VIBE

A Design Space for **VI**sual **B**elief **E**licitation in Data Journalism

Sham Mahajan
Independent

Bonnie Chen
Emory University

Alireza Karduni
Northwestern University

Yea-seul Kim
University of Wisconsin, Madison

Emily Wall
Emory University

EUROVIS 2022 ME

What are beliefs?

Mental representations related to a phenomenon that can be characterized as **numerical** or **categorical values**

Belief elicitation in Data Journalism

The New York Times   FiveThirtyEight   Vox
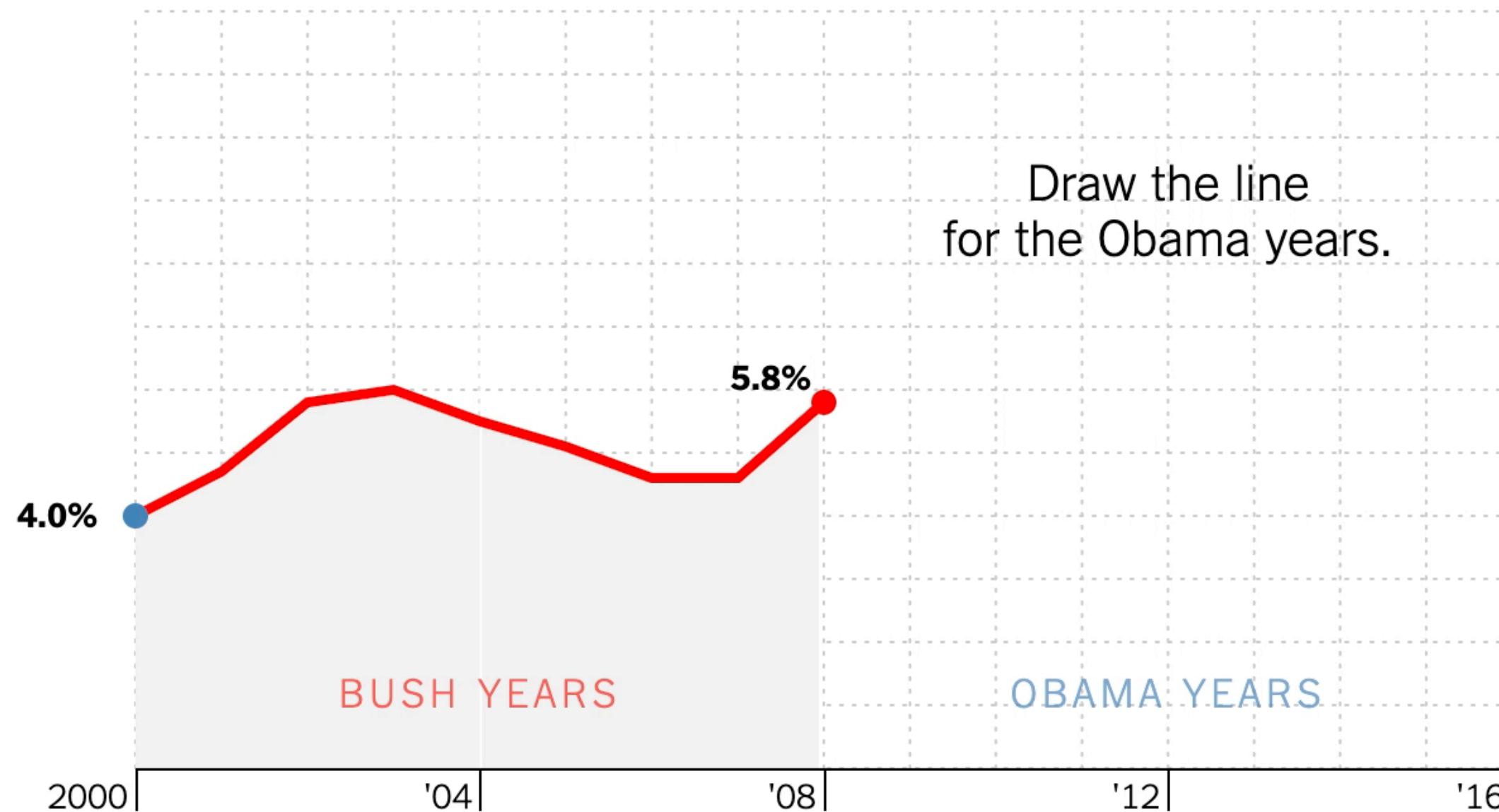
You Draw It: What Got Better or Worse During Obama's Presidency

By **LARRY BUCHANAN**, **HAEYOUN PARK** and **ADAM PEARCE**   JAN. 15, 2017

Draw your guesses on the charts below to
see if you're as smart as you think you are.

Under President Obama, the **unemployment rate** ...

Draw the line
for the Obama years.

5.8%

4.0%

BUSH YEARS

OBAMA YEARS

2000   '04   '08   '12   '16

Show me how I did.

Do You Know Where America Stands On Guns?

