

Somewhere Over the Rainbow:

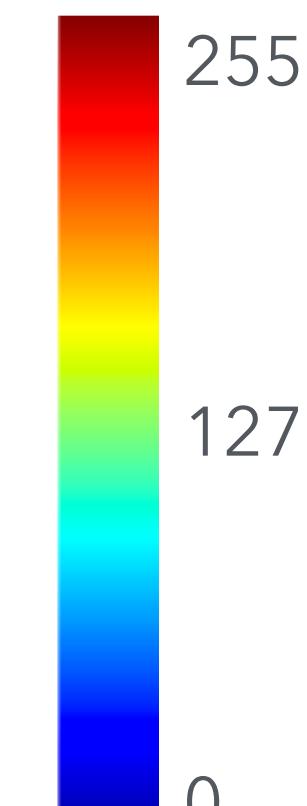
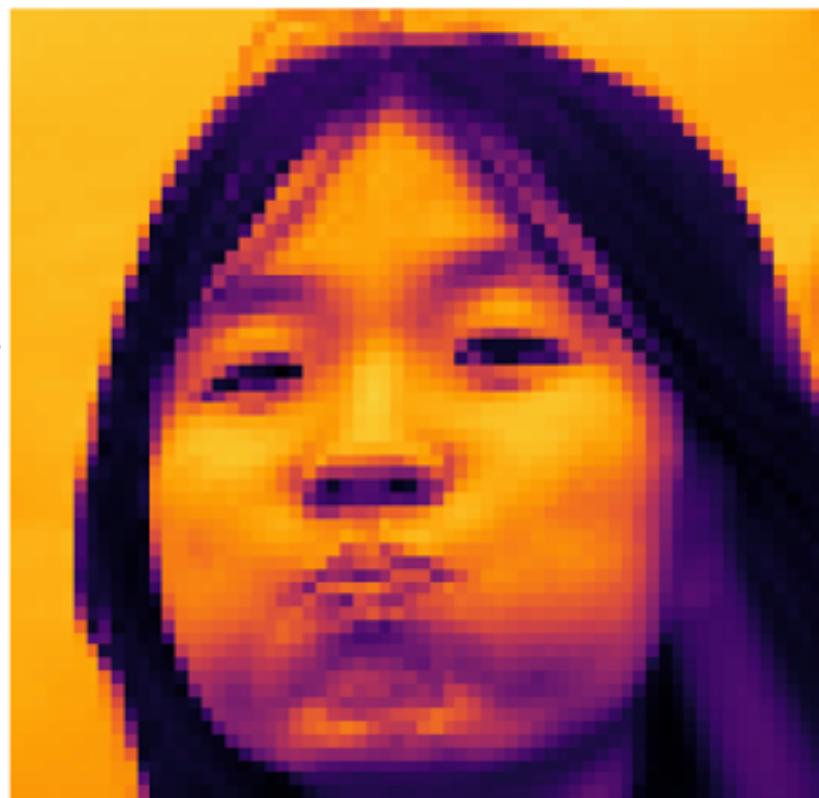
An Empirical Assessment of Quantitative Colormaps

Yang Liu, Jeffrey Heer
University of Washington
Interactive Data Lab
yliu0@cs.washington.edu, jheer@uw.edu



Quantitative Colormaps

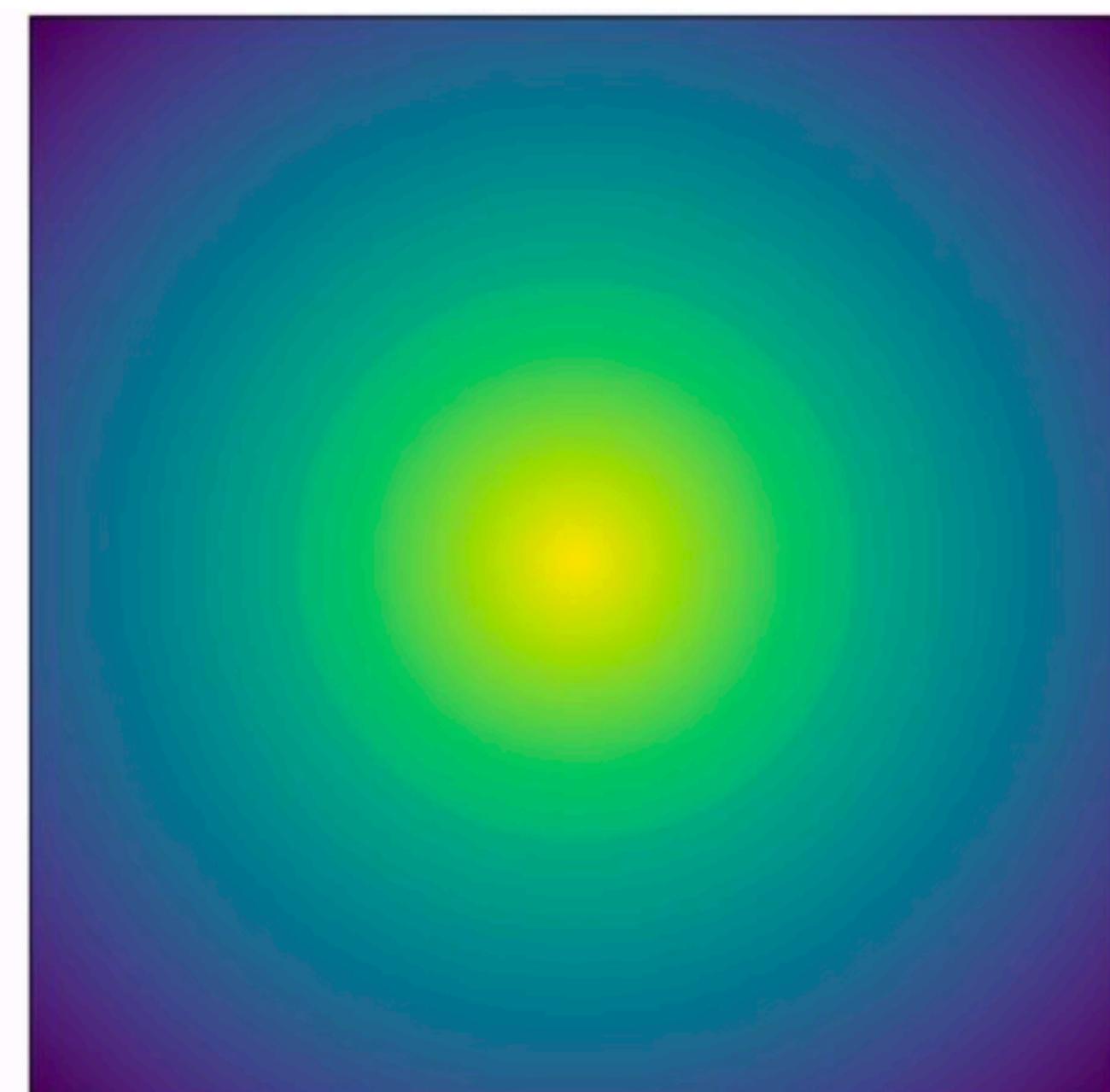
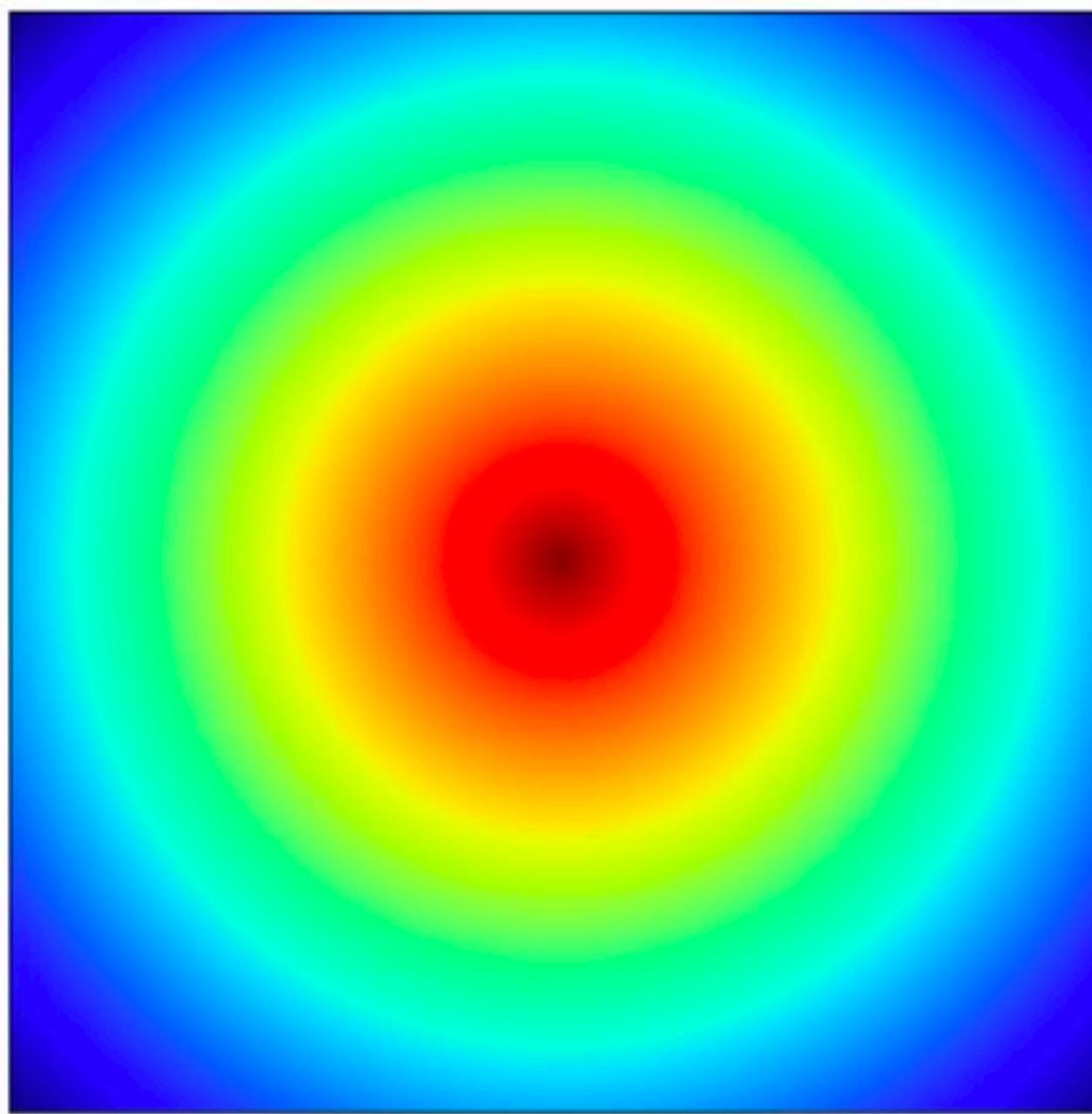
19	14	14	15	15	16	16	18	18	56	56	100	100	123	123	124	124	156	156	172	172	175	175	163	163		
19	14	14	15	15	16	16	18	18	56	56	100	100	123	123	124	124	156	156	172	172	175	175	163	163		
23	14	14	14	14	14	19	19	34	34	117	117	124	124	148	148	181	181	199	199	203	203	201	201	189	189	
23	14	14	14	14	14	19	19	34	34	117	117	124	124	148	148	181	181	199	199	203	203	201	201	189	189	
21	15	15	15	15	15	13	13	83	83	117	117	150	150	176	176	191	191	187	187	186	186	187	187	185	185	
21	15	15	15	15	15	13	13	83	83	117	117	150	150	176	176	191	191	187	187	186	186	187	187	185	185	
15	13	13	19	19	29	29	114	114	141	141	163	163	166	166	170	170	156	156	109	109	78	78	87	87	87	
15	13	13	19	19	29	29	114	114	141	141	163	163	166	166	170	170	156	156	109	109	78	78	87	87	87	
13	16	16	26	26	59	59	119	119	144	144	161	161	144	144	66	66	25	25	18	18	13	13	2	2	2	
13	16	16	26	26	59	59	119	119	144	144	161	161	144	144	66	66	25	25	18	18	13	13	2	2	2	
14	16	16	47	47	74	74	130	130	153	153	120	120	21	21	16	16	58	58	44	44	48	48	92	92	92	
14	16	16	47	47	74	74	130	130	153	153	120	120	21	21	16	16	58	58	44	44	48	48	92	92	92	
14	11	11	66	66	97	97	151	151	127	127	66	66	94	94	136	136	153	153	153	153	131	131	135	135	135	
14	11	11	66	66	97	97	151	151	127	127	66	66	94	94	136	136	153	153	153	153	131	131	135	135	135	
16	16	16	73	73	126	126	176	176	173	173	188	188	183	187	187	177	177	153	153	147	147	147	170	170	189	189
16	16	16	73	73	126	126	176	176	173	173	188	188	183	187	187	177	177	153	153	147	147	147	170	170	189	189
18	21	21	79	79	162	162	197	197	206	206	210	210	208	208	205	205	206	206	207	207	204	204	205	205	205	
18	21	21	79	79	162	162	197	197	206	206	210	210	208	208	205	205	206	206	207	207	204	204	205	205	205	
15	22	22	97	97	188	188	209	209	215	215	218	218	219	219	218	218	217	217	213	213	214	214	211	211	211	
15	22	22	97	97	188	188	209	209	215	215	218	218	219	219	218	218	217	217	213	213	214	214	211	211	211	
12	18	18	128	128	201	201	212	212	218	218	219	219	219	219	220	220	220	220	219	219	214	214	212	212	212	
12	18	18	128	128	201	201	212	212	218	218	219	219	219	219	220	220	220	220	219	219	214	214	212	212	212	
13	12	12	164	164	207	207	212	212	210	210	215	215	220	220	221	221	220	220	217	217	209	209	200	200	200	
13	12	12	164	164	207	207	212	212	210	210	215	215	220	220	221	221	220	220	217	217	209	209	200	200	200	
13	18	18	180	180	206	206	205	205	209	209	214	214	215	215	215	215	213	213	206	206	200	200	177	177	177	
13	18	18	180	180	206	206	205	205	209	209	214	214	215	215	215	215	213	213	206	206	200	200	177	177	177	
13	25	25	186	186	201	201	204	204	206	206	208	208	203	204	204	207	207	205	205	202	202	192	192	173	173	173
13	25	25	186	186	201	201	204	204	206	206	208	208	203	204	204	207	207	205	205	202	202	192	192	173	173	173
12	25	25	183	183	195	195	198	198	201	201	200	200	200	200	203	203	199	199	198	198	195	195	189	189	189	
12	25	25	183	183	195	195	198	198	201	201	200	200	200	200	203	203	199	199	198	198	195	195	189	189	189	
12	20	20	176	176	187	187	192	192	194	194	195	195	197	197	197	197	201	201	204	204	206	206	206	206	206	



What makes jet ineffective?

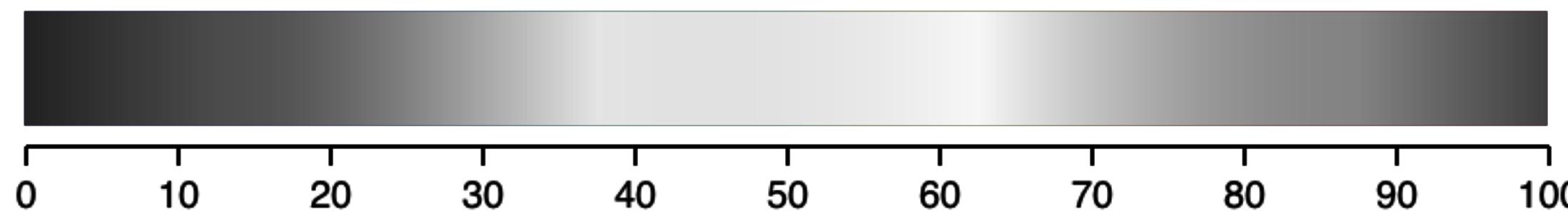


Criticisms: Gradients



Banding effects at **hue boundaries** make it hard to observe **gradients**

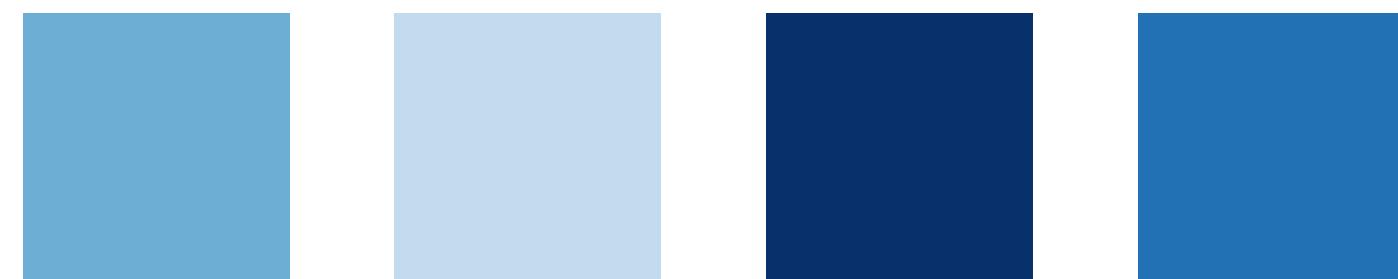
Criticisms: Black-and-White



To a lesser degree, *jet* is unfriendly to other types of **Color Vision Deficiency**

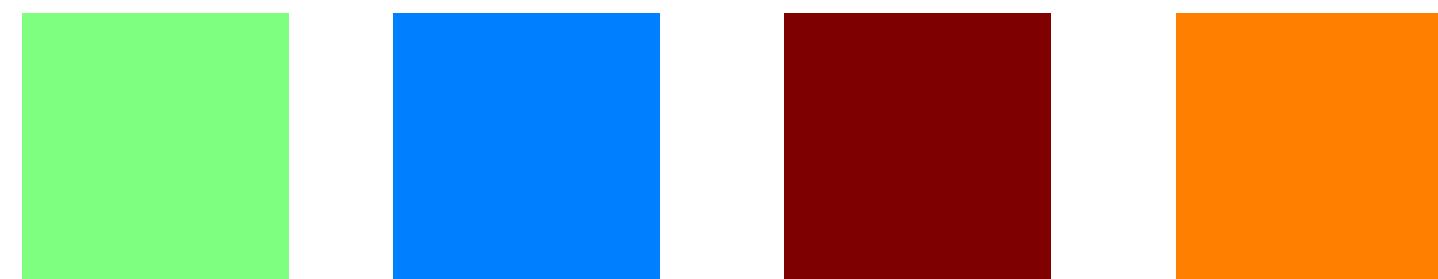
Criticisms: Ordering

Order?



Luminance is a strong
perceptual cue for **order**

Order?



However, *jet* ramps
primarily in **hue**

Design Guidelines

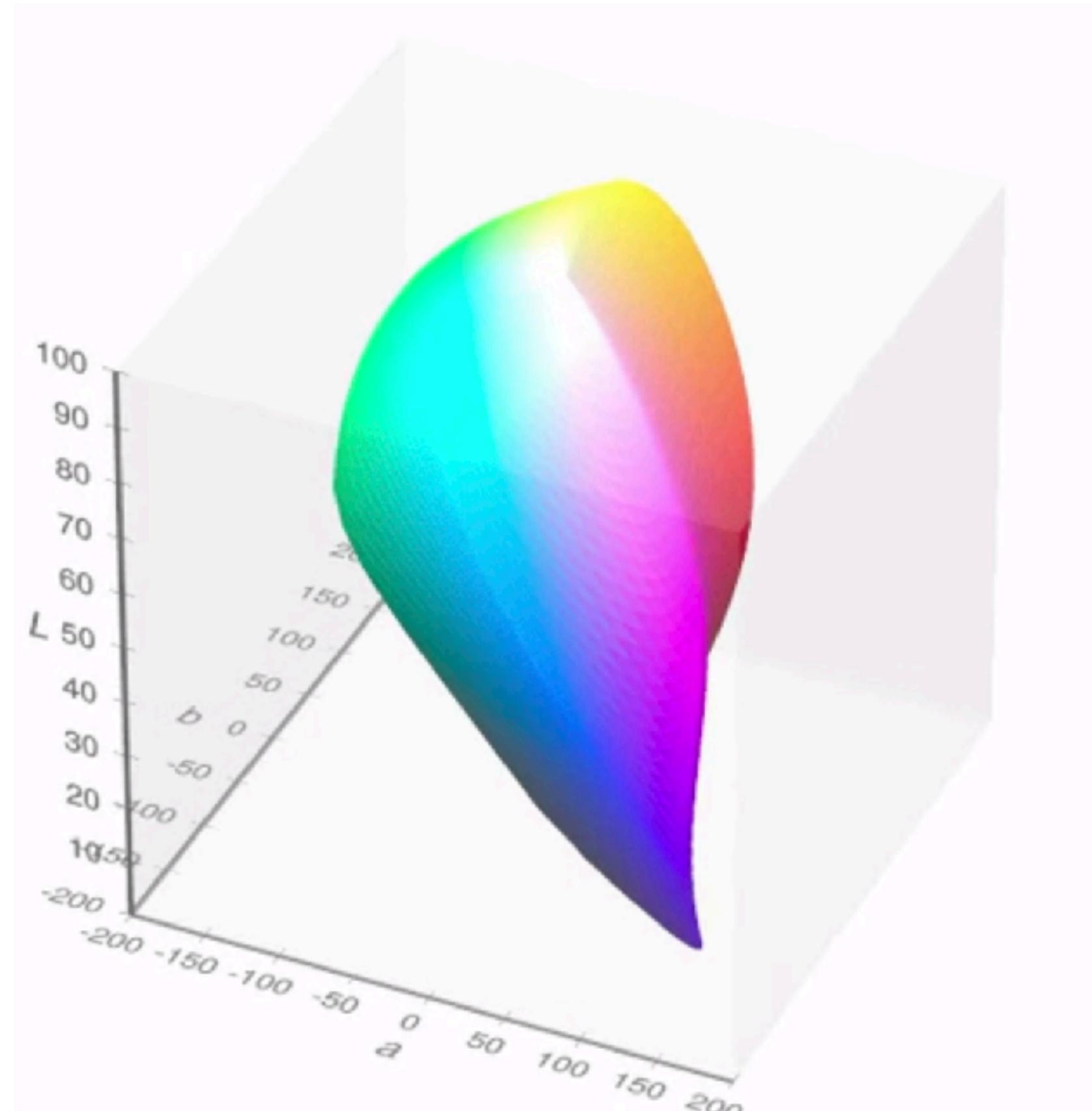
- 1 Ramp in luminance (dark to light, or vice versa)
- 2 Limit the use of hues?

