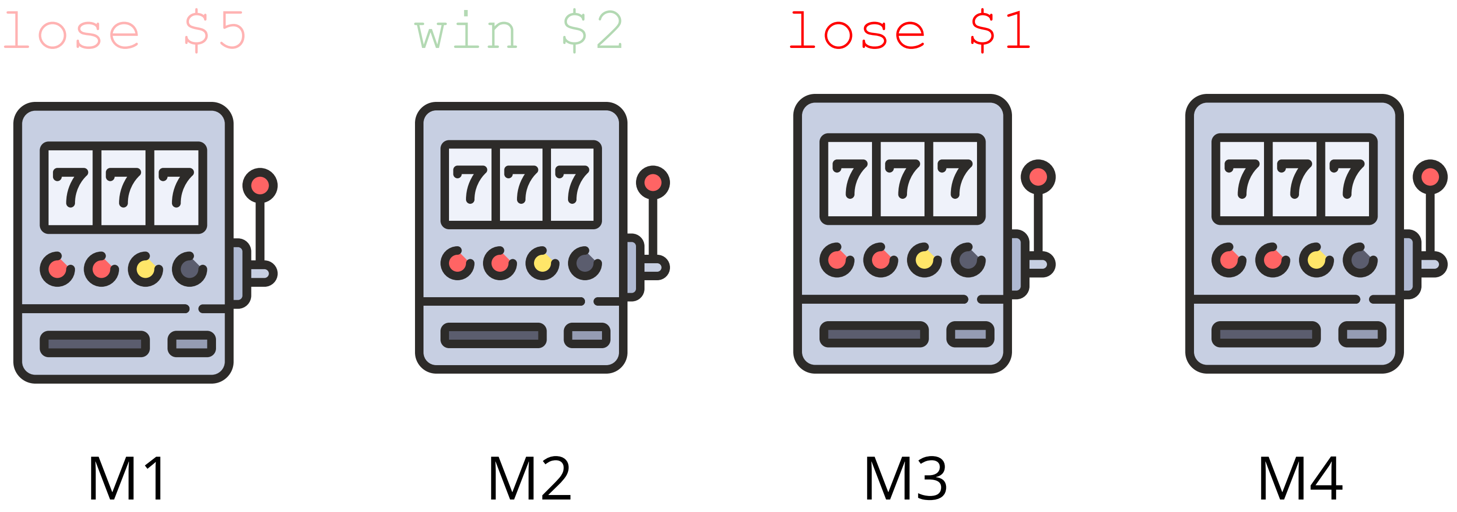
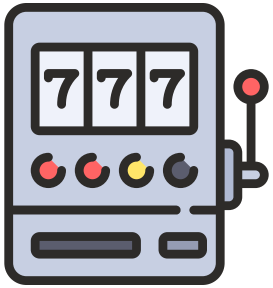
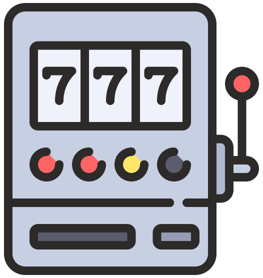
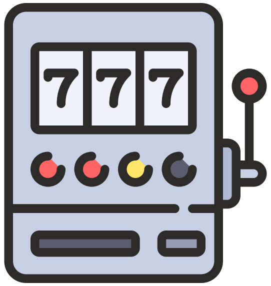
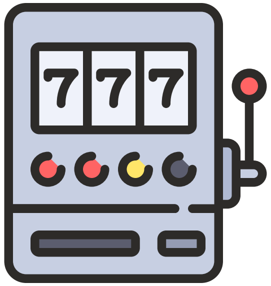
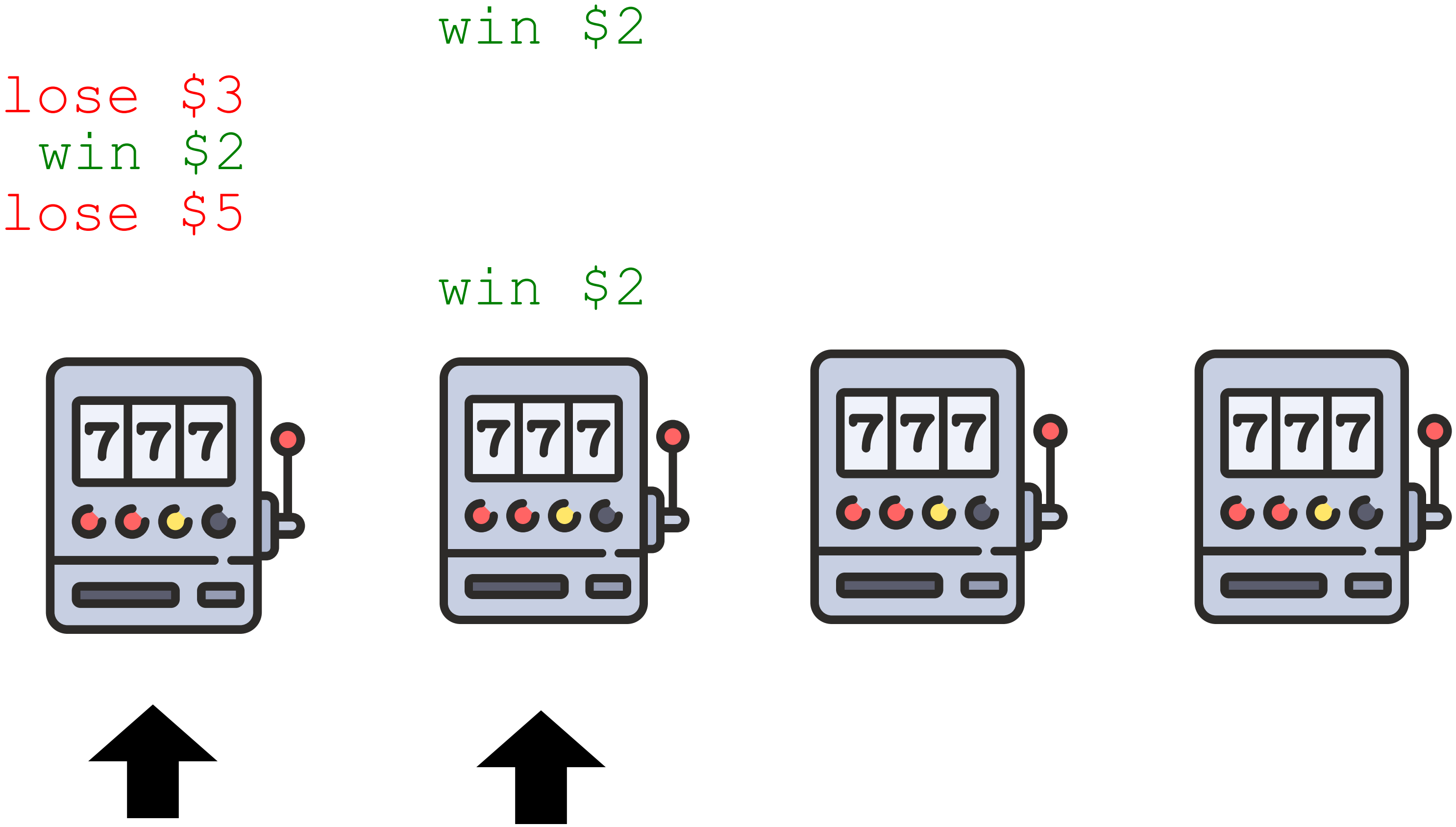


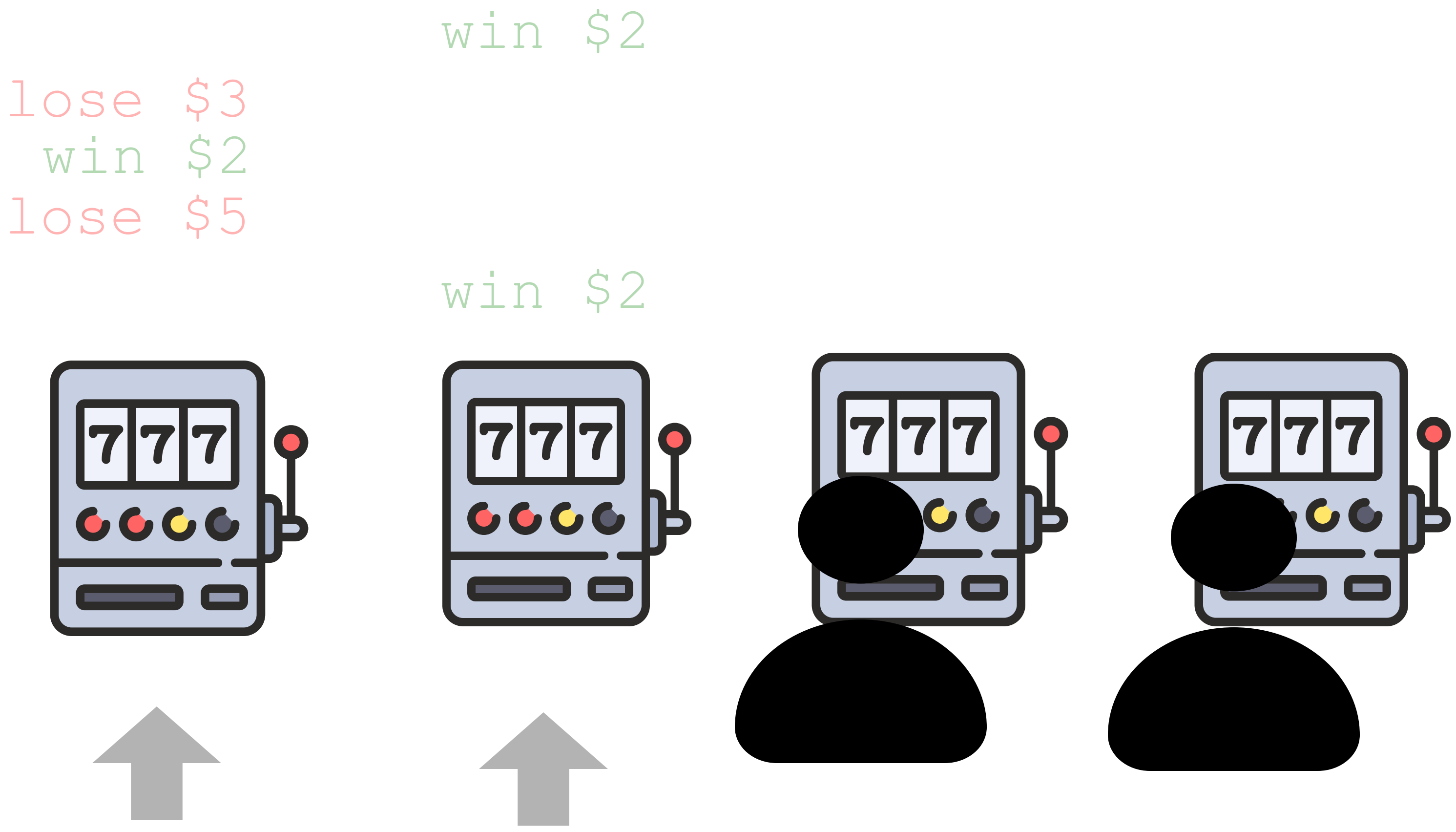
image source: flaticon

image source: flaticon


777

777

777

777



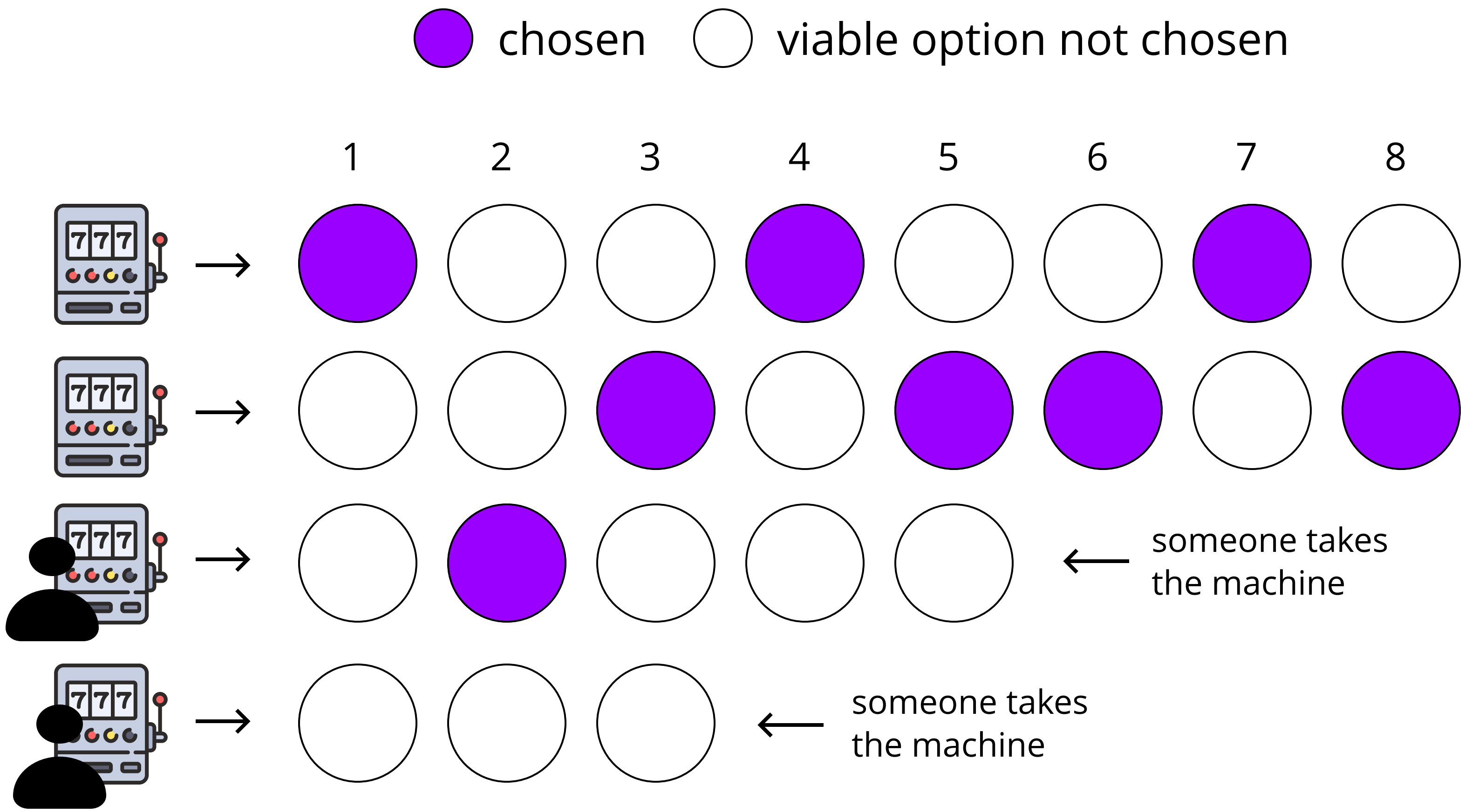
I've concentrated recent
bets on these machines



I've concentrated recent bets on these machines

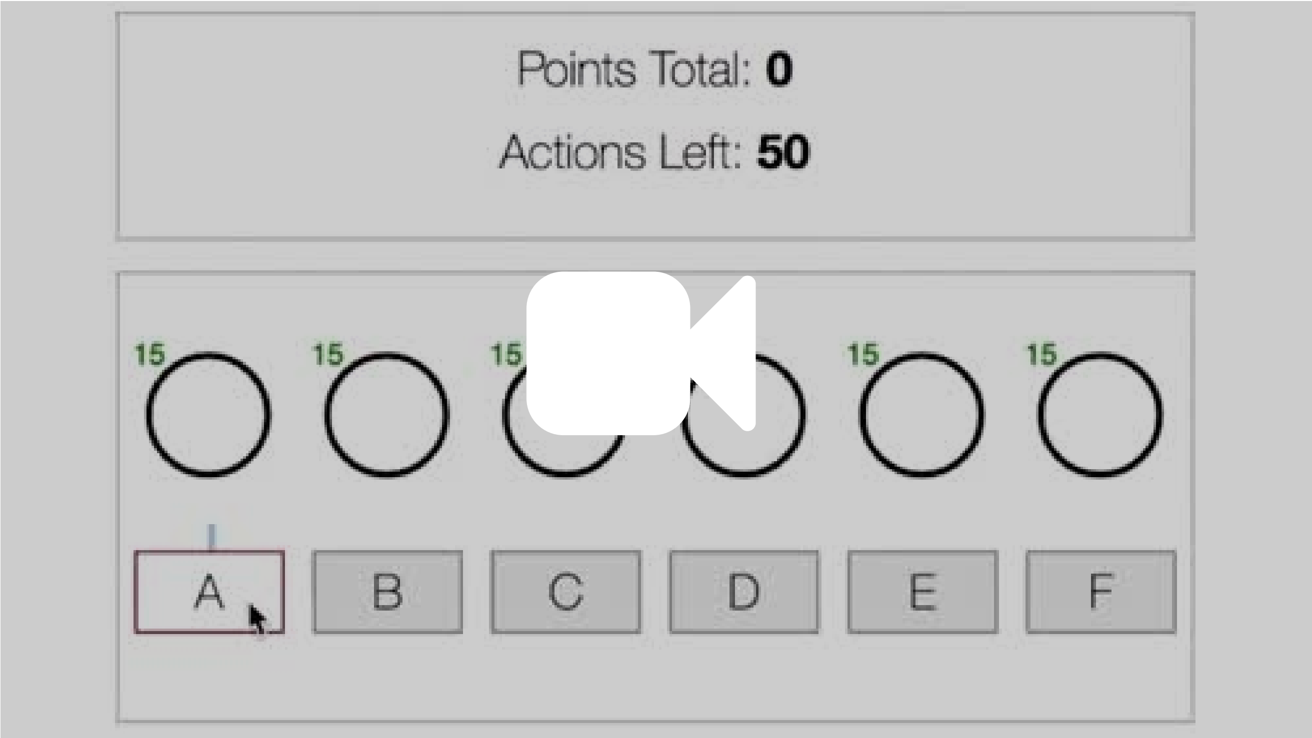
I've not used these machines recently, and someone else has taken them

RL approximation: options not pursued for N trials vanish



Experimental task

Experimental task



Method details

Method details

- Task:
  - Six armed bandit
  - Horizon: 50 trials (x3)
  - Feedback between games

