Prevalence-induced perceptual shift due to prior change

Peiling Jiang

Bayesian Modeling of Behavior Spring 2021

Modeling of

Levari, D. E., Gilbert, D. T., Wilson, T. D., Sievers, B., Amodio, D. M., & Wheatley, T. (2018). Prevalence-induced concept change in human judgment. Science, 360(6396), 1465-1467.

"When blue dots became rare, participants began to see purple dots as blue; when threatening faces became rare, participants began to see neutral faces as threatening; and when unethical requests became rare, participants began to see innocuous requests as unethical."

Levari et al. (2018)

Method


1000 trials
...
50
Blue (0, 0, 255)
Blue spectrum
~ (50, 0, 205)
0
Purple spectrum
~ (51, 0, 204)
-50
Purple (100, 0, 155)

Method

Study 1 - Stable condition / Decrease condition

Study 2 - Some offered verbal instructions

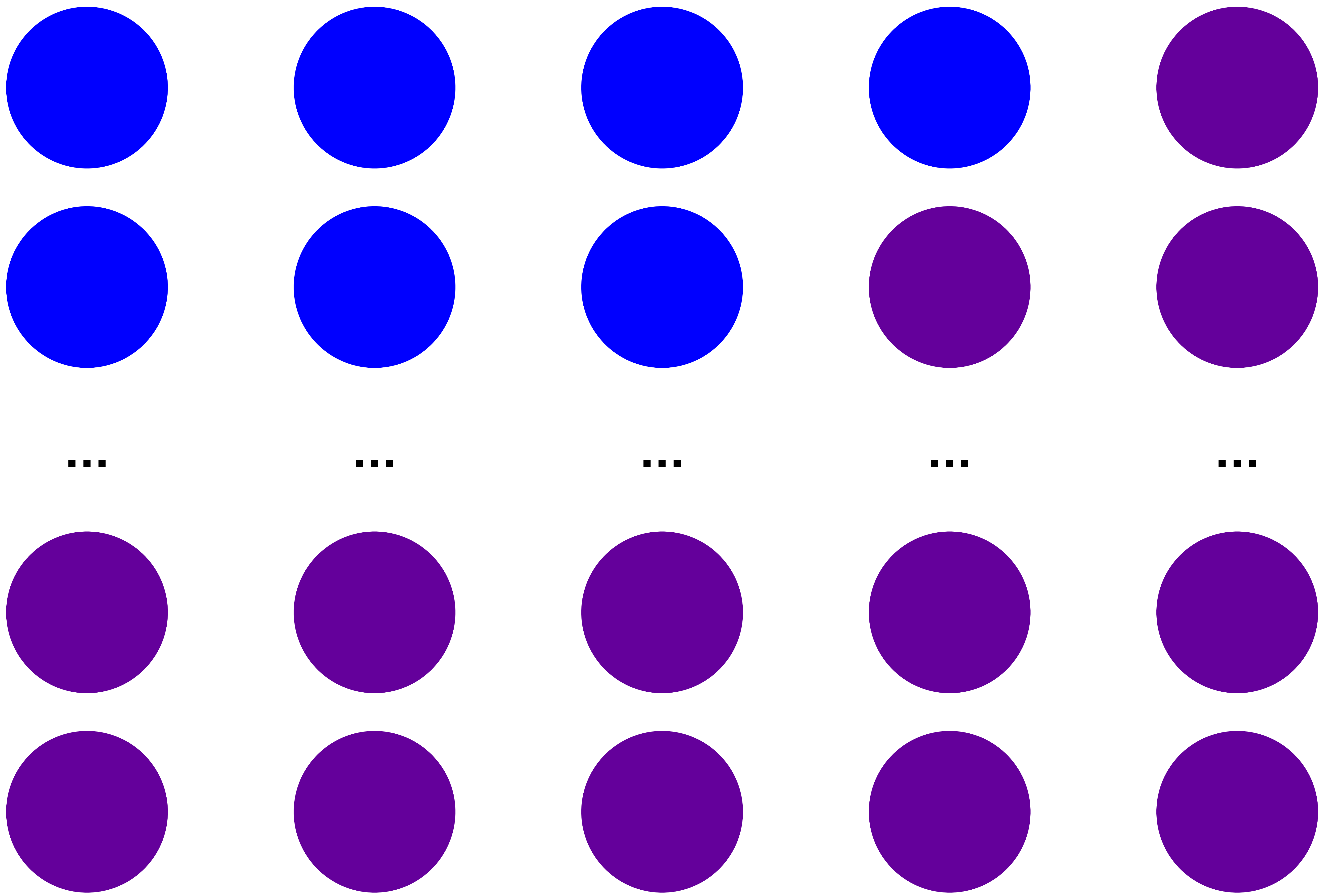
Study 3 - Some offered verbal instructions and monetary reward