By **Dhrumil Mehta** and **Julia Wolfe**
Get the data on **GitHub**

It seems almost routine. There's a high-profile mass shooting, followed by a week or two of fierce debate around gun policy. (And usually no legislative change.)

When it comes to the specific policy debates, where does the American public stand? This is normally where we'd answer that question for you, but why should we do all the work?

How well do you know how Americans feel about guns? Let's start with the basics. (Keep in mind: This all comes from polling[1] conducted after the school shooting at Stoneman Douglas High School in Parkland, Florida.)

What share of Americans support **stricter gun laws**?

| 0 | 25 | 50% | 75 | 100 |

◀ LESS SUPPORT   MORE SUPPORT ▶

Guess

Source: fivethirtyeight.com

Why elicit beliefs?

Why elicit beliefs?

Belief elicitation can improve recall and comprehension of data

[Kim et. al, CHI 2017]

Commonly used to increase engagement in data journalism

Beliefs can form into biases

[Wall et. al, VIS 2017]

Potentially mitigate biases

Bayesian cognitive models depend on measuring beliefs at the beginning of data exploration and again in response to some stimuli (new data)

[Karduni et. al, TVCG 2020]

Enable future research directions related to cognition, decision-making

Our contribution
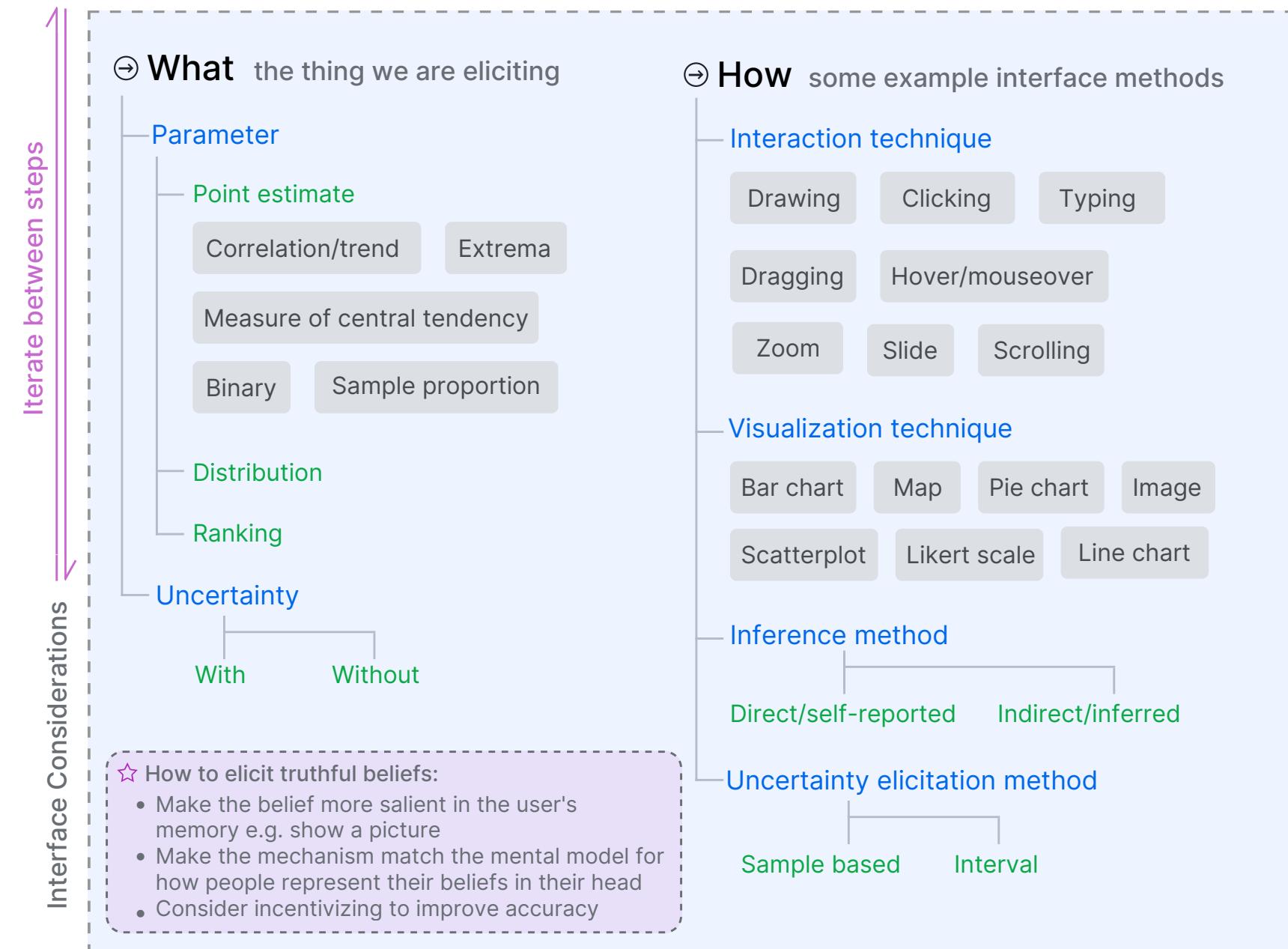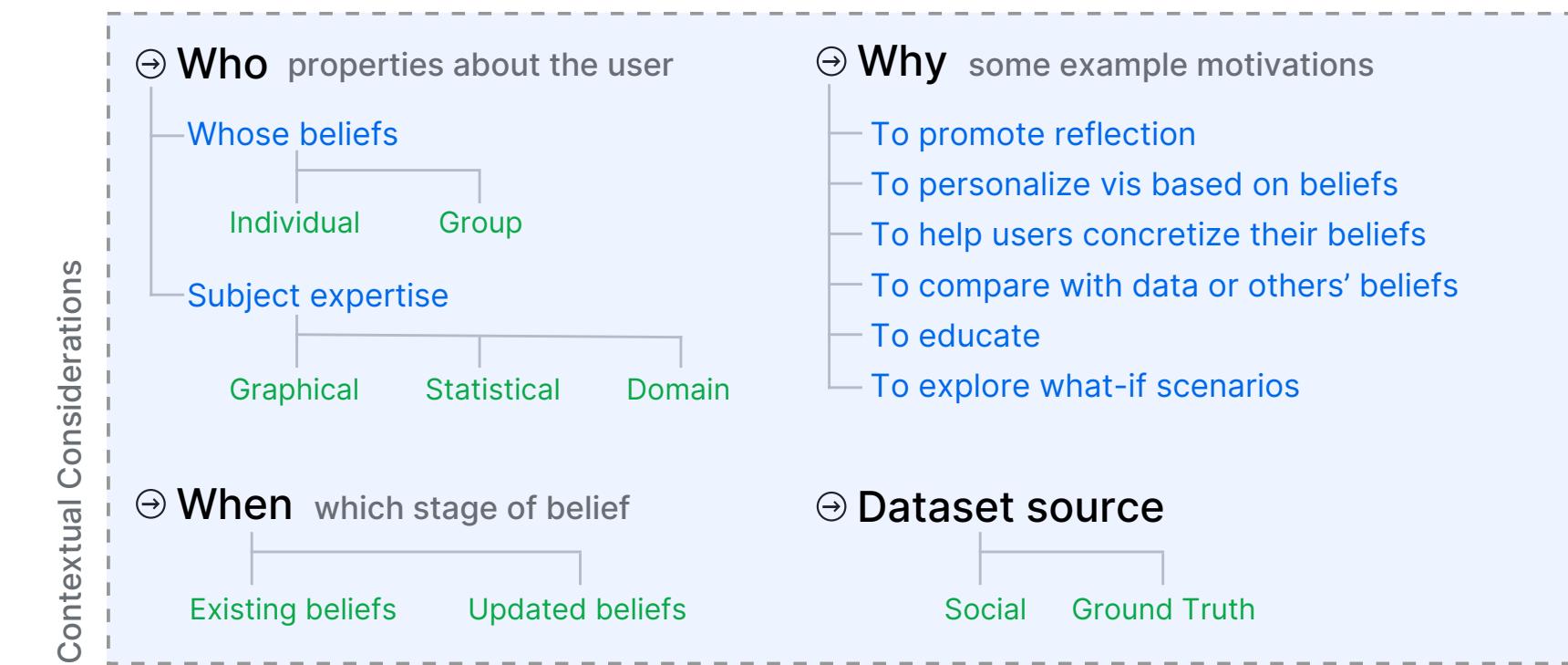
What is VIBE?