Method details

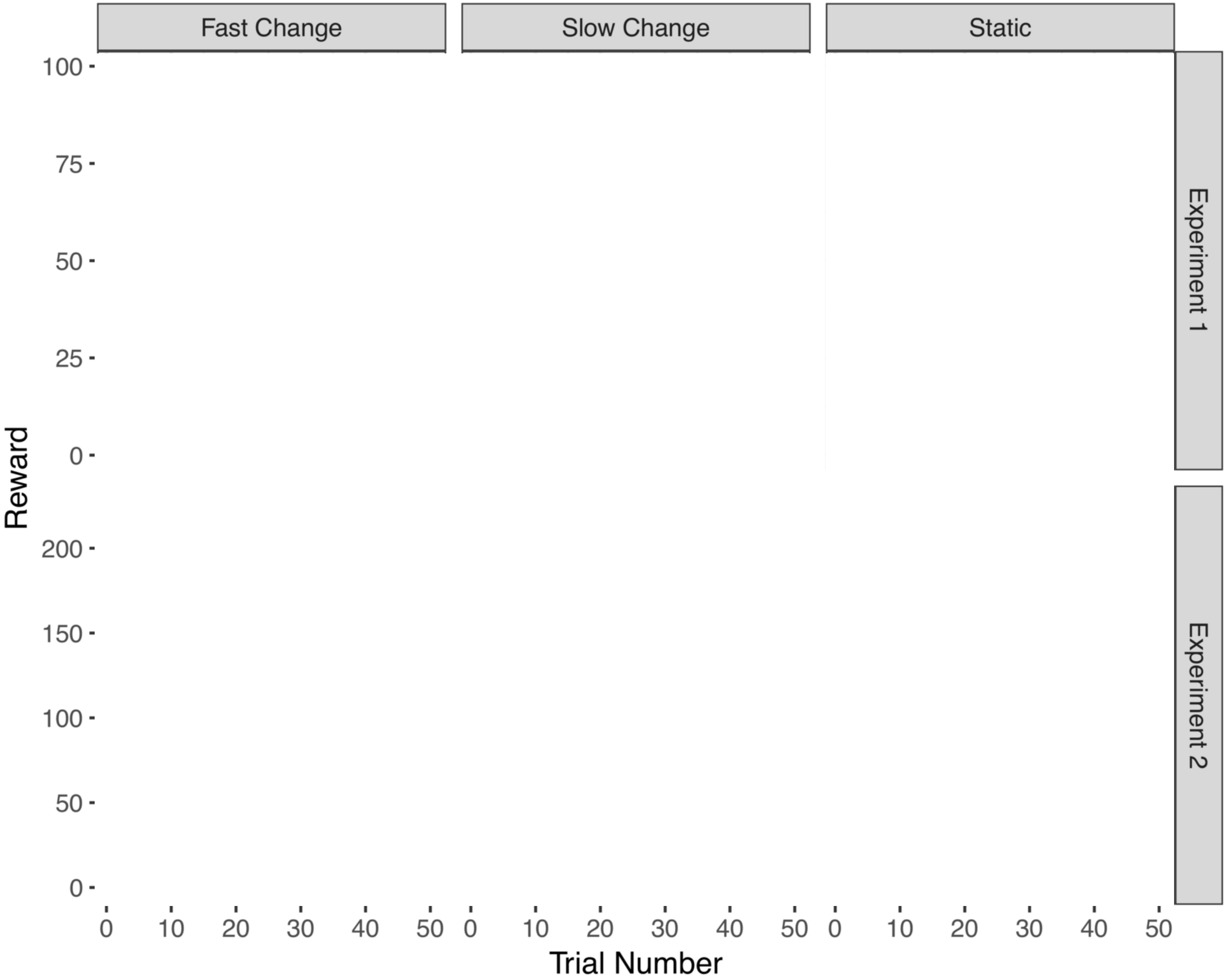
- Task:
  - Six armed bandit
  - Horizon: 50 trials (x3)
  - Feedback between games

- Other details:
  - Experiments run on Amazon Mechanical Turk
  - Expt 1: N = 400, Expt 2: N = 300, Pay: US$10/hr
  - Instructions had short "test" to check understanding

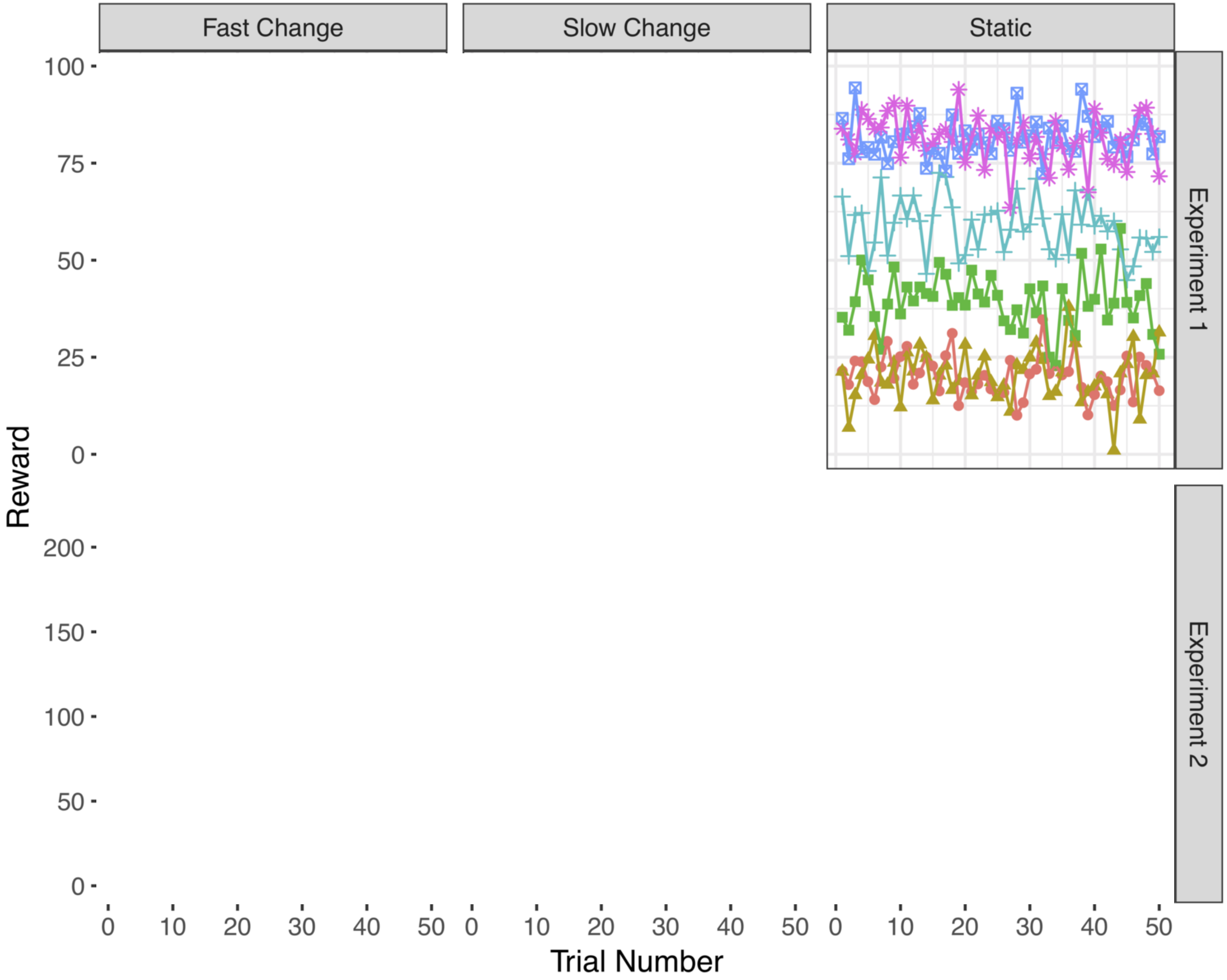
Method details

- Task:
  - Six armed bandit
  - Horizon: 50 trials (x3)
  - Feedback between games
- Manipulations:
  - Availability (const., threat)
  - Change (static, slow, fast)
  - Drift (none, biased)
- Other details:
  - Experiments run on Amazon Mechanical Turk
  - Expt 1: N = 400, Expt 2: N = 300, Pay: US$10/hr
  - Instructions had short "test" to check understanding