The Rise of *Viridis*

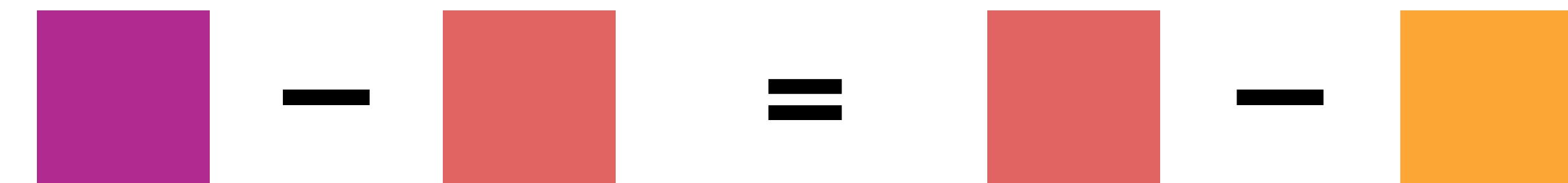


Designed using Uniform Color Space

Viridis: Uniform Color Space

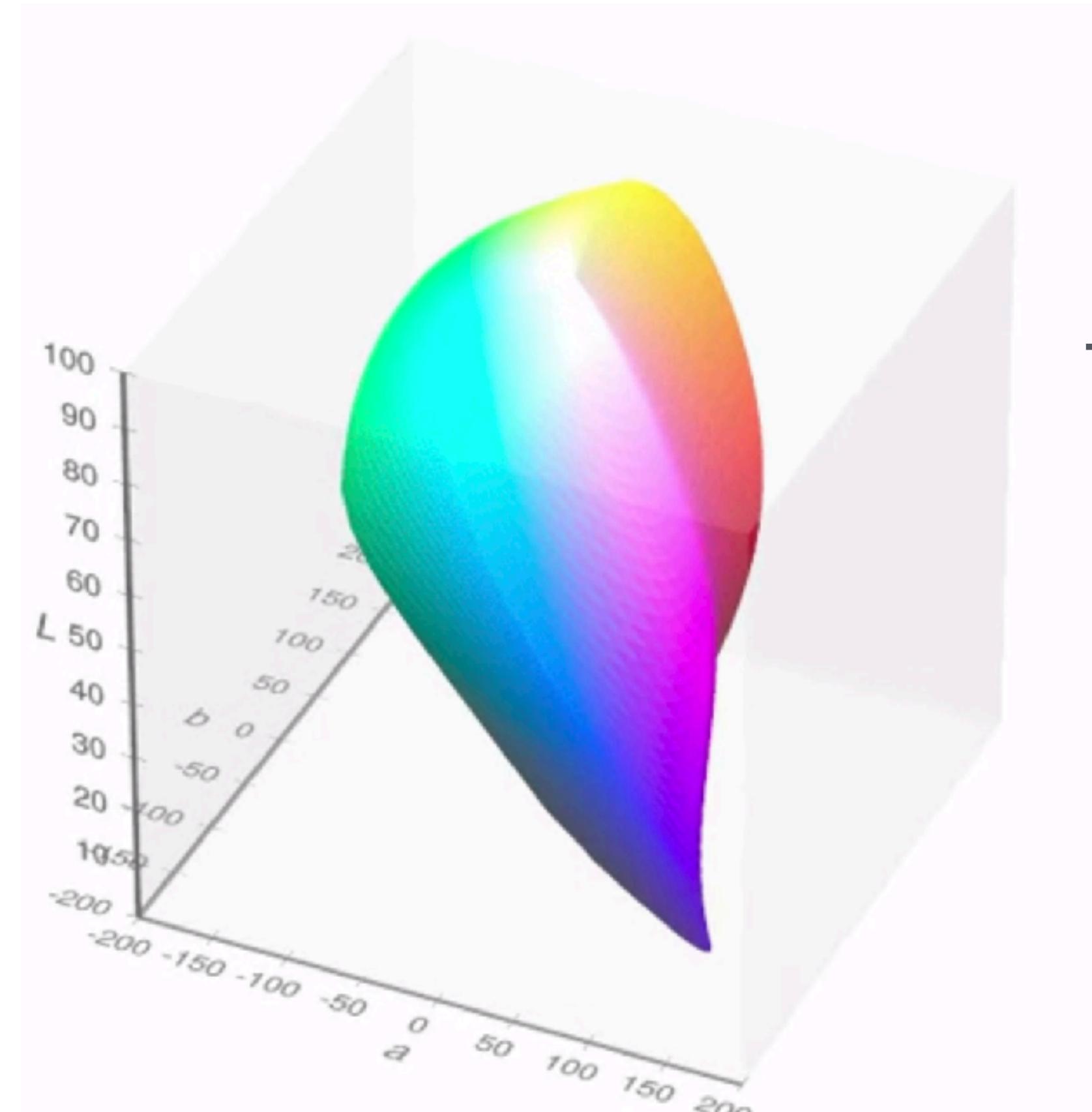


Equal distances in the space approximate
equal perceptual differences

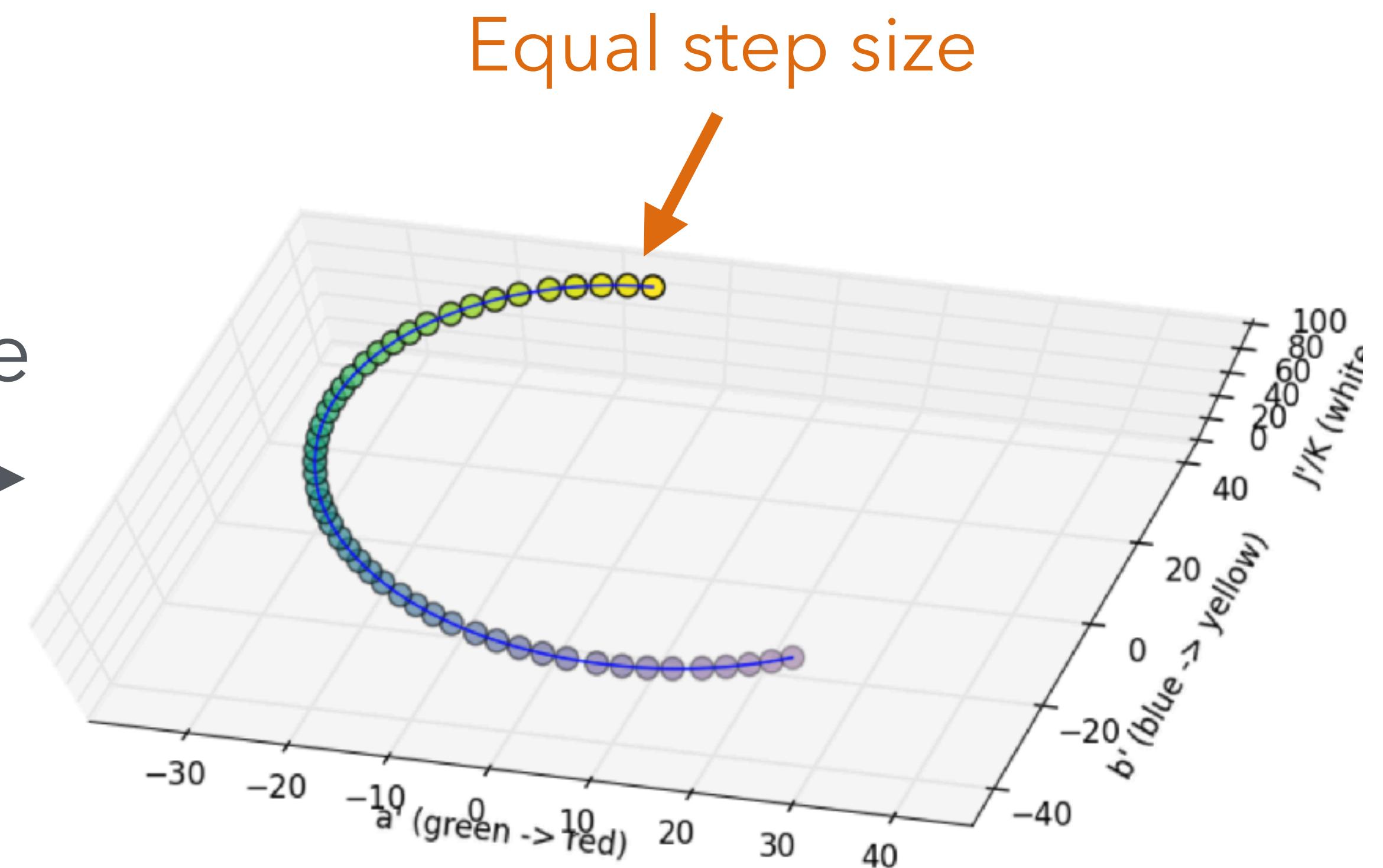


Source: https://en.wikipedia.org/wiki/Lab_color_space

Viridis: Uniform Color Space



Trace Curve
→



Source: https://en.wikipedia.org/wiki/Lab_color_space

S. Walt and N. Smith, <http://bids.github.io/colormap/>

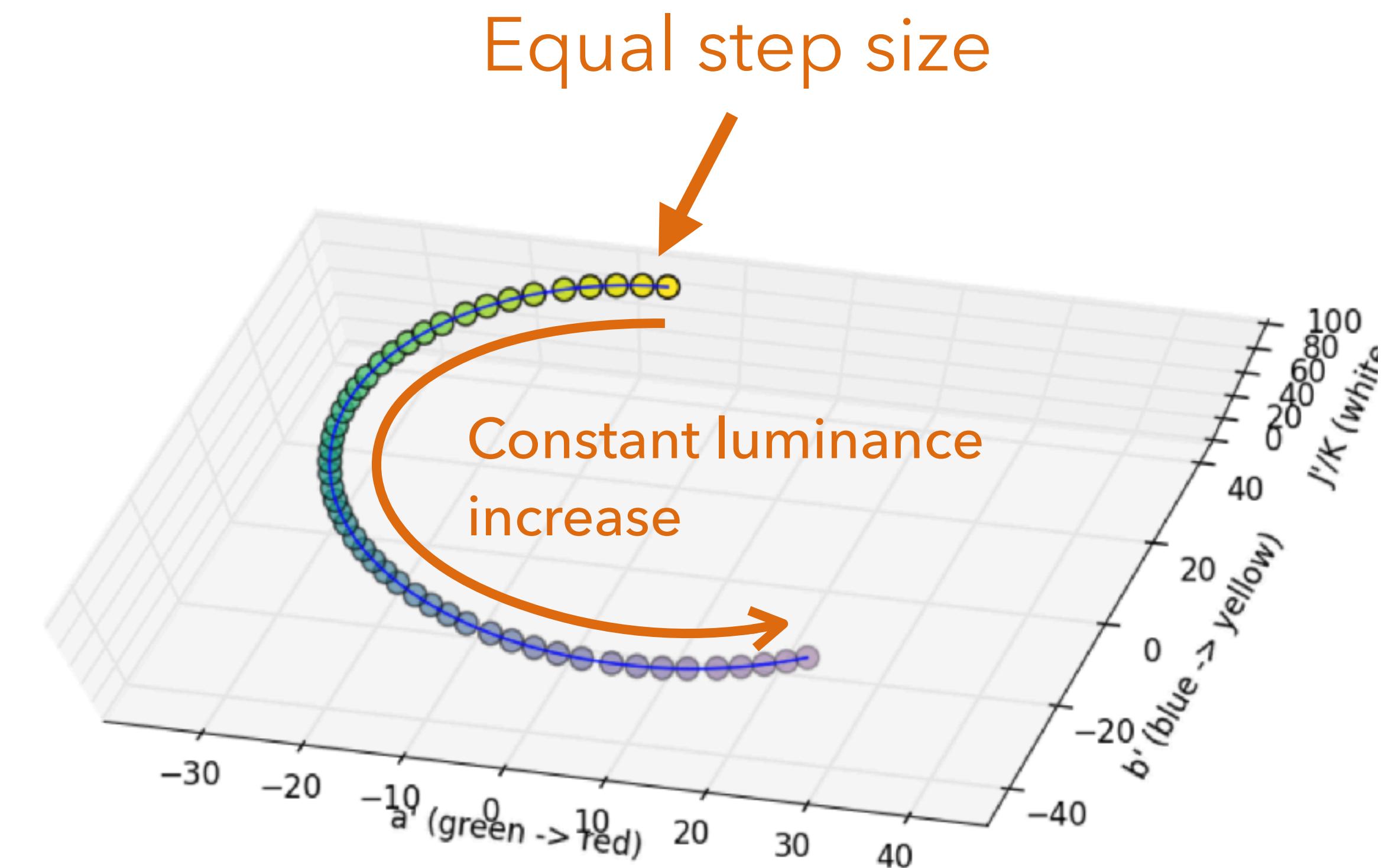
Viridis: Uniform Color Space

Design Strategies

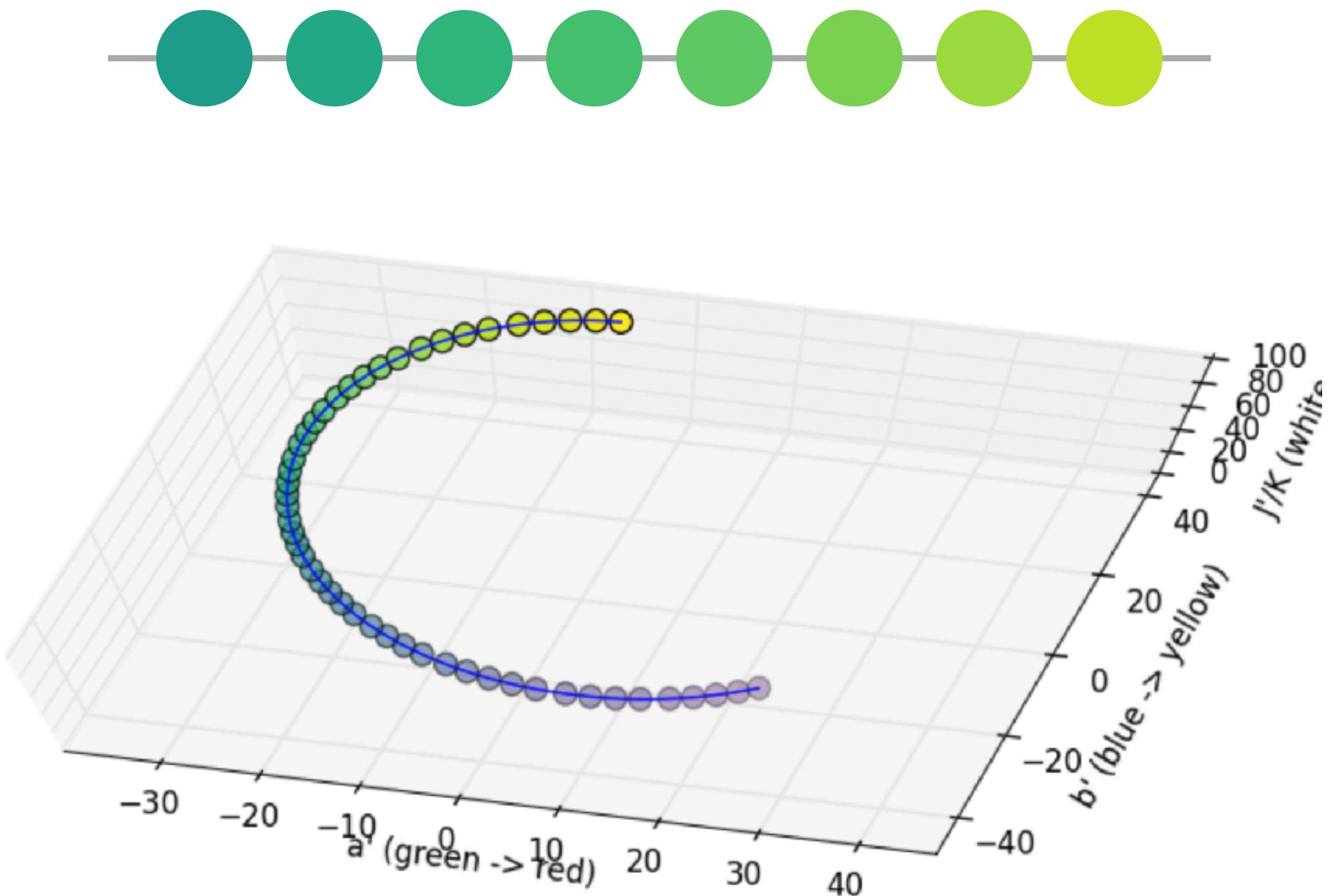
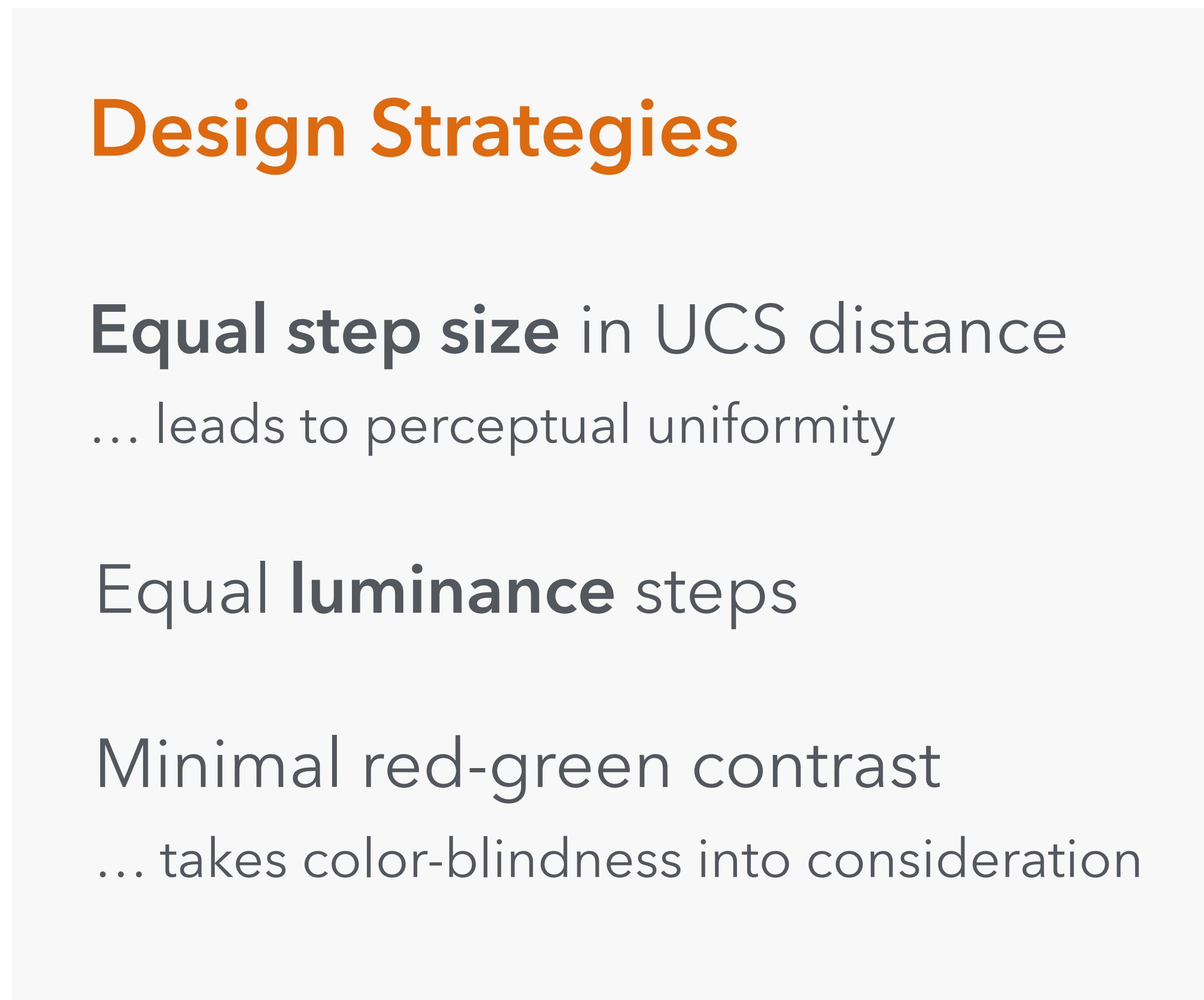
Equal step size in UCS distance

... leads to perceptual uniformity

Equal luminance steps



Viridis: Uniform Color Space



S. Walt and N. Smith, <http://bids.github.io/colormap/>

Questions

- 1 Can we use **hue** in addition to **luminance**?
- 2 Might we gain **insights** for designing more effective colormaps?

Somewhere Over the Rainbow

An Empirical Assessment of
Quantitative Colormaps

Introduction

Hues vs. luminance: can we use both?

Methods

A suite of experiments comparing 9 colormaps

Results

Global statistics and interesting special cases

Conclusion

Our contributions, limitations and takeaways

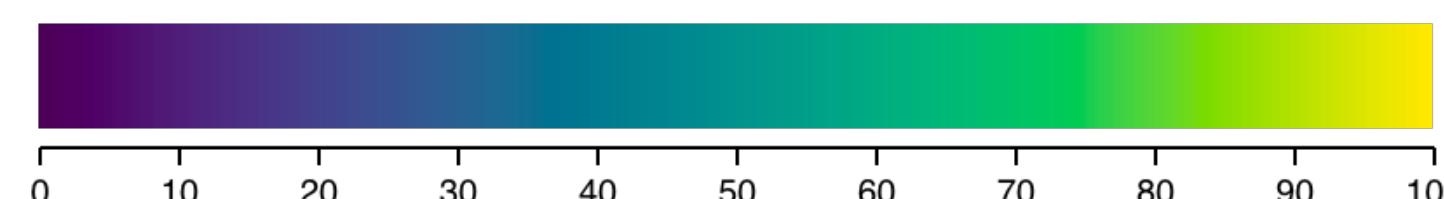
Objective

Assess the **effectiveness** of the colormaps in
encoding scalar information

Task

Value Estimation

What value does the color stand for?

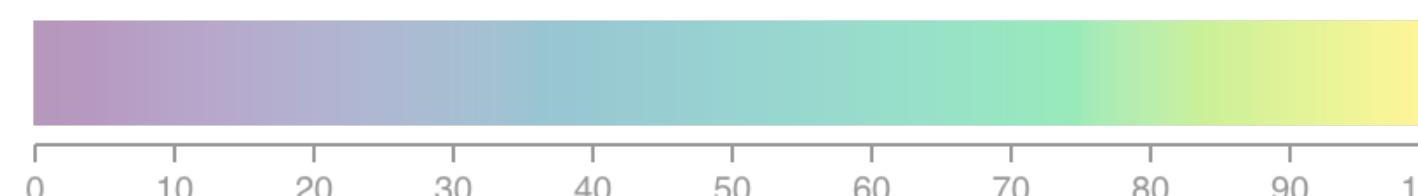


Prior work: color is a **poor** visual channel for **direct value reading**

Task

Value Estimation

What value does the color stand for?



Ordinal Judgment

Which is more similar to the reference, A or B?



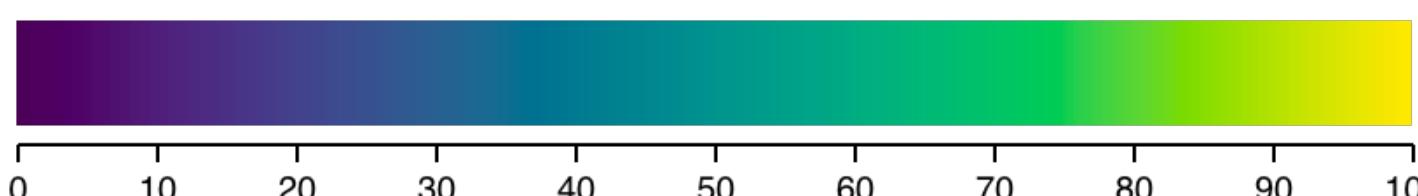
A



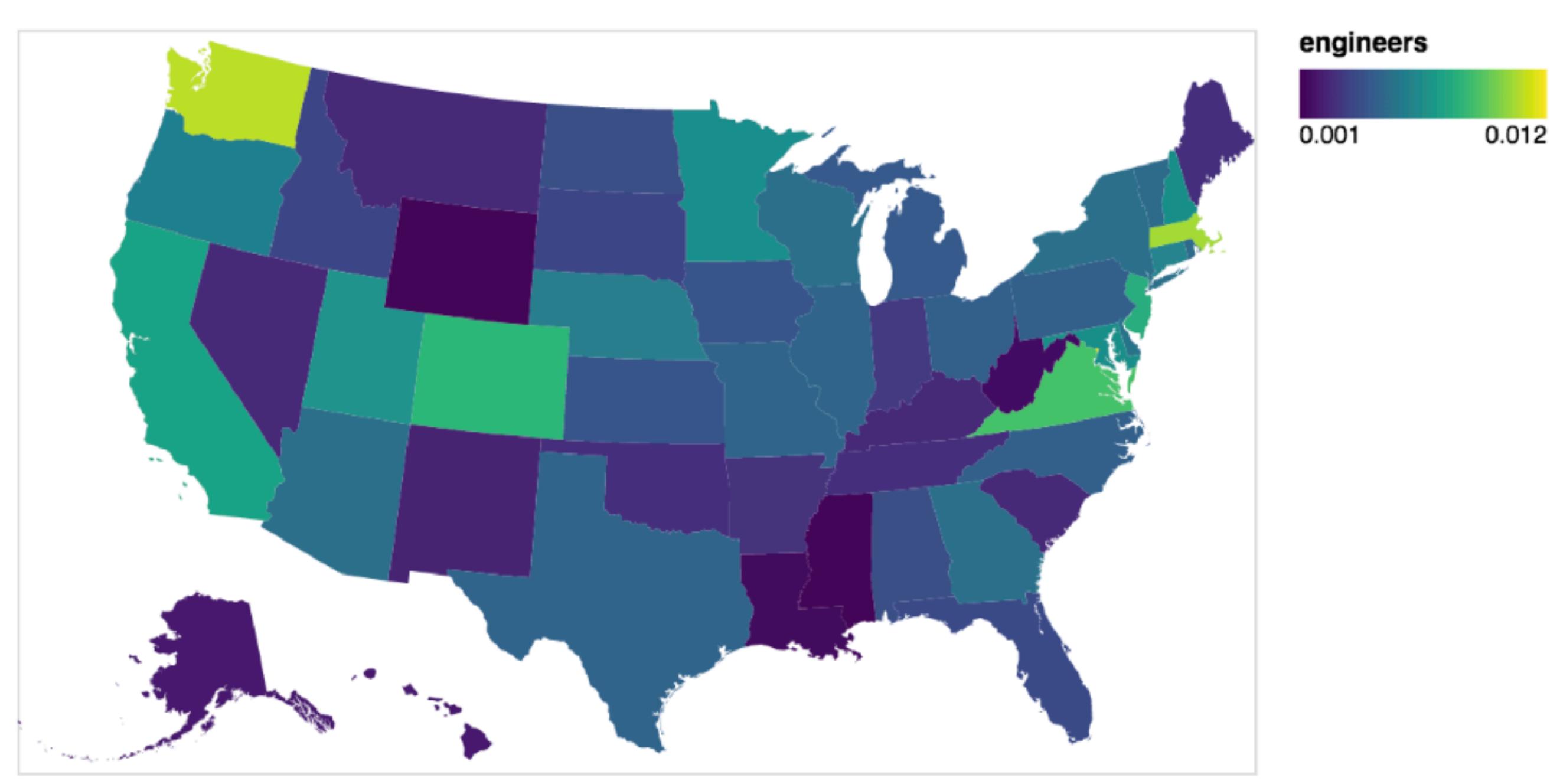
Ref



B

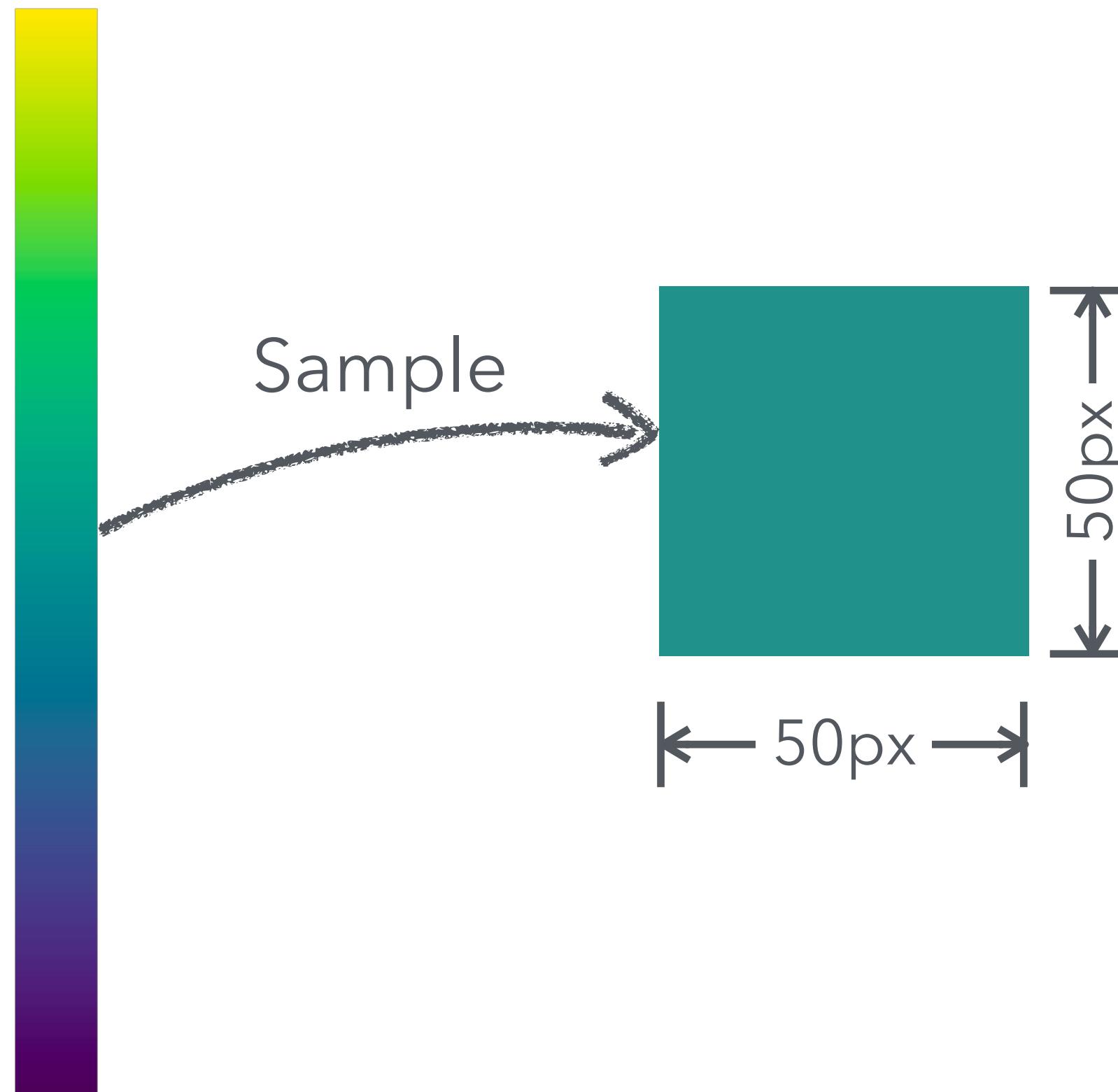


Task: Realistic Example



Which U.S. state has a rate more similar to
California: Texas or Washington?

Stimuli



Choice of Stimuli

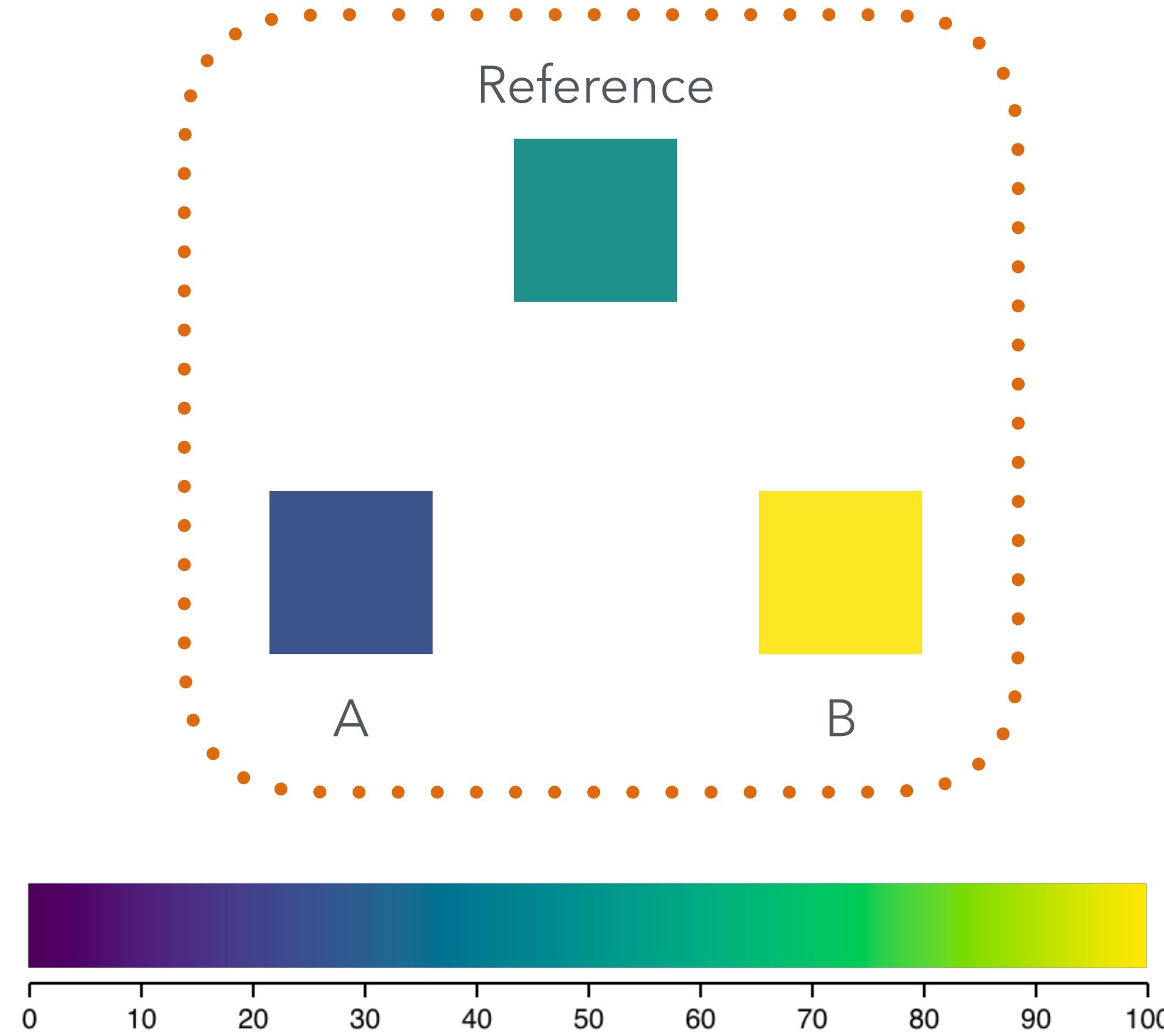
Stay closer to the conditions for which
color models are defined

Mitigate potential **confounds**:

- mark size
- simultaneous contrast
- spatial distance

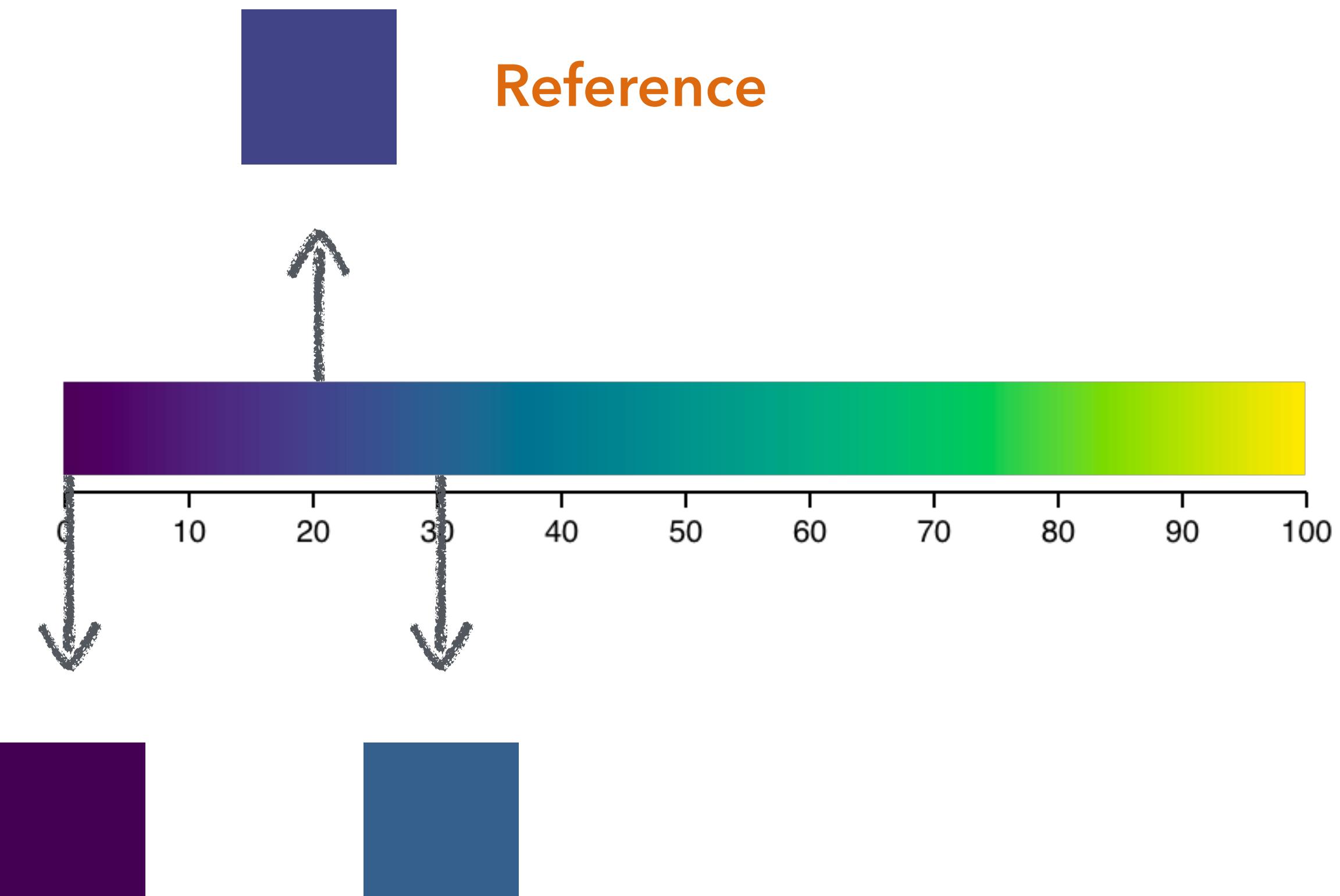
Stimuli: Triplets

Which is closer to the reference, A or B?