A **design space** for creating visualization-driven interfaces

for **visualization creators** like data journalists, designers, or visualization experts

to **elicit people's beliefs**

What is VIBE?

1. Belief formation

2. Belief elicitation

Contextual Considerations

⊕ **Who** properties about the user

Whose beliefs

Individual   Group

Subject expertise

Graphical   Statistical   Domain

⊕ **Why** some example motivations

- To promote reflection
- To personalize vis based on beliefs
- To help users concretize their beliefs
- To compare with data or others' beliefs
- To educate
- To explore what-if scenarios

⊕ **When** which stage of belief

Existing beliefs   Updated beliefs

⊕ **Dataset source**

Social   Ground Truth

Iterate between steps

Interface Considerations

⊕ **What** the thing we are eliciting

Parameter

Point estimate

Correlation/trend   Extrema

Measure of central tendency

Binary   Sample proportion

Distribution

Ranking

Uncertainty

With   Without

☆ How to elicit truthful beliefs:
- Make the belief more salient in the user's memory e.g. show a picture
- Make the mechanism match the mental model for how people represent their beliefs in their head
- Consider incentivizing to improve accuracy

⊕ **How** some example interface methods

Interaction technique

Drawing   Clicking   Typing

Dragging   Hover/mouseover

Zoom   Slide   Scrolling

Visualization technique

Bar chart   Map   Pie chart   Image

Scatterplot   Likert scale   Line chart

Inference method

Direct/self-reported   Indirect/inferred

Uncertainty elicitation method

Sample based   Interval

3. Seeing new data

⊕ **Feedback**

Veracity   Deviation   Explanation

4. Belief updating

What if I asked you to elicit people's beliefs for:
How does family income predict children's college chances?