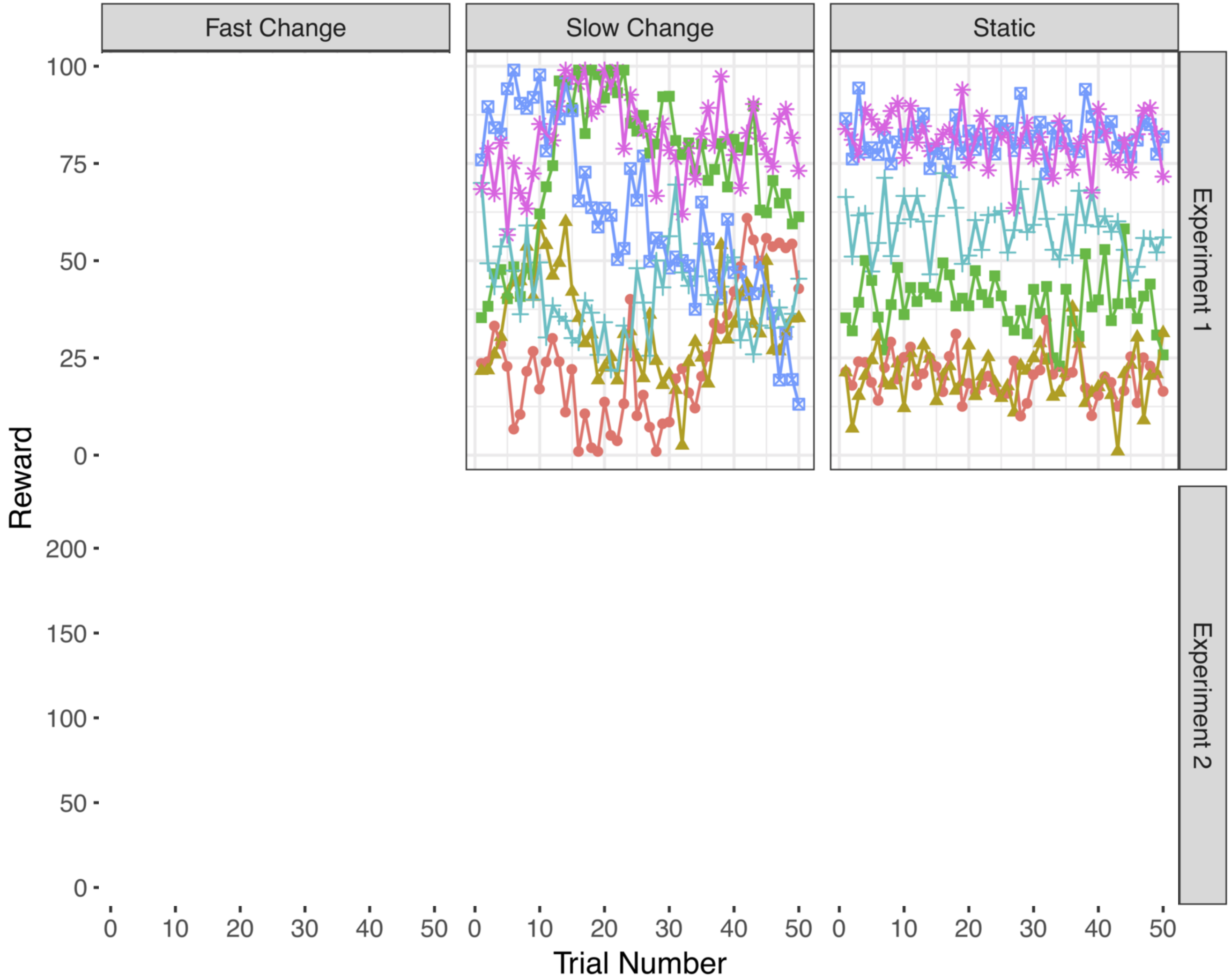
Environments



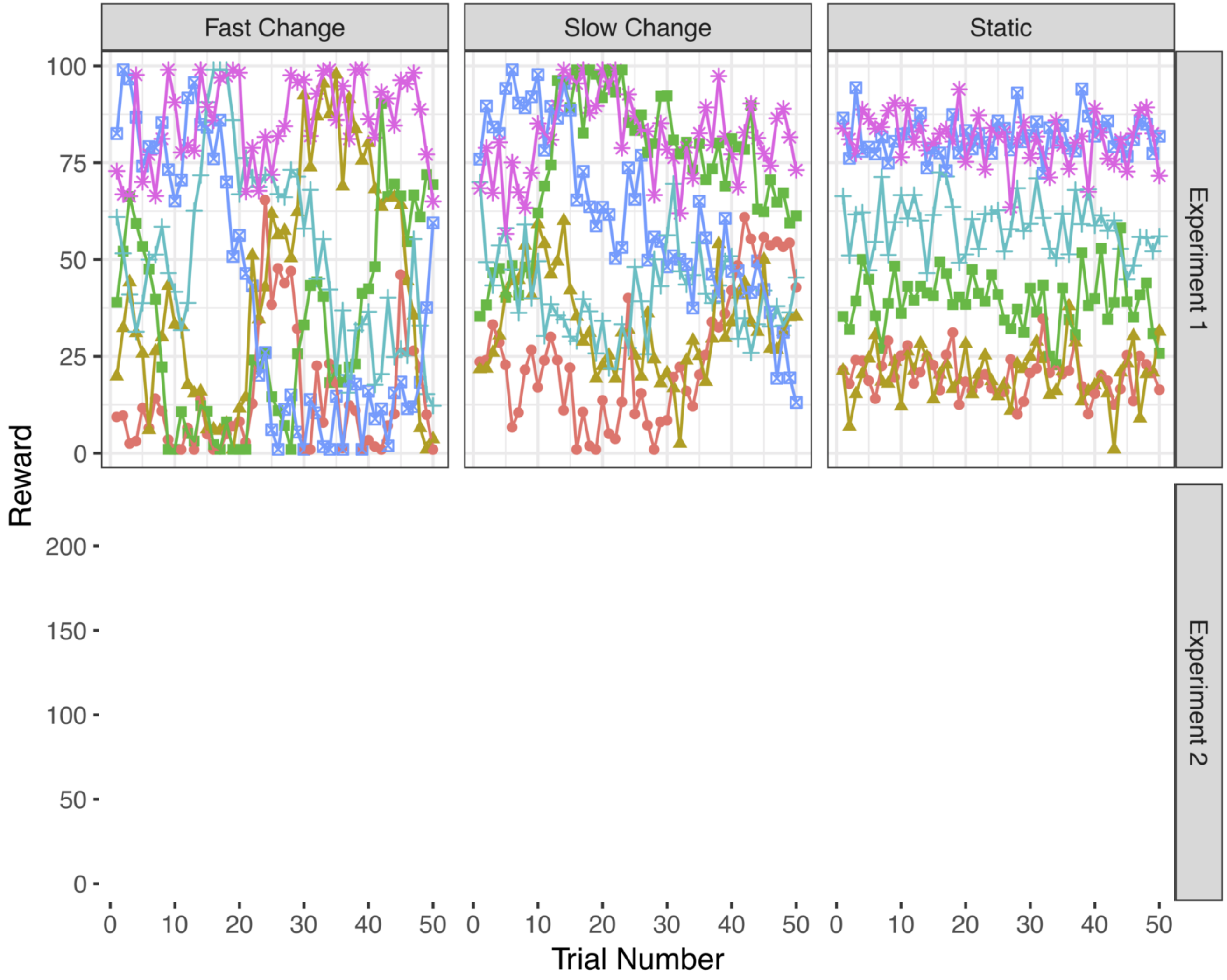
Environments



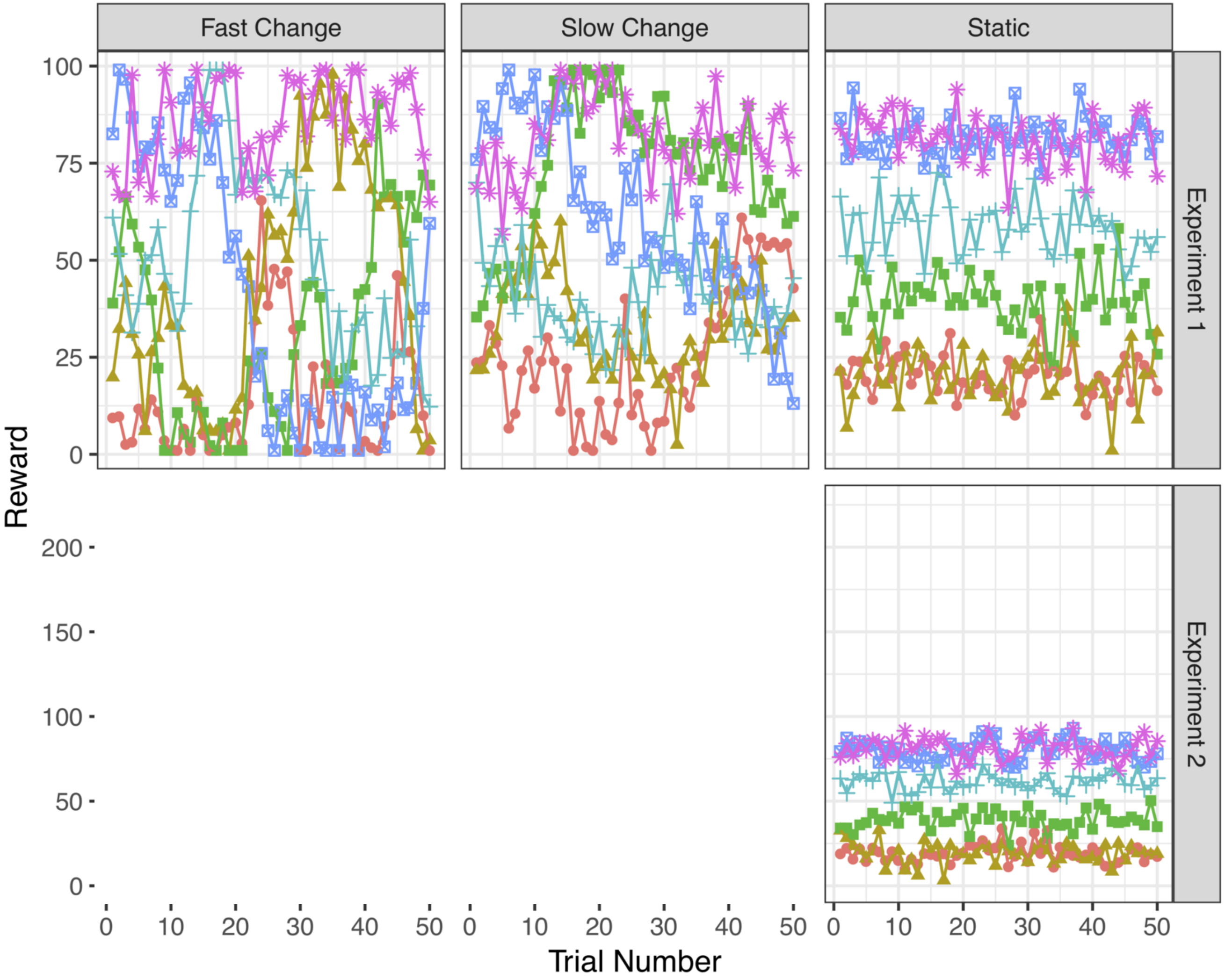
Environments



Environments

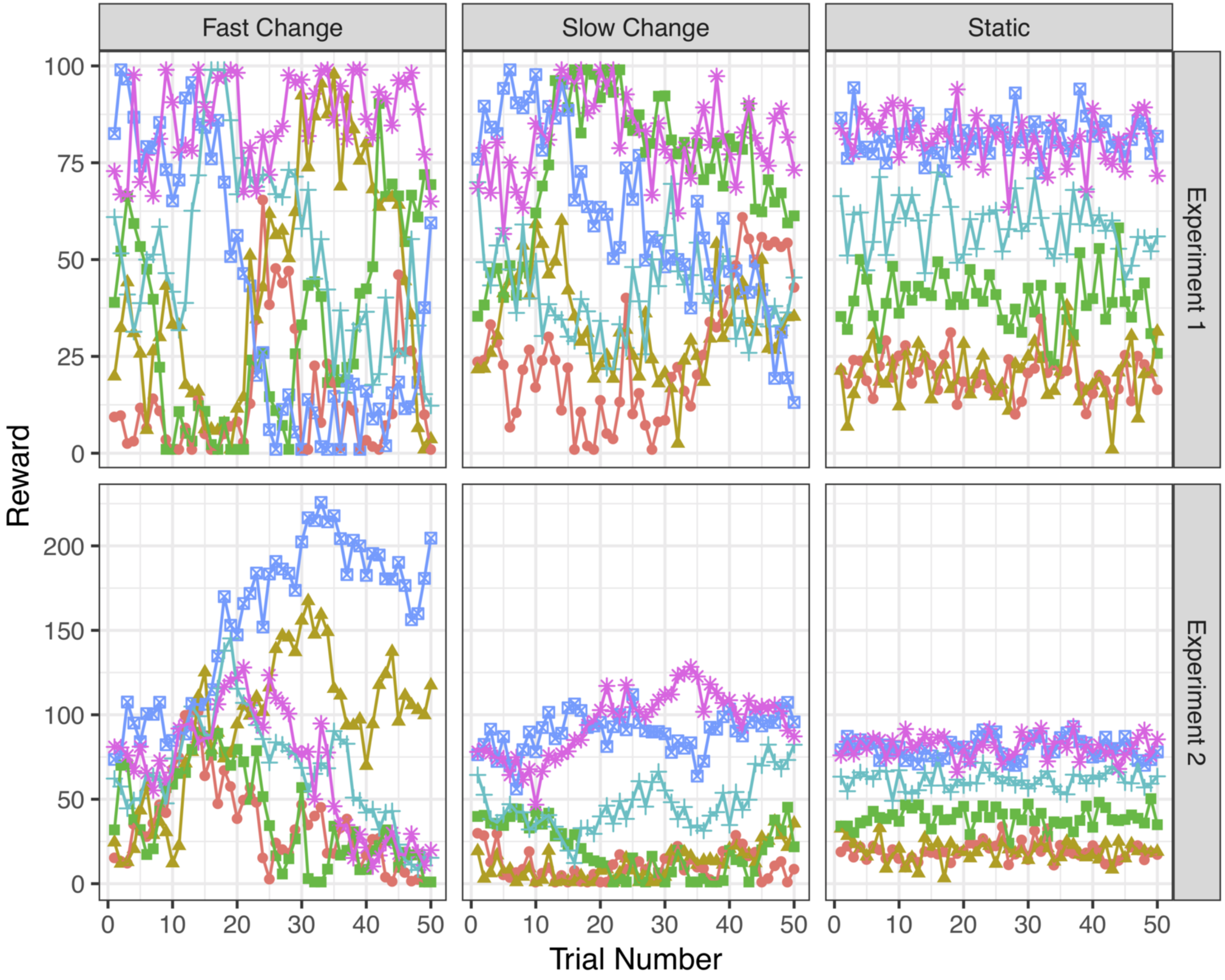


Environments



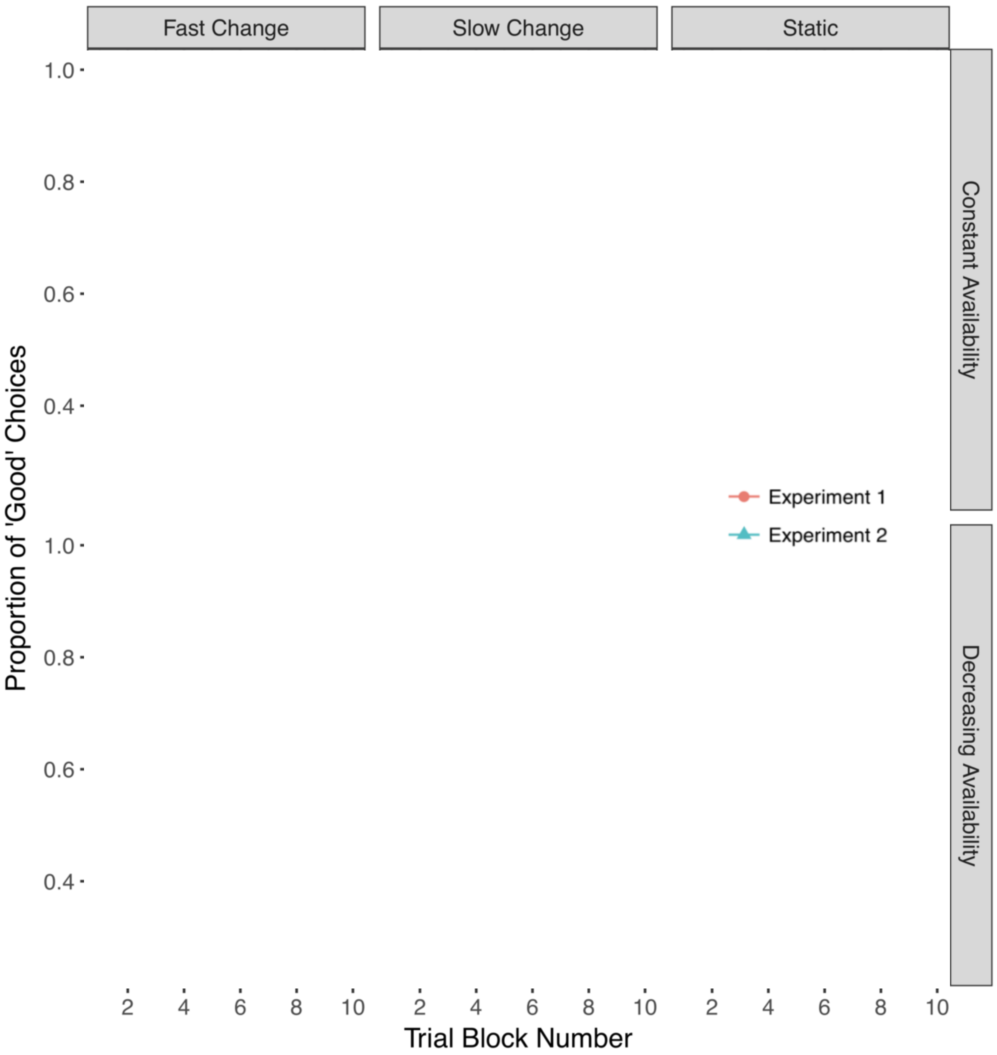
Environments

Environments

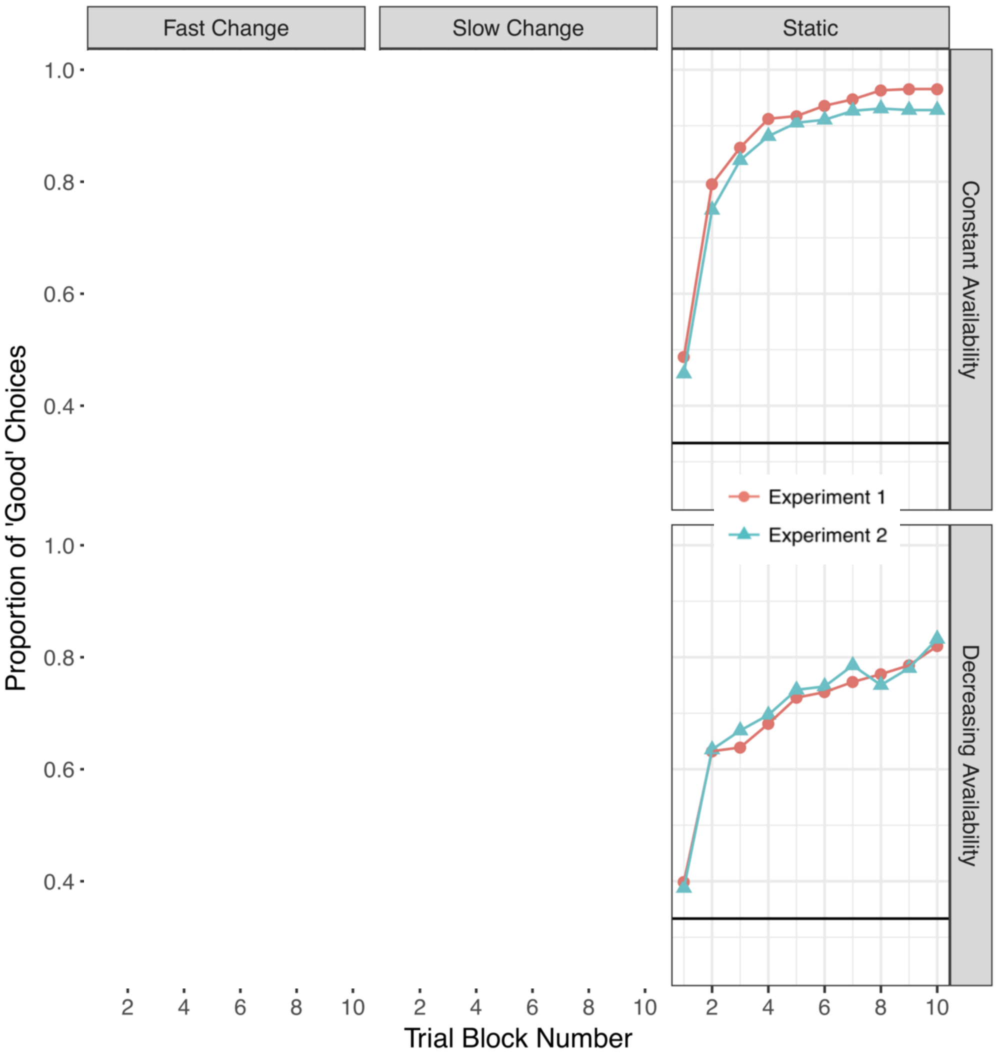


Results

Learning curves

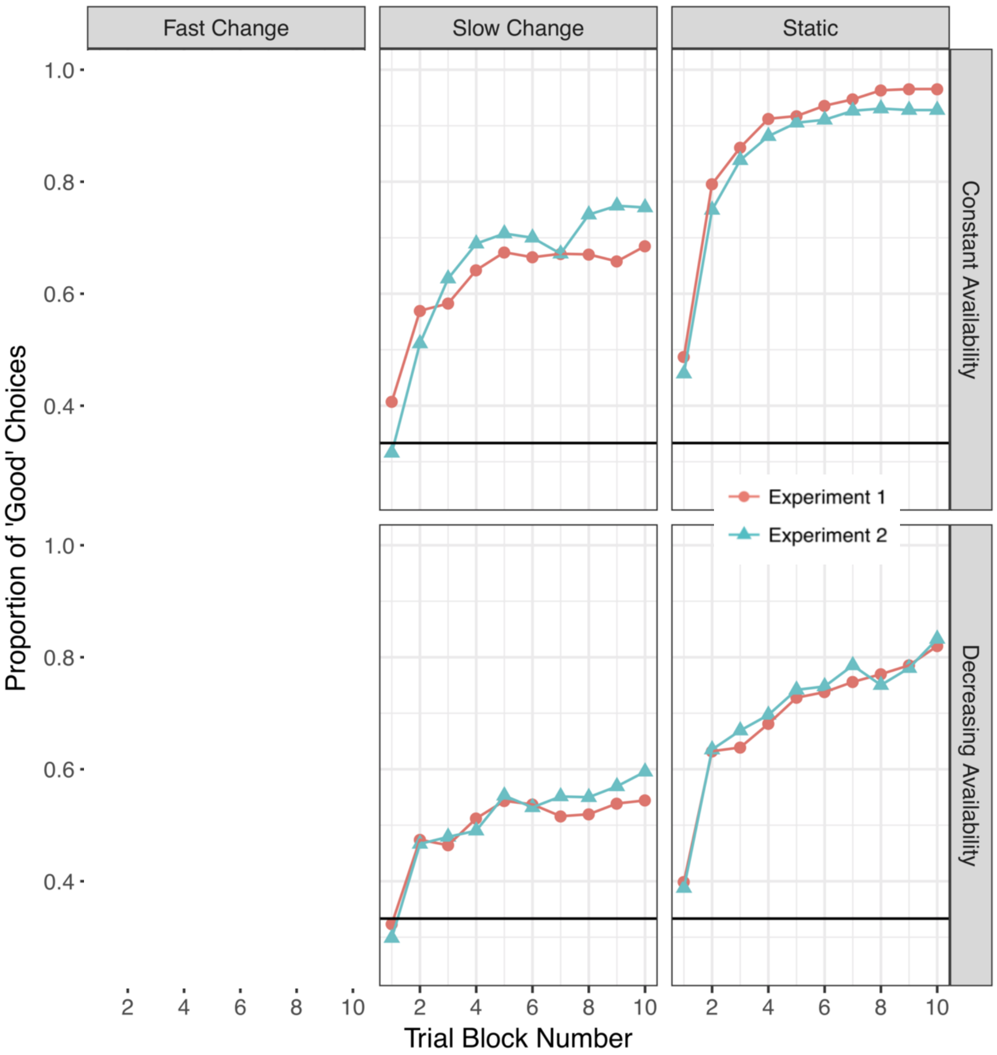


Learning curves



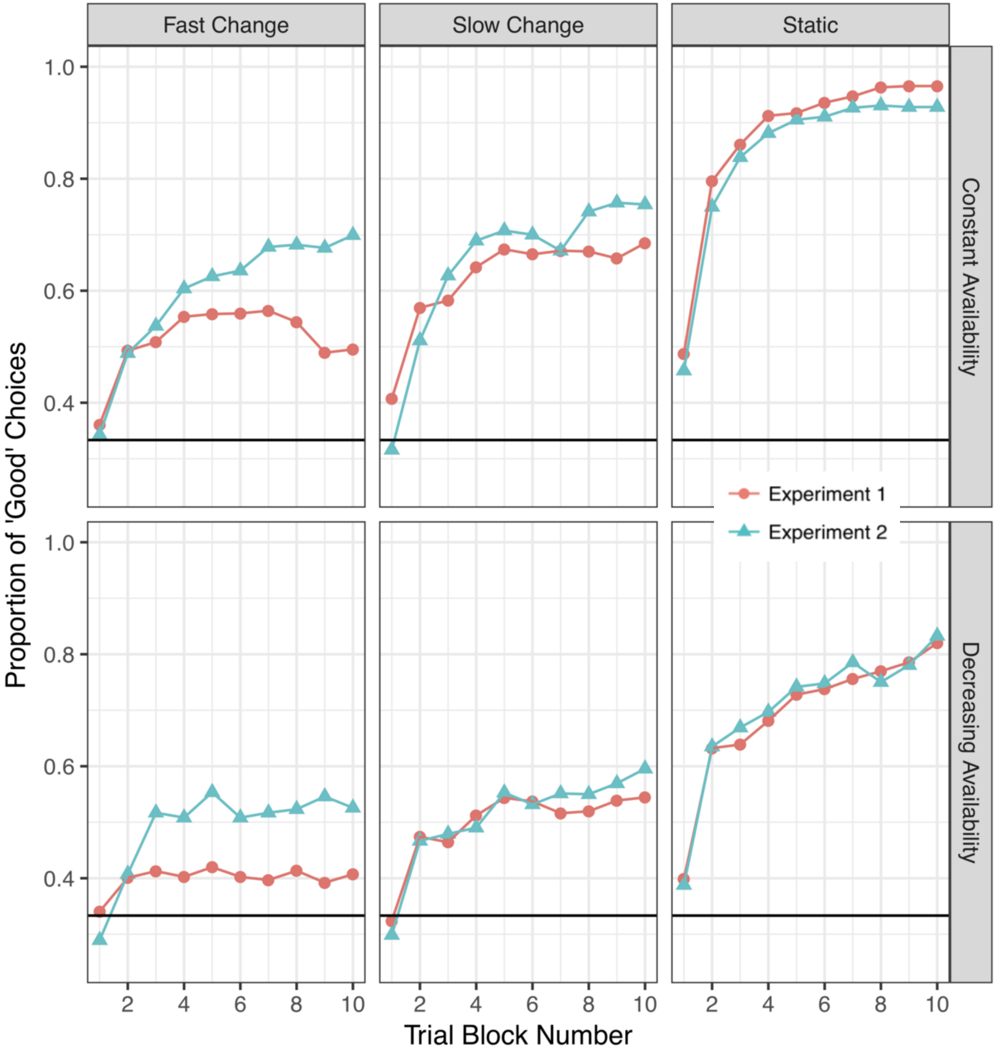
- Good = "top 2" option
- People learn quickly