Study 4 - Stable condition / Gradually decrease condition / Abruptly de. Condition

Study 5 - Stable condition / Increase condition

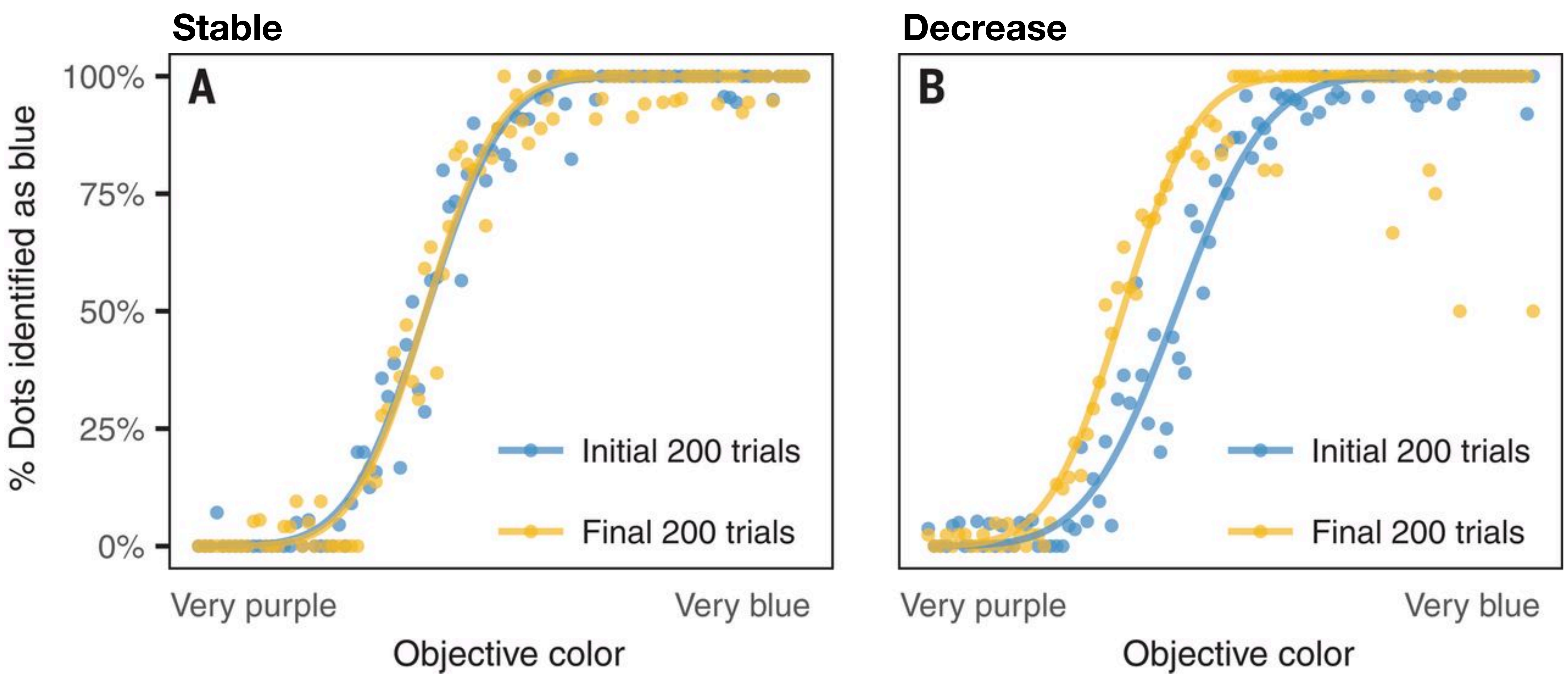
Method

Trials	1-200	201-250	251-300	301-350	351-1000



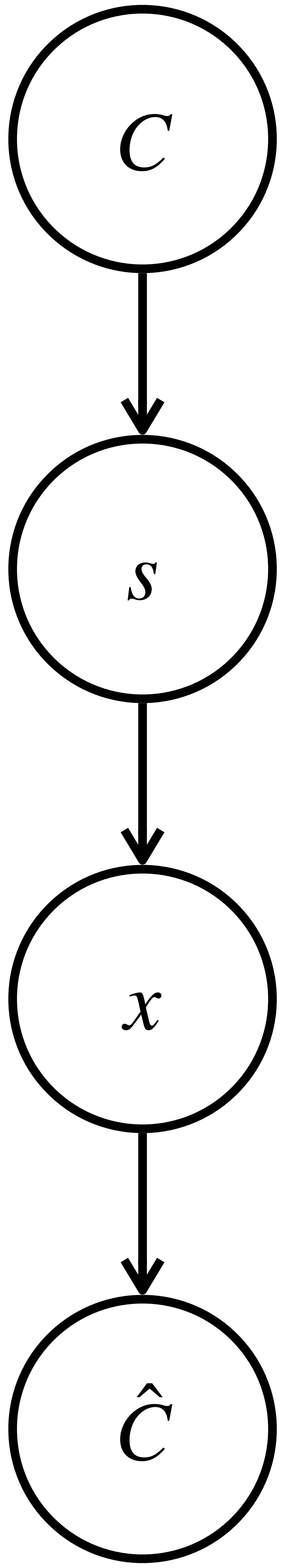
Probability color drawn from blue spectrum	50%	40%	28%	16%	6%

Results



Study 1

Modeling



$p(C)$

$p(s \mid C)$

$p(x \mid s)$

$p(\hat{C} \mid x)$

$$
\begin{aligned}
p(\hat{C}|x) &= p(C)p(x|C) \\
&= p(C)\int p(x|s)p(s|C)ds \\
\because p(x|s) &\sim \mathcal{N}(s,\sigma^2) \\
p(s|C) &= \frac{1}{a} \text{ when } \operatorname{sign}(s) = C \text{ and } |s| < a \\
\therefore d &= \log\frac{p(C=1)}{p(C=-1)} + \log\frac{\int p(x|s)p(s|C=1)ds}{\int p(x|s)p(s|C=-1)ds} \\
&= \boxed{\log\frac{p(C=1)}{p(C=-1)}} + \log\frac{\int_0^a \exp-\frac{(s-x)^2}{2\sigma^2}ds}{\int_{-a}^0 \exp-\frac{(s-x)^2}{2\sigma^2}ds}
\end{aligned}
$$

Hypothesis: the prior was updated continuously based on the previous observations towards a certain type of stimuli

Modeling

Rule A - Normal Dirichlet

i.e. (1, 1) + (1, 0) if reports blue / (0, 1) if reports purple

$$\log \frac{p(C=1)}{p(C=-1)} = \log \frac{\frac{\alpha'_1}{\sum_i \alpha'}}{\frac{\alpha'_{-1}}{\sum_i \alpha'}} = \log \frac{\alpha_1 + c_1}{\alpha_{-1} + c_{-1}}$$