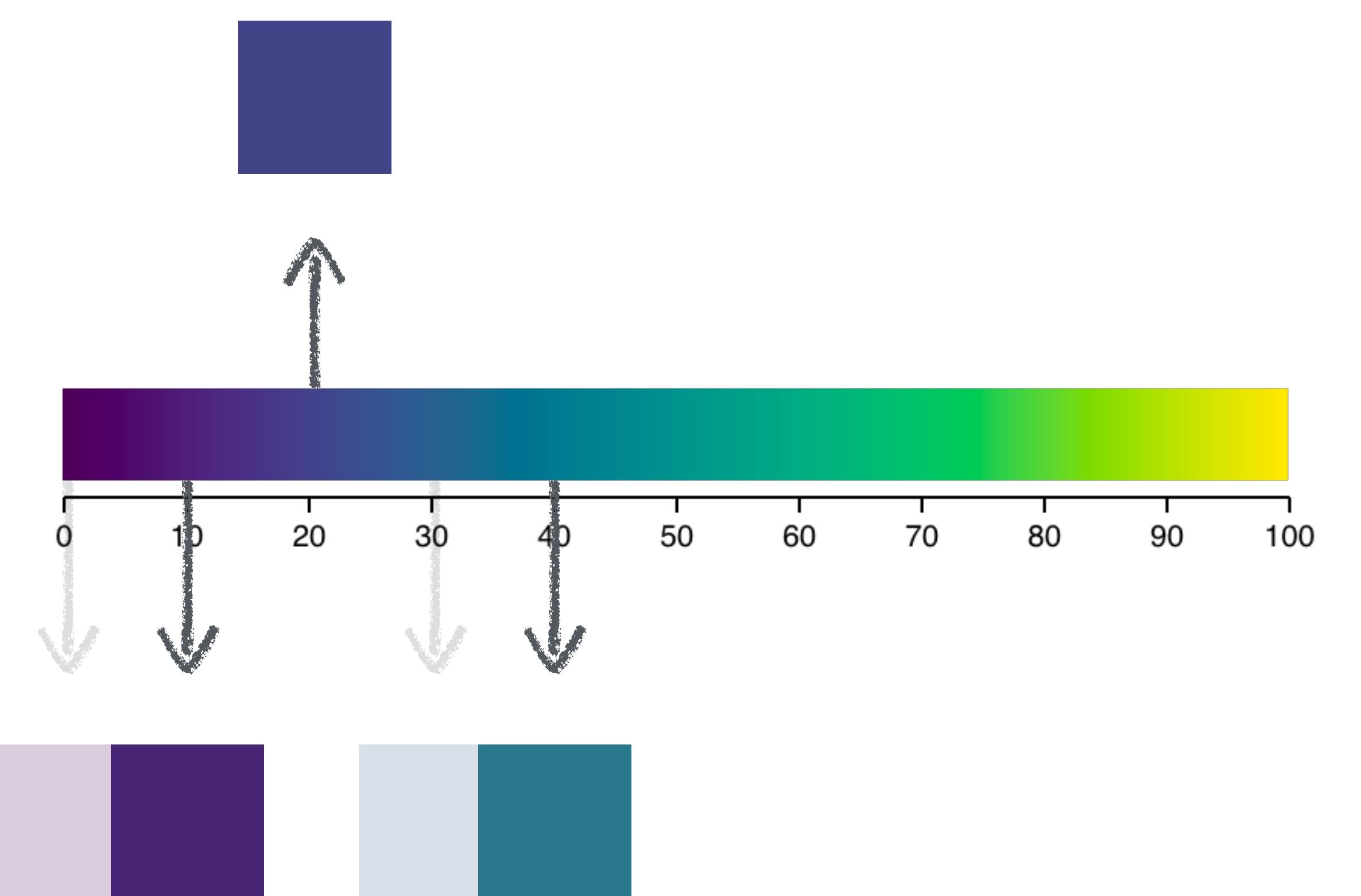


How did we generate
the triplets?

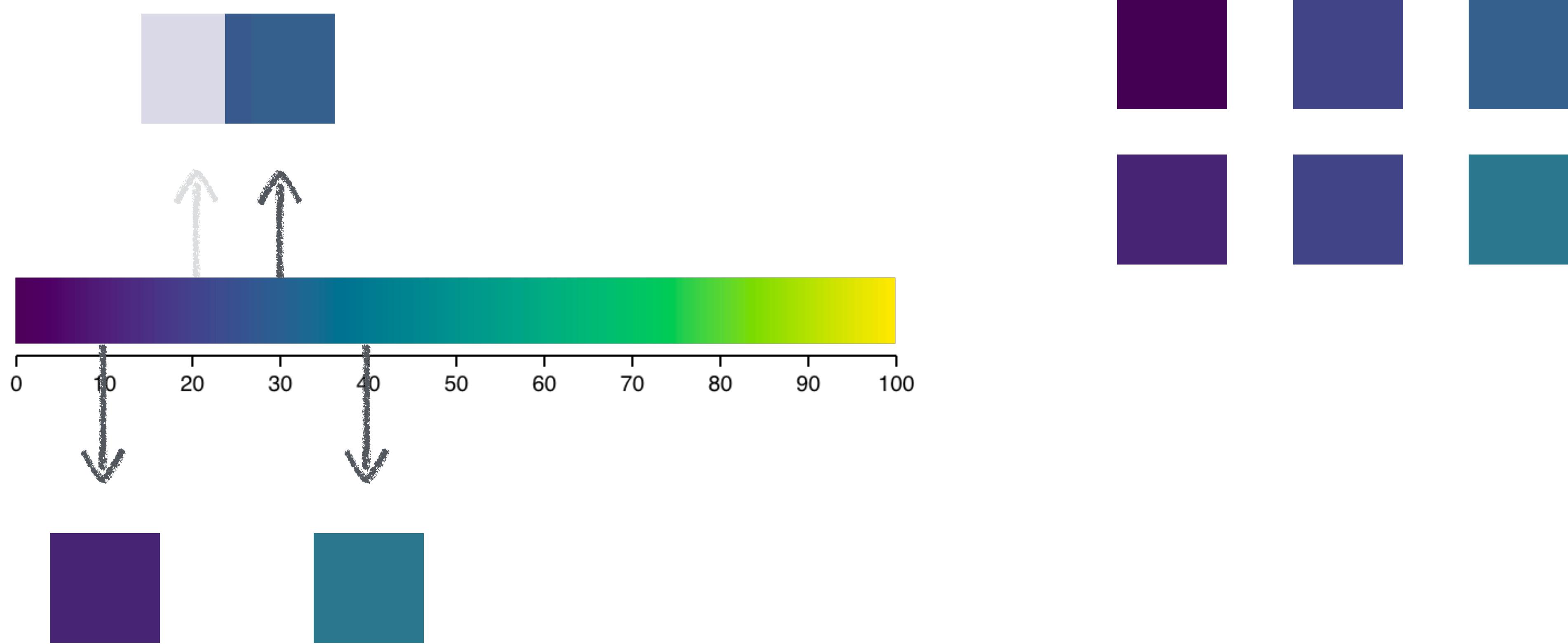
Stimuli: Triplets



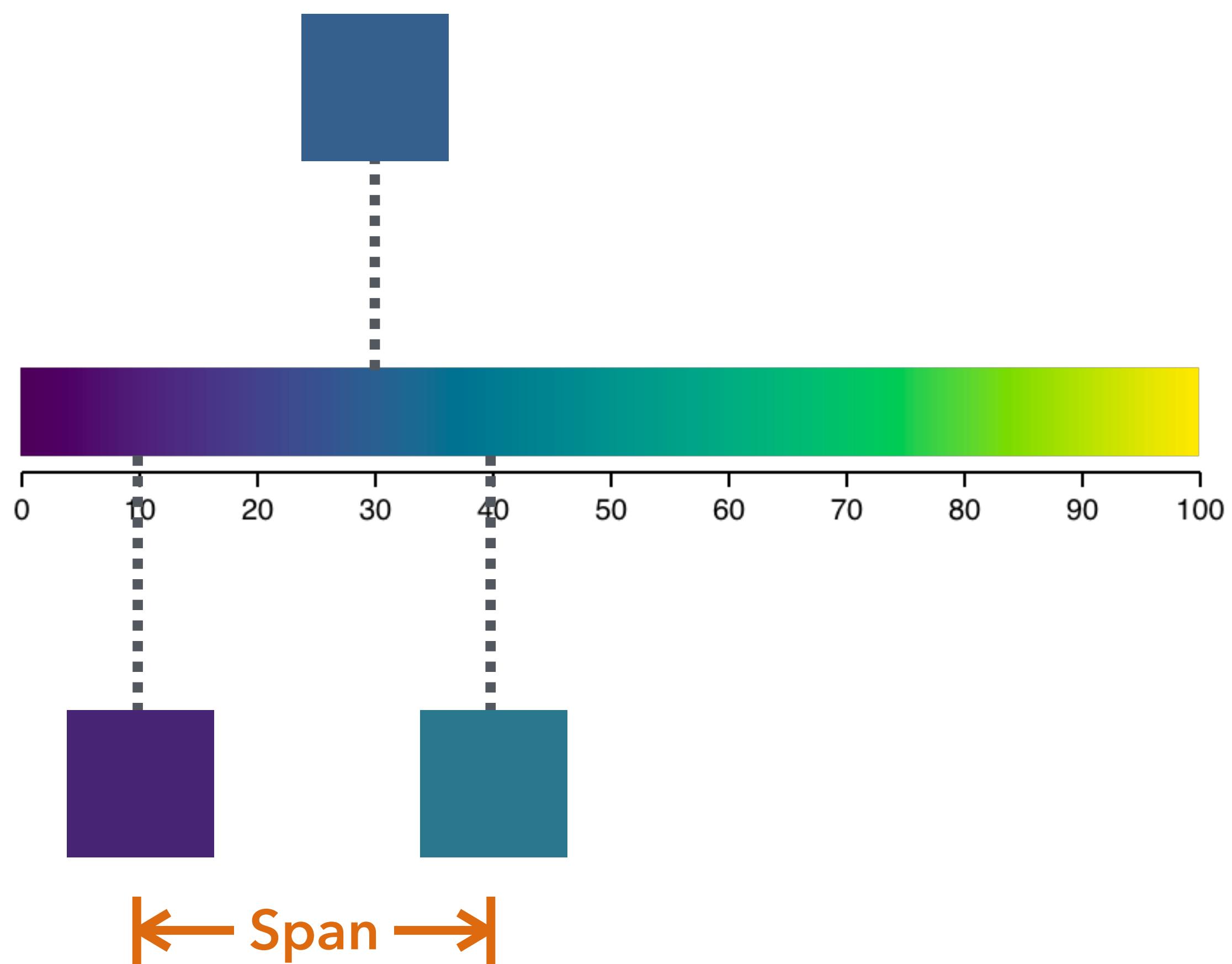
Stimuli: Triplets



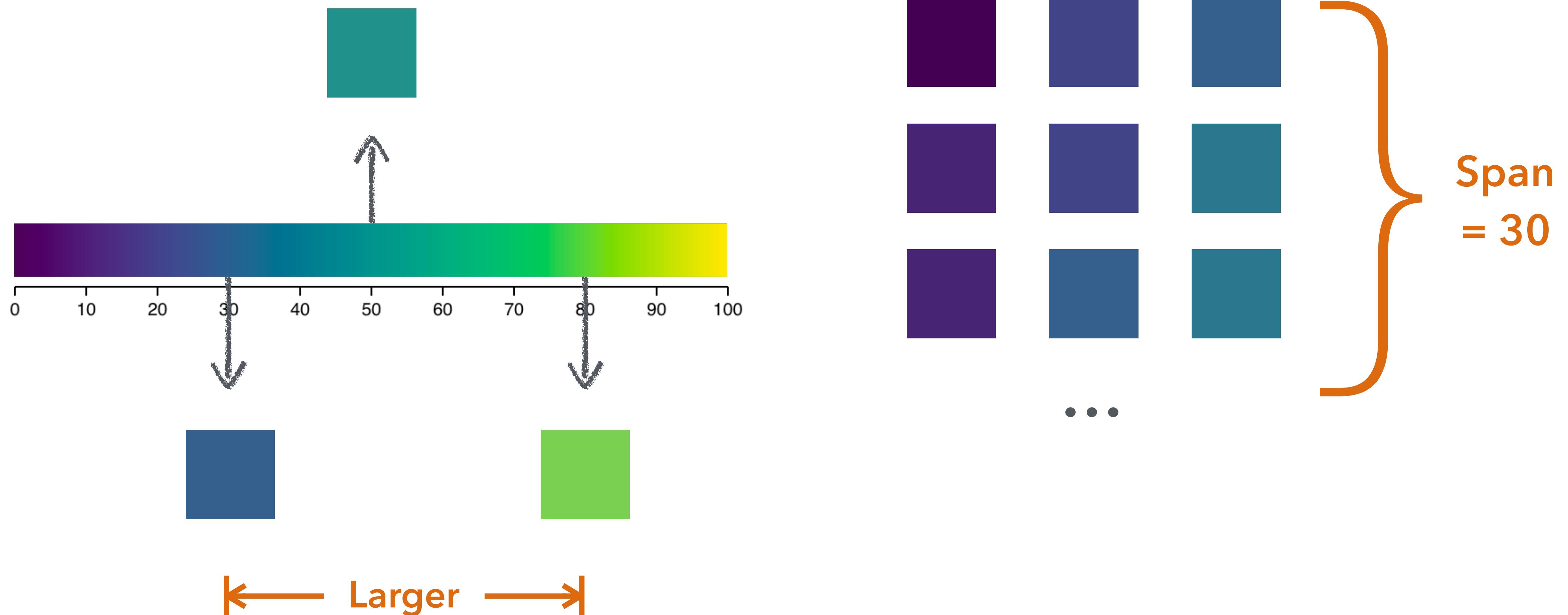
Stimuli: Triplets



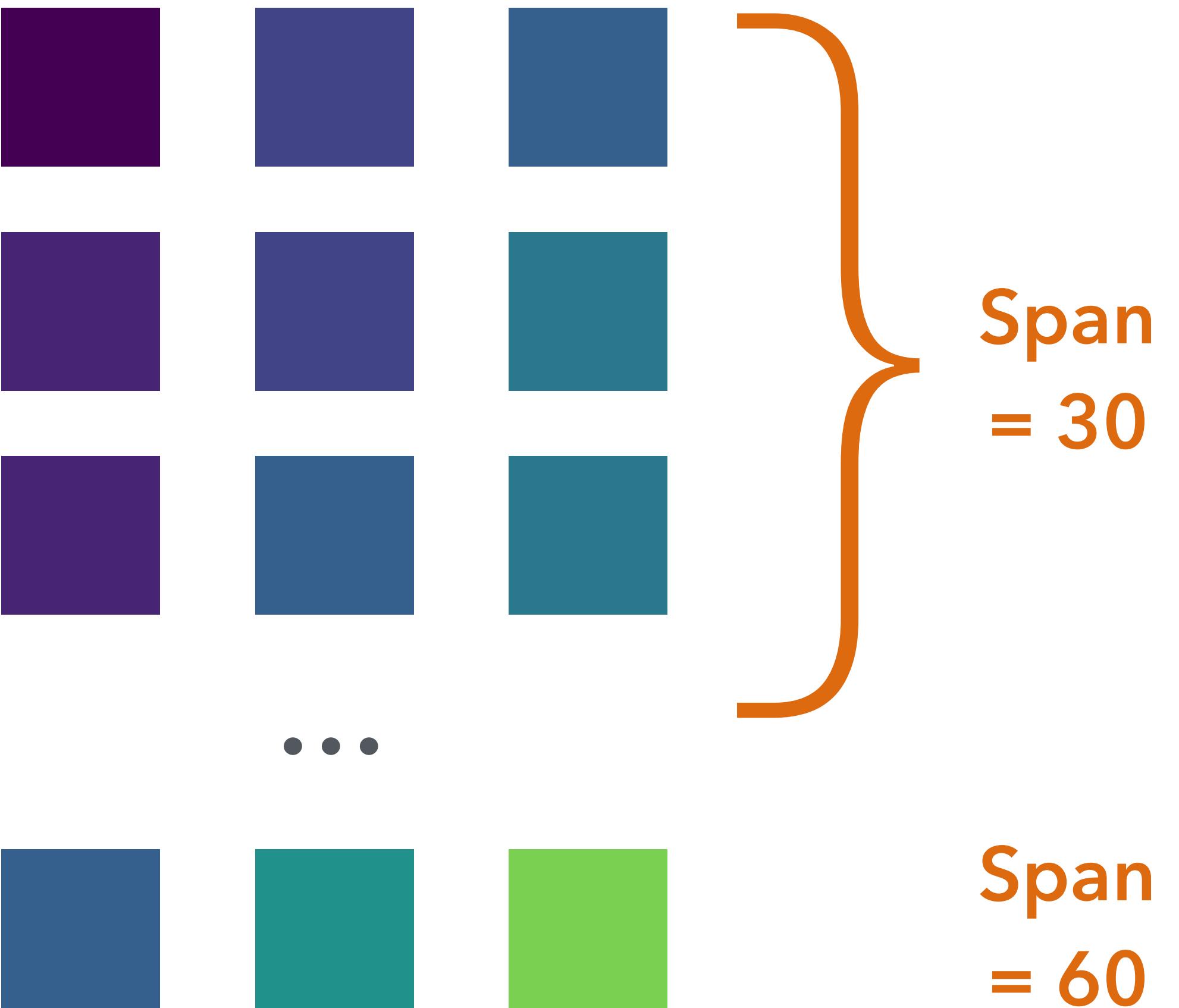
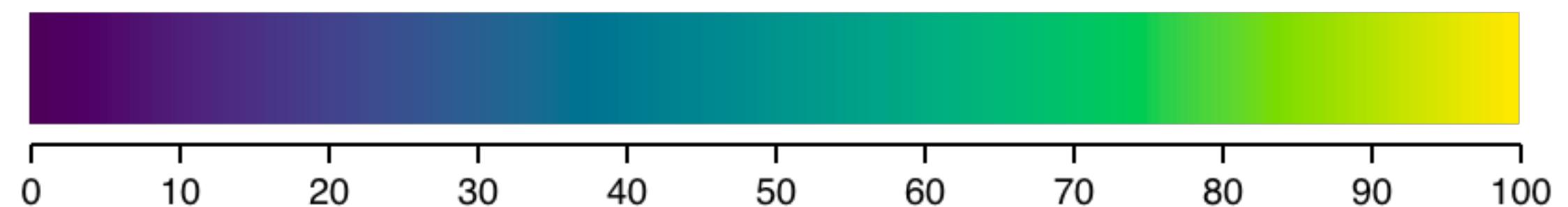
Stimuli: Triplets