Learning curves



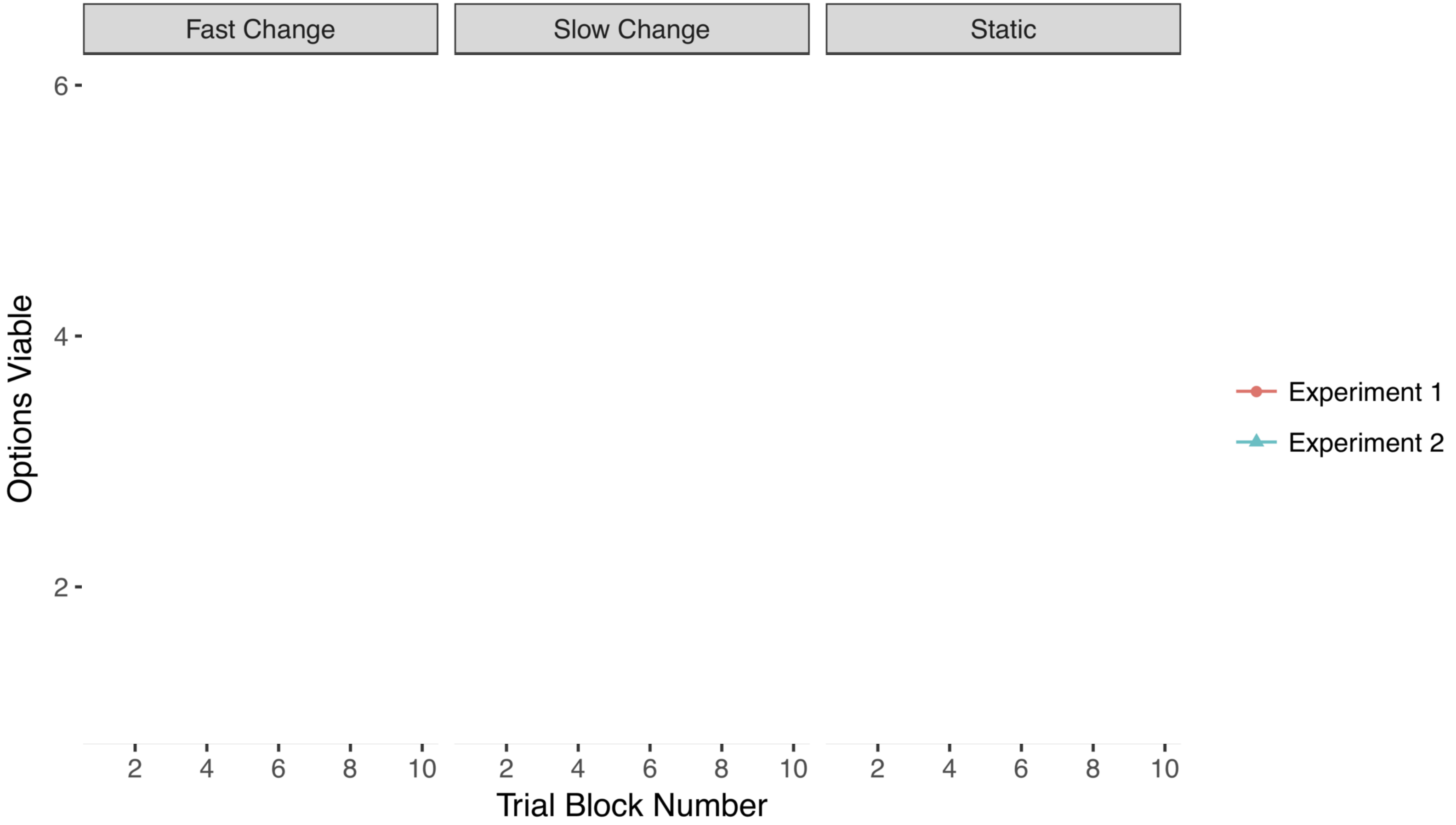
- Good = "top 2" option
- People learn quickly
- Fewer good choices when:
  - option threat exists
  - environment changes

Learning curves

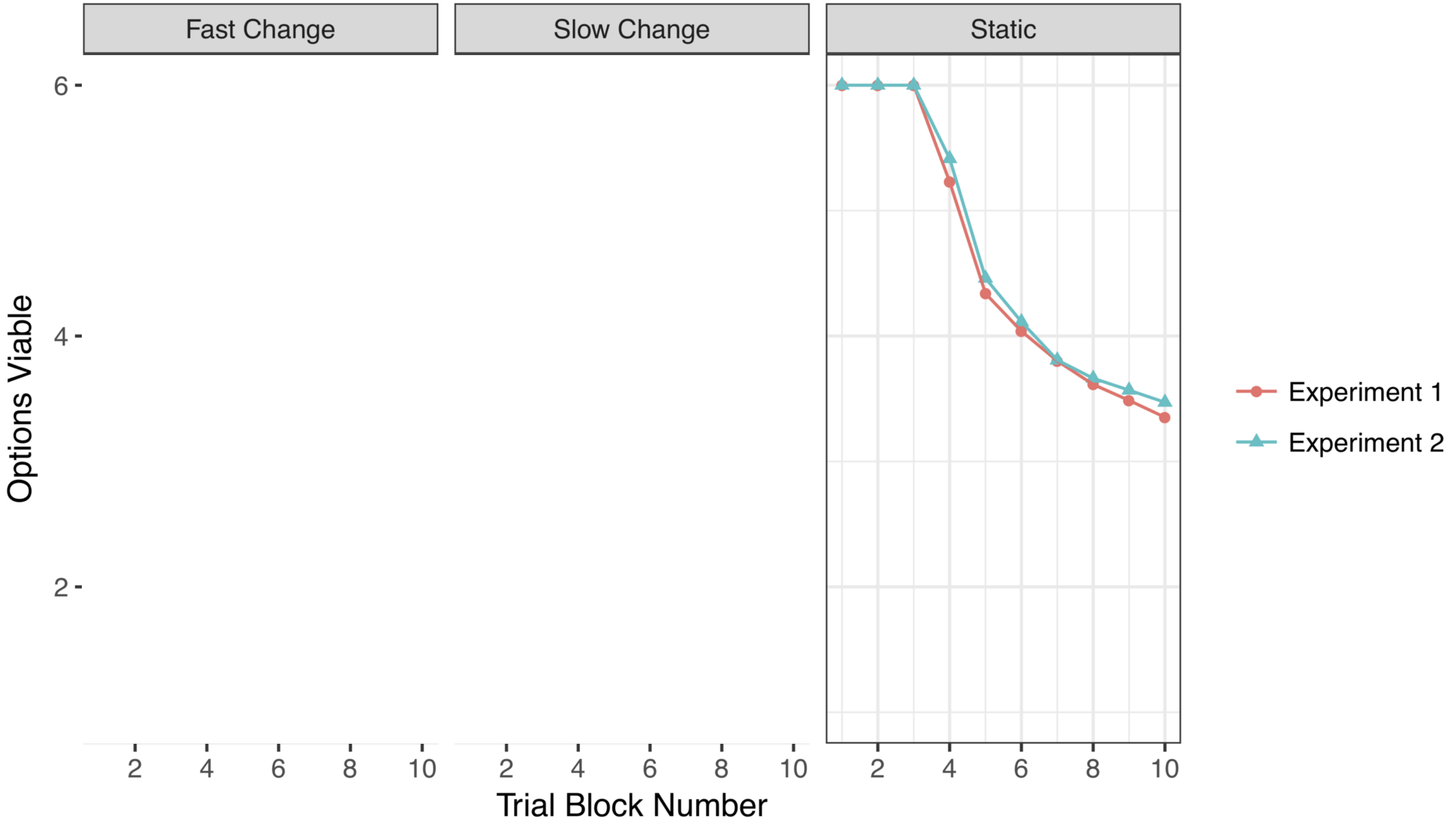


- Good = "top 2" option
- People learn quickly
- Fewer good choices when:
  - option threat exists
  - environment changes
- (Runaway winner effect in Exp 2 fast change)

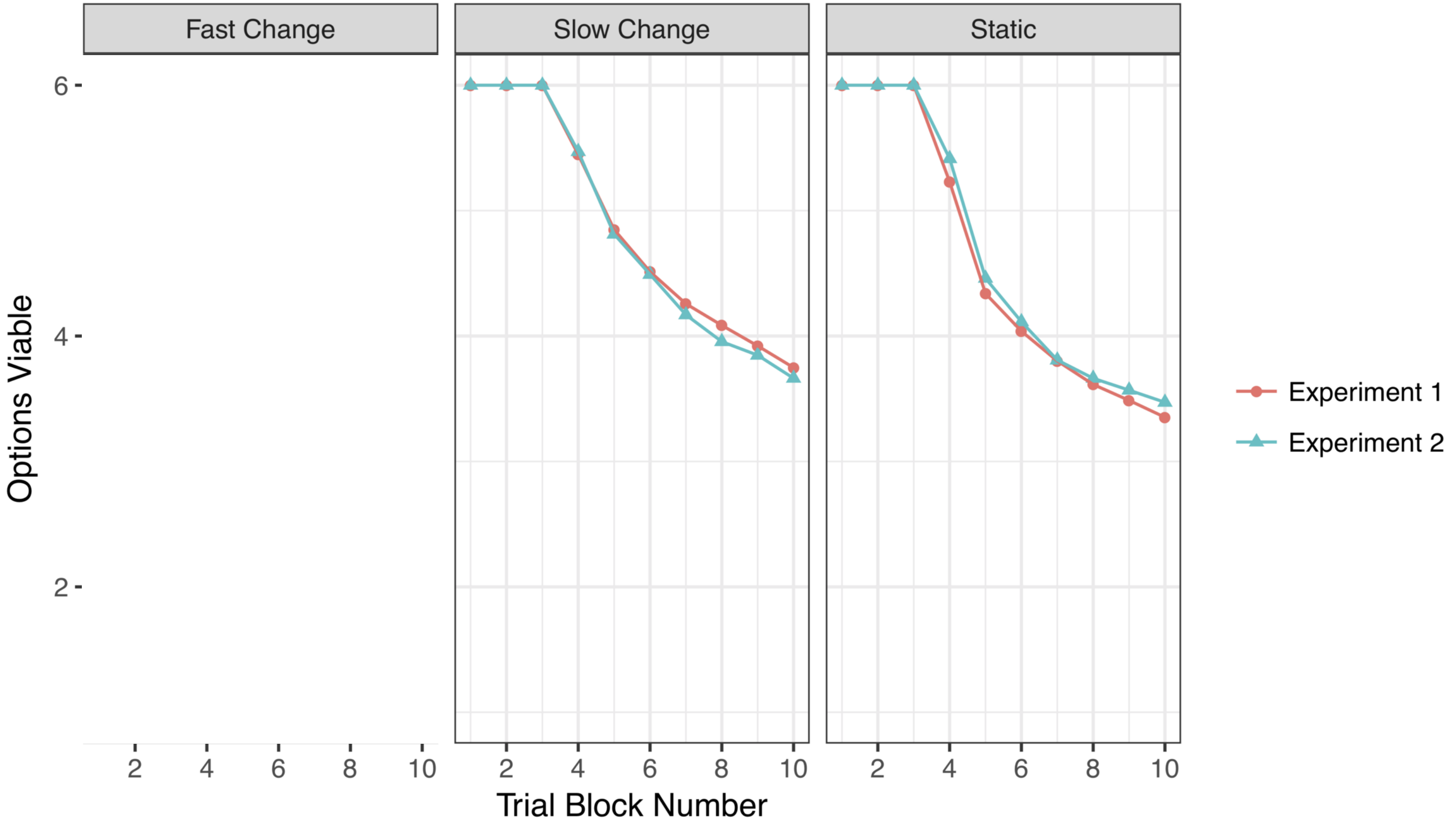
How many options do we keep?



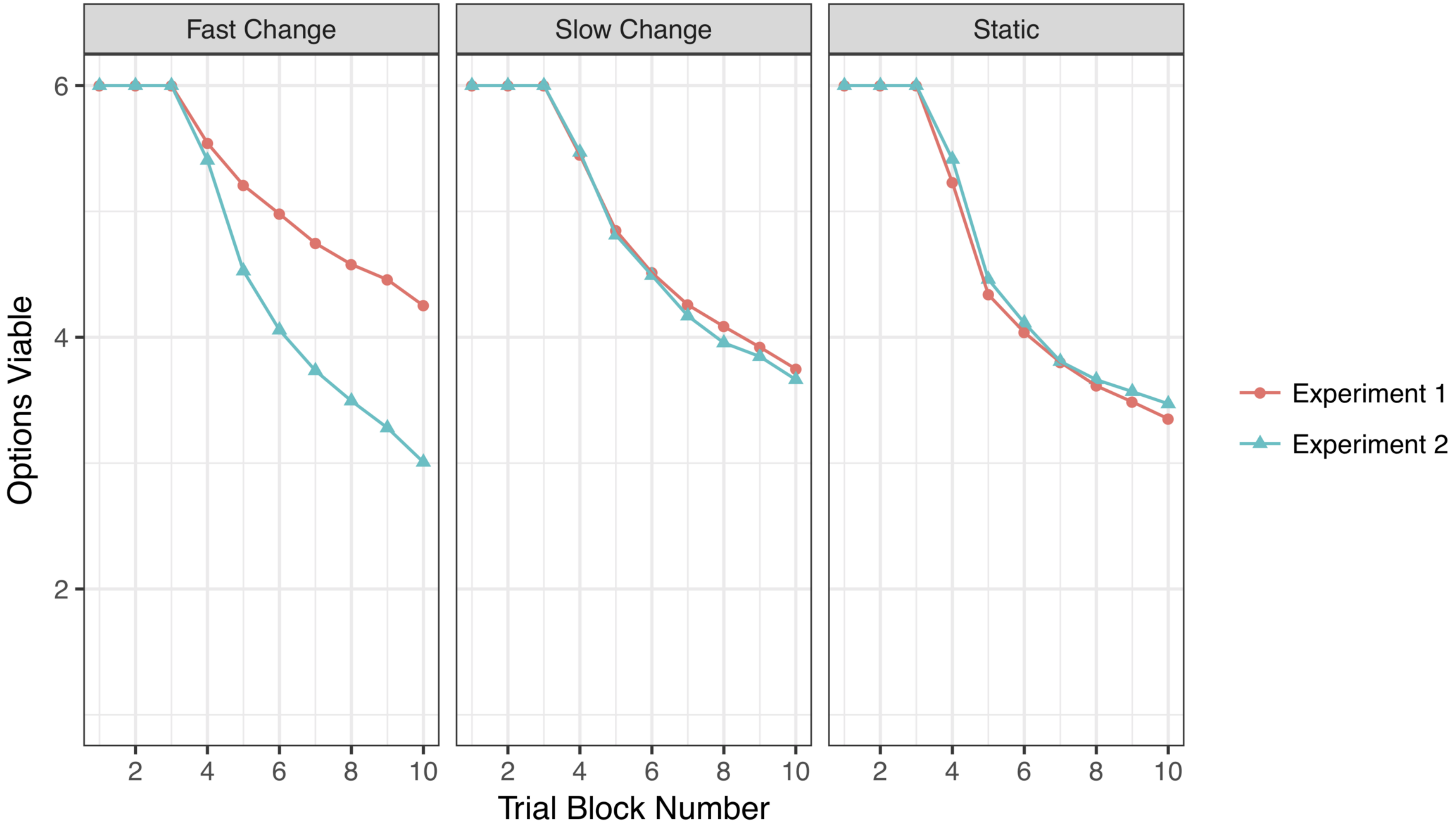
How many options do we keep?



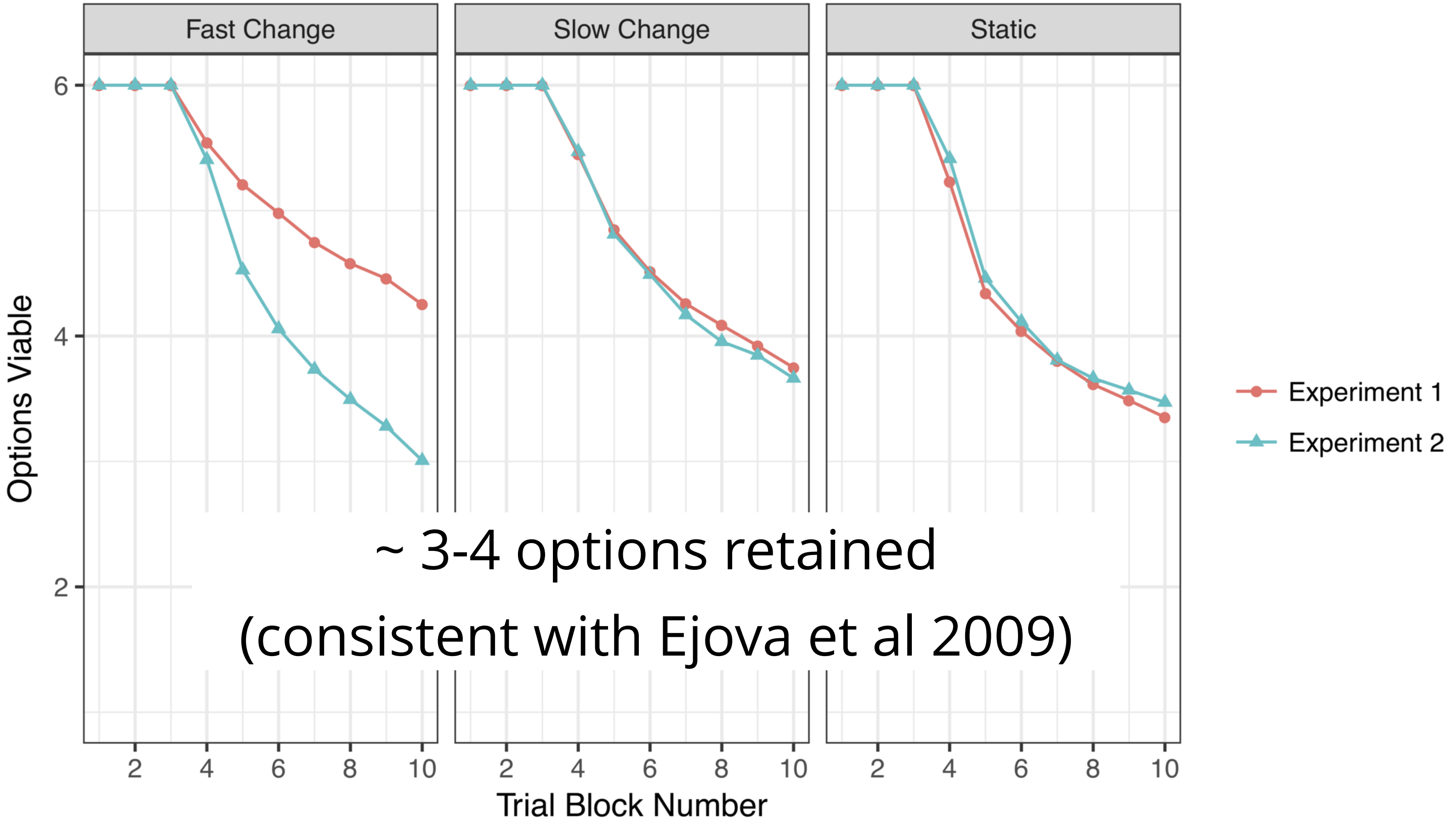
How many options do we keep?



How many options do we keep?



How many options do we keep?