Rule B - Signal search

i.e. (1, 1) + (1, 0) if reports blue / (0, 0) if reports purple

$$\log \frac{p(C=1)}{p(C=-1)} = \log \frac{\alpha_1 + c_1}{\alpha_{-1}}$$

Rule C - Twisted mind

i.e. (1, 1) + (0, 1) if reports blue / (1, 0) if reports purple

$$\log \frac{p(C=1)}{p(C=-1)} = \log \frac{\alpha_1 + c_{-1}}{\alpha_{-1} + c_1}$$

Rule D - Observation count

i.e. (1000, 1000) + (-1, 0) if reports blue / (0, -1) if reports purple

$$\log \frac{p(C=1)}{p(C=-1)} = \log \frac{\alpha_1 - c_1}{\alpha_{-1} - c_{-1}}$$

Rule A
Normal Dirichlet

	Reports	α_1	α_{-1}
Initial		1	1
1	P	0	1
2	B	1	0
3	P	0	1
4	P	0	1
5	B	1	0
6	P	0	1
7	P	0	1
8	B	1	0
Total		4	6

$$\frac{p(C=1)}{p(C=-1)} \qquad \frac{2}{3}$$

Rule B
Signal search

	Reports	α_1	α_{-1}
Initial		1	1
1	P	0	0
2	B	1	0
3	P	0	0
4	P	0	0
5	B	1	0
6	P	0	0
7	P	0	0
8	B	1	0
Total		4	1

$$\frac{4}{1}$$

Rule C
Twisted mind

	Reports	α_1	α_{-1}
Initial		1	1
1	P	1	0
2	B	0	1
3	P	1	0
4	P	1	0
5	B	0	1
6	P	1	0
7	P	1	0
8	B	0	1
Total		6	4