Stimuli: Triplets



Stimuli: Triplets



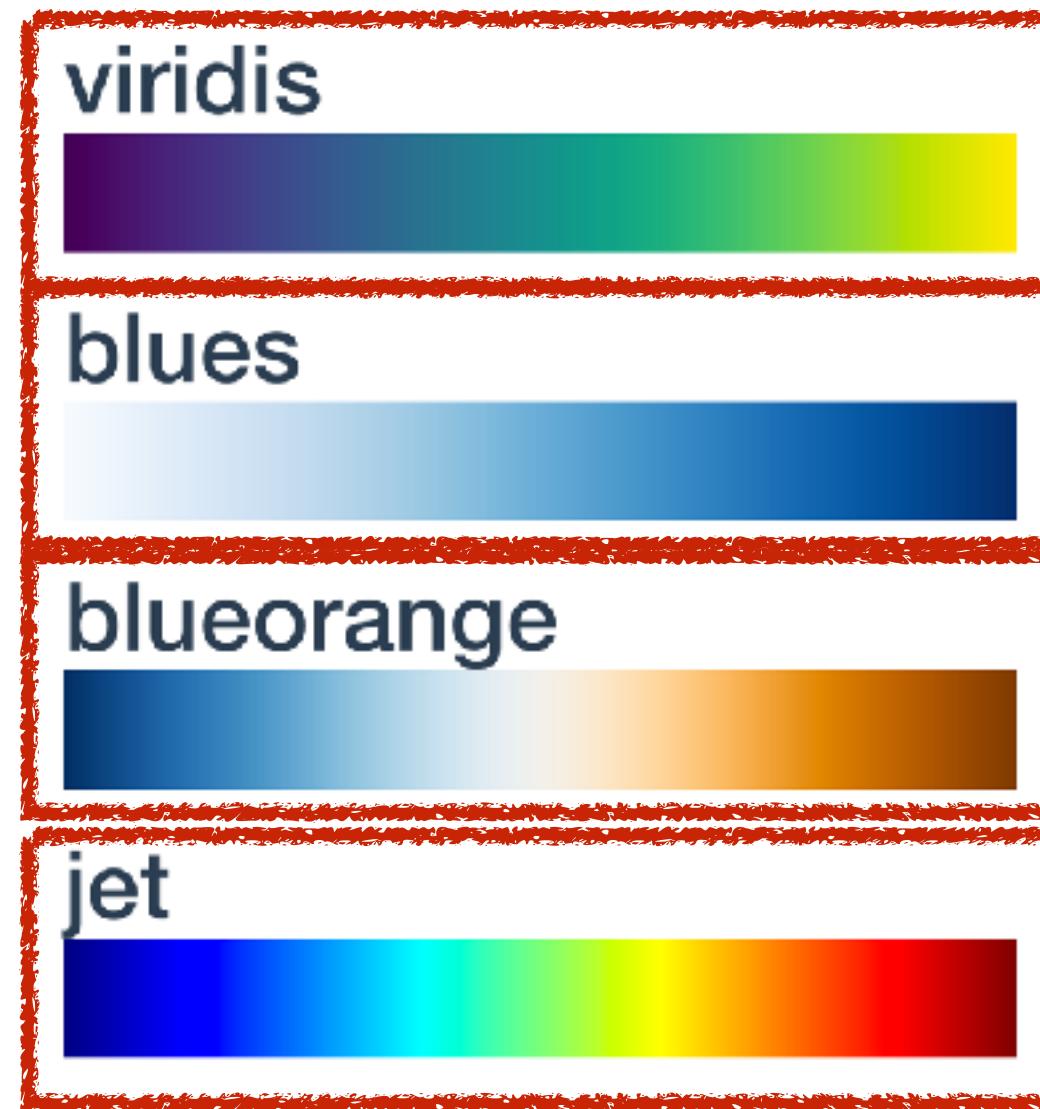
Stimuli: Triplets

3 span conditions: **15, 30, 60**

42 trials per colormap

Note: $42 = (9 + 7 + 5) \times 2$

Stimuli: Colormaps



Multi-hue sequential type

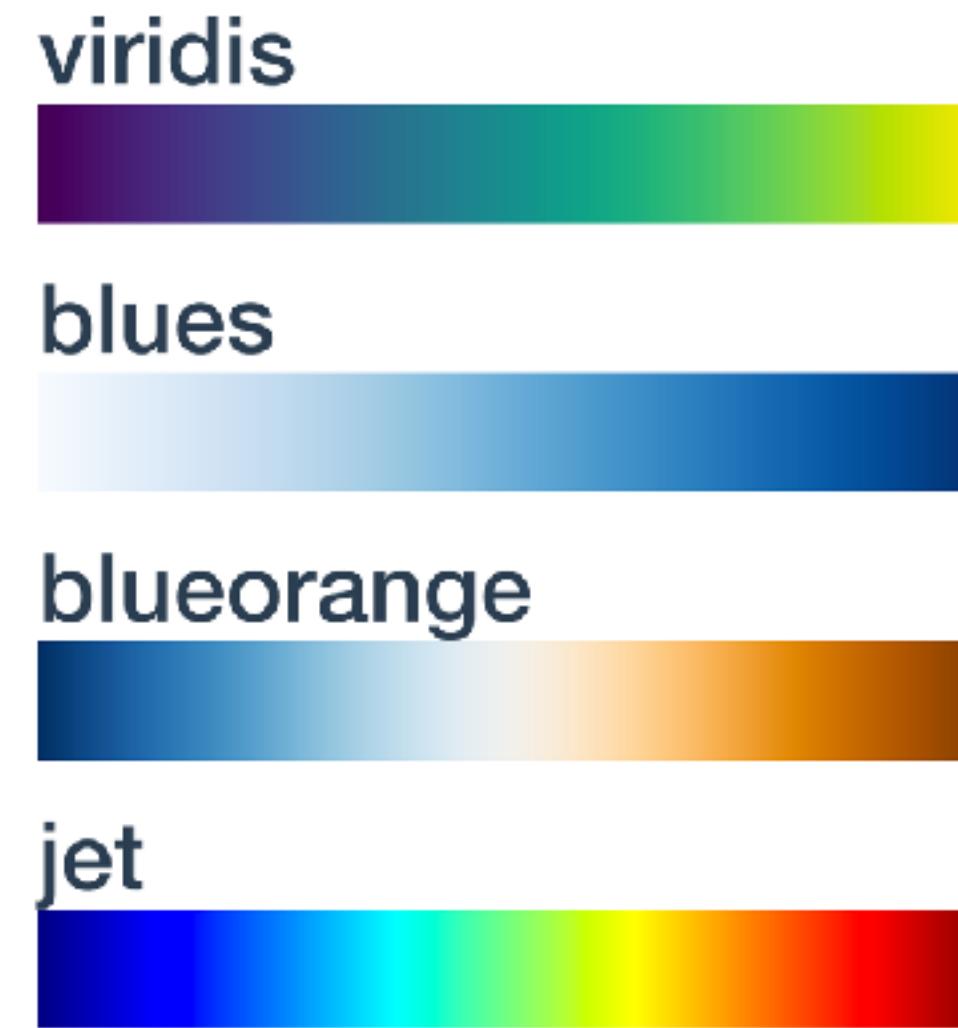
Single-hue sequential type

Diverging type

Rainbow

Assorted
Group

Stimuli: Colormaps



Assorted
Group

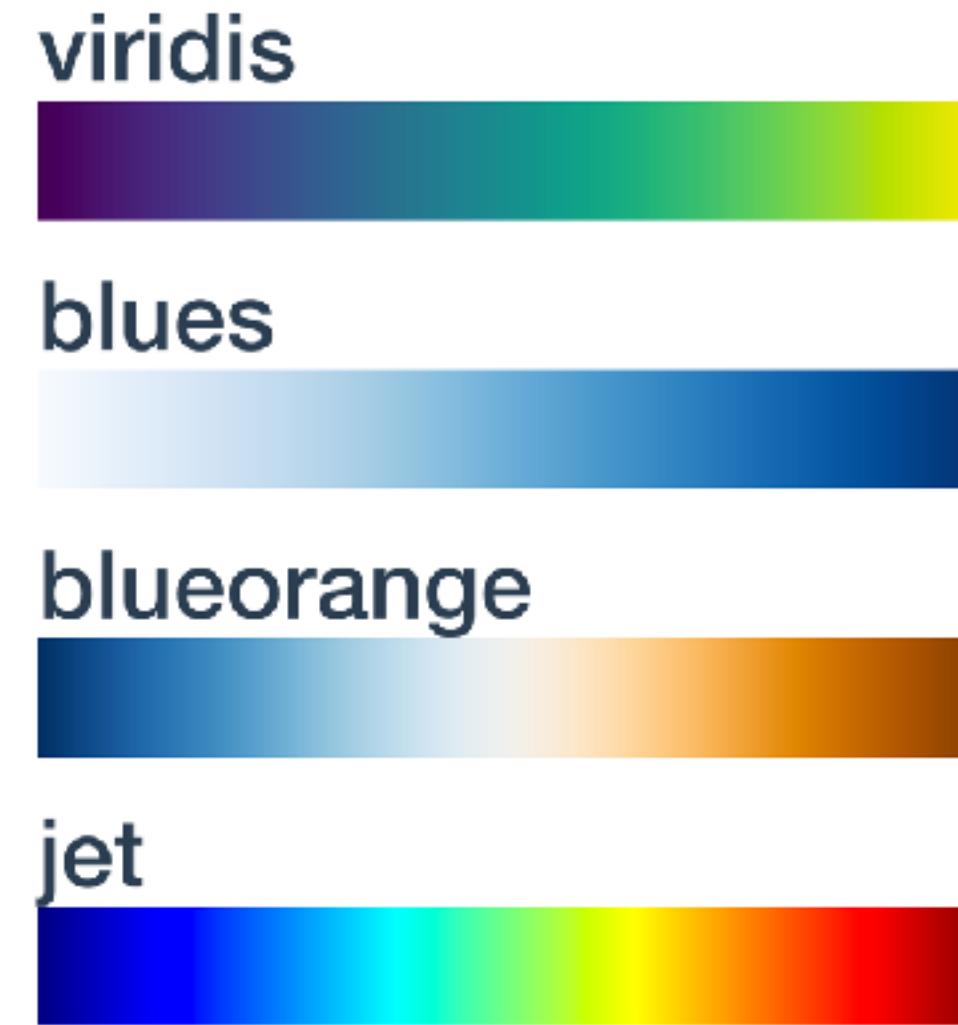


Single-hue
Group

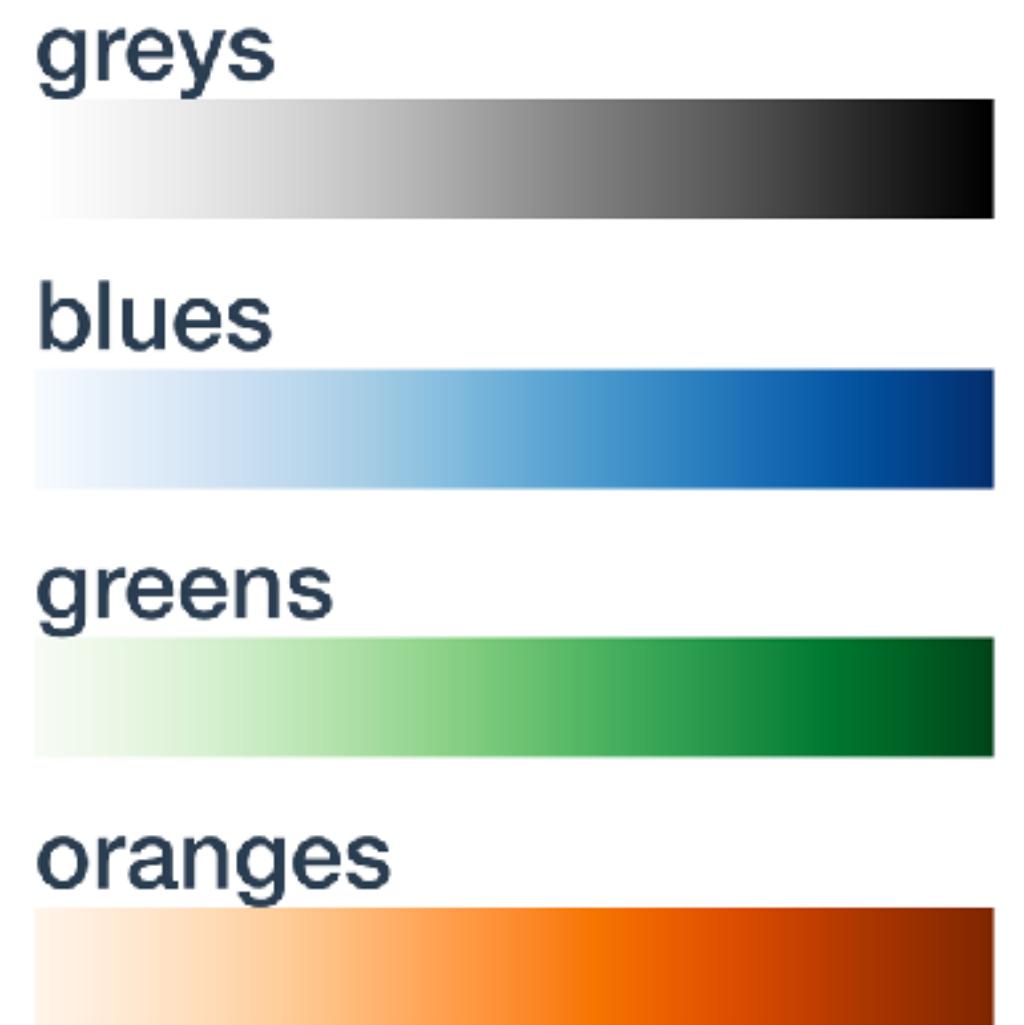
Purely Achromatic

Relatively opposing
hues in LAB space

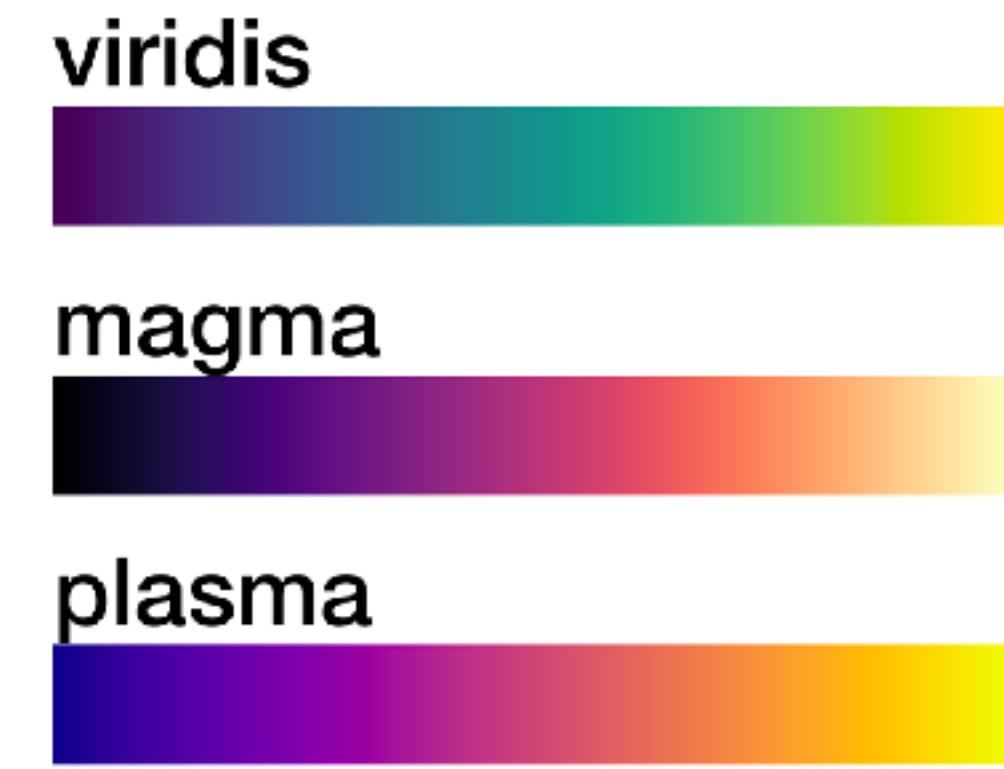
Stimuli: Colormaps



Assorted
Group



Single-hue
Group



Multi-hue UCS
Group

Procedure

Recruit ▷ Screen ▷ Intro ▷ Practice ▷ Judge

The screenshot shows the Amazon Mechanical Turk worker dashboard. At the top, it displays "Worker ID: A1R139QYQ21GYB" and "Hello, Smirzen | Sign Out". The navigation bar includes "HITs", "Dashboard", and "Qualifications", with a search bar for "Search All HITs". Below the navigation, there are tabs for "All HITs" and "Your HITs Queue". The main content area is titled "HIT Groups (1-20 of 1343)" and features a large banner for "Amazon Mechanical Turk" with the sub-instruction "Locate unboxed fashion items in the image (DON'T FORGET BELT!!)". A table lists 13 HIT groups, each with a title, description, number of workers, reward, creation time, preview link, and an "Accept & Work" button. The first HIT group is for "Parts Availability and Pricing" by "Ray Tsai". Other groups include "Opinion Survey" by "UnSpun Opinions", "Choose the best photo" by "TripAdvisor Machine Learning", and various damage assessment tasks from "RS-MT-Requestor".

Rank	HIT Group	Description	Workers	Reward	Created	Actions
1	Ray Tsai	Parts Availability and Pricing	11,604	\$0.02	1d ago	Preview Accept & Work
2	UnSpun Opinions	Opinion Survey	11,066	\$0.02	2h ago	Preview Accept & Work
3	TripAdvisor Machine Learning	Choose the best photo	9,951	\$0.50	7m ago	Preview Accept & Work
4	Crowdsurf Support	Transcribe up to 35 Seconds of Media to Text - Earn up to \$0.17 per HIT!!	9,776	\$0.01	6h ago	Preview Accept & Work
5	CCMP5	Which Product Type is the most relevant? (v3.1D)	6,902	\$0.05	4m ago	Preview Accept & Work
6	RS-MT-Requestor	Building Damage Assessment; Batch:FM_40k_2_2_2018-04-12_14:28:55	5,172	\$0.05	7h ago	Preview Accept & Work
7	RS-MT-Requestor	Building Damage Assessment; Batch:FM_40k_2_1_2018-04-12_14:28:55	5,000	\$0.04	10h ago	Preview Accept & Work
8	RS-MT-Requestor	Building Damage Assessment; Batch:FM_40k_1_2_2018-04-12_13:51:41	5,000	\$0.04	10h ago	Preview Accept & Work
9	UnSpun Opinions	Opinion Survey	2,921	\$1.00	27m ago	Preview Accept & Work
10	f8b64e4e-b7c8-47a8-9ee3-b161727f1ca2	Judge the reputation polarity of Article Clips	2,583	\$0.08	6h ago	Preview Accept & Work
11	Usability Tester	Quick survey - approx. 4 minutes	1,627	\$0.80	2h ago	Preview Accept & Work
12	Noah	Write a paragraph describing events in each video, then label the time range for each event.	1,558	\$1.30	21h ago	Preview Accept & Work
13	DAEHO KIM	Drawing Bounding Box and Finding Relations Between Objects	1,231	\$0.03	2d ago	Preview Accept & Work

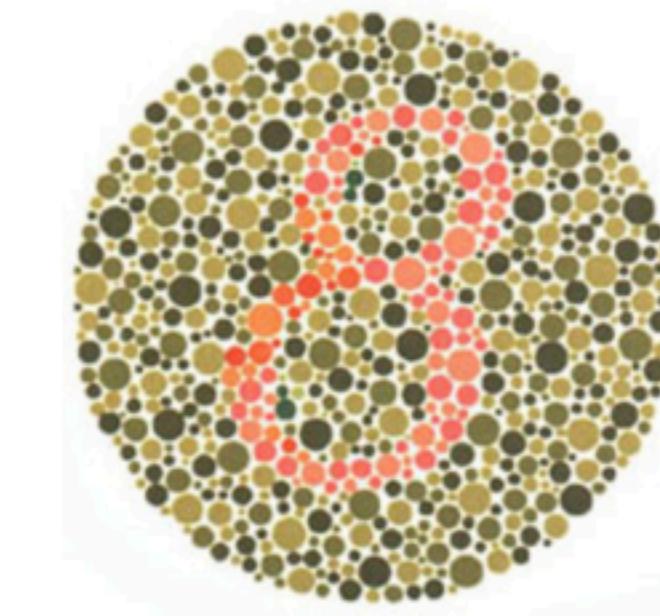
Procedure

Recruit ▷ Screen ▷ Intro ▷ Practice ▷ Judge

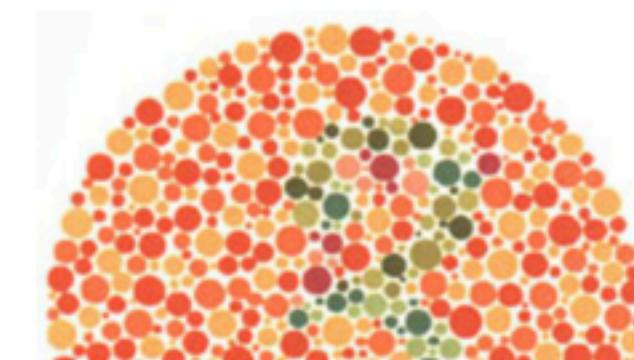
Inclusion Criteria

Before we proceed to the experiment, we want to make sure that you have normal color vision. The questions below aim to test if you have color vision deficiency.

Below, each image contains a number. Type the corresponding number in the box to pass the test.



What is the number?

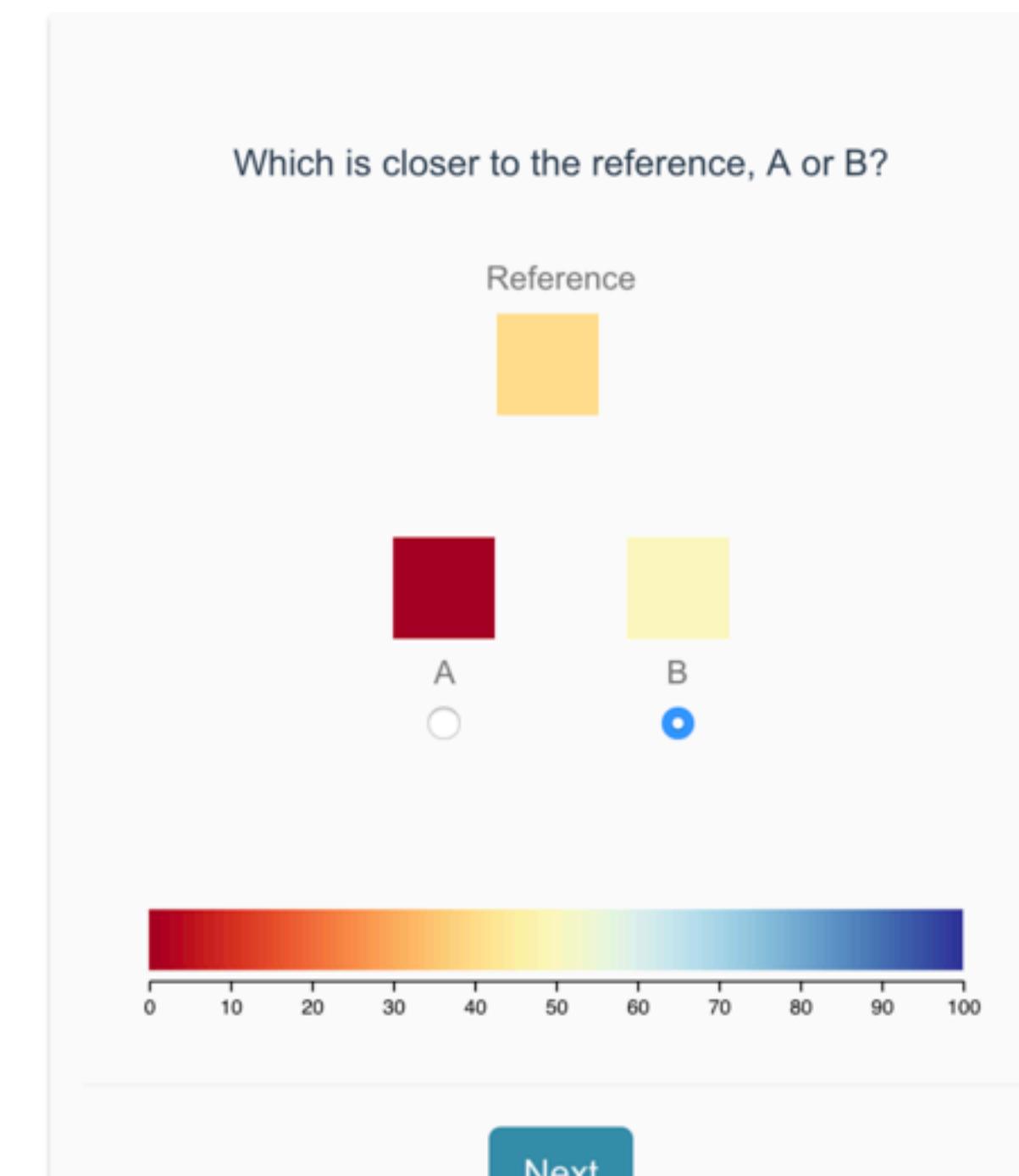


Procedure

Recruit ▷ Screen ▷ Intro ▷ Practice ▷ Judge

Tutorial

In each question, you will be presented with a triplet of colors. Your task is to judge which of the two colors, A or B, is closer to the reference. For example, in the question below, the answer is B:



Procedure

Recruit ▷ Screen ▷ Intro ▷ Practice ▷ Judge

Which is closer to the reference, A or B?

Reference

A color calibration task. At the top, the text "Which is closer to the reference, A or B?" is displayed. Below this, the word "Reference" is followed by a blue square. Below the reference are two other blue squares labeled "A" and "B". Square "A" is positioned at approximately 40% on a color scale, and square "B" is positioned at approximately 70%. Below the squares is a horizontal color bar with a gradient from red to blue, with numerical labels from 0 to 100. At the bottom of the slide is a teal button labeled "Next".

0 10 20 30 40 50 60 70 80 90 100

A B

Next

Procedure

Recruit ▷ Screen ▷ Intro ▷ Practice ▷ Judge

Which is closer to the reference, A or B?

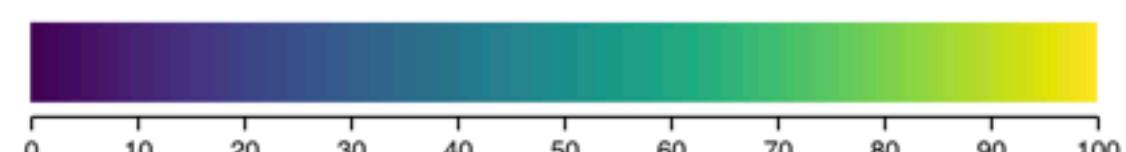
Reference



A



B



Next

Somewhere Over the Rainbow

An Empirical Assessment of
Quantitative Colormaps

Introduction

Hues vs. luminance: can we use both?

Methods

A suite of experiments comparing 9 colormaps

Results

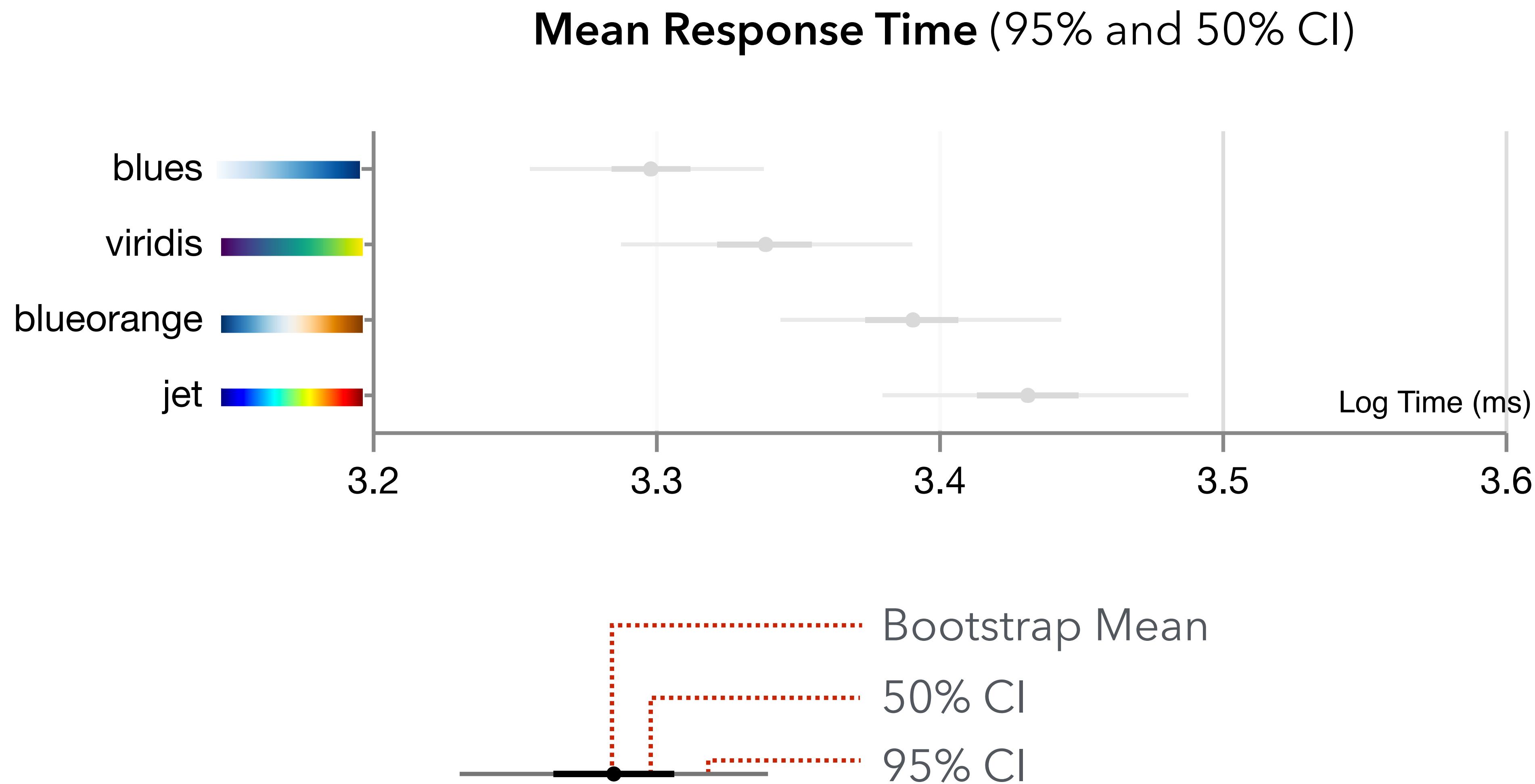
Global statistics and interesting special cases

Conclusion

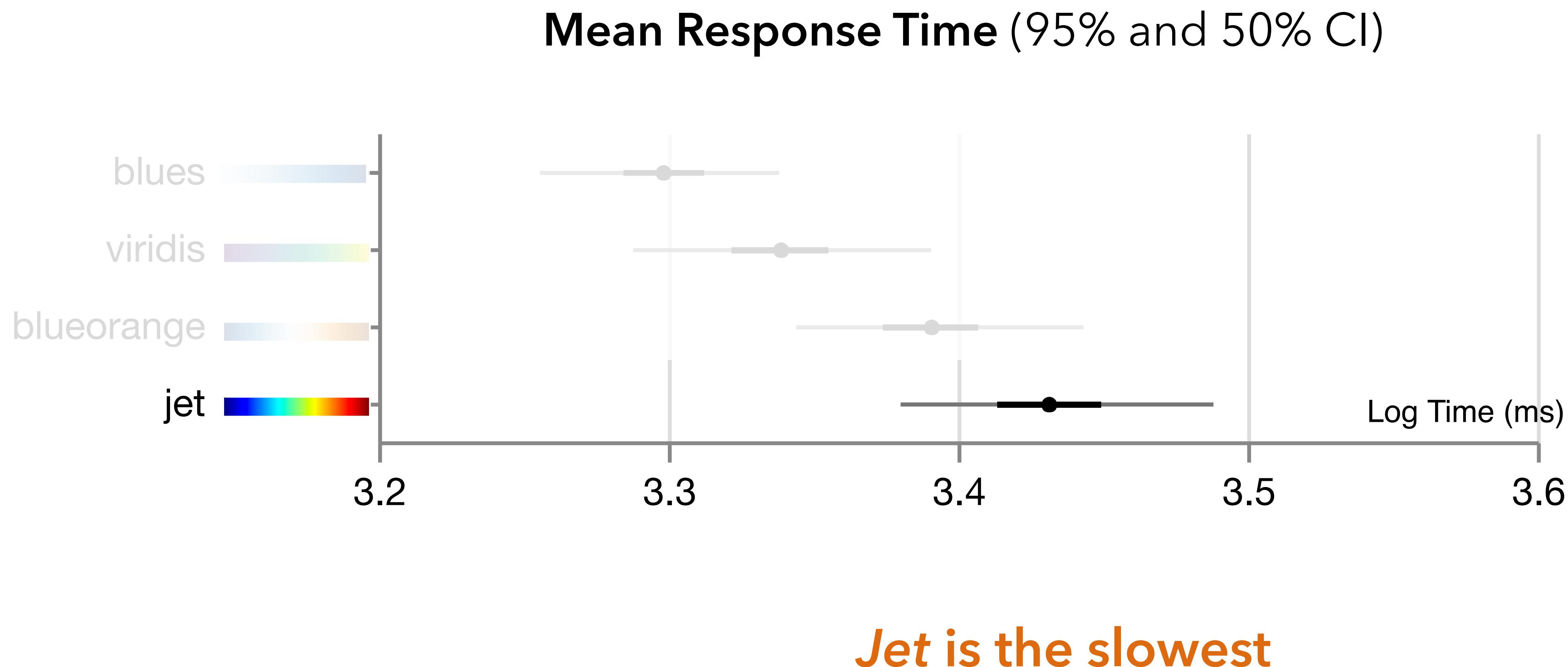
Our contributions, limitations and takeaways