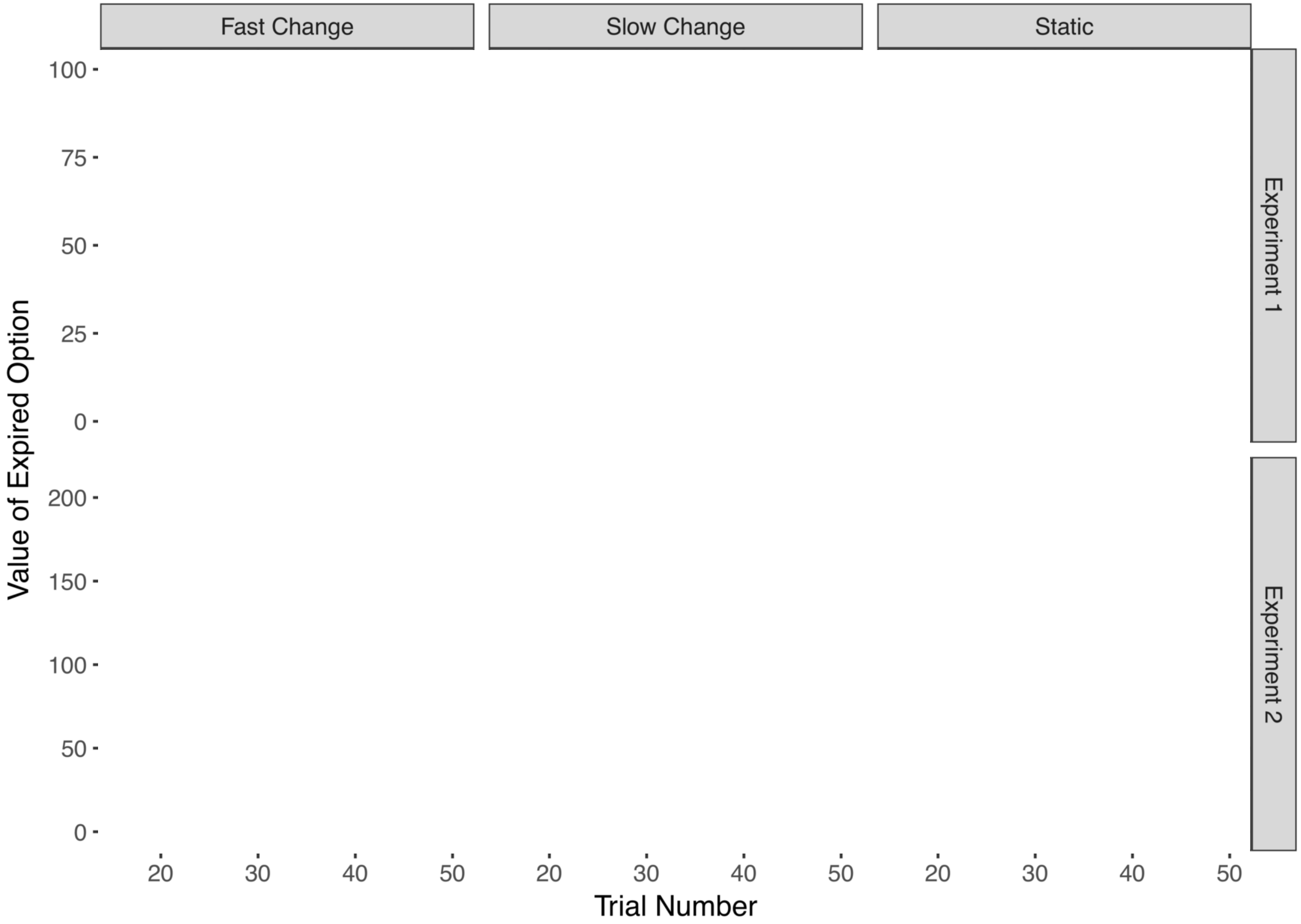


~ 3-4 options retained

(consistent with Ejova et al 2009)

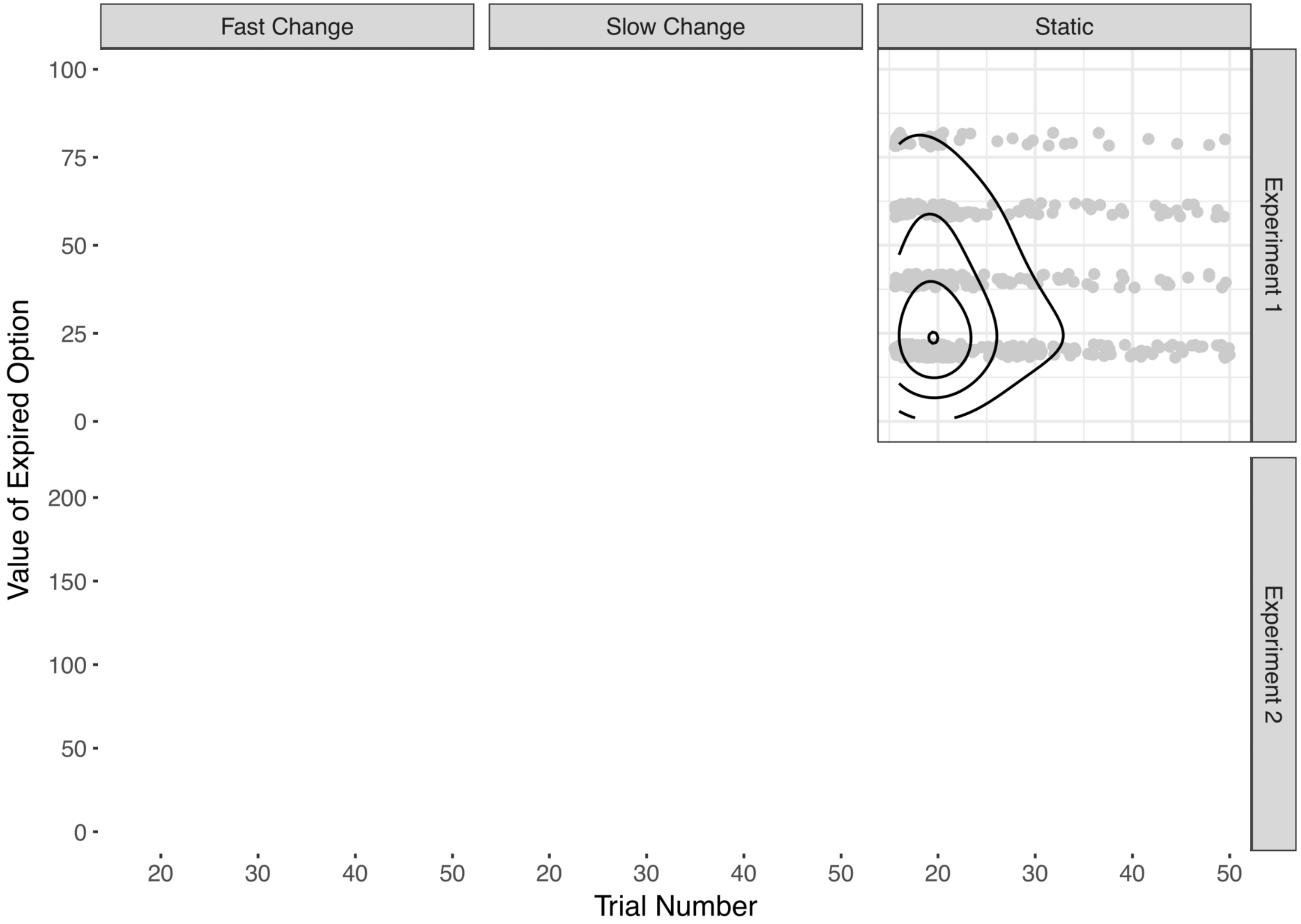
When do we let options go?

(an unapologetically exploratory analysis!)



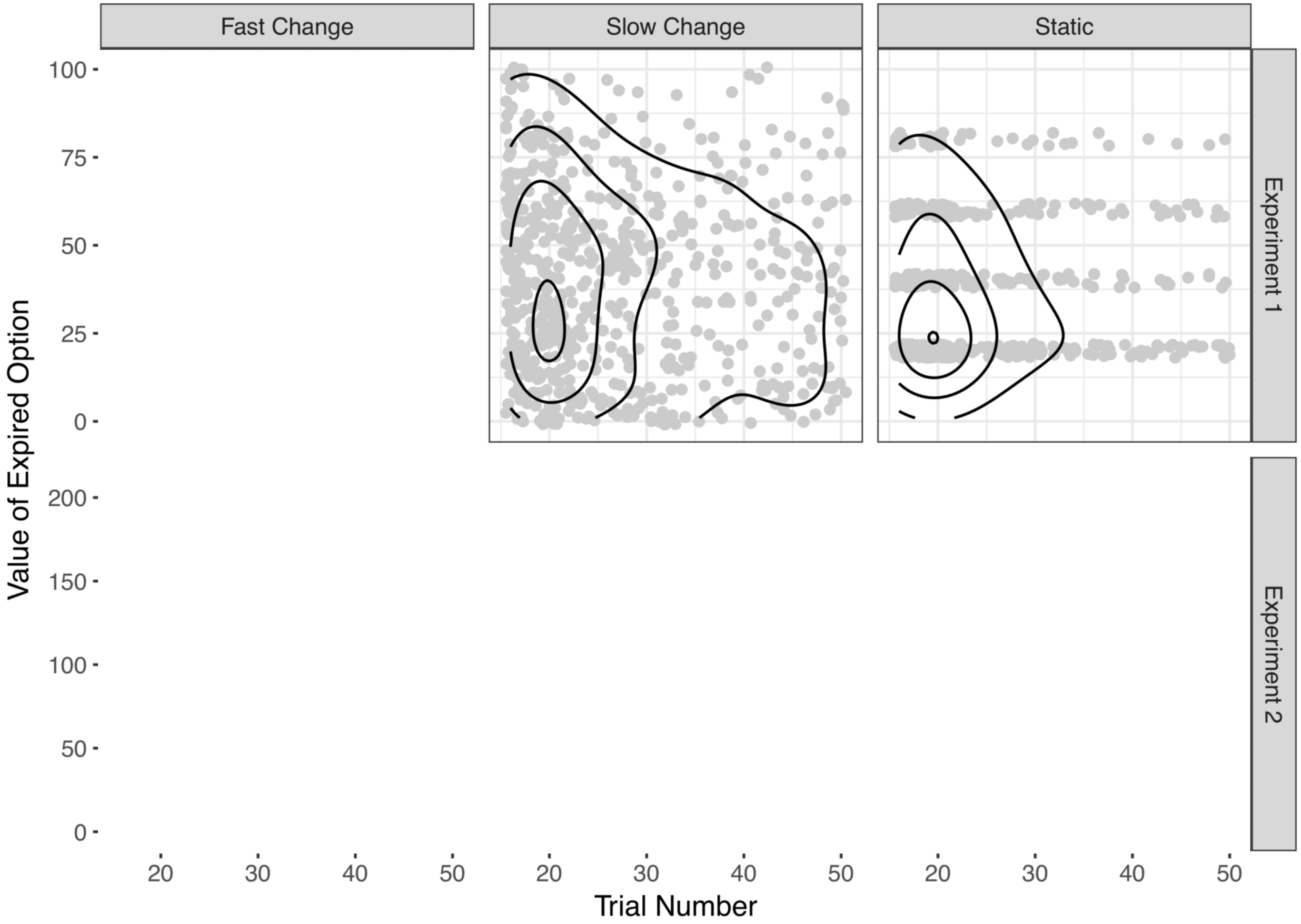
When do we let options go?

(an unapologetically exploratory analysis!)



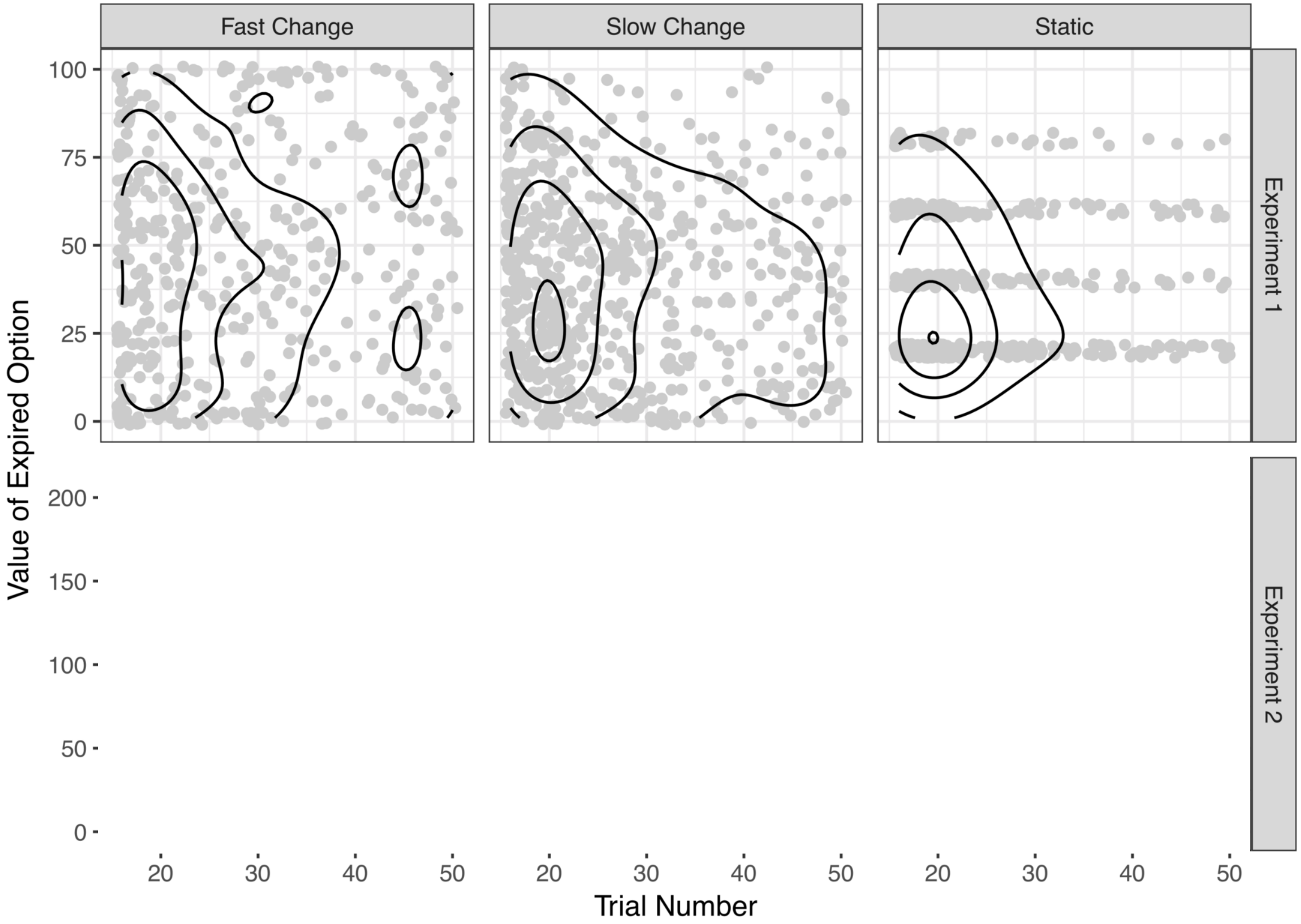
When do we let options go?

(an unapologetically exploratory analysis!)



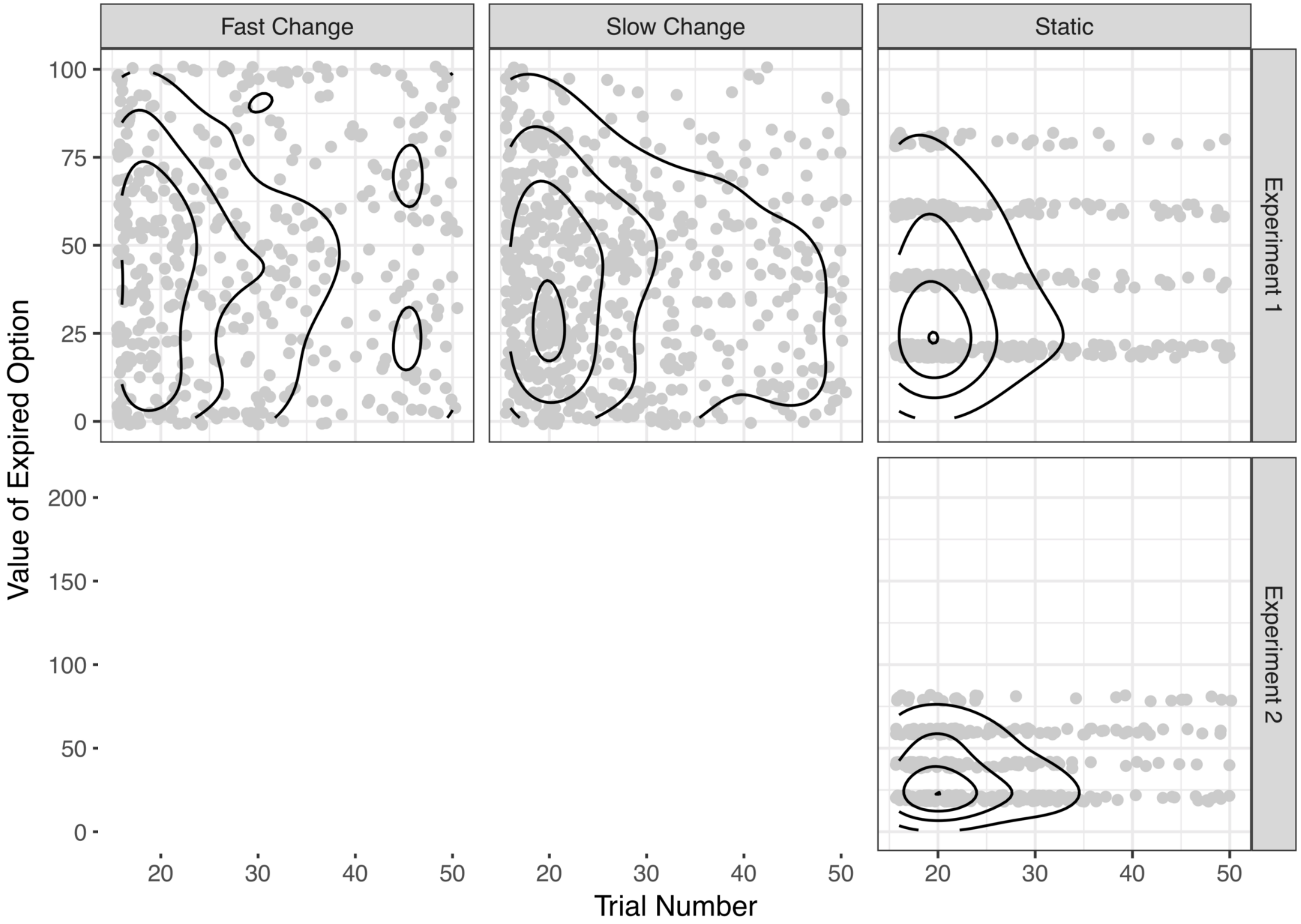
When do we let options go?

(an unapologetically exploratory analysis!)