$$\frac{3}{2}$$

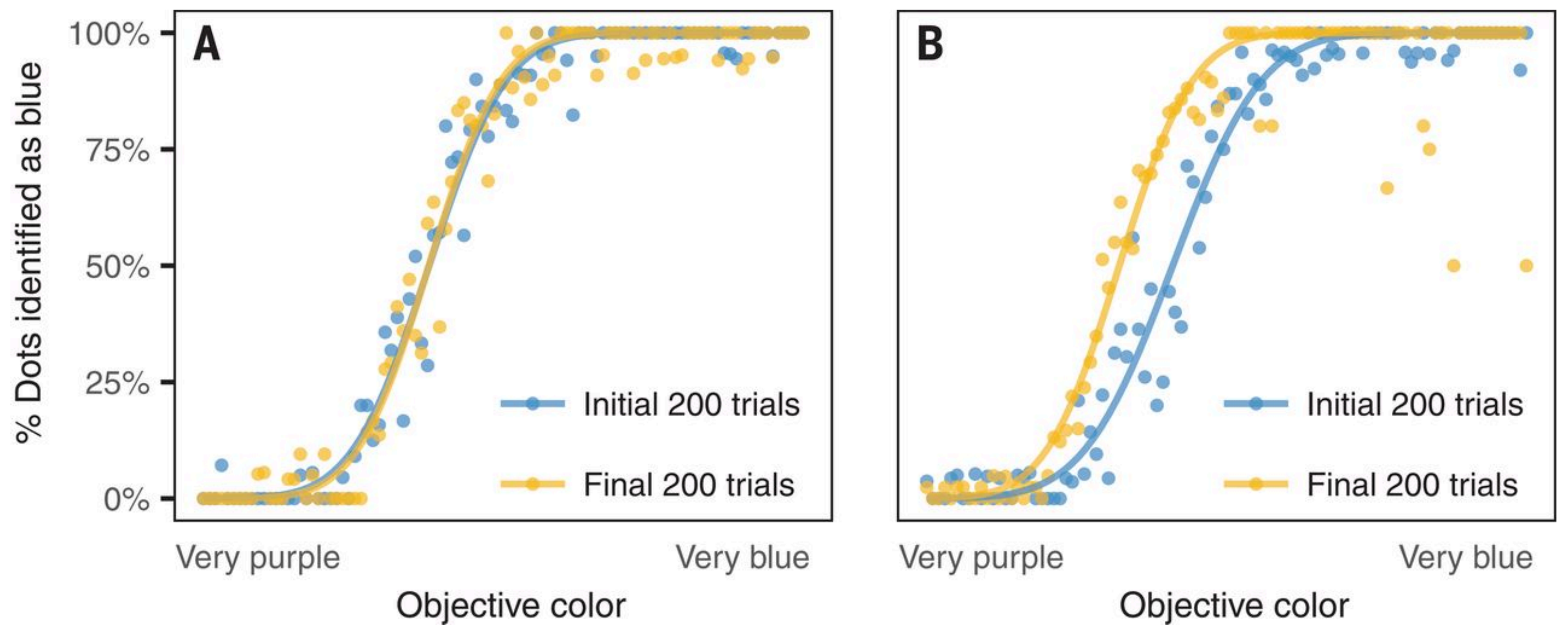
Rule D
Observation count

	Reports	α_1	α_{-1}
Initial		8	8
1	P	0	-1
2	B	-1	0
3	P	0	-1
4	P	0	-1
5	B	-1	0
6	P	0	-1
7	P	0	-1
8	B	-1	0
Total		5	3