Belief Elicitation

How Does Family Income Predict Children's College Chances?

| Almost No relationship | Weakly Influences | Moderately Influences | Strongly Influences | Almost Completely Influences |

Belief Elicitation

How Does Family Income Predict Children's College Chances?

Almost No Relationship · Moderately Influences · Almost Completely Influences

Percent of children who attended college

Choose the intercept → Choose the intercept uncertainty

Choose the slope uncertainty ← Choose slope

Parents' income percentile · Poorest · Richest

Methodology

Scoping and defining belief

Scoping: Stages of Belief

Belief formation

Belief elicitation

Seeing new data

Belief updating

Kim et. al, CHI 2019
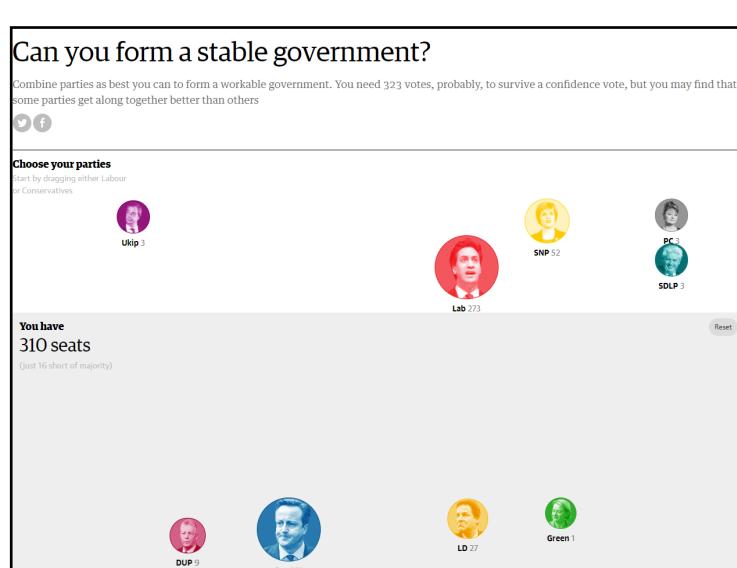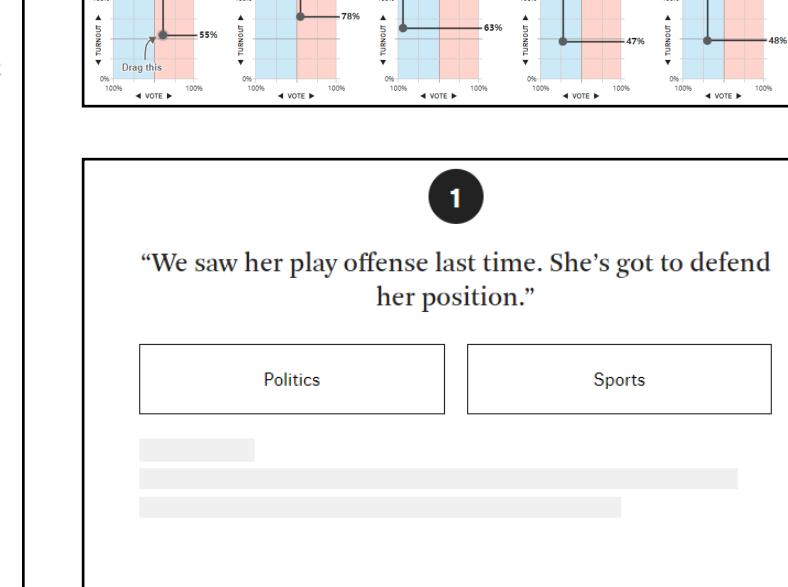