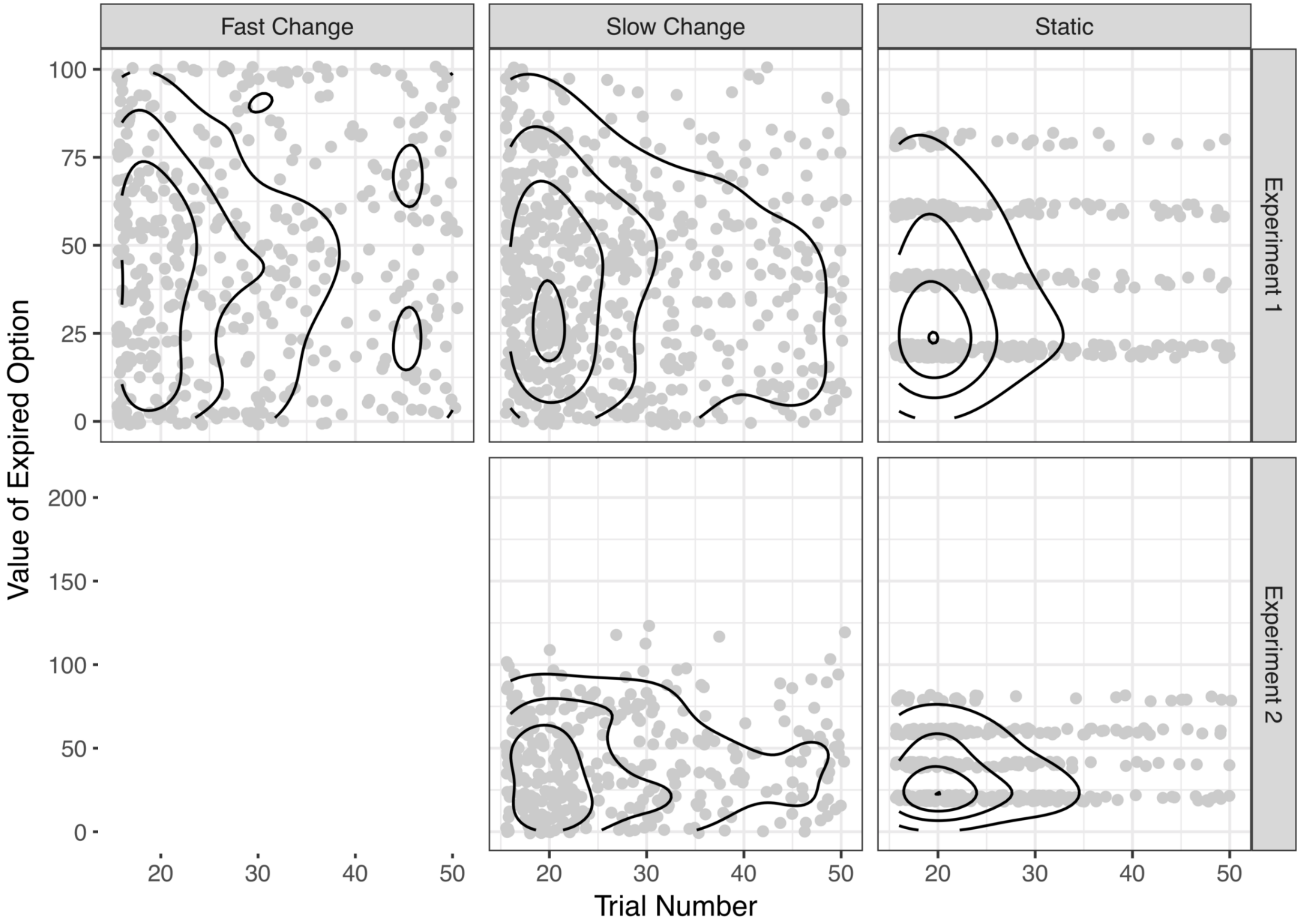
When do we let options go?

(an unapologetically exploratory analysis!)



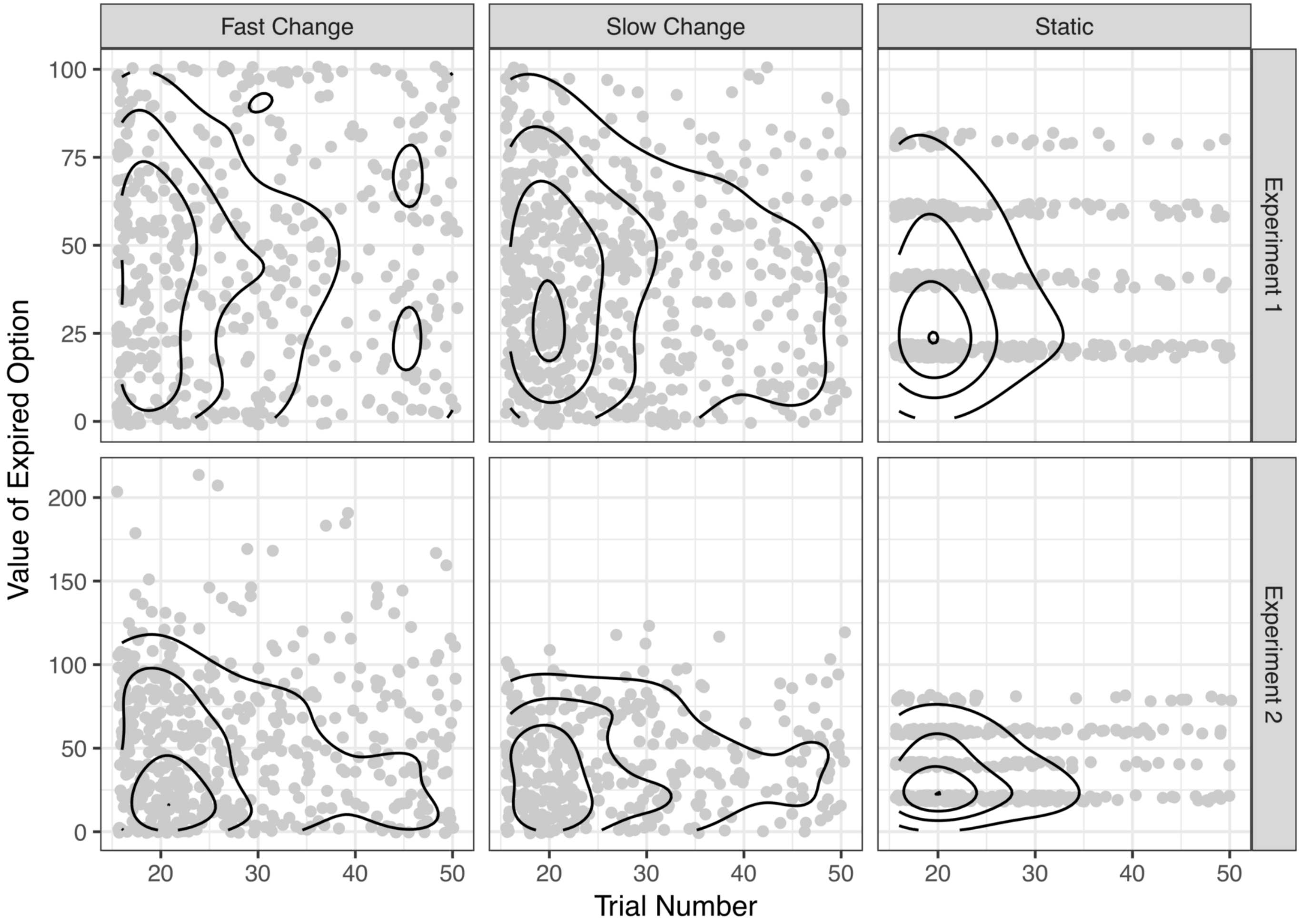
When do we let options go?

(an unapologetically exploratory analysis!)

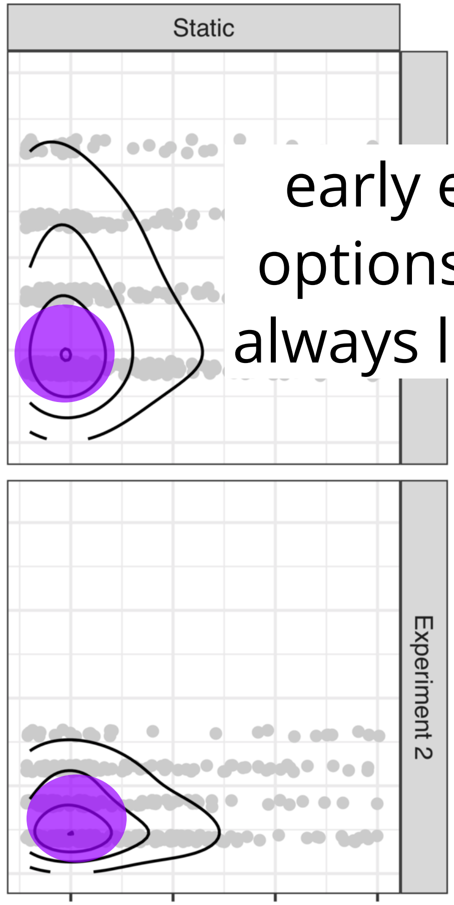


When do we let options go?

(an unapologetically exploratory analysis!)

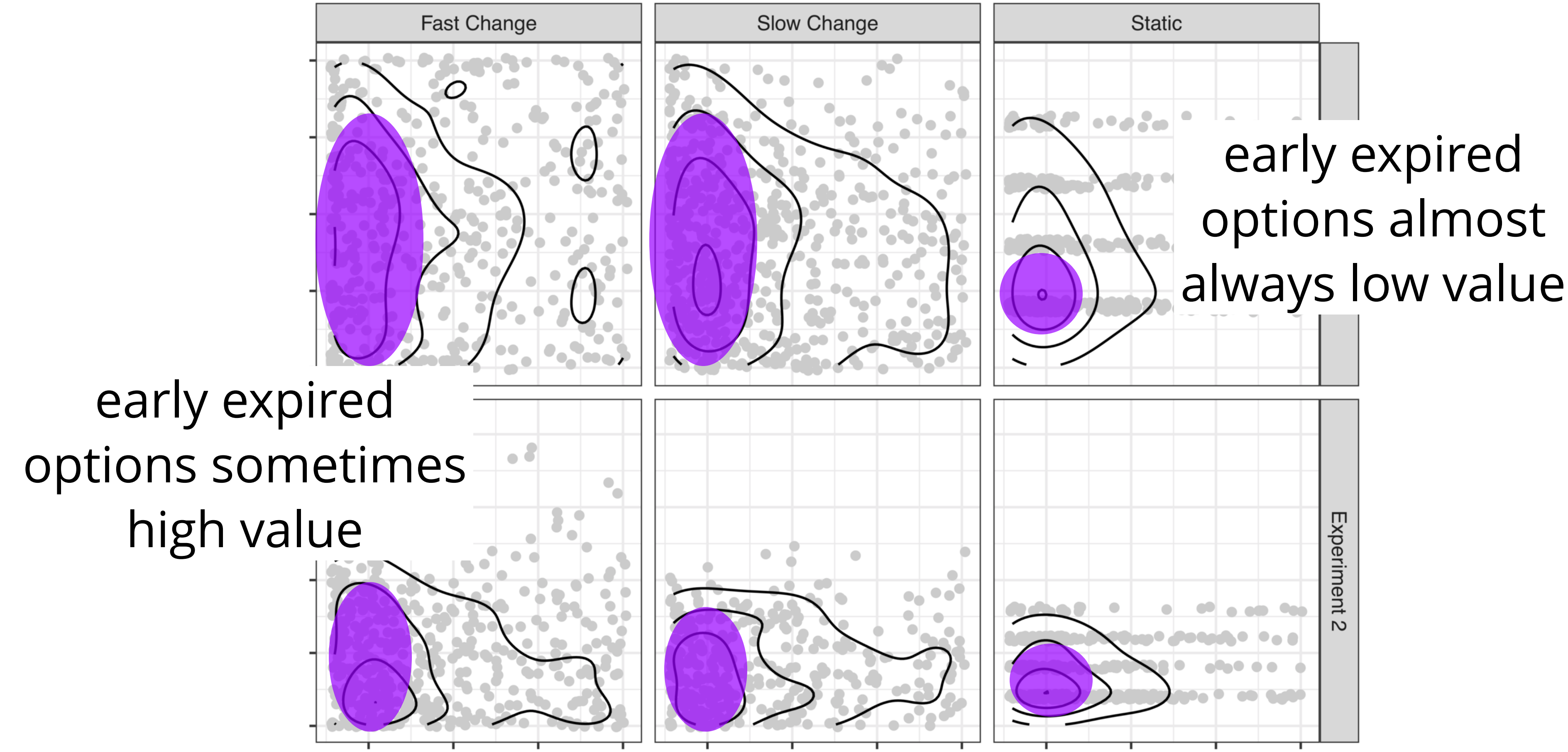


1. Poorer discrimination in volatile environments?

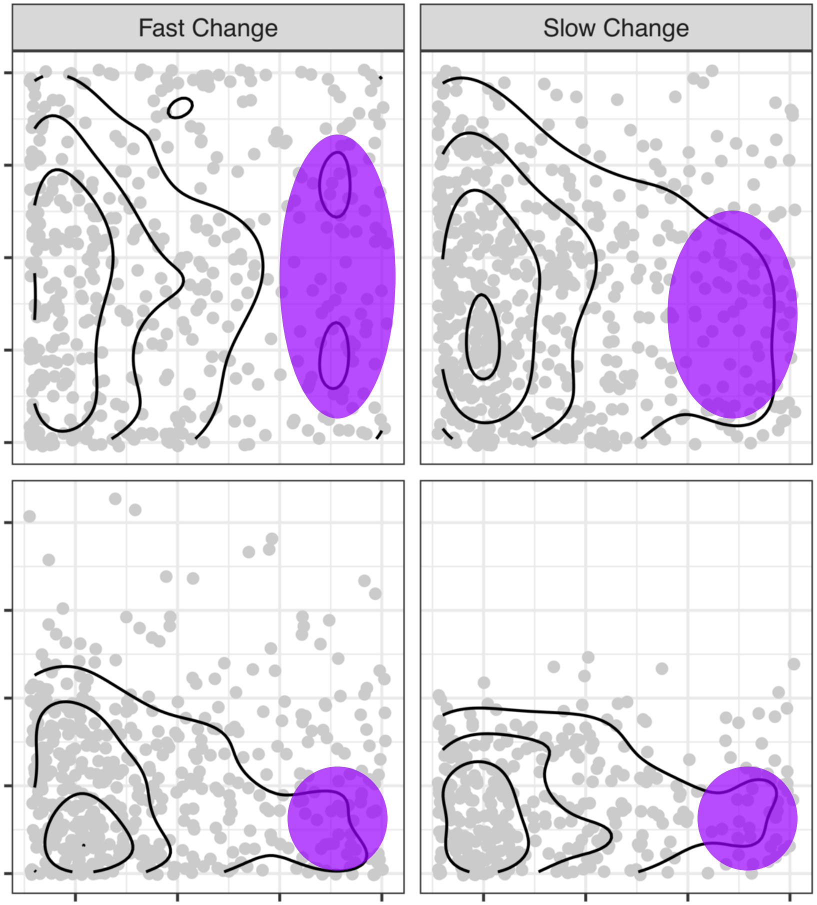


early expired options almost always low value

1. Poorer discrimination in volatile environments?



2. Letting go near the deadline?



Looks like people are "clinging" to a few suboptimal options only to let them expire right before the deadline?

Interim summary

Interim summary

- People mostly make good choices, but it is hard in extremely volatile environments (not surprisingly)

Interim summary

- People mostly make good choices, but it is hard in extremely volatile environments (not surprisingly)
- People <u>do</u> let options expire but are perhaps reluctant: agrees with Ejova et al (2009), Neth et al (2014), possibly also with Shin & Ariely (2004)

Interim summary

- People mostly make good choices, but it is hard in extremely volatile environments (not surprisingly)
- People do let options expire but are perhaps reluctant: agrees with Ejova et al (2009), Neth et al (2014), possibly also with Shin & Ariely (2004)
- There appears to be systematicity to how and when we allow options to expire

Open questions

Open questions

- Is the expiry any different to what we'd expect from a standard RL model (e.g. Kalman filter)

Open questions

- Is the expiry any different to what we'd expect from a standard RL model (e.g. Kalman filter)
- If there are differences, what pattern do they take?

Open questions

- Is the expiry any different to what we'd expect from a standard RL model (e.g. Kalman filter)
- If there are differences, what pattern do they take?
- Do the differences in responding across volatility levels reflect a strategy change, or the same approach expressed differently because the environment is different?

Computational Modelling

Kalman filter

Expected reward for option j on last trial

$$E_{j,t-1}$$

$$S_{j,t-1}$$

Uncertainty about reward for option j on last trial

Kalman filter

Expected reward for option j on this trial

$$E_{j,t-1} \longrightarrow E_{jt}$$

$$S_{j,t-1} \longrightarrow S_{jt}$$

Uncertainty about reward for option j on this trial

Kalman filter

$$\begin{tikzcd}
E_{j,t-1} \arrow[rr] & & E_{jt} \\
& K_{jt} & \\
S_{j,t-1} \arrow[rr] \arrow[ur, "\text{Uncertainty drives Kalman gain}"] & & S_{jt}
\end{tikzcd}$$

Kalman filter

$$
\begin{tikzcd}
E_{j,t-1} \arrow[rr] & & E_{jt} \\
& K_{jt} \arrow[ur] \arrow[dr] & \\
S_{j,t-1} \arrow[ur] \arrow[rr] & & S_{jt}
\end{tikzcd}
$$

Kalman gain influences beliefs about expected reward and uncertainty

Kalman filter

$$E_{jt} = E_{j,t-1} + \delta_{jt} K_{jt} \left(r_t - E_{j,t-1} \right)$$

Kalman filter

$$E_{jt} = E_{j,t-1} + \delta_{jt} K_{jt} \left(r_t - E_{j,t-1} \right)$$

Predicted reward for choosing the option

Kalman filter

$$E_{jt} = E_{j,t-1} + \delta_{jt} K_{jt} \left(r_t - E_{j,t-1} \right)$$

Predicted reward for choosing the option

Prediction error

Kalman filter

(only update chosen option)

$$E_{jt} = E_{j,t-1} + \delta_{jt} K_{jt} \left(r_t - E_{j,t-1} \right)$$

Predicted reward for choosing the option

Prediction error

Kalman filter

(only update chosen option)

$$E_{jt} = E_{j,t-1} + \delta_{jt} K_{jt} \left(r_t - E_{j,t-1} \right)$$

Predicted reward for choosing the option

Prediction error

Amount of learning depends on the Kalman gain

Kalman filter

$$E_{jt} = E_{j,t-1} + \delta_{jt} K_{jt} \left(r_t - E_{j,t-1} \right)$$

Kalman filter

$$E_{jt} = E_{j,t-1} + \delta_{jt} K_{jt} \left(r_t - E_{j,t-1} \right)$$

$$S_{jt} = (1 - \delta_{jt} K_{jt})(S_{j,t-1} + {\sigma_w}^2)$$

KF updates uncertainty

Kalman filter

$$E_{jt} = E_{j,t-1} + \delta_{jt} K_{jt} \left(r_t - E_{j,t-1}\right)$$

$$S_{jt} = (1 - \delta_{jt} K_{jt})(S_{j,t-1} + \sigma_w{}^2)$$

KF updates uncertainty

$$K_{jt} = \frac{S_{j,t-1} + \sigma_w{}^2}{S_{j,t-1} + \sigma_n{}^2 + \sigma_w{}^2} \leftarrow$$