Global Time and Error

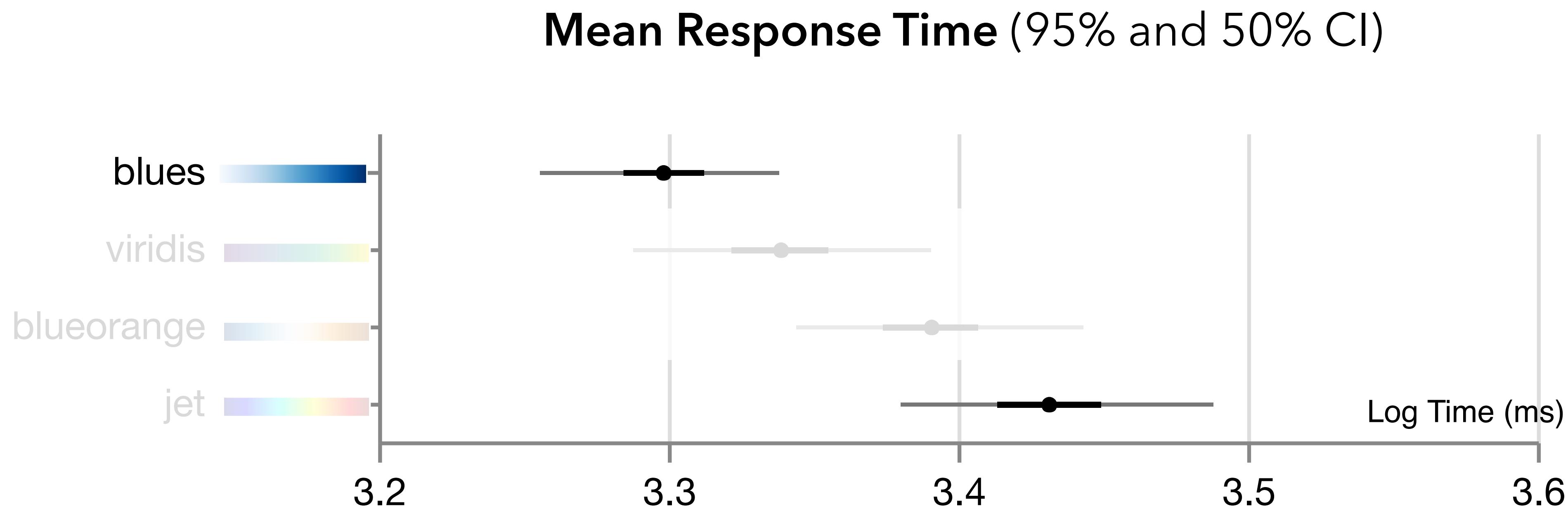
Assorted Colormaps: Time



Assorted Colormaps: Time

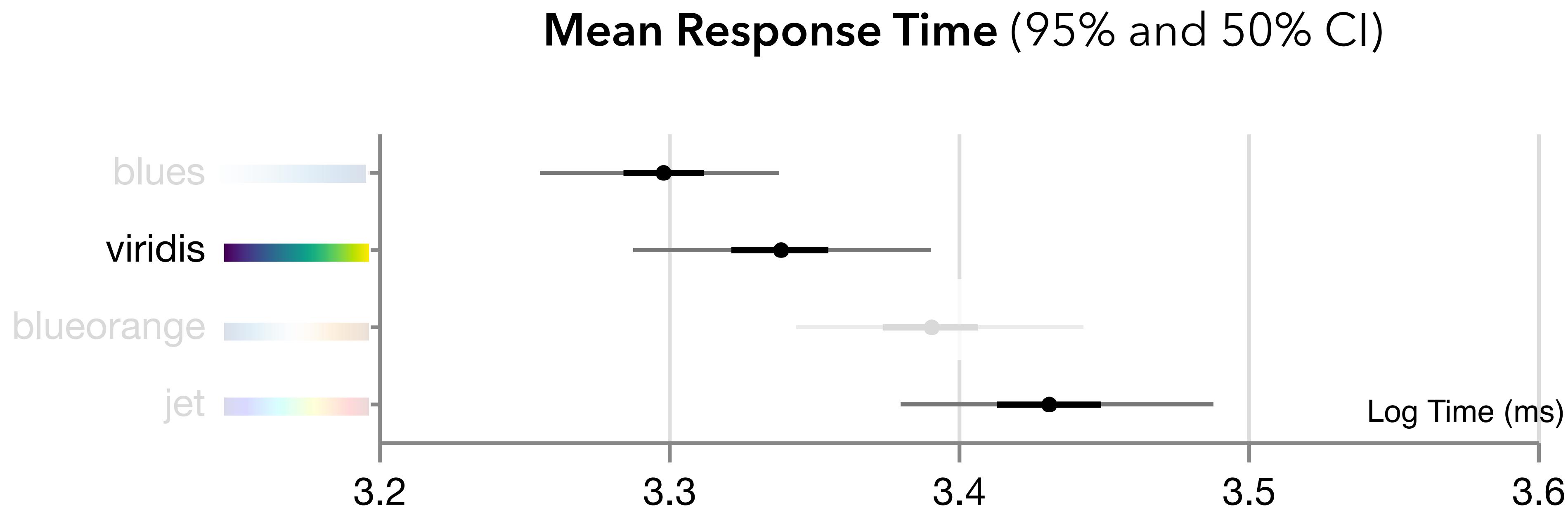


Assorted Colormaps: Time



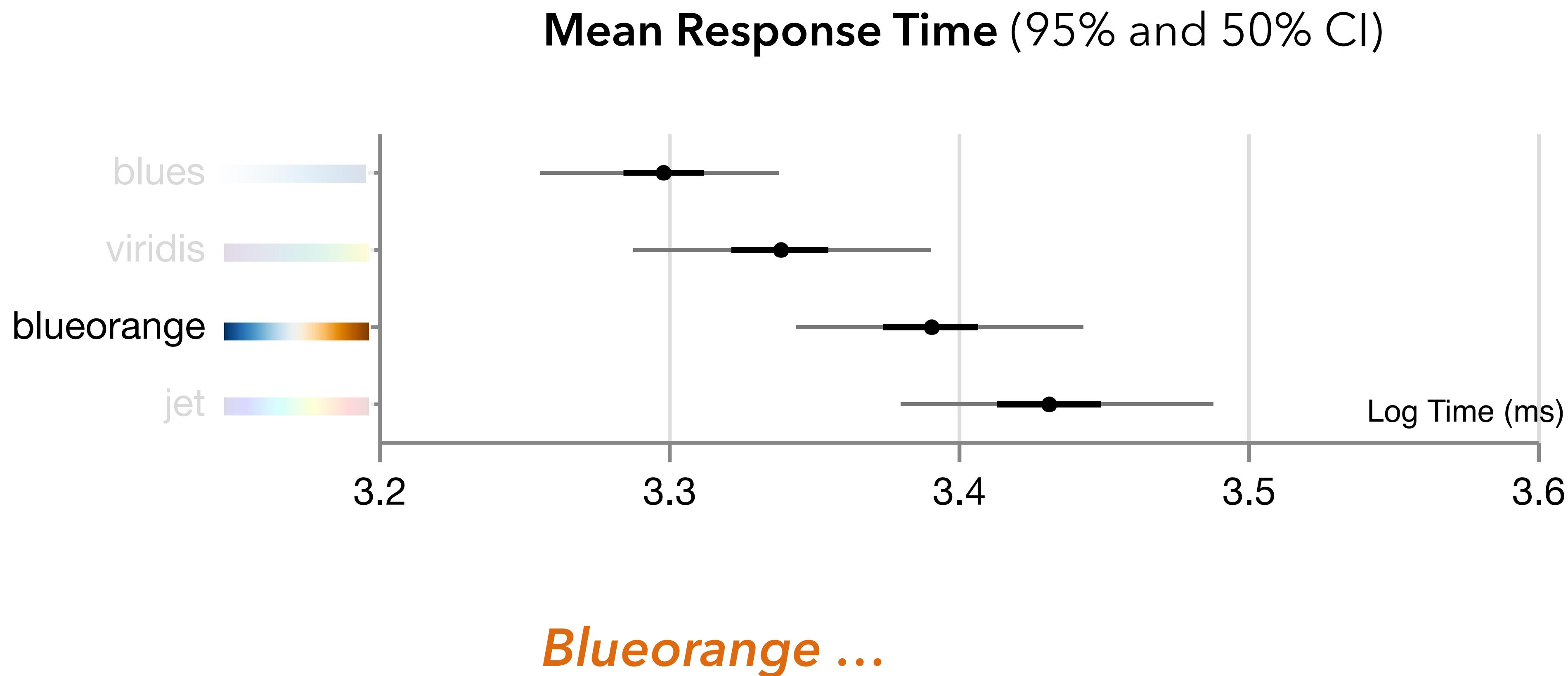
Blues is the fastest

Assorted Colormaps: Time

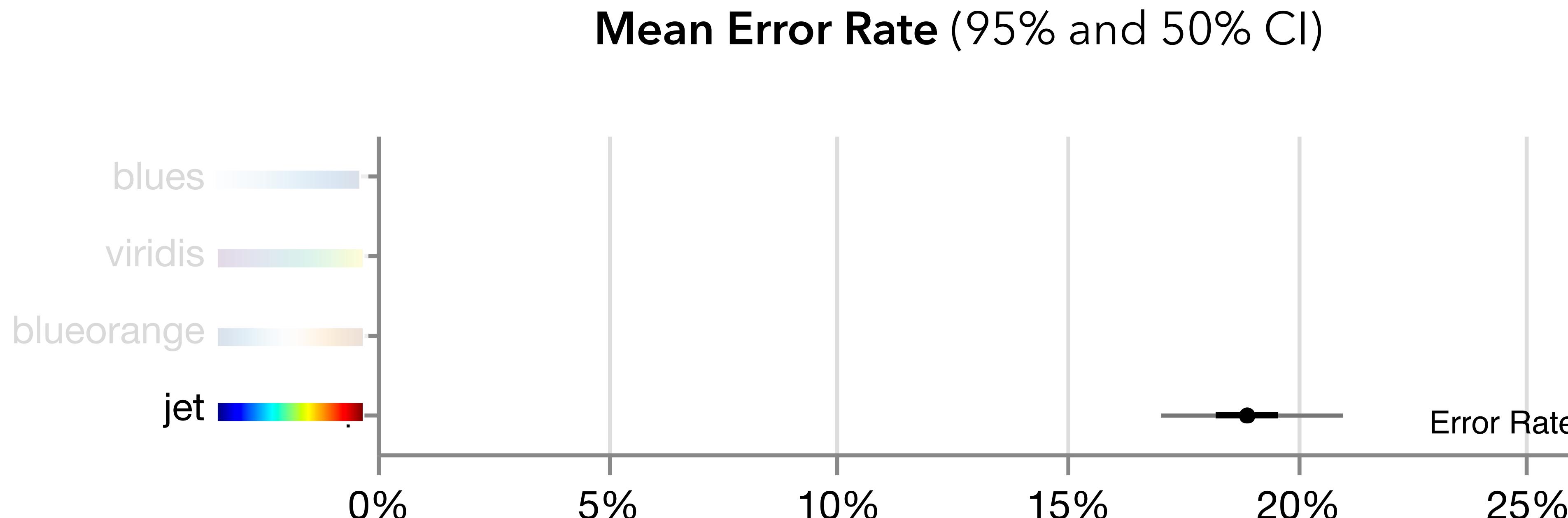


Blues is the fastest, followed by viridis

Assorted Colormaps: Time

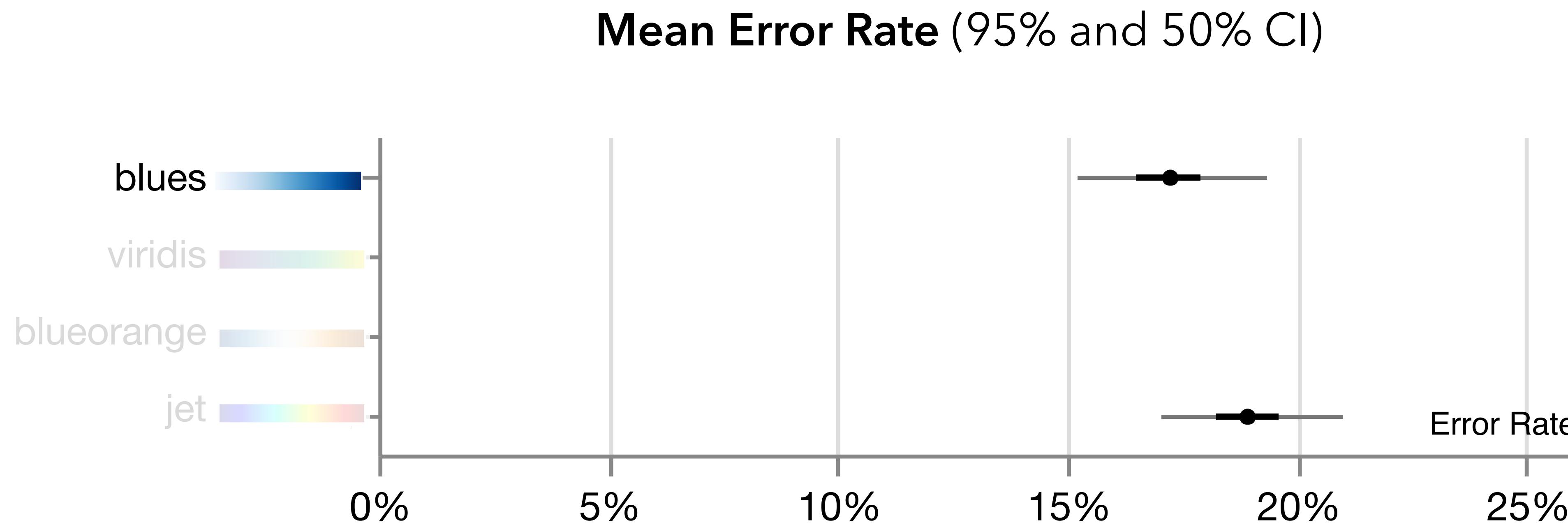


Assorted Colormaps: Error



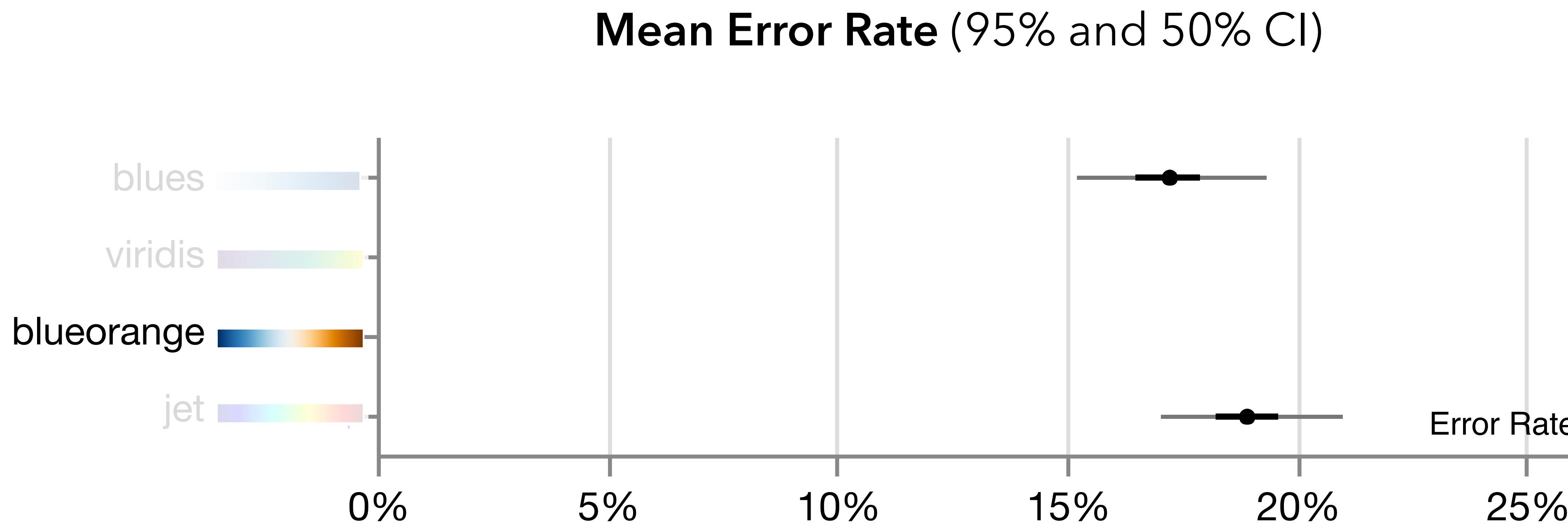
Participants made ~18% errors using *jet*

Assorted Colormaps: Error



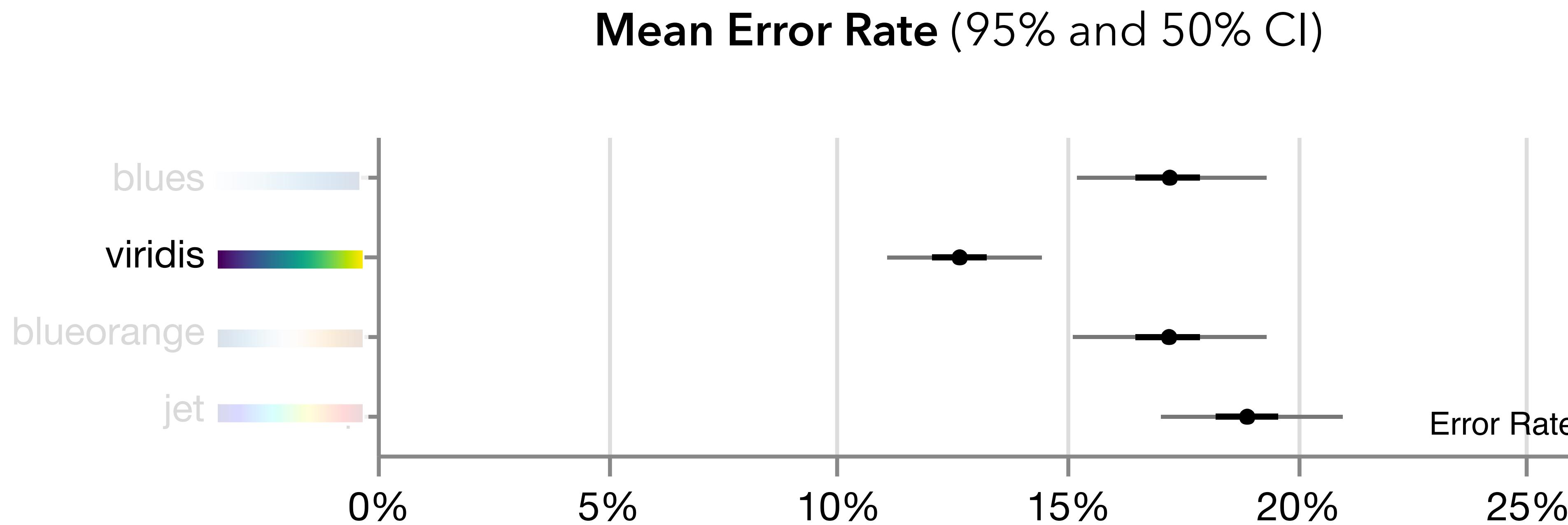
Blues is slightly less error-prone

Assorted Colormaps: Error



Blues and blueorange is slightly less error-prone

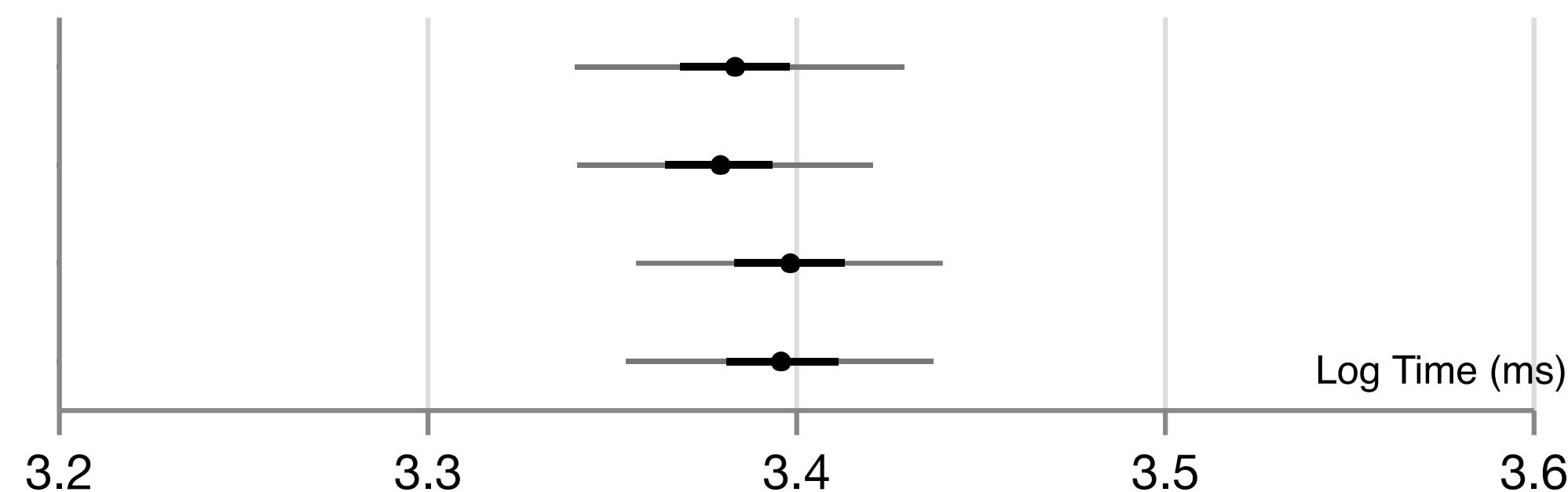
Assorted Colormaps: Error



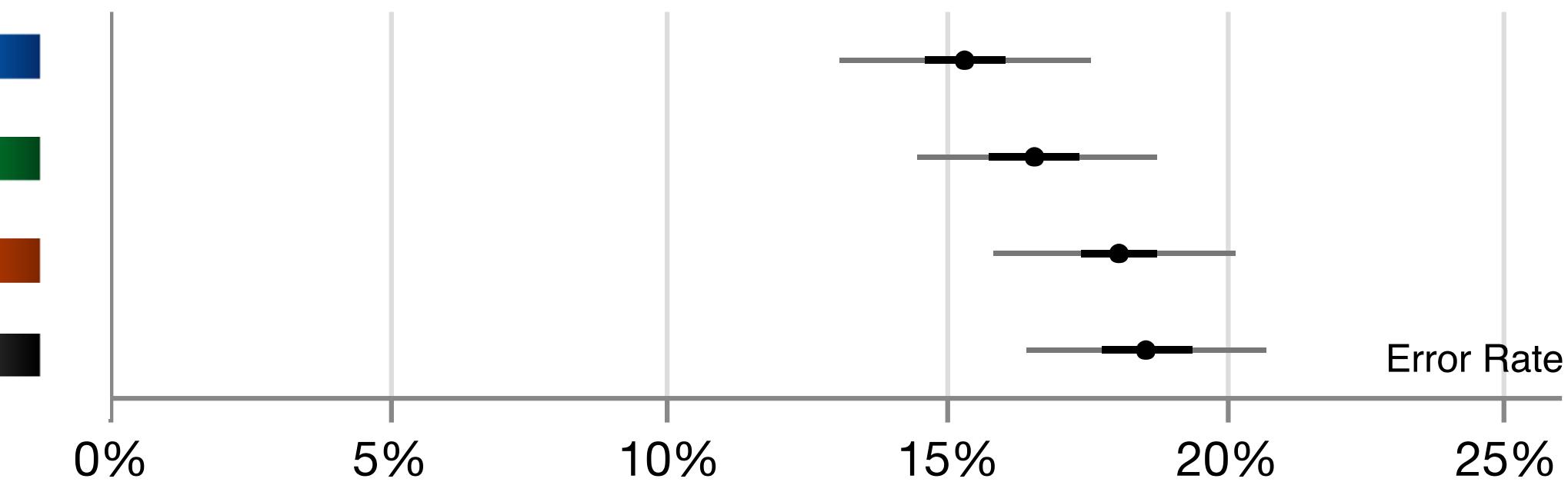
Viridis is the winner in accuracy!

Single-hue Colormaps

Mean Response Time (95% and 50% CI)



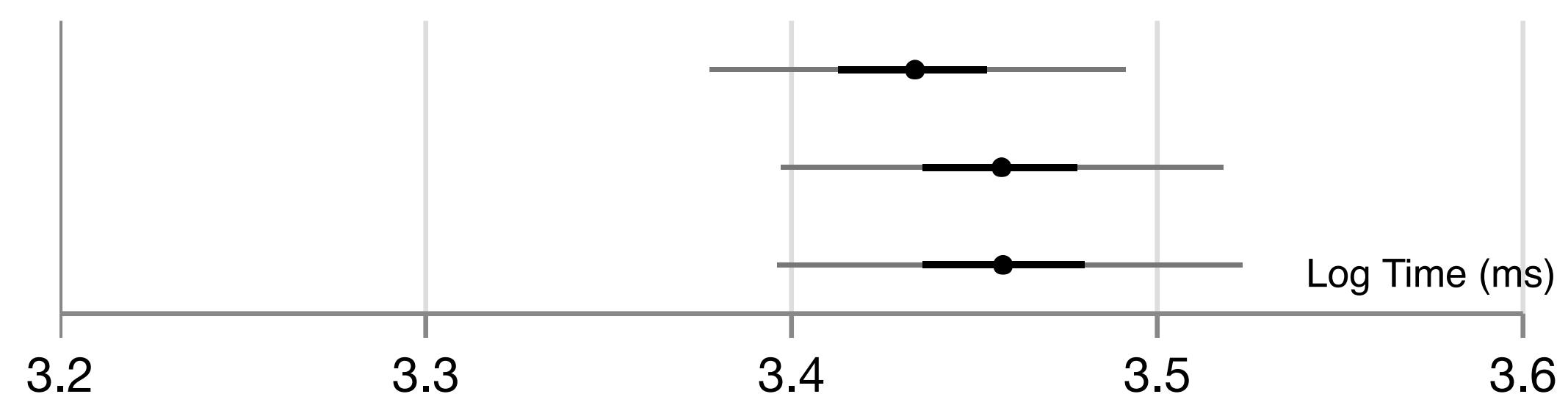
Mean Error Rate (95% and 50% CI)



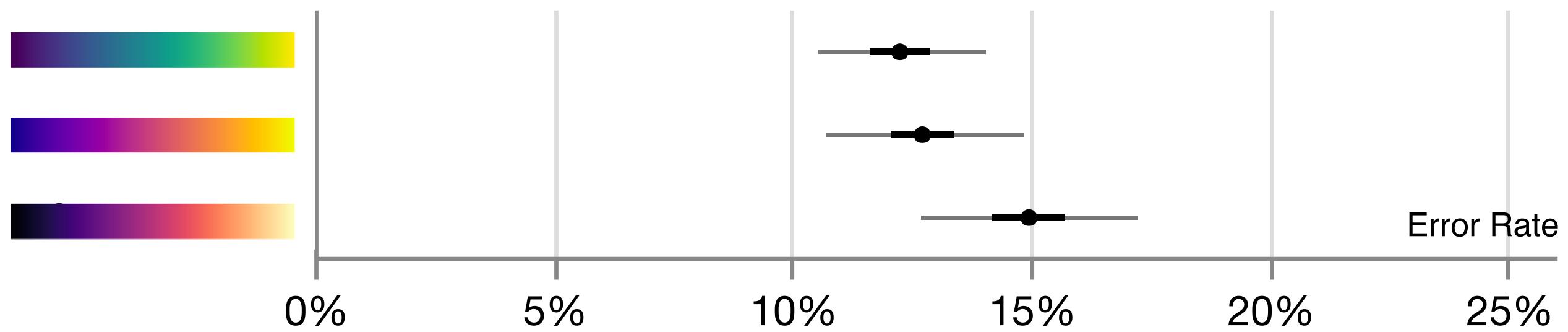
Comparable speed and accuracy
... with nuances

Multi-hue UCS Colormaps

Mean Response Time (95% and 50% CI)



Mean Error Rate (95% and 50% CI)