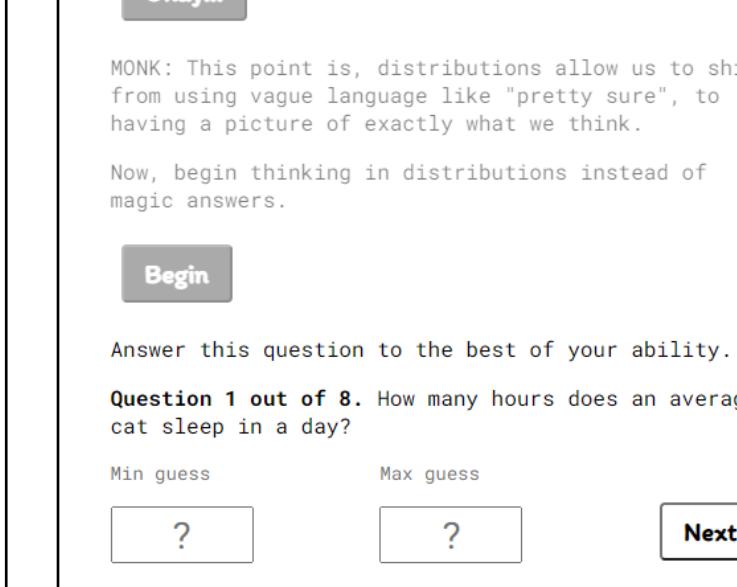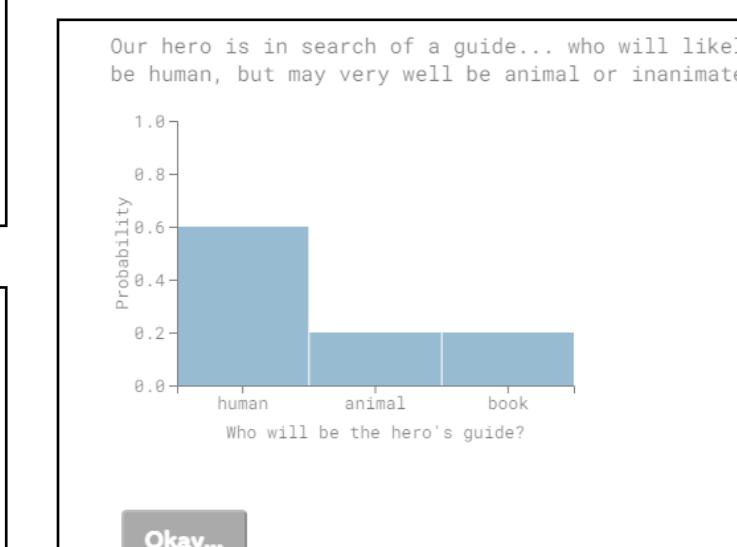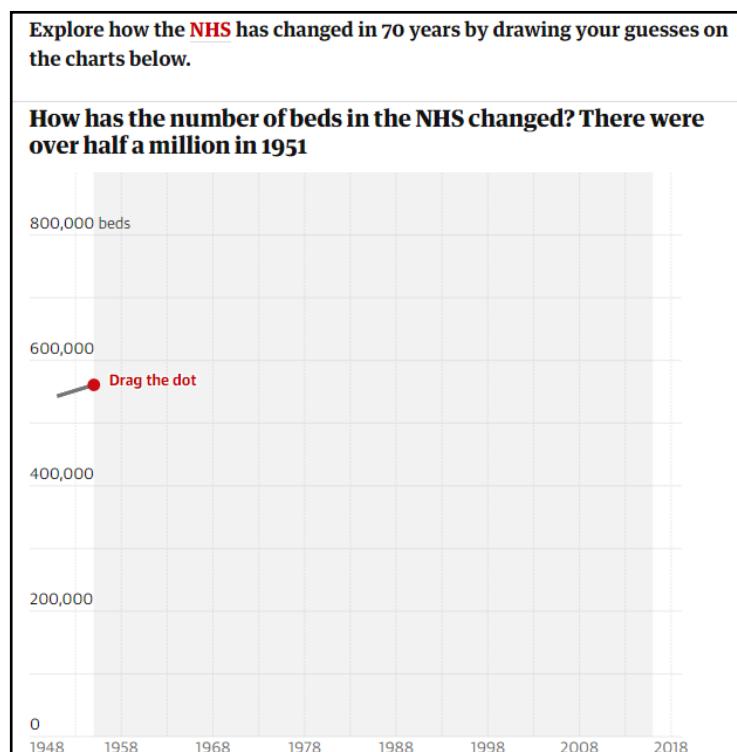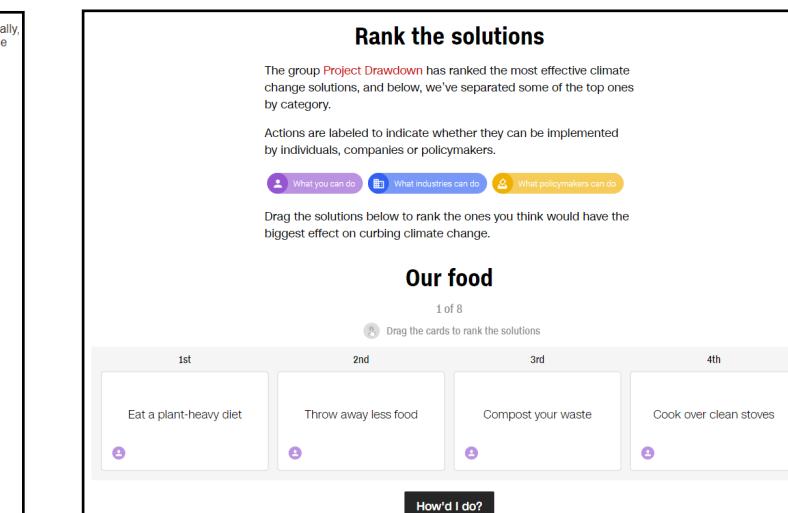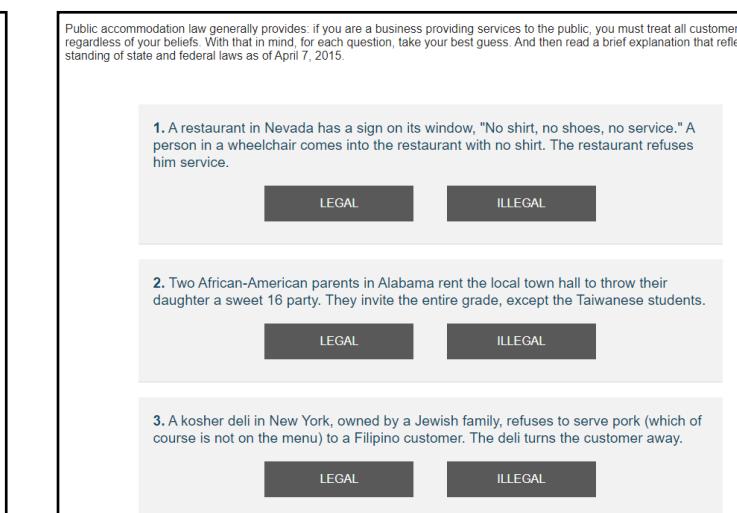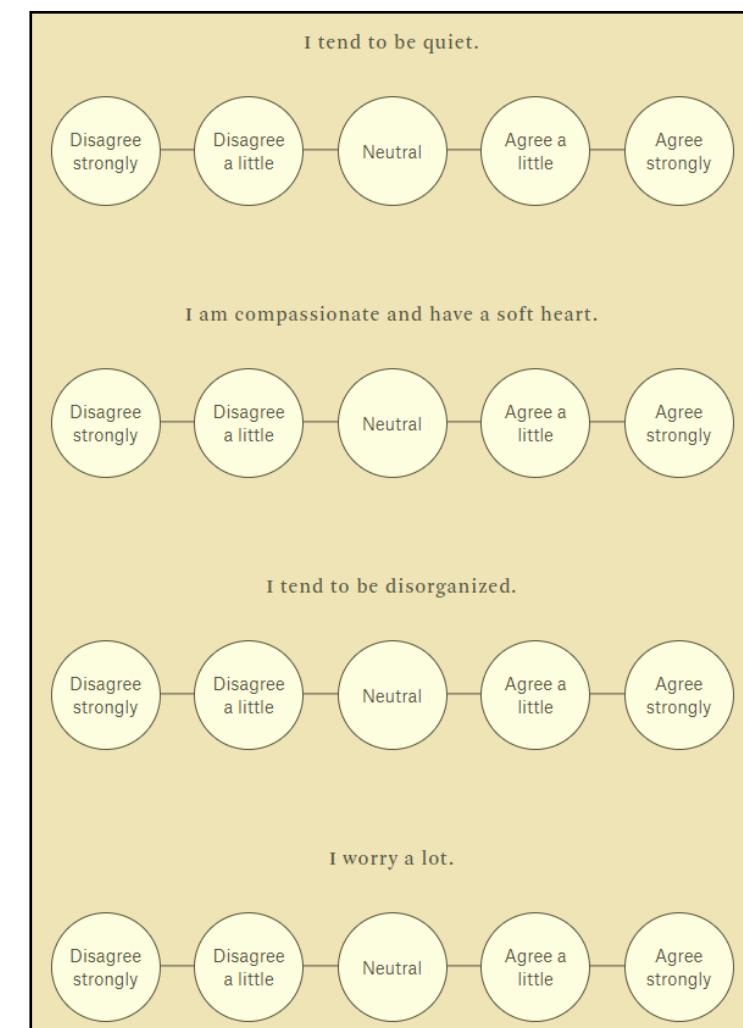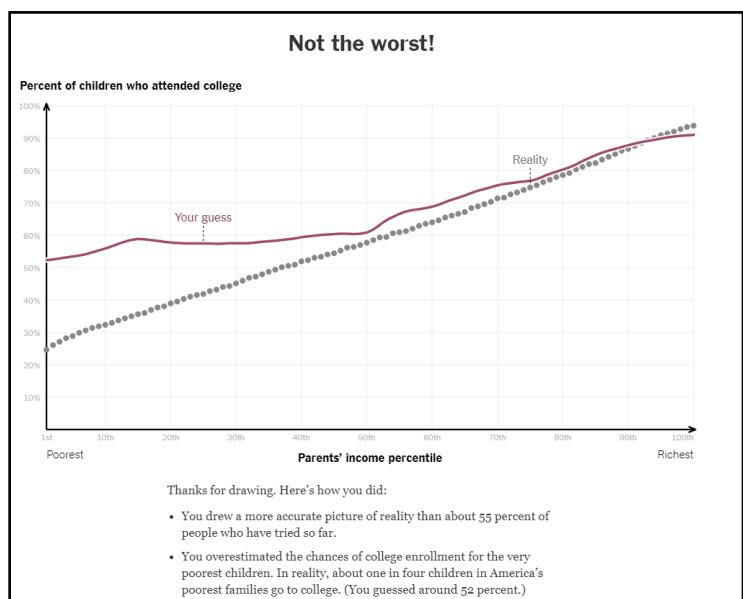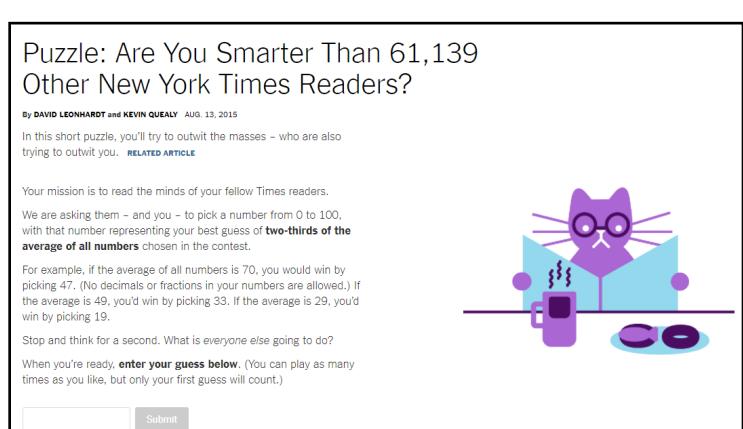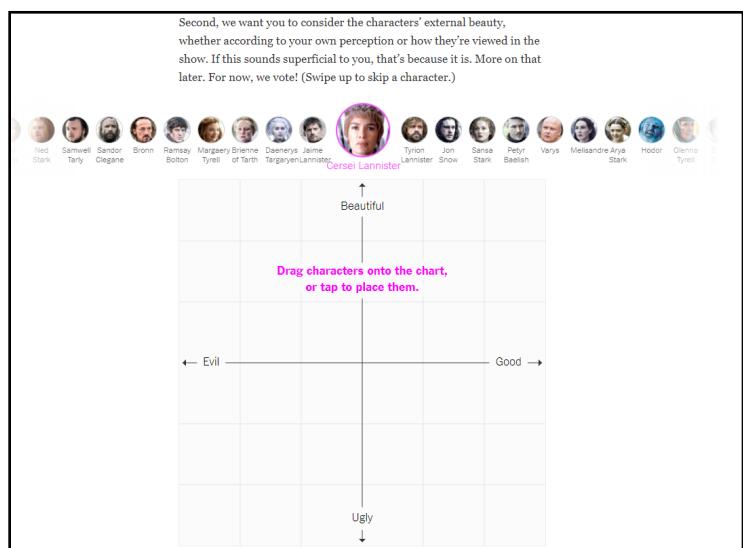
Scoping and defining belief