Gain depends on uncertainty

Kalman filter

$$E_{jt} = E_{j,t-1} + \delta_{jt} K_{jt} \left(r_t - E_{j,t-1} \right)$$

$$S_{jt} = (1 - \delta_{jt} K_{jt})(S_{j,t-1} + \sigma_w{}^2)$$

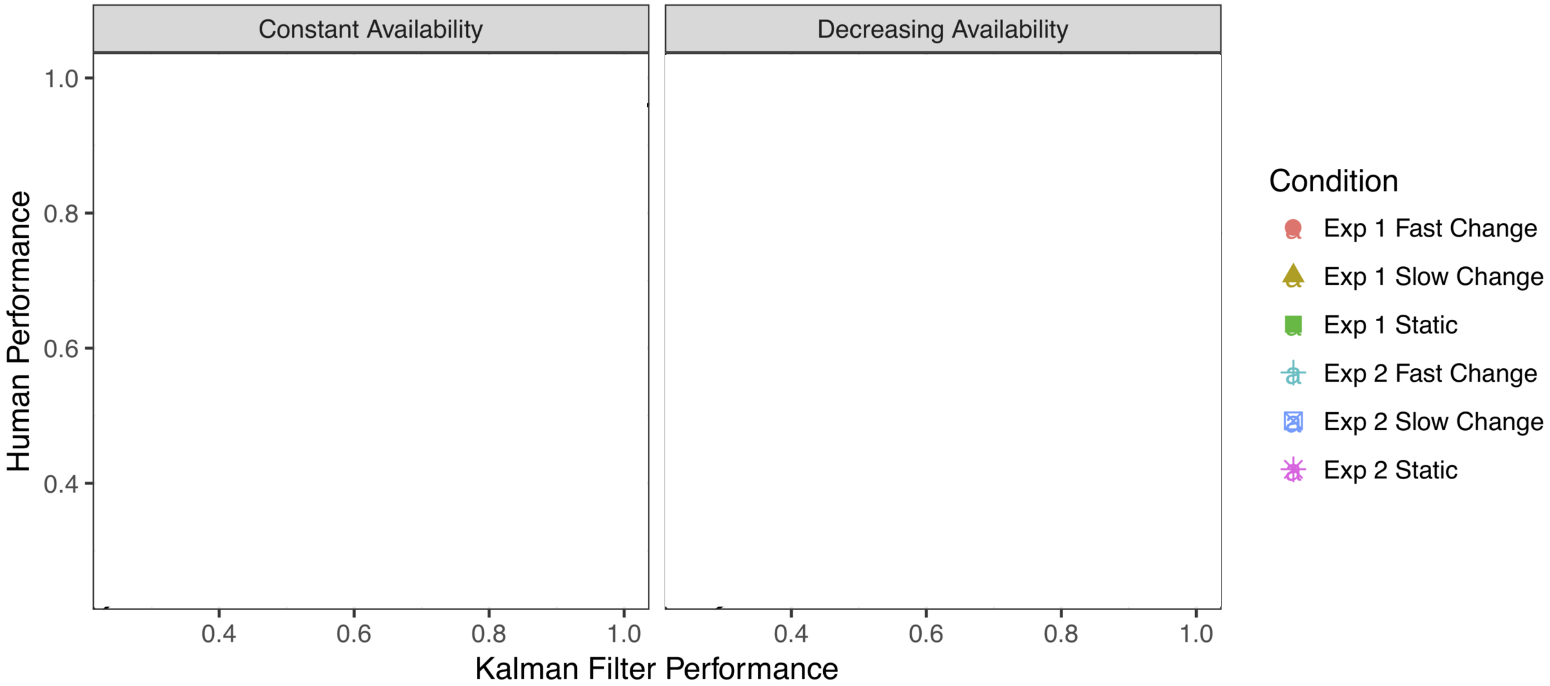
KF updates uncertainty

$$K_{jt} = \frac{S_{j,t-1} + \sigma_w{}^2}{S_{j,t-1} + \sigma_n{}^2 + \sigma_w{}^2}$$

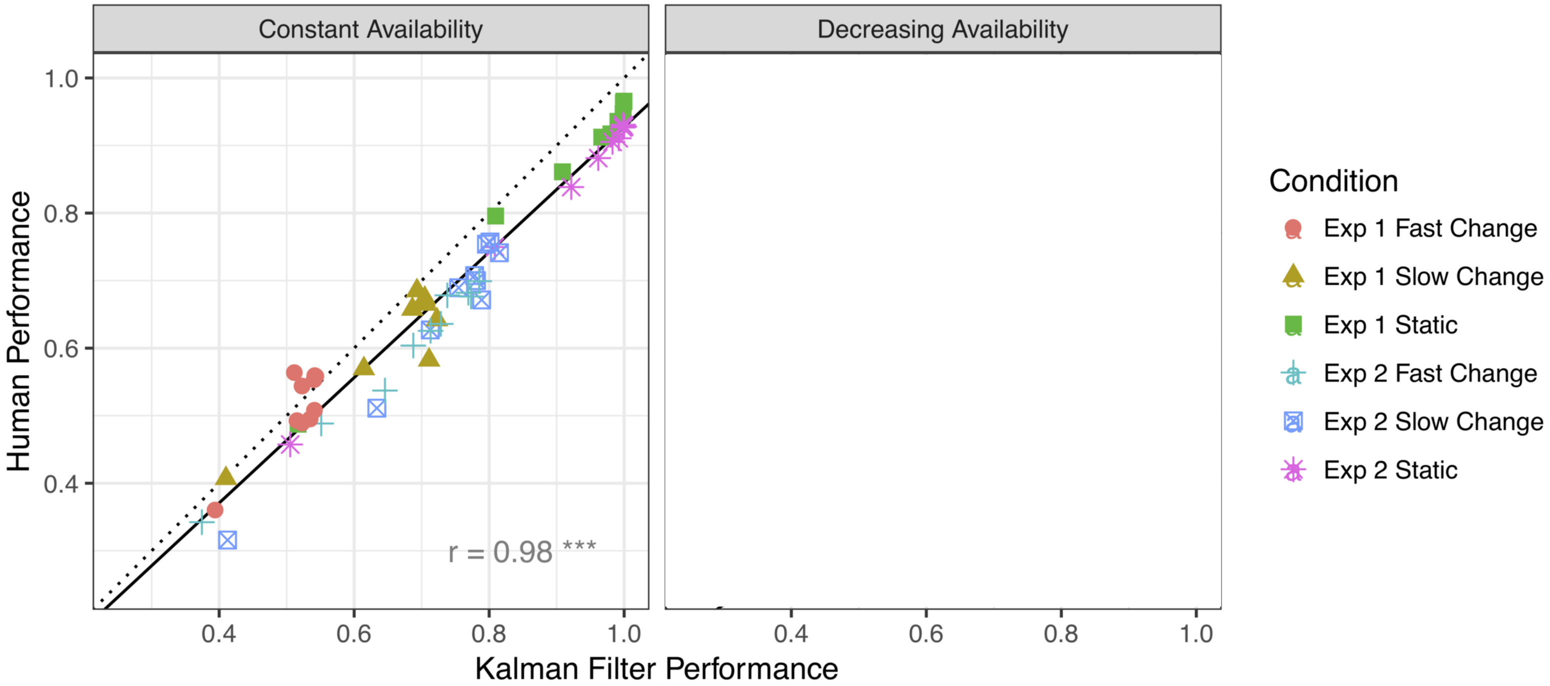
Gain depends on uncertainty

- Volatility σ_w and noise σ_n fixed at veridical values
- Initial values E_{j0} and S_{j0} reflect diffuse prior
- Model not yoked to participant: purely predictive

Choice probabilities

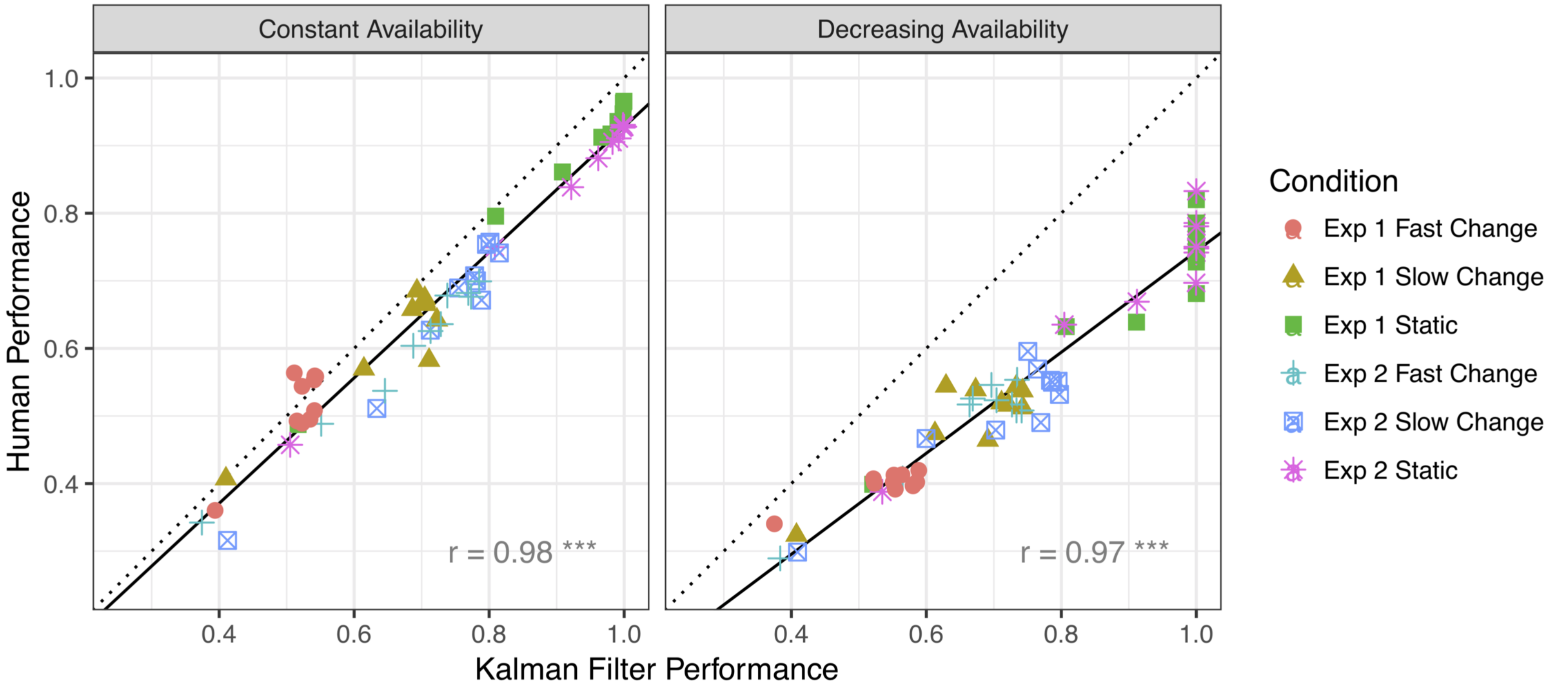


Choice probabilities



KF model provides an excellent account of choice behaviour when options do not expire

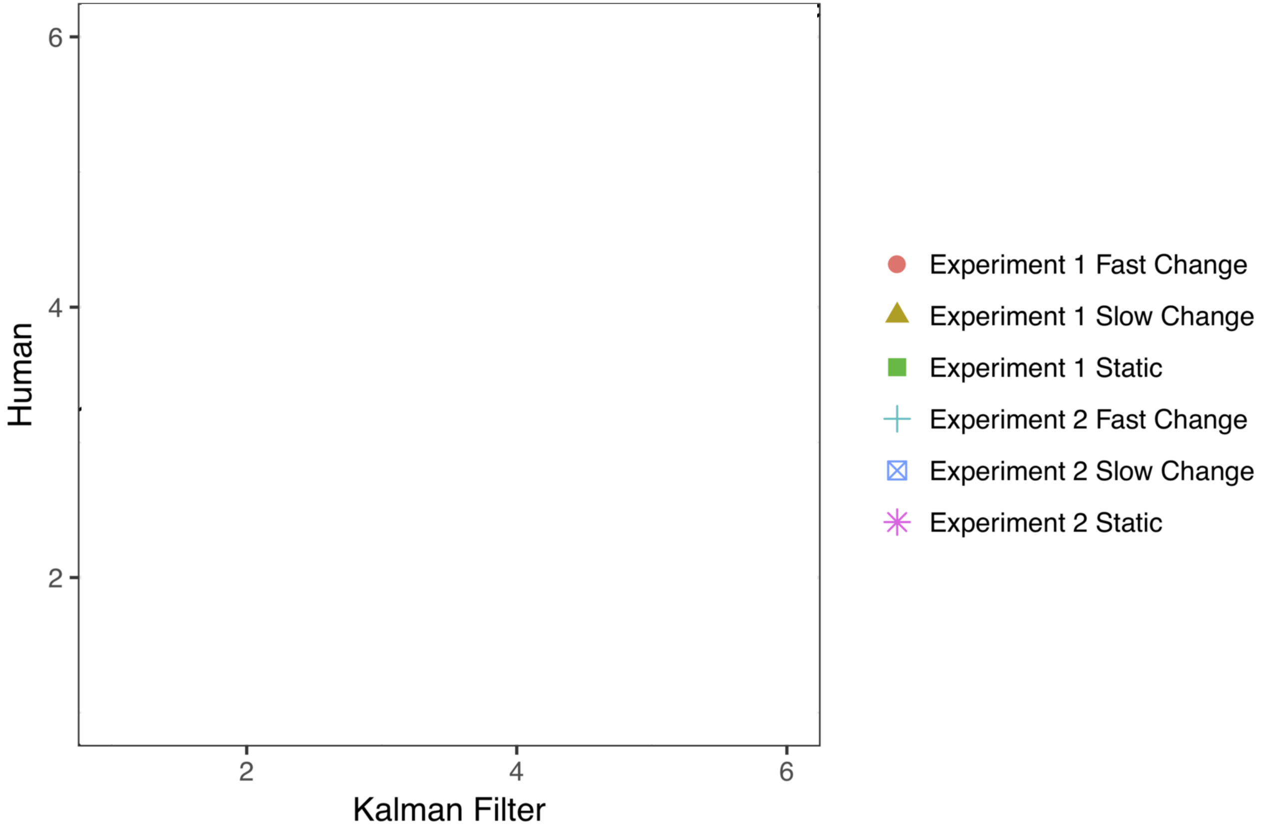
Choice probabilities



KF model provides an excellent account of choice behaviour when options do not expire

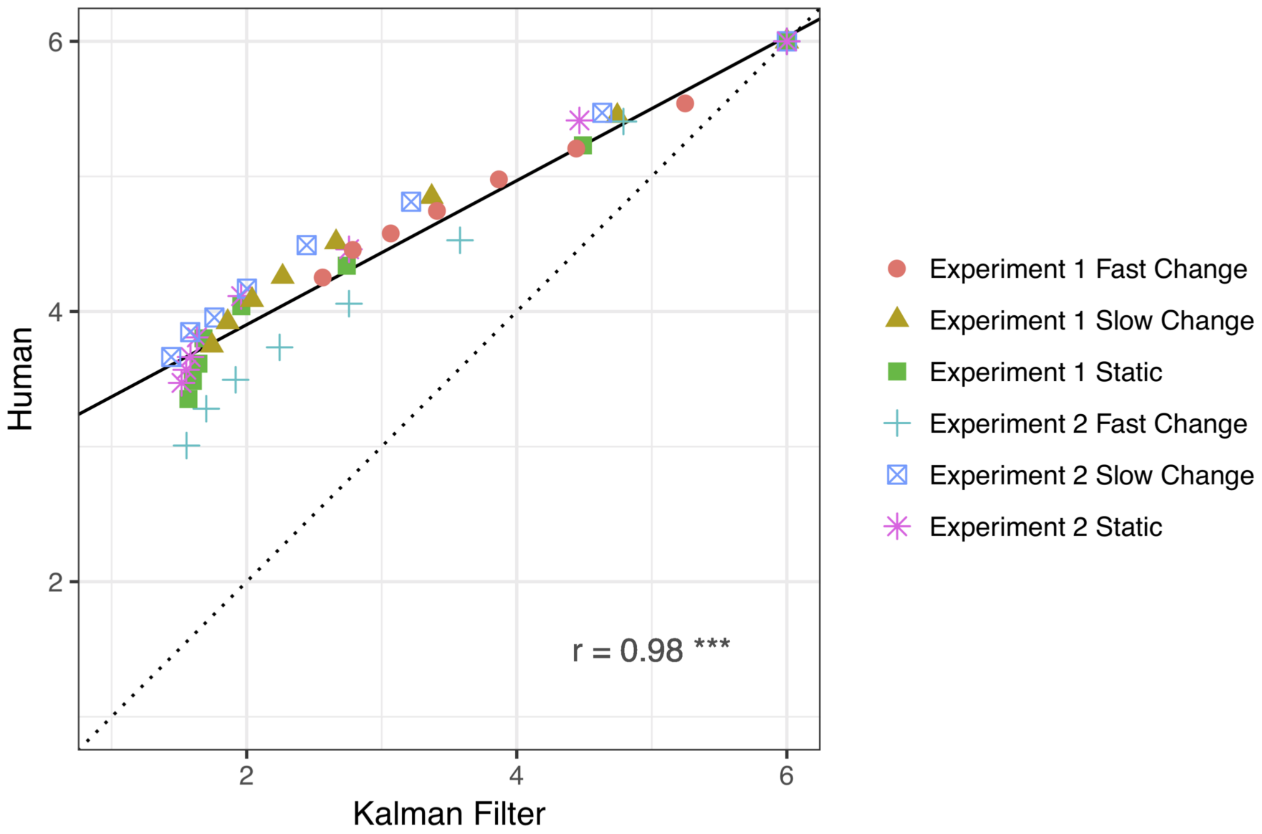
There is a systematic difference when option loss is a possibility

Options retained?