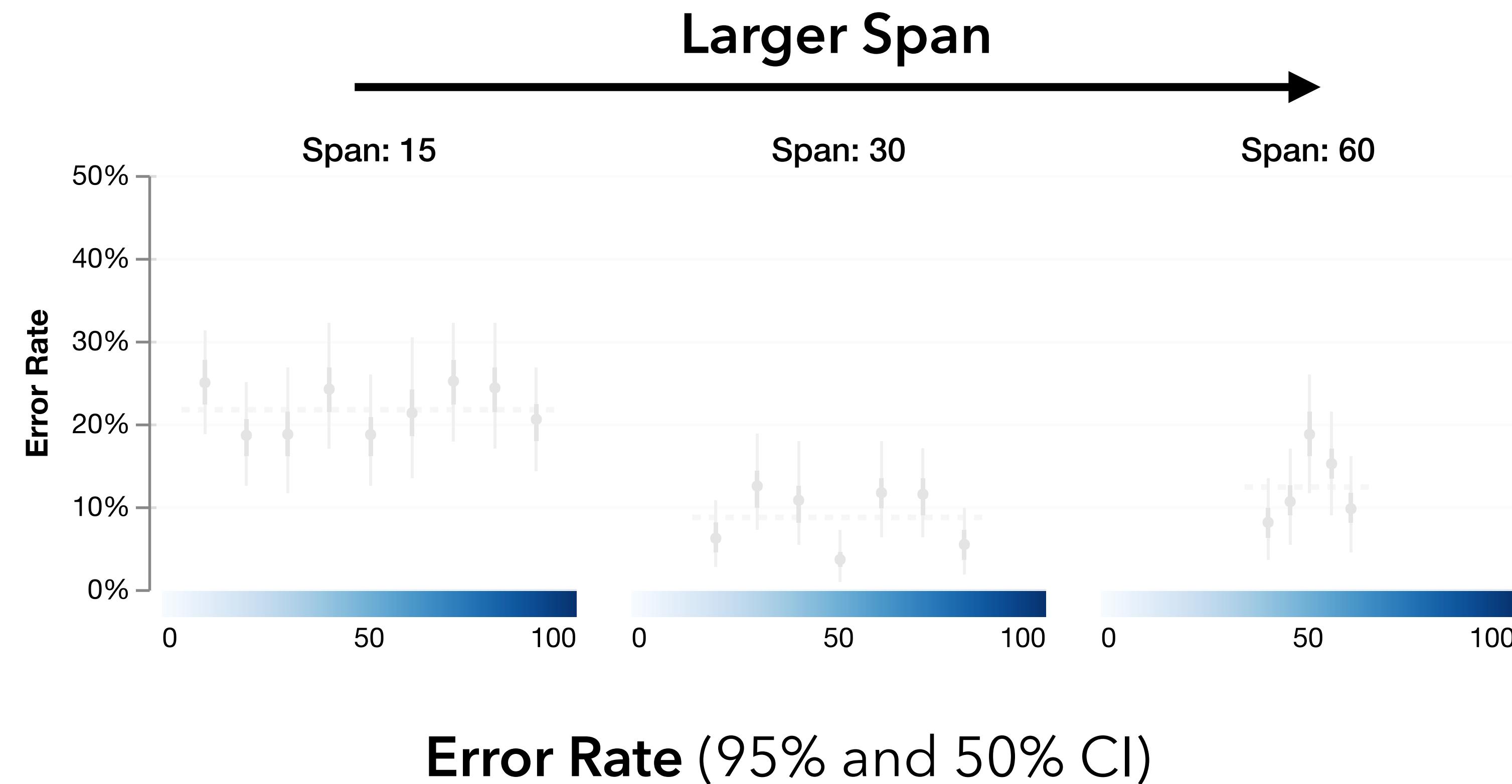


Comparable speed and accuracy
... with nuances

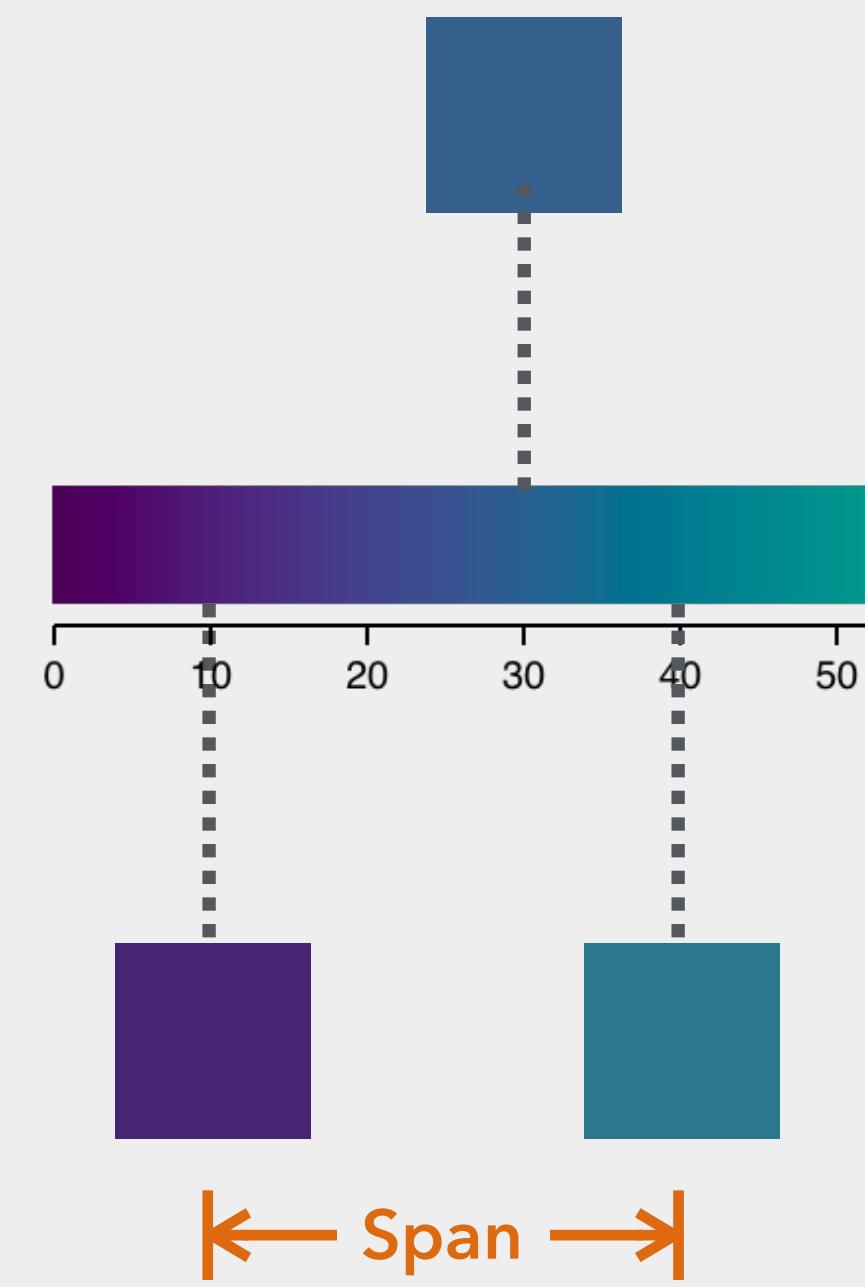
UCS colormaps are the most accurate
... across studies, though with slightly longer response time

Special Cases

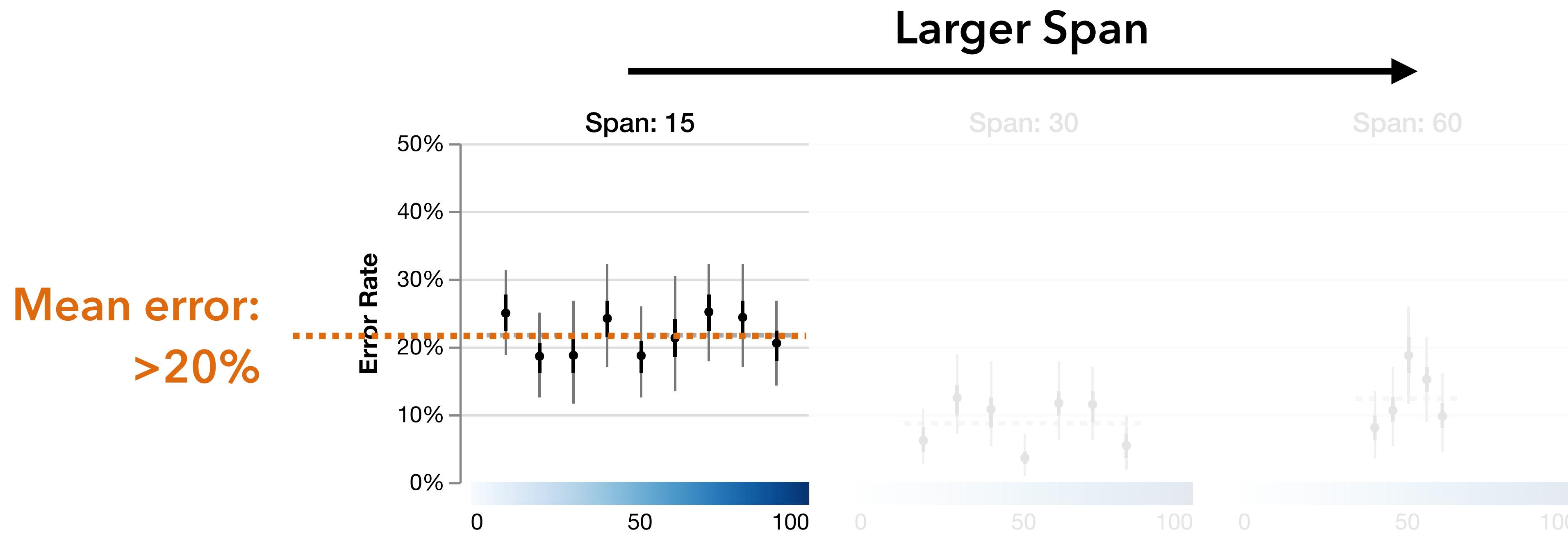
Resolution Issue



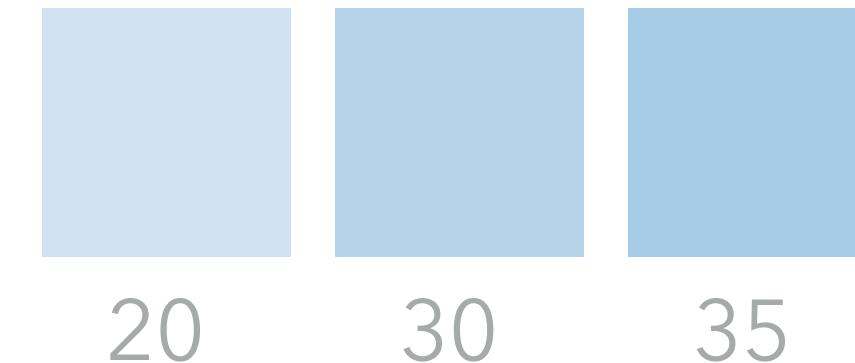
Definition: Span



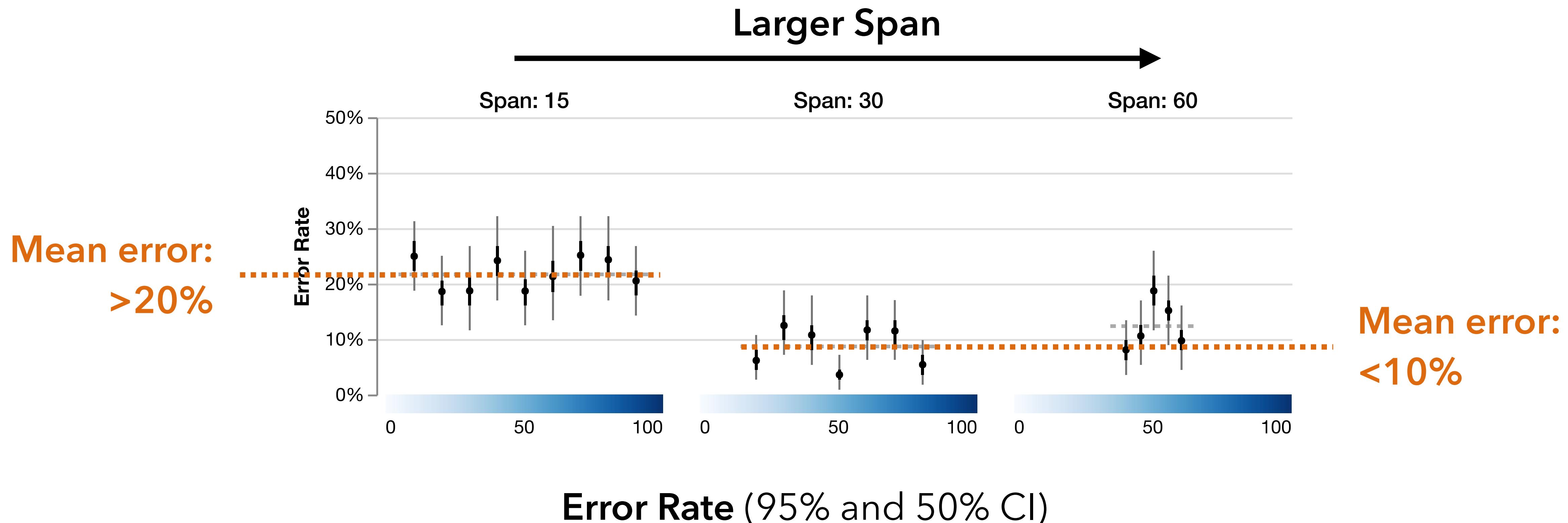
Resolution Issue



Example:



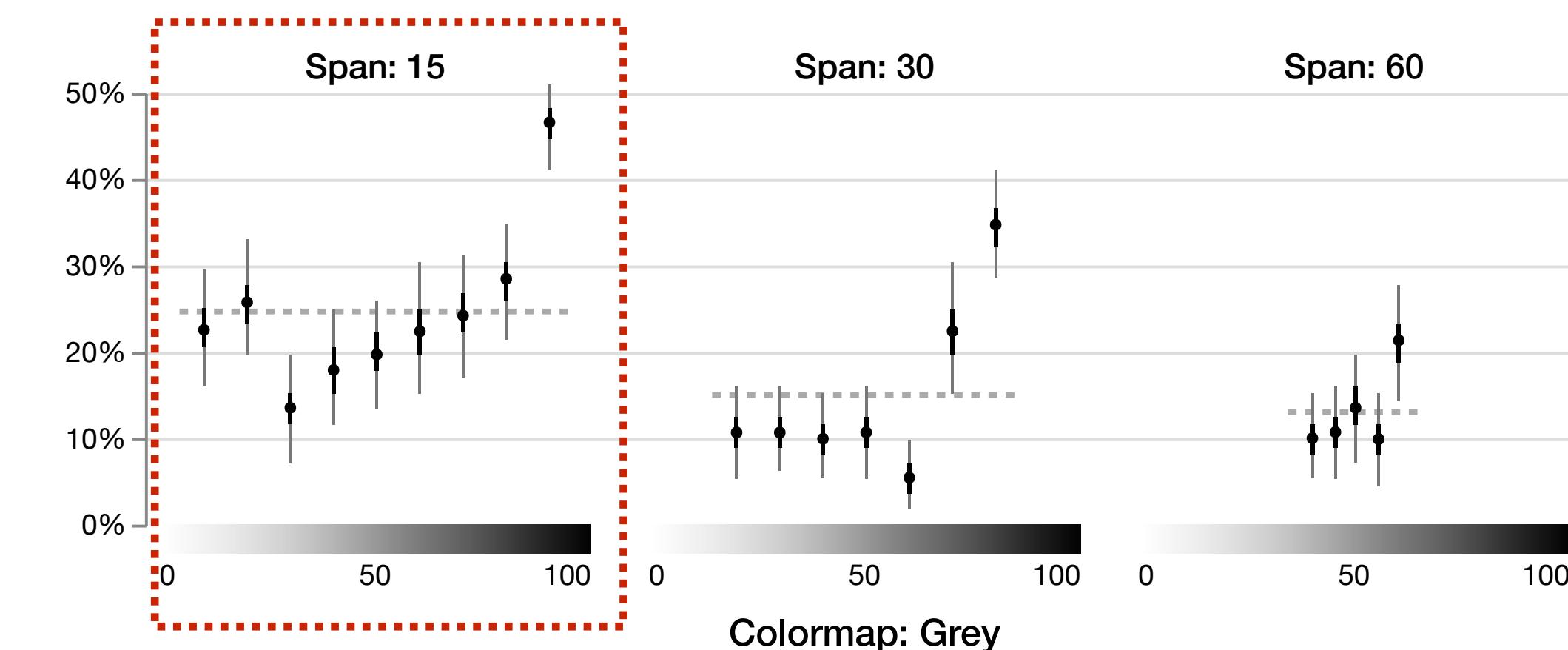
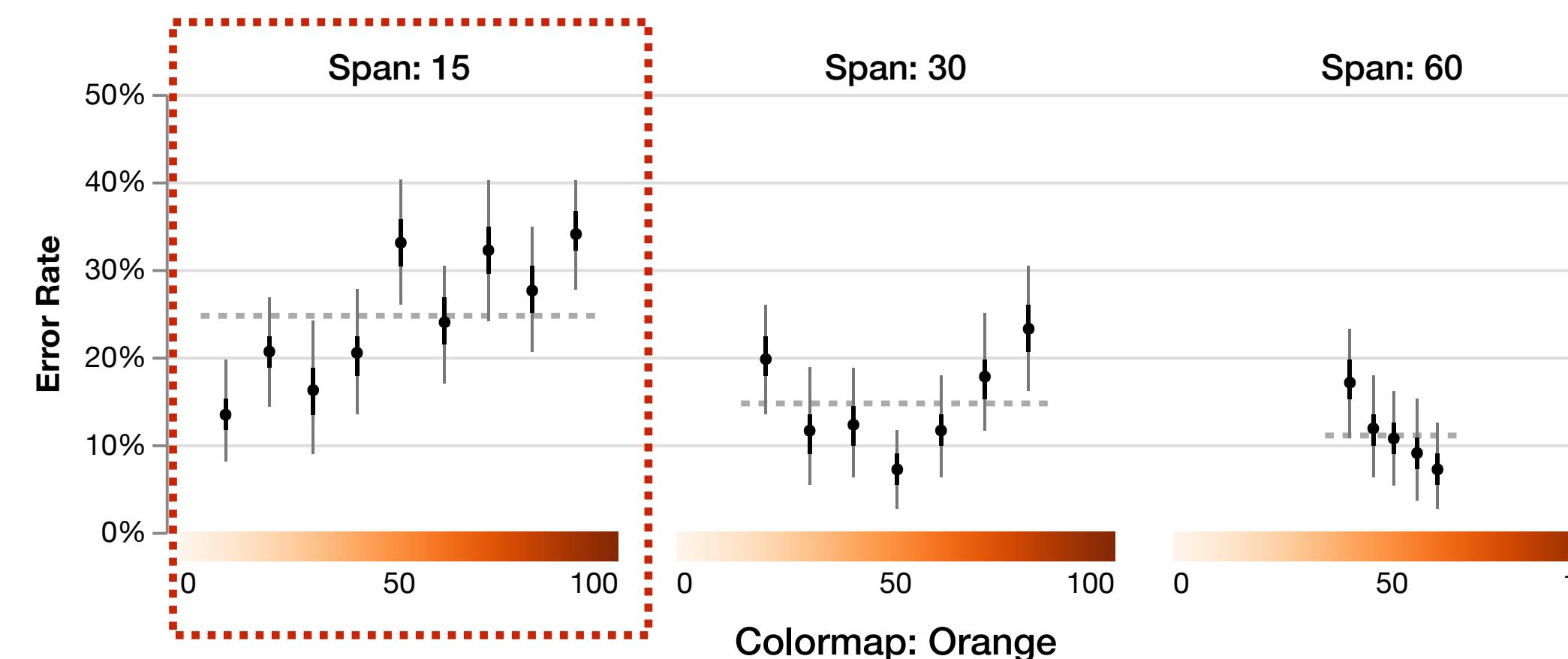
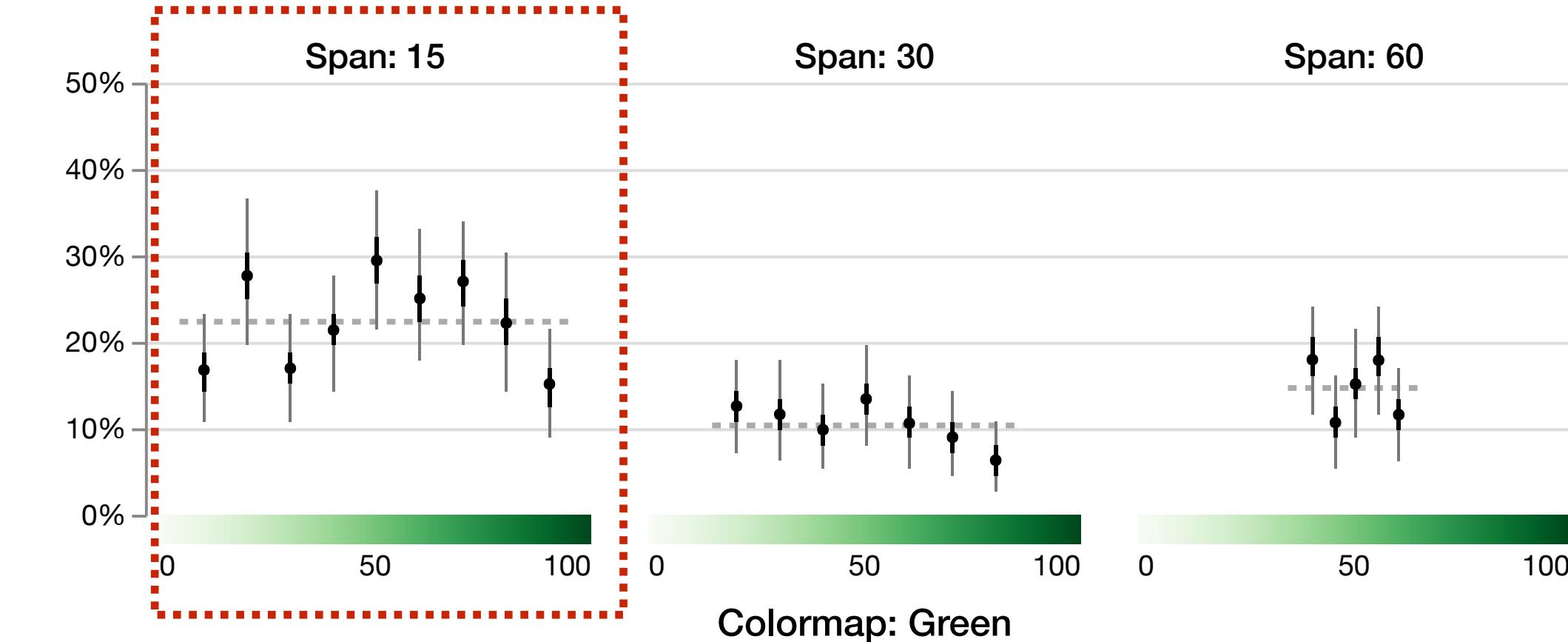
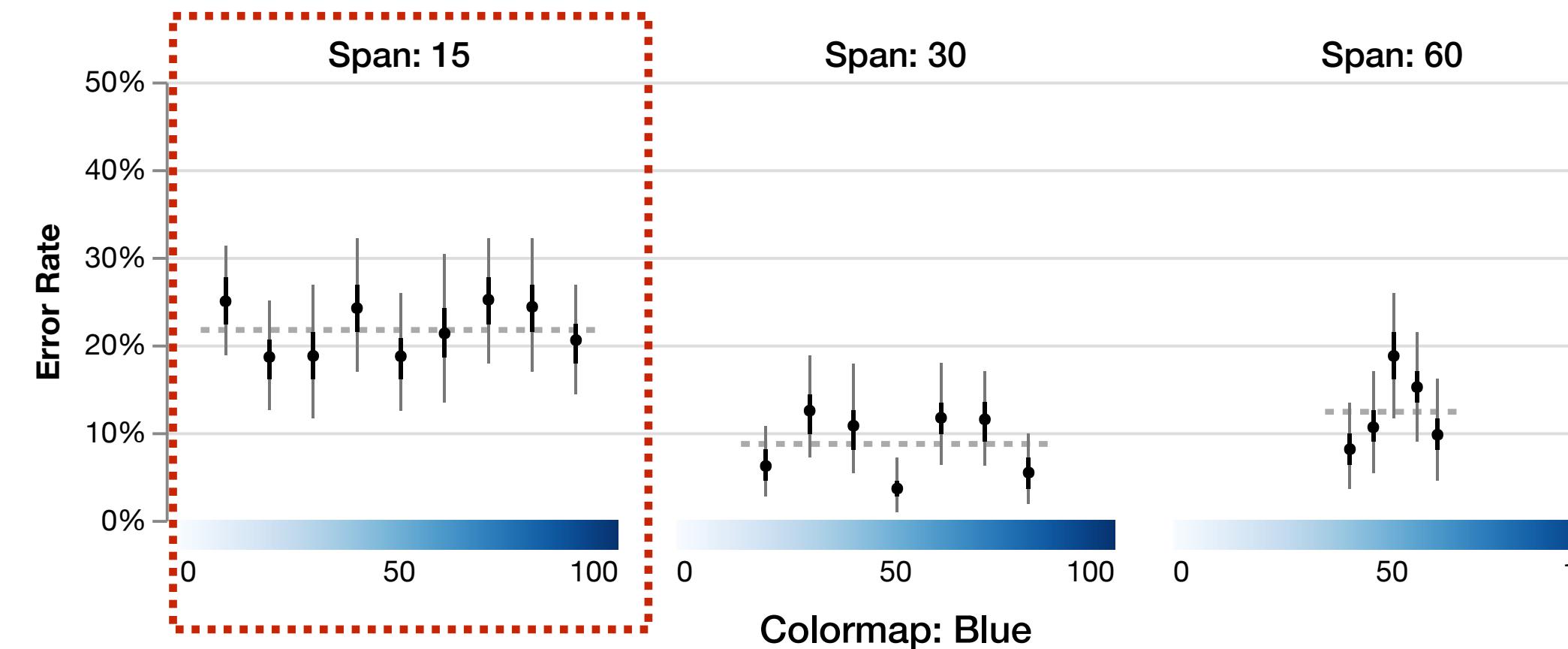
Resolution Issue



Example:



Single-hue Colormaps: Resolution Issue



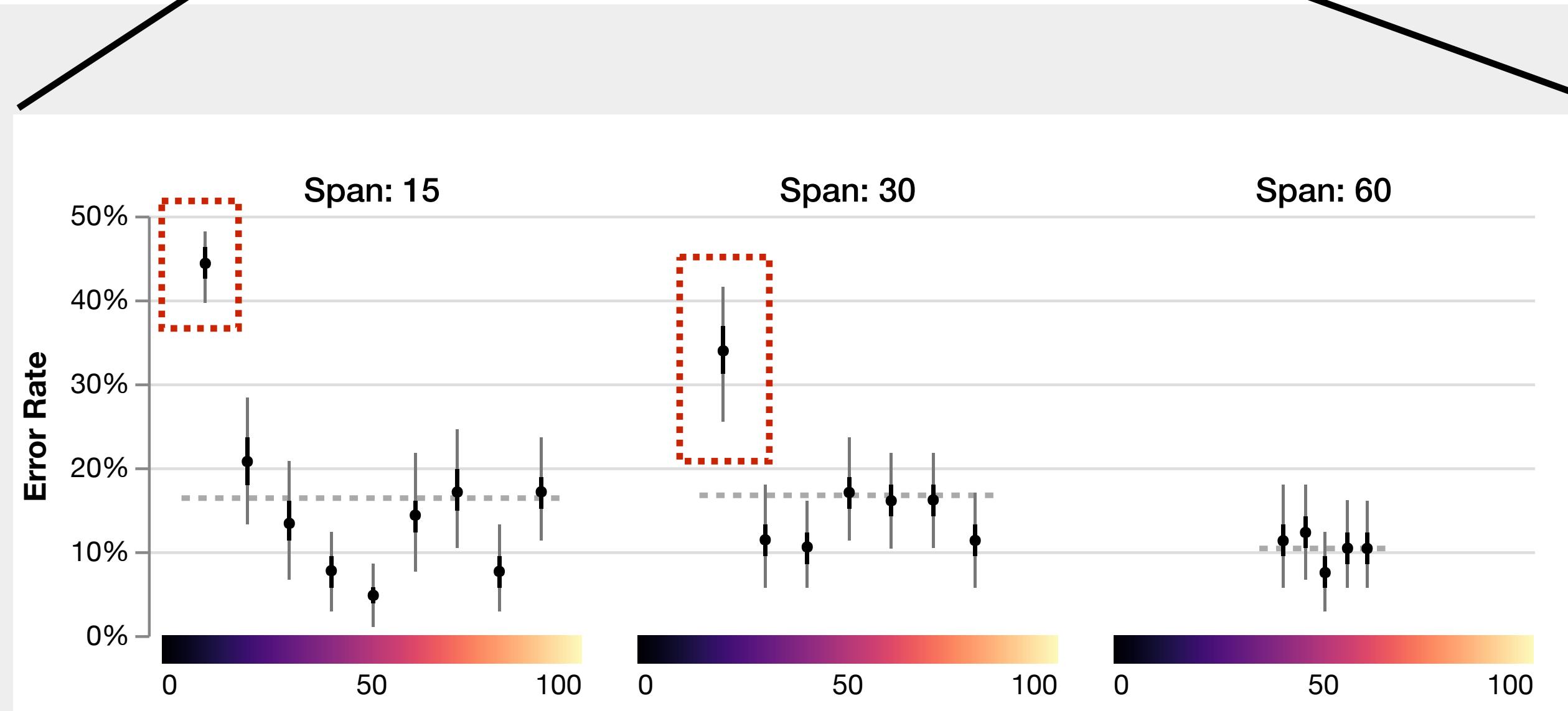
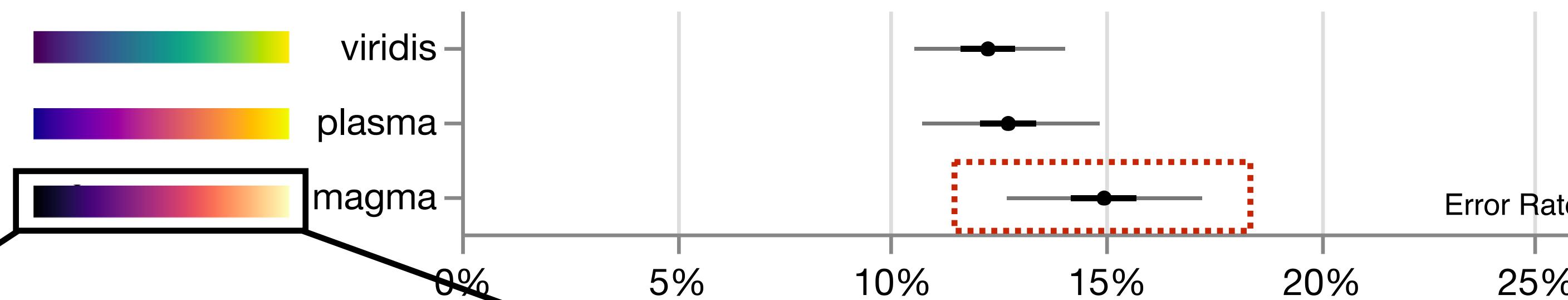
All single-hue colormaps we tested suffer from
low resolution for nearer value comparisons