Collecting surveys from data journalism venues

Collecting surveys

14 examples

Scoping and defining belief

Collecting surveys from data journalism venues
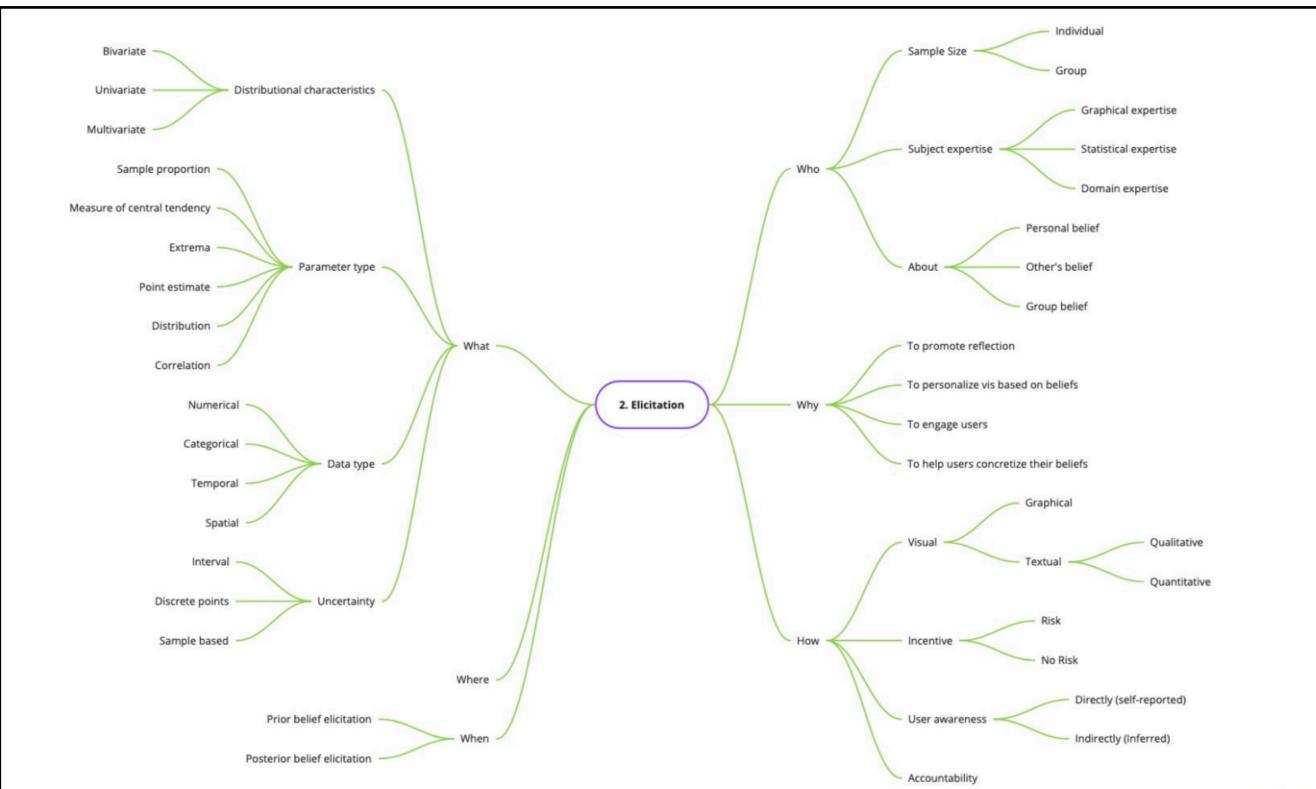
Card sort