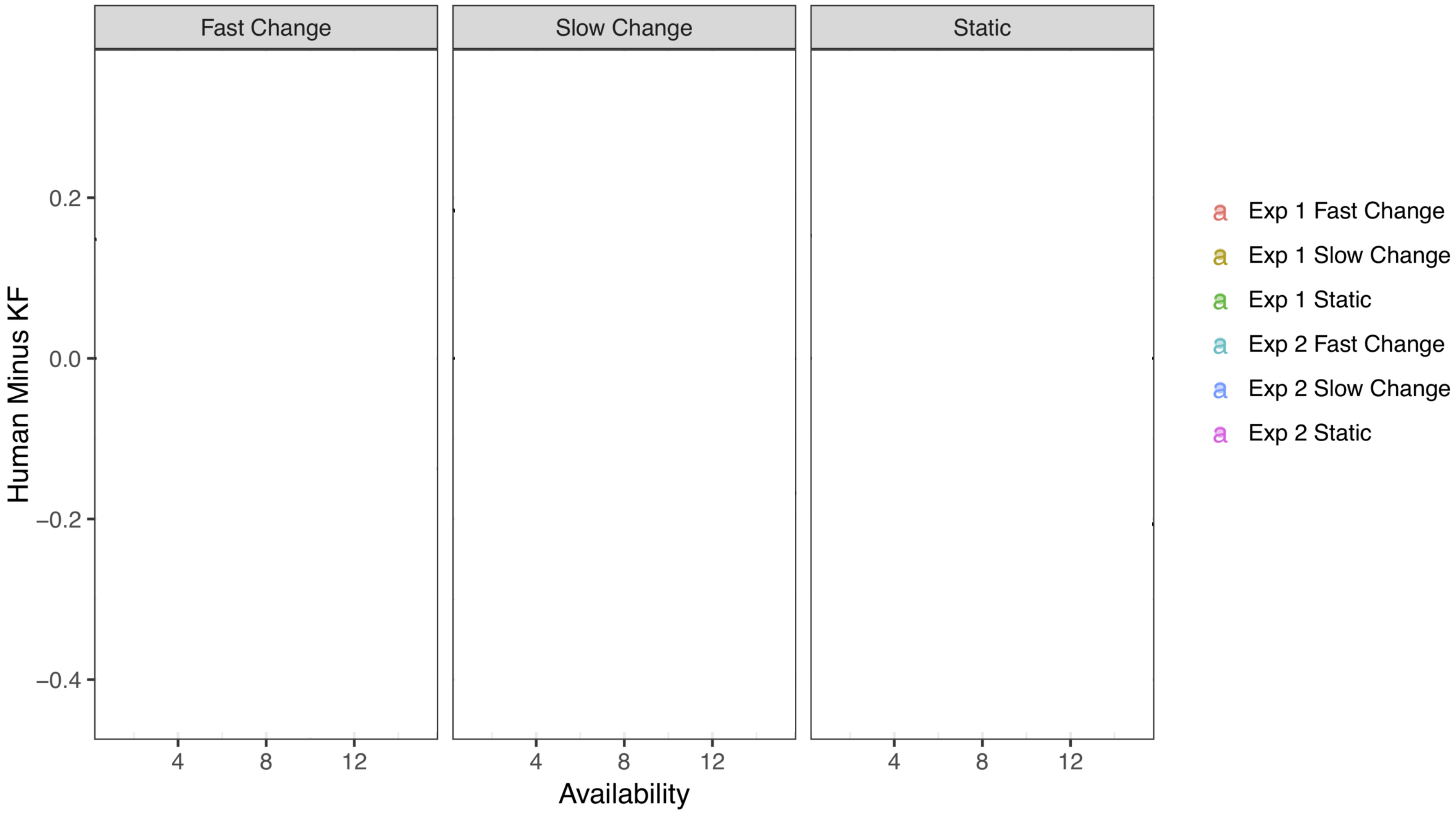


Options retained?

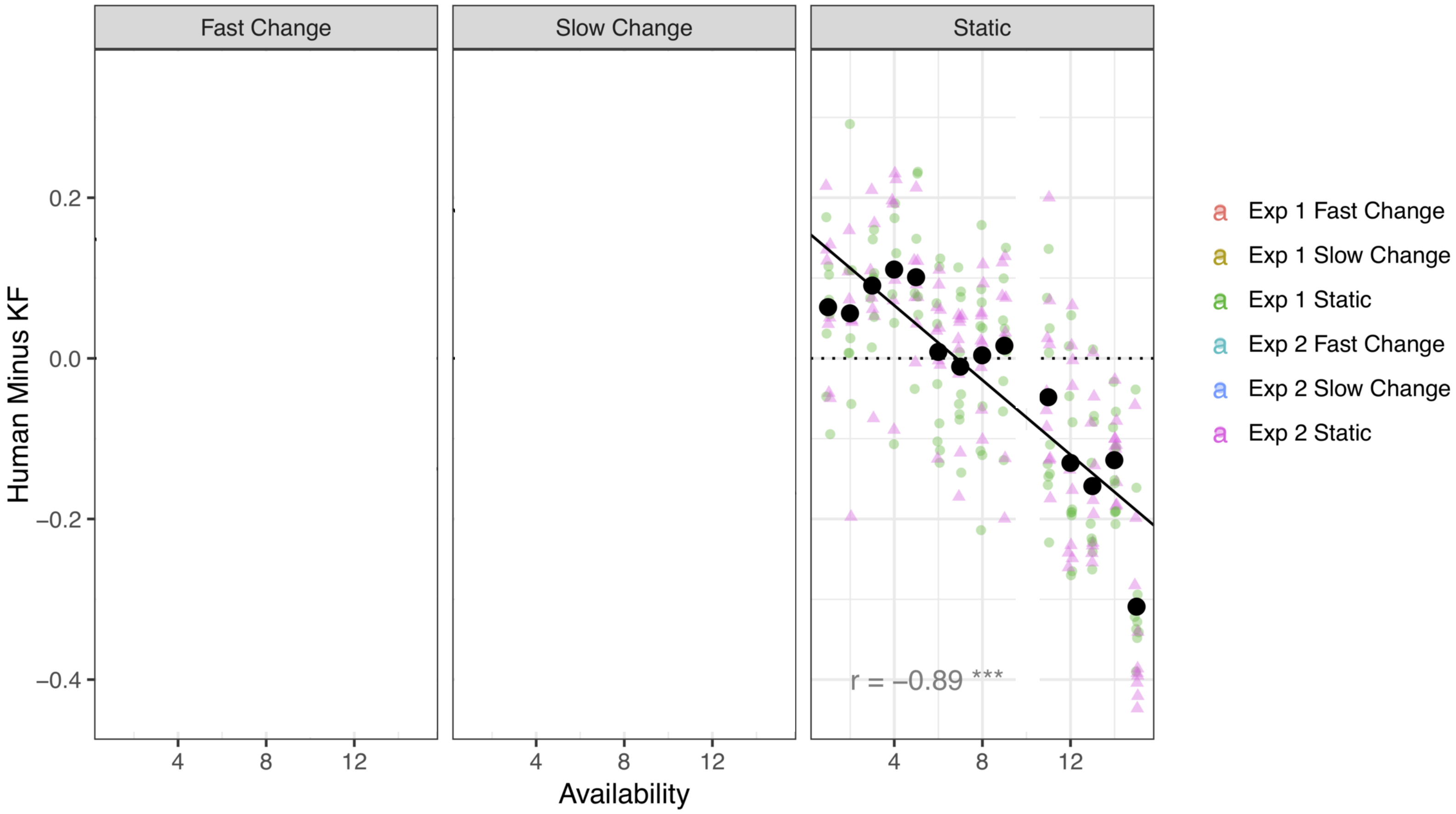
Human decision makers retain more options than the KF model



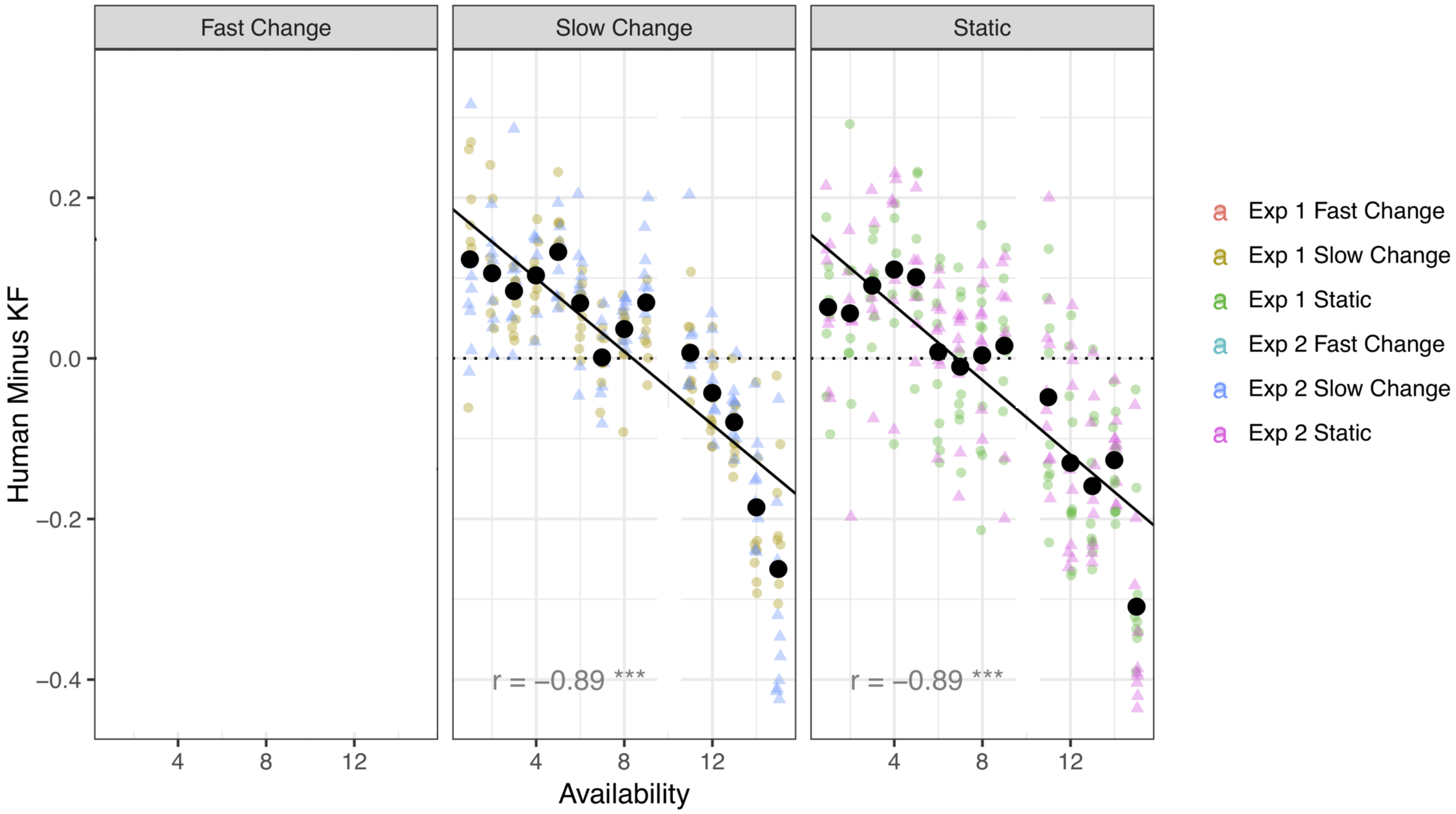
Quantifying loss aversion?



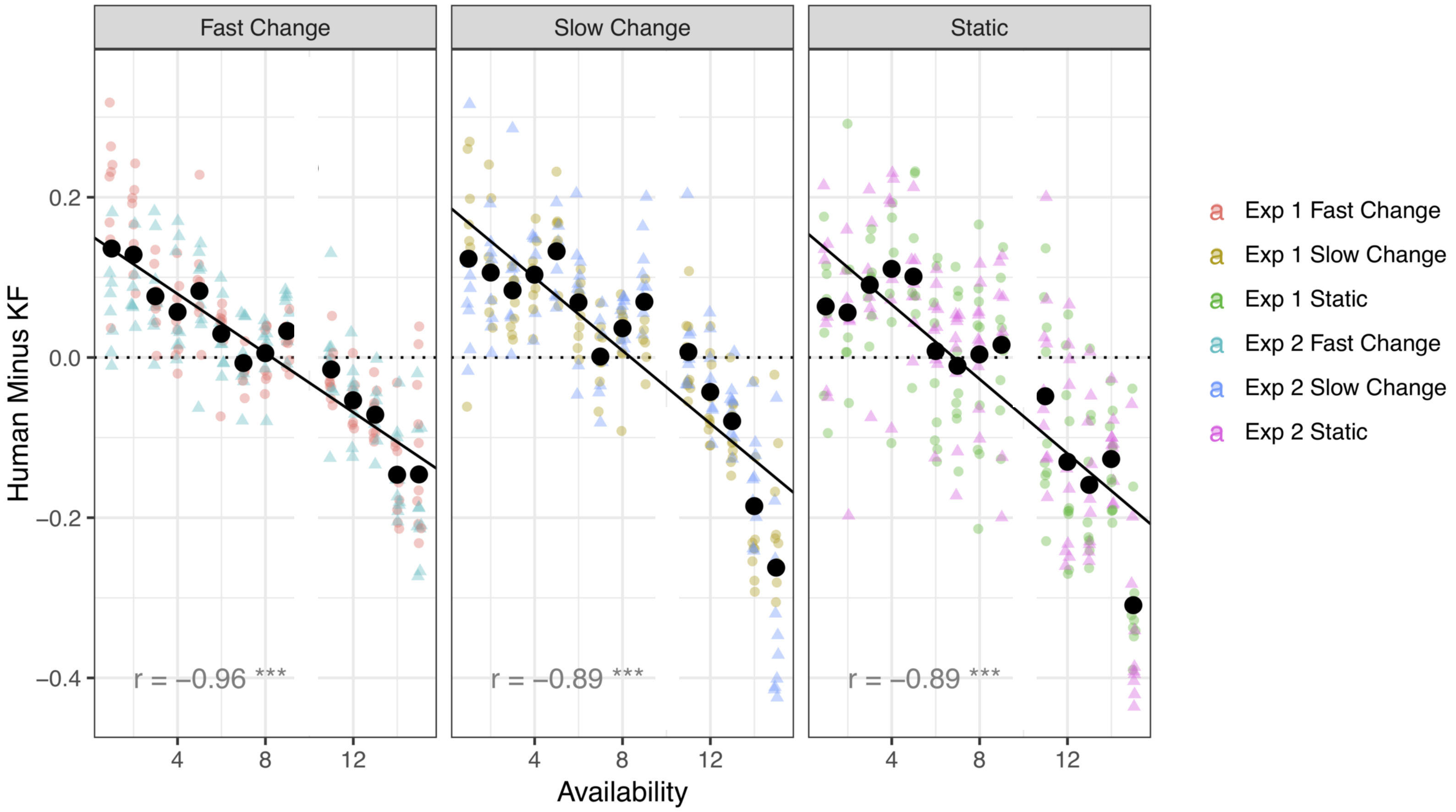
Quantifying loss aversion?



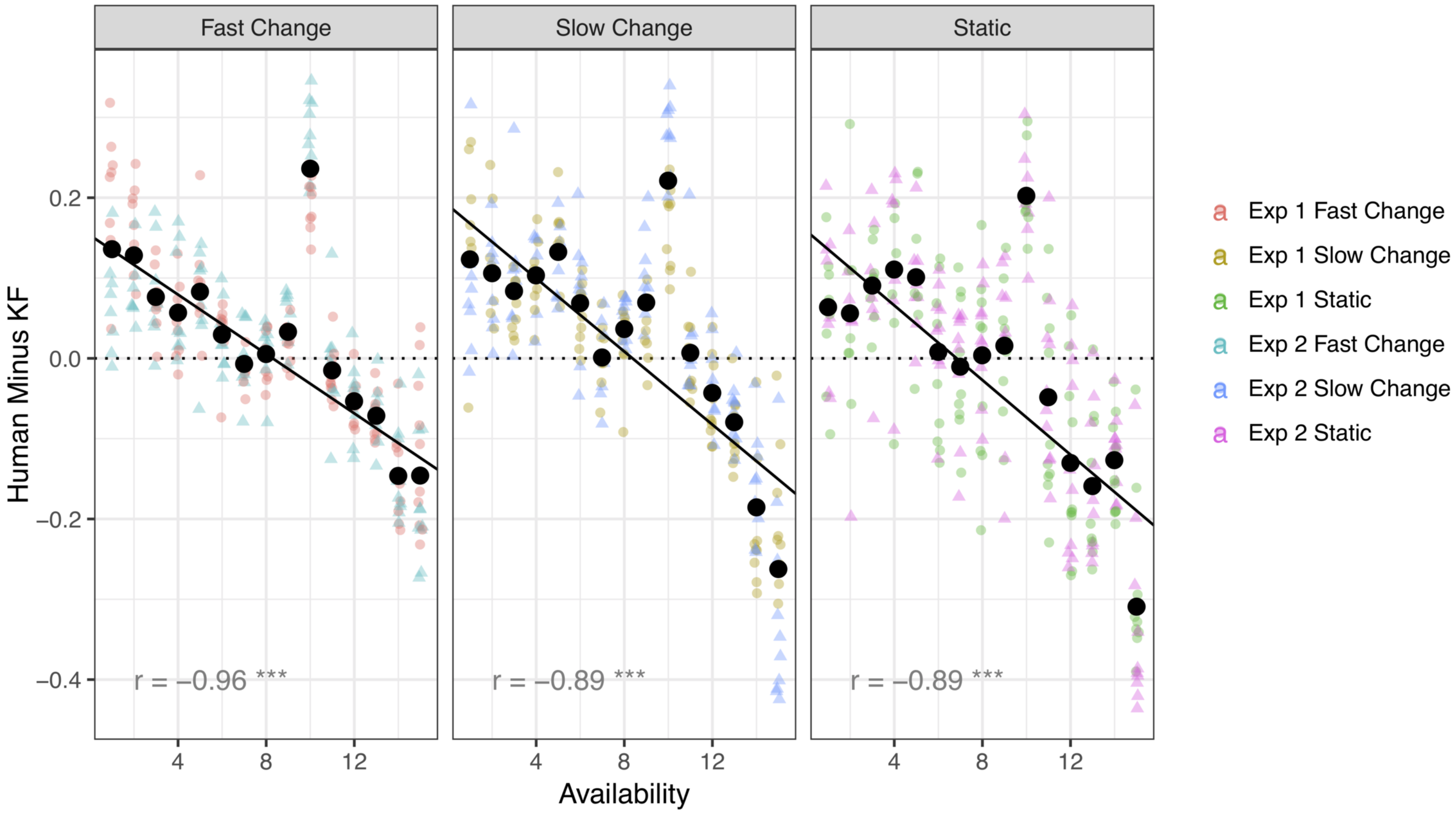
Quantifying loss aversion?



Quantifying loss aversion?



Quantifying loss aversion?



Conclusions?

Main findings

Main findings

- Replicated findings from "doors" tasks in a bandit task framework without explicit switching costs

Main findings

- Replicated findings from "doors" tasks in a bandit task framework without explicit switching costs
- Extension to several dynamic environments

Main findings

- Replicated findings from "doors" tasks in a bandit task framework without explicit switching costs
- Extension to several dynamic environments
- Computational modelling to measure the shift in decision policy when option loss exists

Main findings

- Replicated findings from "doors" tasks in a bandit task framework without explicit switching costs
- Extension to several dynamic environments
- Computational modelling to measure the shift in decision policy when option loss exists
- Attempted to quantify the loss aversion signal

Main findings

- Replicated findings from "doors" tasks in a bandit task framework without explicit switching costs
- Extension to several dynamic environments
- Computational modelling to measure the shift in decision policy when option loss exists
- Attempted to quantify the loss aversion signal

Follow up?

- Covariates? Does anxiety play a particular role here?

Main findings

- Replicated findings from "doors" tasks in a bandit task framework without explicit switching costs
- Extension to several dynamic environments
- Computational modelling to measure the shift in decision policy when option loss exists
- Attempted to quantify the loss aversion signal

Follow up?

- Covariates? Does anxiety play a particular role here?
- Why the "gradual rising" pattern? Hazard in the task is abrupt (cliff) not smooth (lion). Why do people treat a "cliff" task like a "lion" threat?

Thanks!

Contact:

- compcogscisydney.org
- d.navarro@unsw.edu.au
- twitter.com/djnavarro
- github.com/djnavarro

Support:

Australian Government
Australian Research Council

UNSW SYDNEY

Project:

- Preprint: psyarxiv.com/3g4p5
- OSF: osf.io/nzvqp

(gifeed.com)