Special Cases

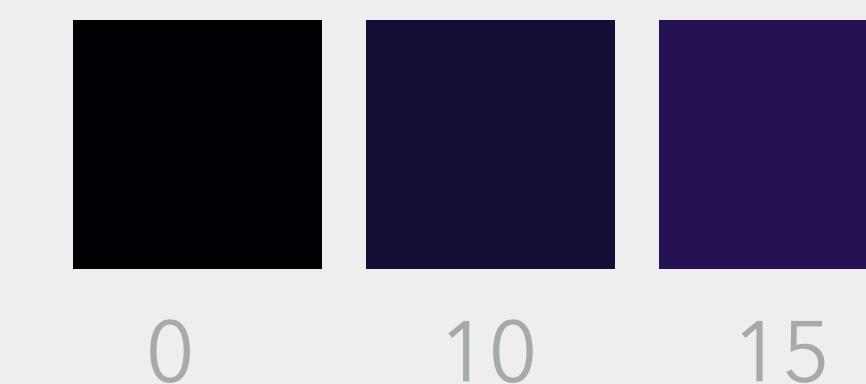
- 1 Single-hue colormaps suffer from low resolution

Dark Regions

Mean Error Rate (95% and 50% CI)

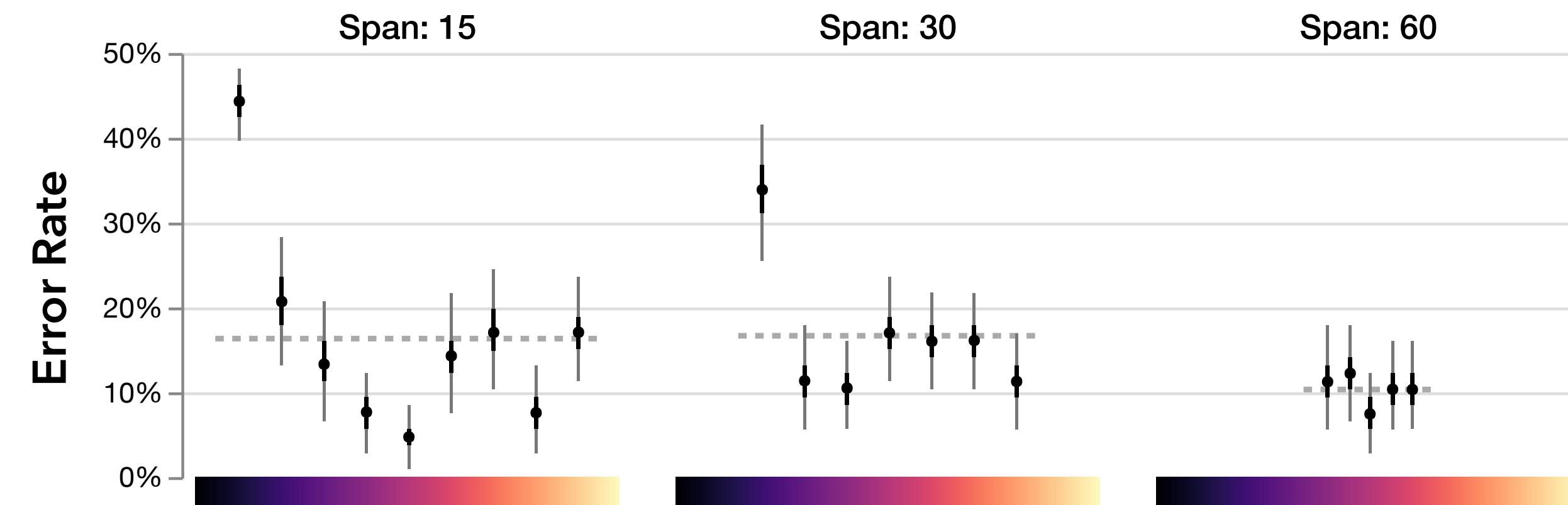


Example Triplets:



Dark Regions

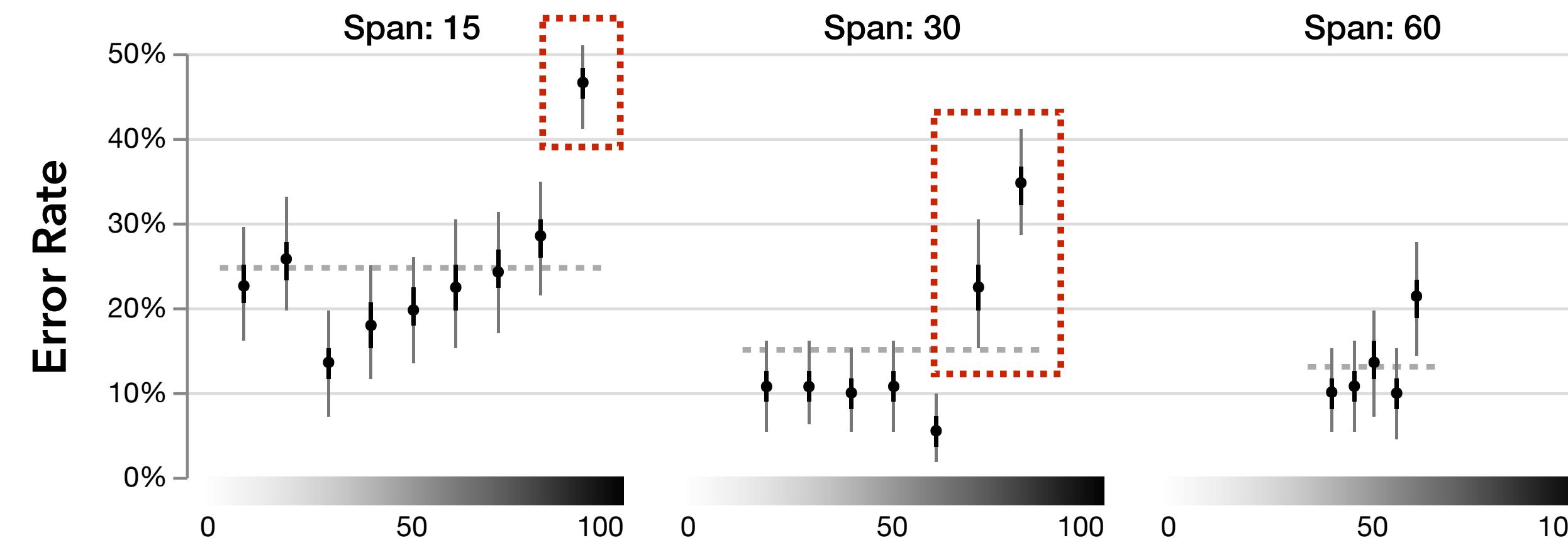
Example Triplets:



0 10 15



0 20 30



85 90 100



70 80 100

Special Cases

- 1 Single-hue colormaps suffer from low resolution
- 2 Performance degrades in low luminance regions

Blueorange: Blues and Oranges

Blueorange



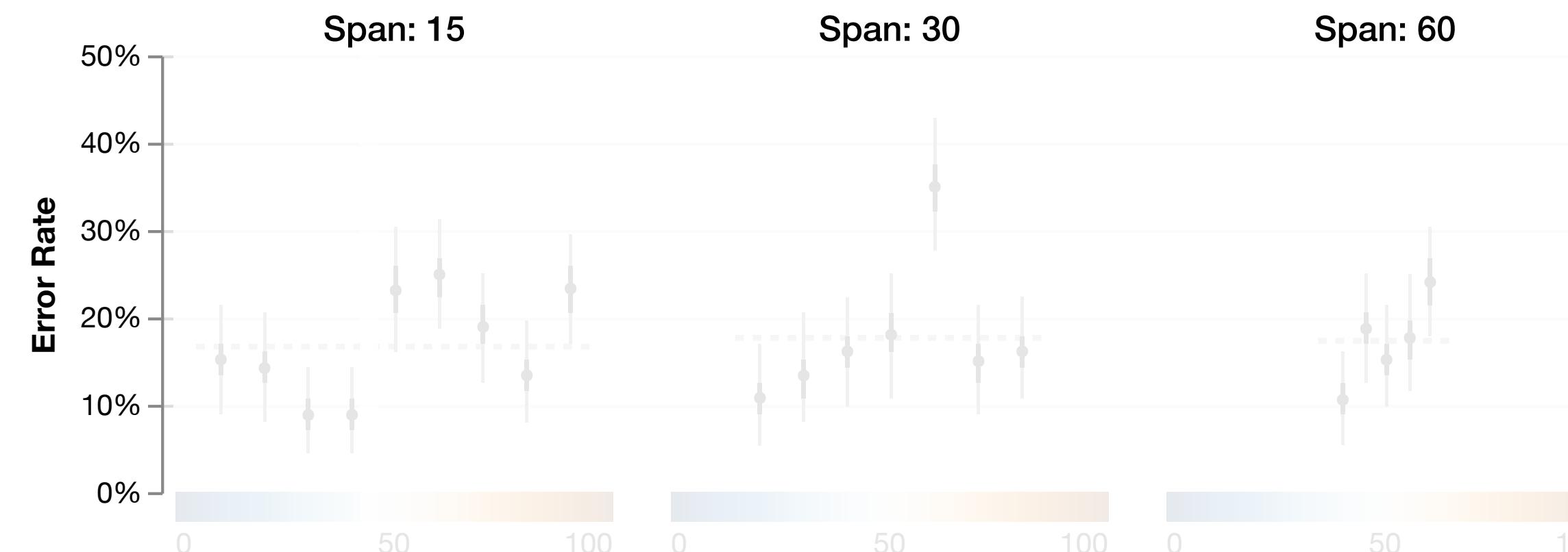
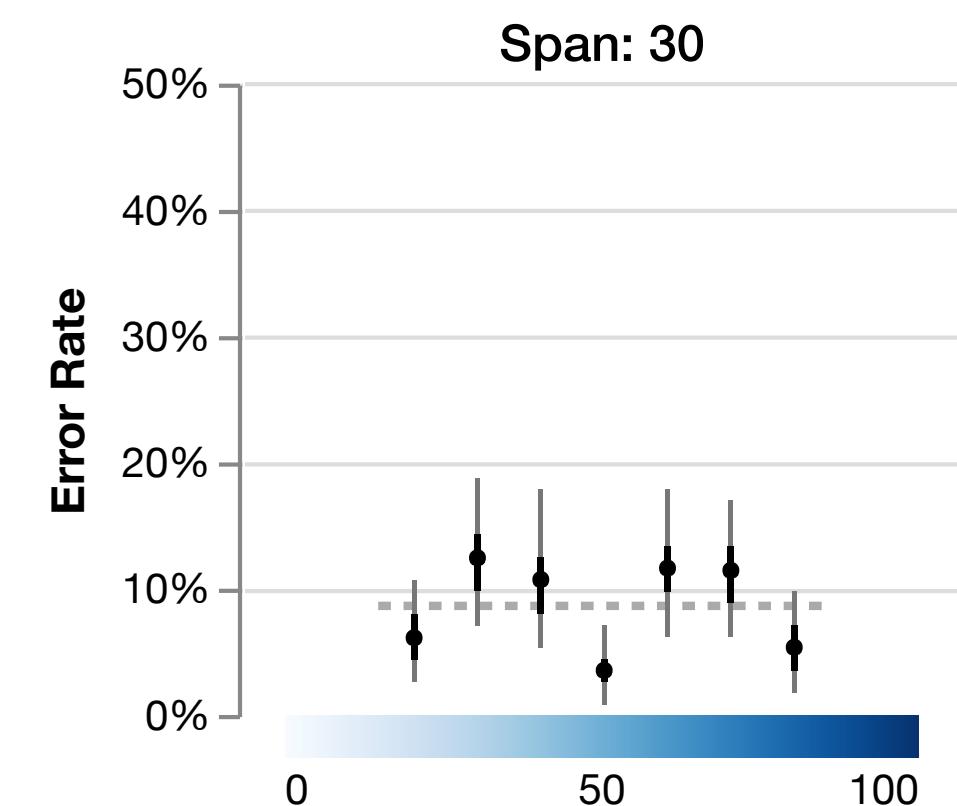
Blues



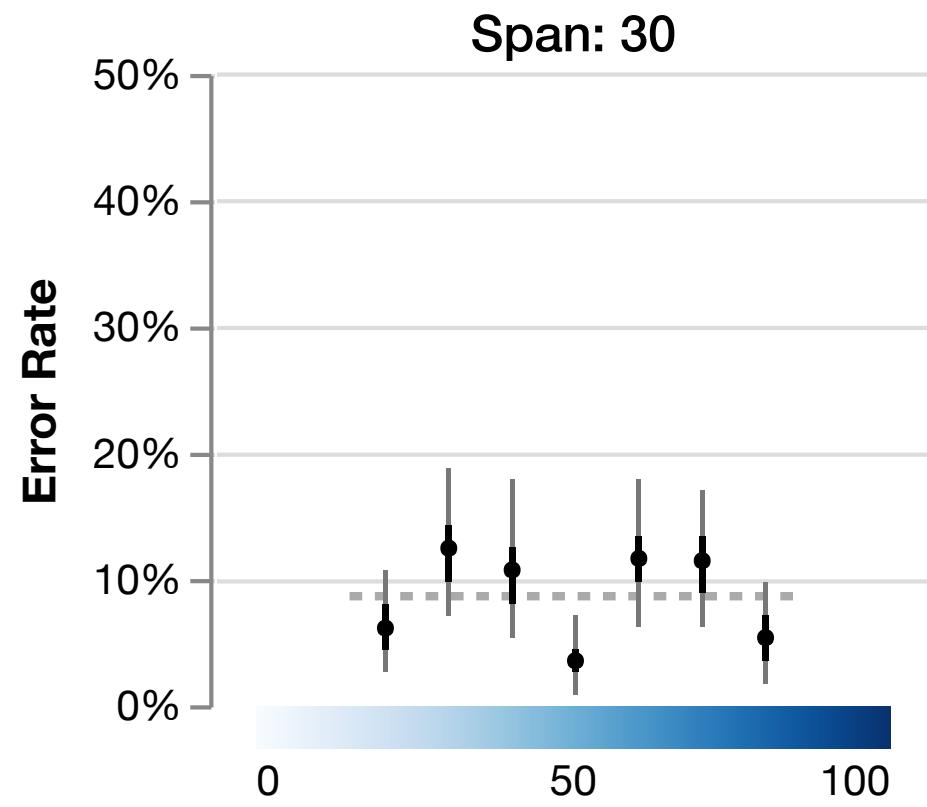
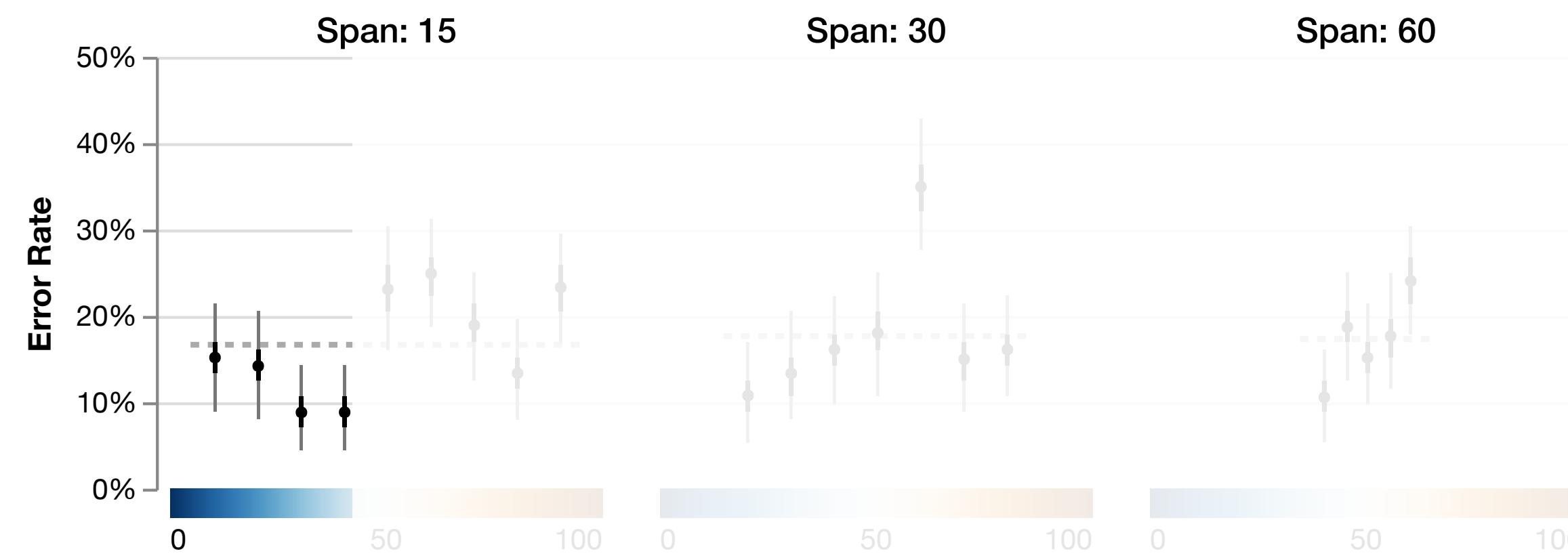
Oranges



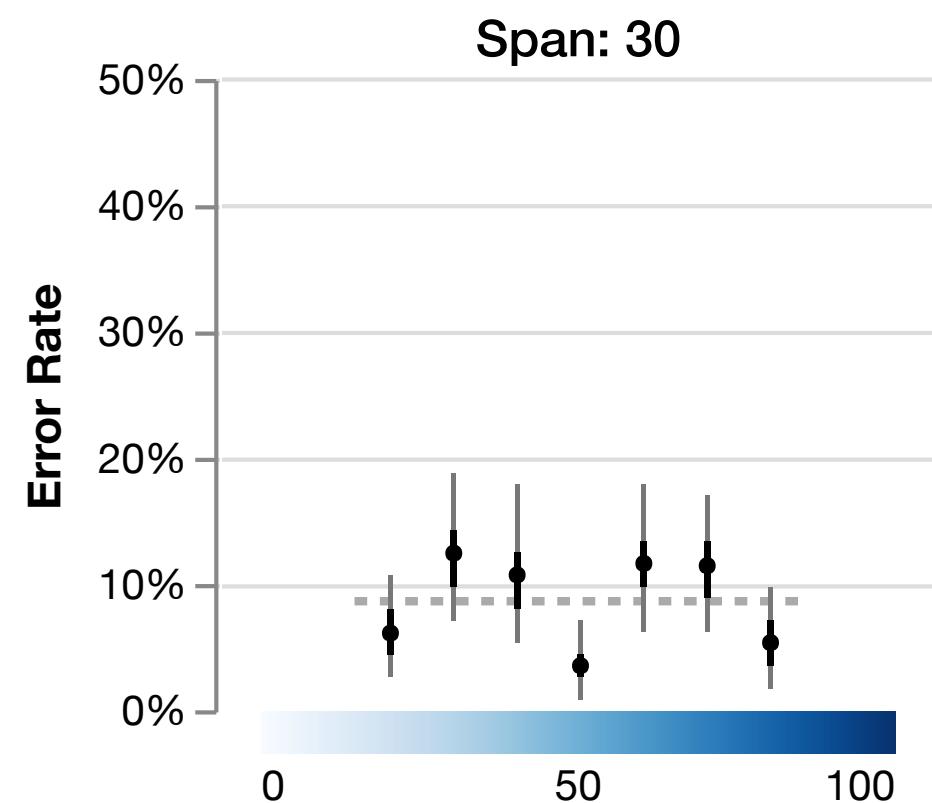
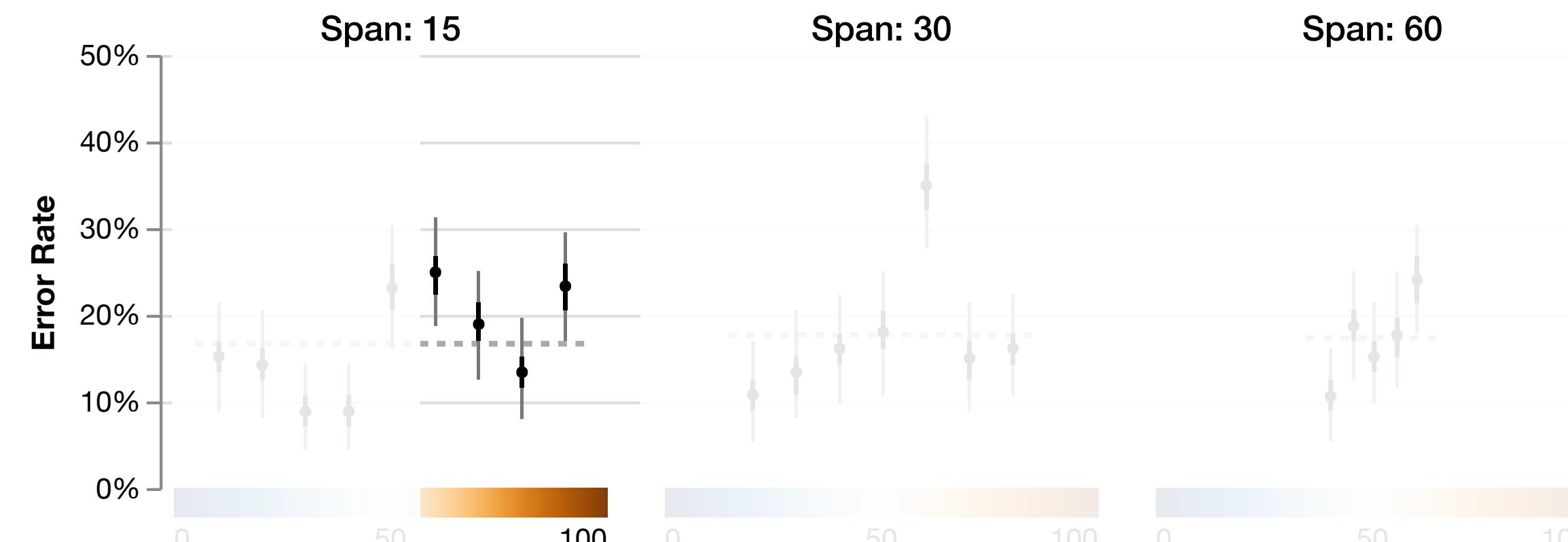
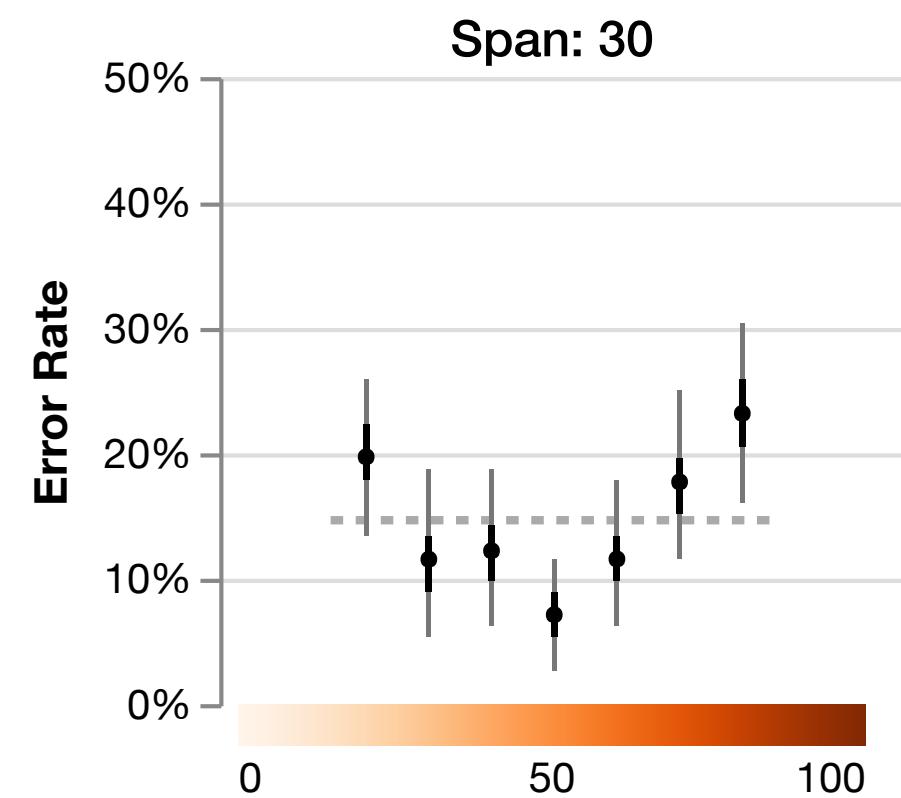
Blueorange: Blues and Oranges



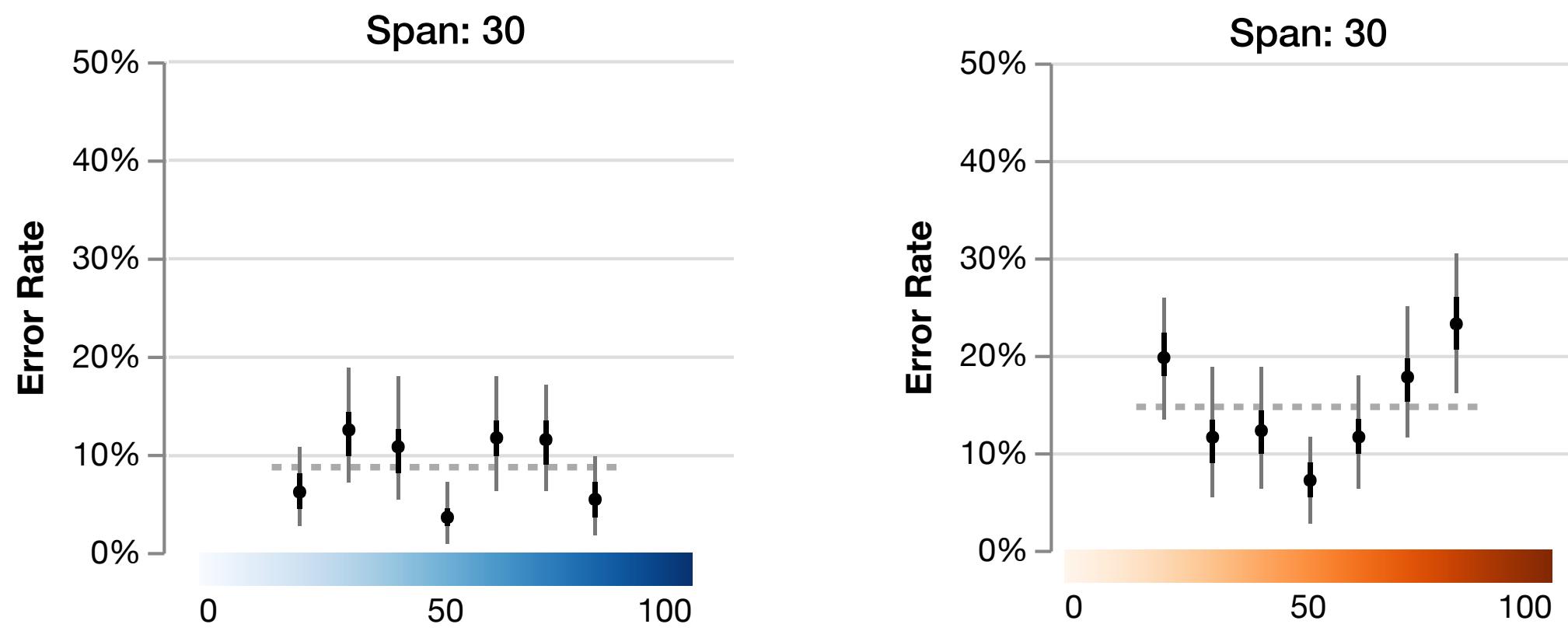
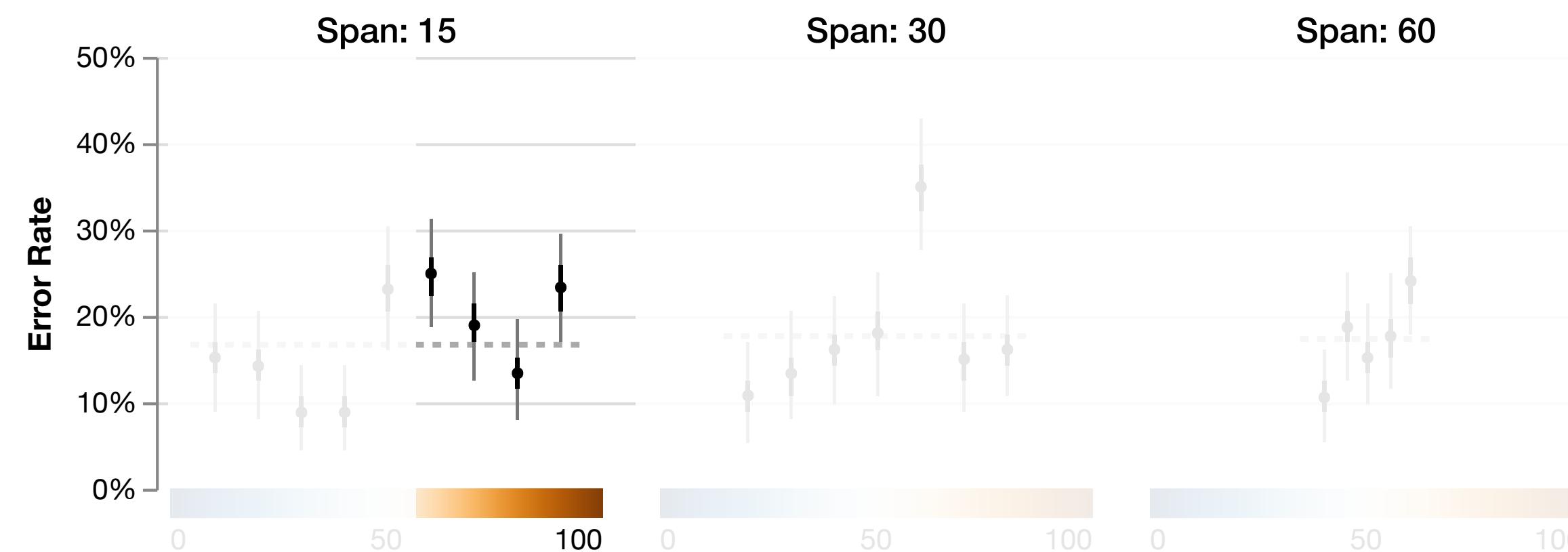
Blueorange: Blues and Oranges