$$\frac{5}{3}$$

Modeling Results

Study 1

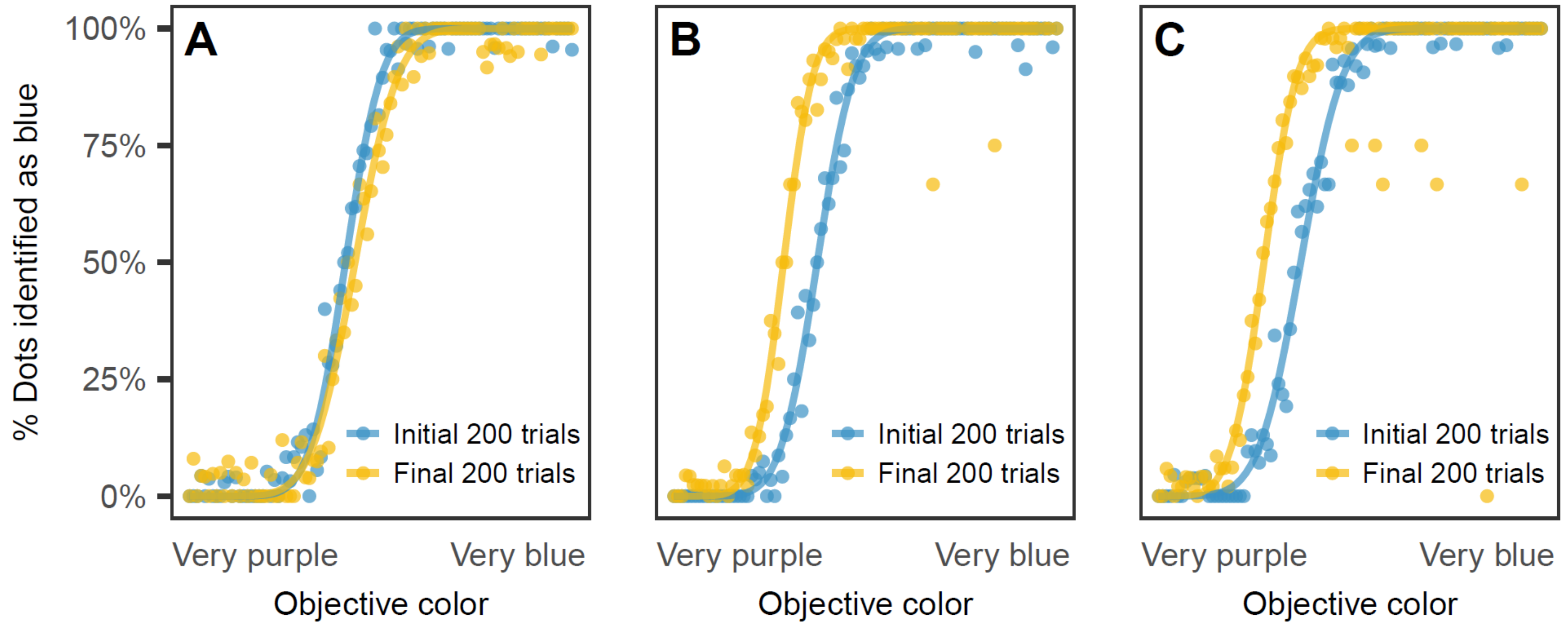


Stable

Decrease

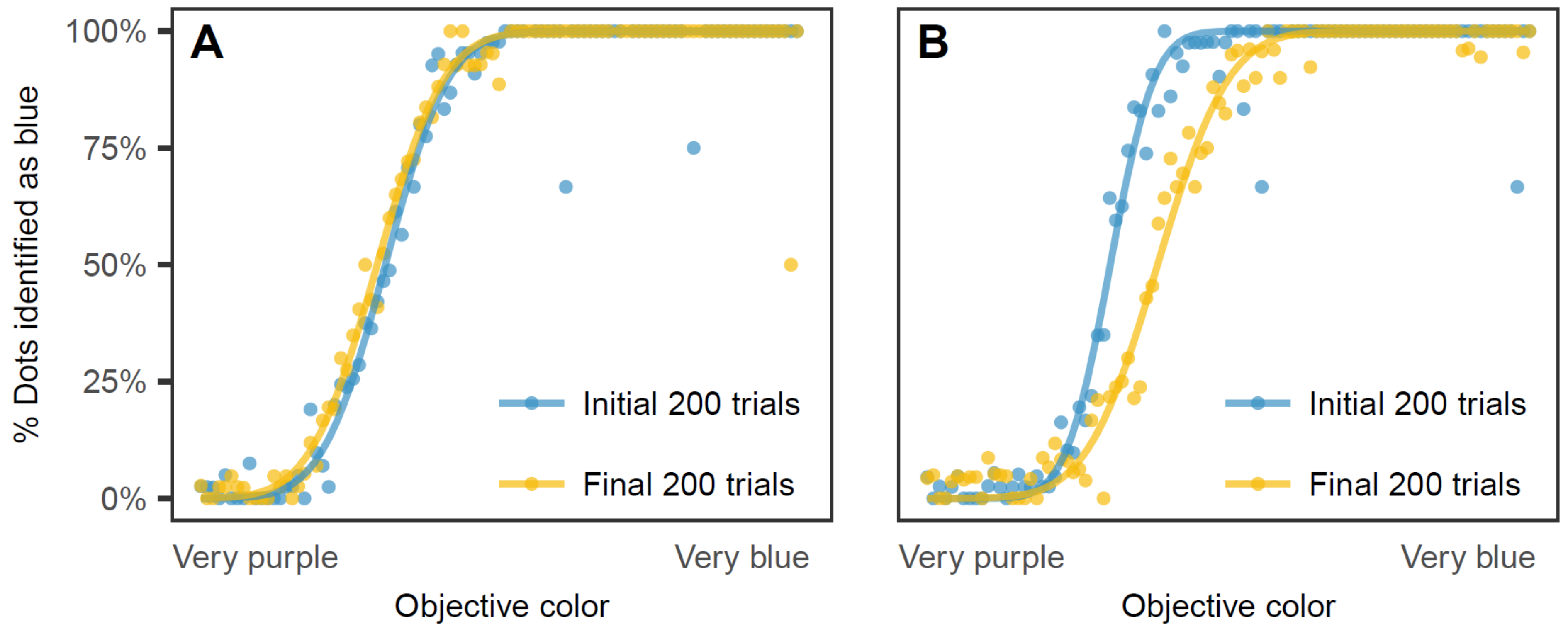
Modeling Results

Study 4



Modeling Results

Study 5



Stable

Increase

Modeling Results
New Study

Stable condition: stays at 20% instead of 50% for 1000 trials

Future Works

Statistical analysis of the simulation data.
Fit real experiment data.

Thank you.