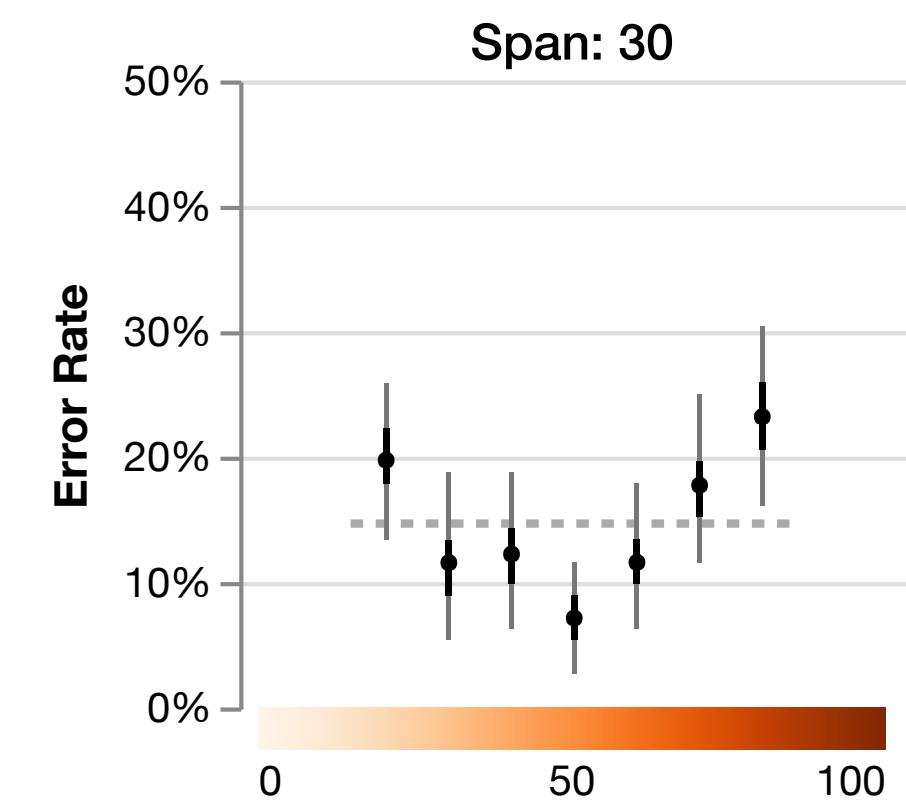
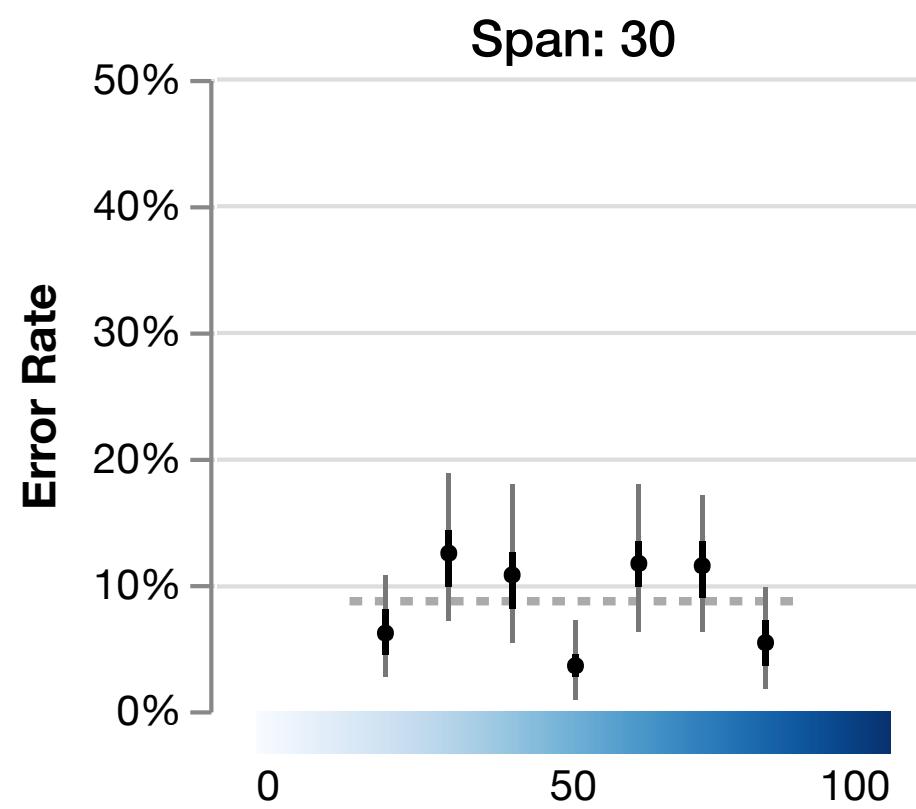
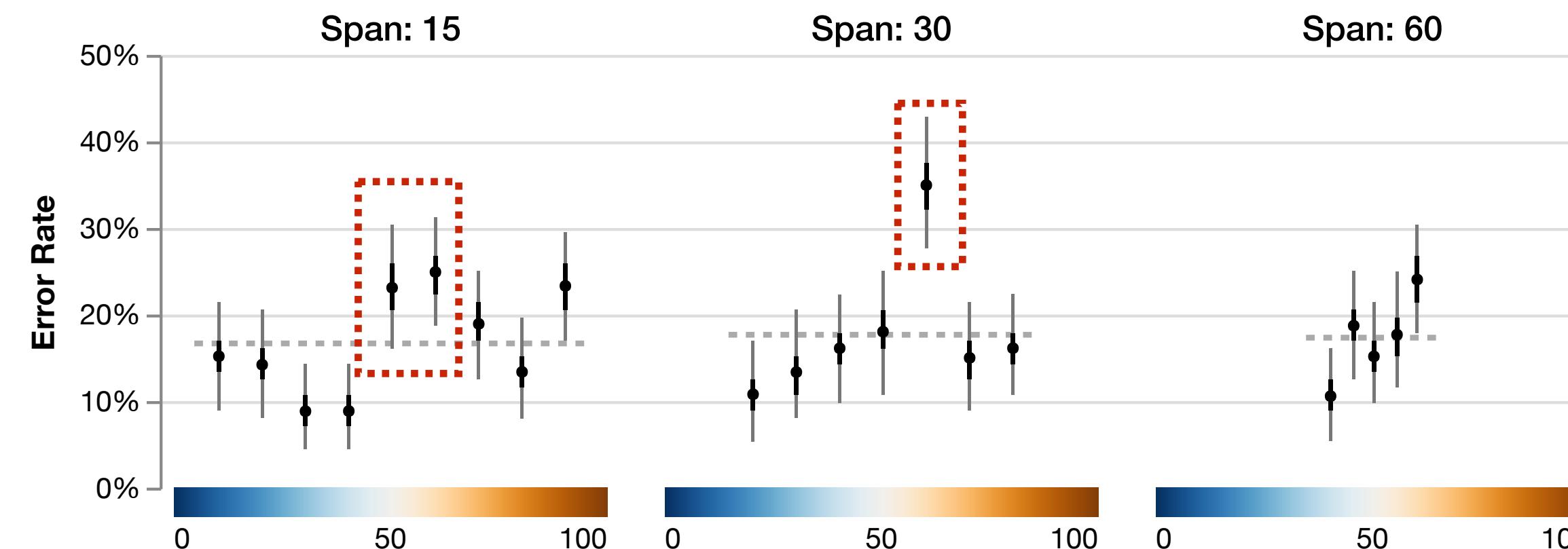
Blueorange: Blues and Oranges



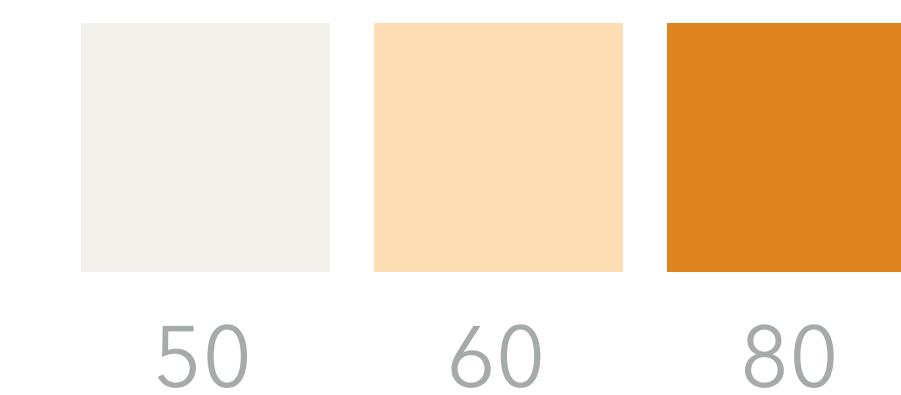
Blueorange: the Midpoint



Blueorange: the Midpoint



Example Triplet:

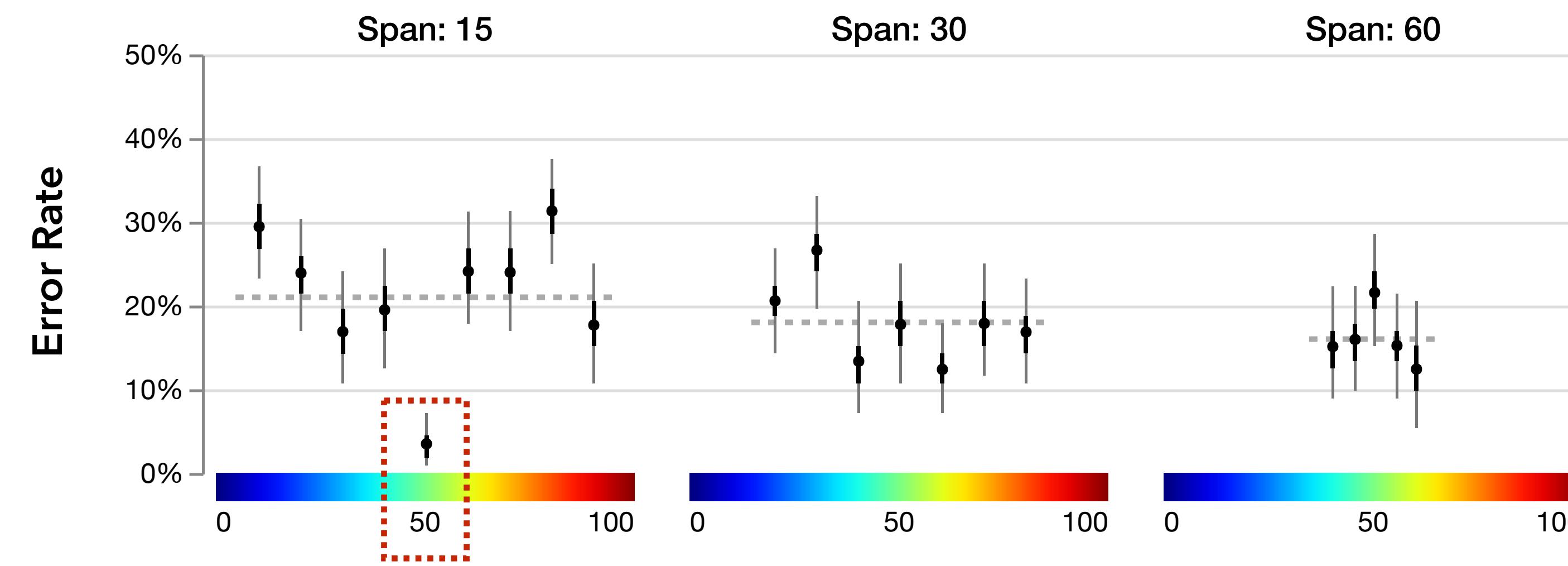


Participants are prone to mistakes when values straddle the **midpoint**

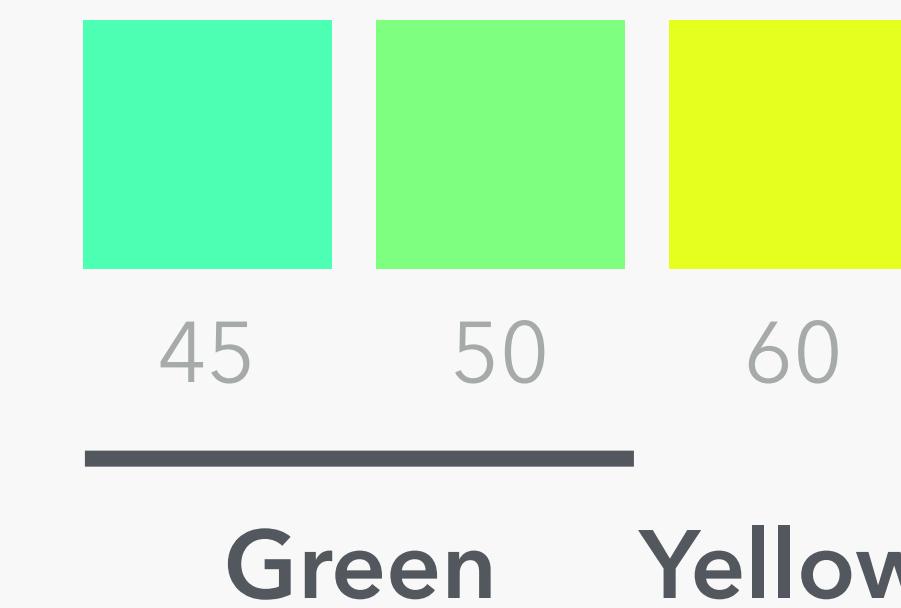
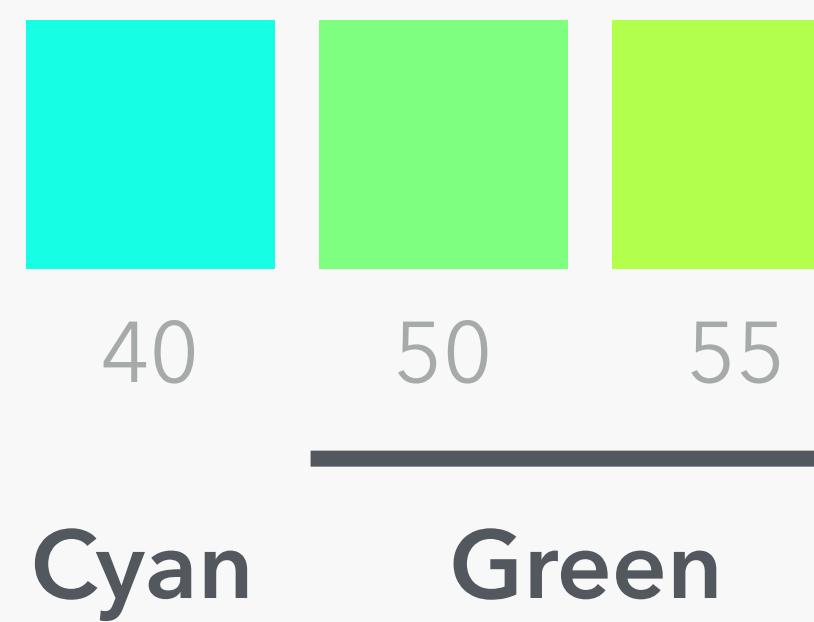
Special Cases

- 1 Single-hue colormaps suffer from low resolution
- 2 Performance degrades in low luminance regions
- 3 *Blueorange* suffers when values straddle the midpoint

Where the Rainbow Shines



Corresponding Triplets:



Roughly **isoluminant**
Color names align with value differences

Special Cases

- 1 Single-hue colormaps suffer from low resolution
- 2 Performance degrades in low luminance regions
- 3 *Blueorange* suffers when values straddle the midpoint
- 4 Where the rainbow shines: color name association

Color Model Analysis

Details in the paper!

Given an untested colormap, might
we predict its performance?

Perceptual Color Model + Color Naming Model



Higher predictive accuracy than either alone
... though the overall accuracy is poor

Somewhere Over the Rainbow

An Empirical Assessment of
Quantitative Colormaps

Introduction

Hues vs. luminance: can we use both?

Methods

A suite of experiments comparing 9 colormaps

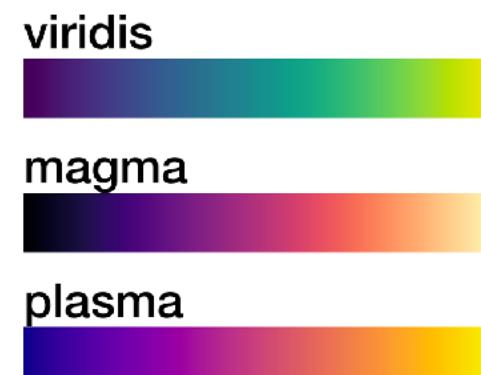
Results

Global statistics and interesting special cases

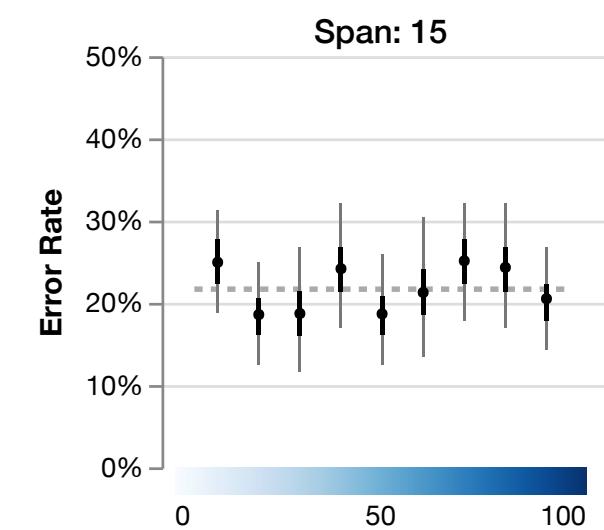
Conclusion

Our contributions, limitations and takeaways

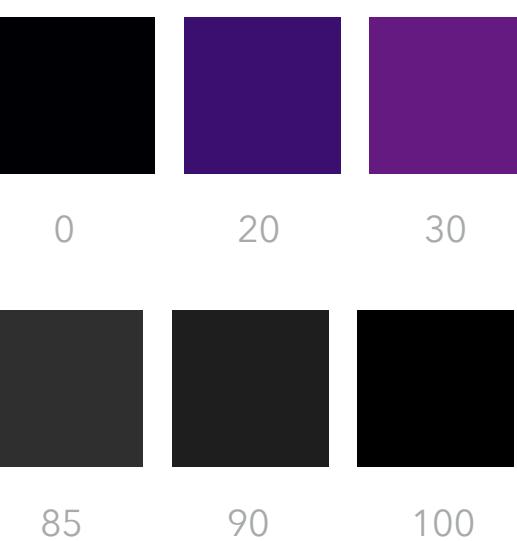
Contributions



Judiciously designed **multi-hue** colormaps **perform well** in terms of time and error



Single-hue colormaps similarly **perform well**, but have **resolution issue** for nearer value comparisons

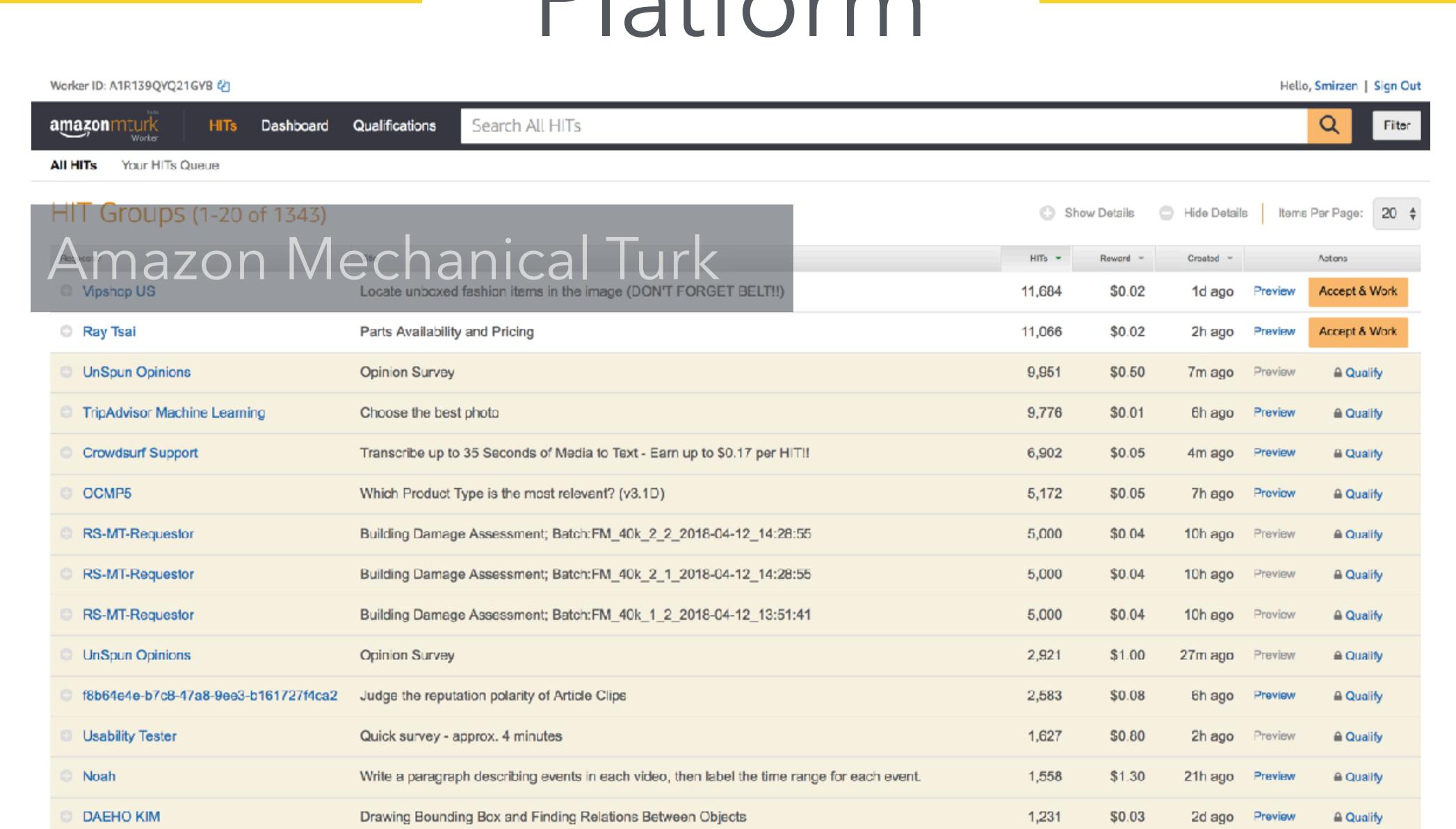


We identify recurring poor performances, notably the degradation in the **low luminance** regions



Limitations

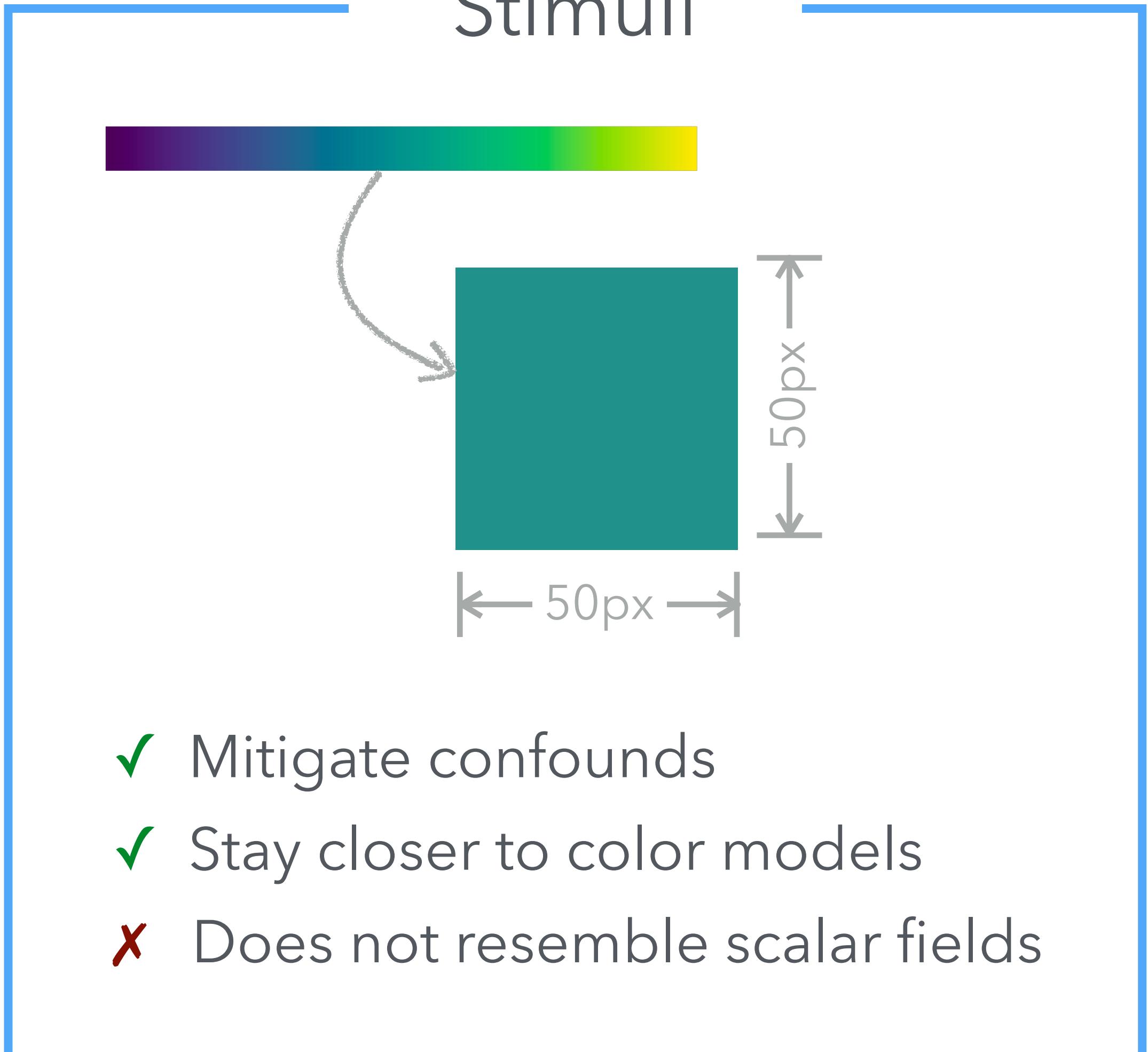
Platform



The screenshot shows the Amazon Mechanical Turk worker interface. At the top, there's a navigation bar with 'Amazon Mechanical Turk' logo, 'HITs', 'Dashboard', 'Qualifications', and a search bar 'Search All HITs'. Below the navigation is a section titled 'HIT Groups (1-20 of 1343)' with a sub-section 'Amazon Mechanical Turk'. It includes a note 'Locate unboxed fashion items in the image (DON'T FORGET BELT!!)'. The main area displays a table of HIT groups with columns: 'HITs', 'Reward', 'Created', 'Actions', and 'Accept & Work'. The table lists various HITs from different workers like Ray Tsai, UnSpun Opinions, TripAdvisor Machine Learning, etc. Each row shows the number of HITs, reward amount, creation time, preview link, and an 'Accept & Work' button.

↓ Experimental control
↑ Ecological validity

Stimuli



Takeaways

- 1 If judiciously designed, more hues improves resolution

Continuous data:



Discrete data:



- 2 Be careful about low luminance colors against high luminance backgrounds

Colormap designers:

Caution against dark ranges

Visualization designers:

Caution against