Iterate≈7

Apply to use case (coding)
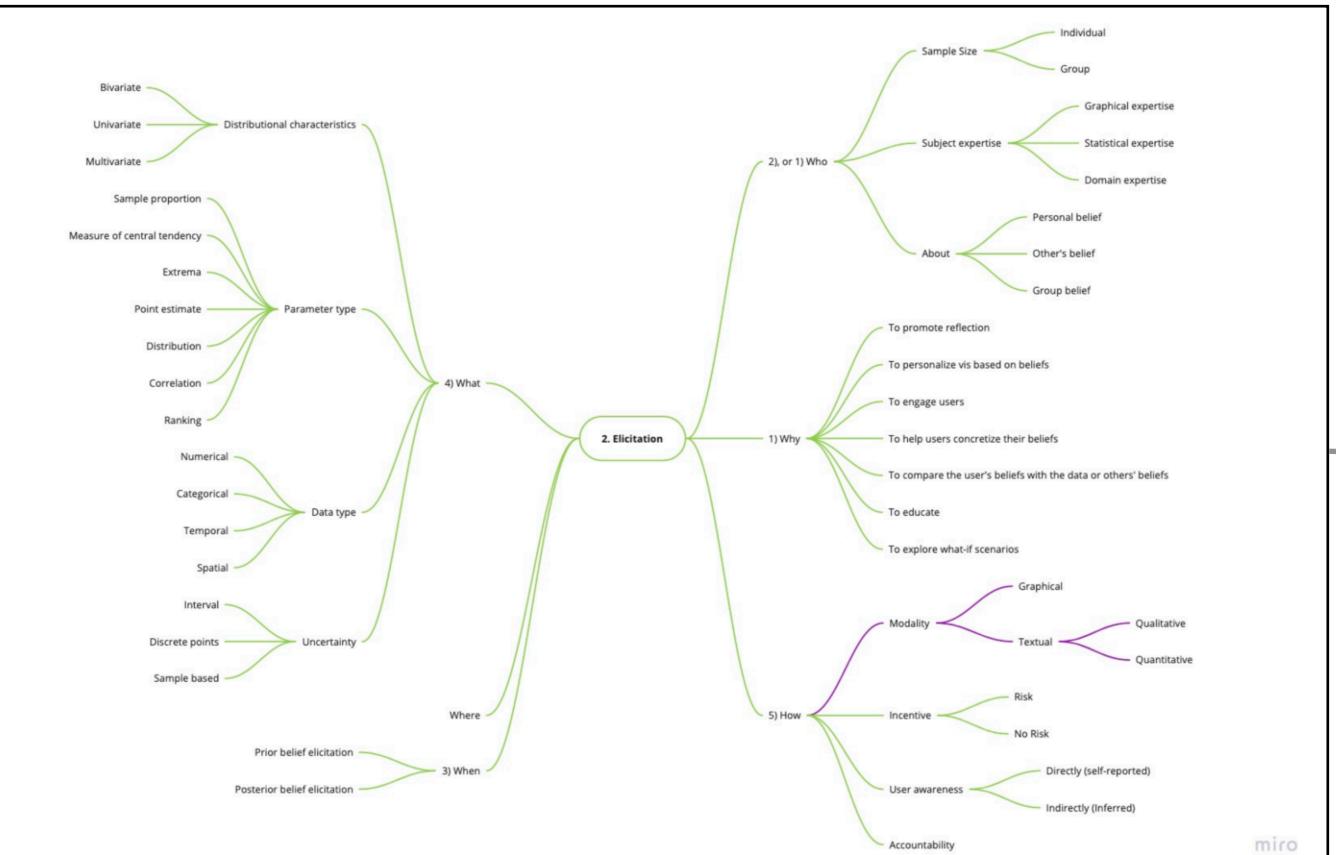
Card sort (5W+H framework)
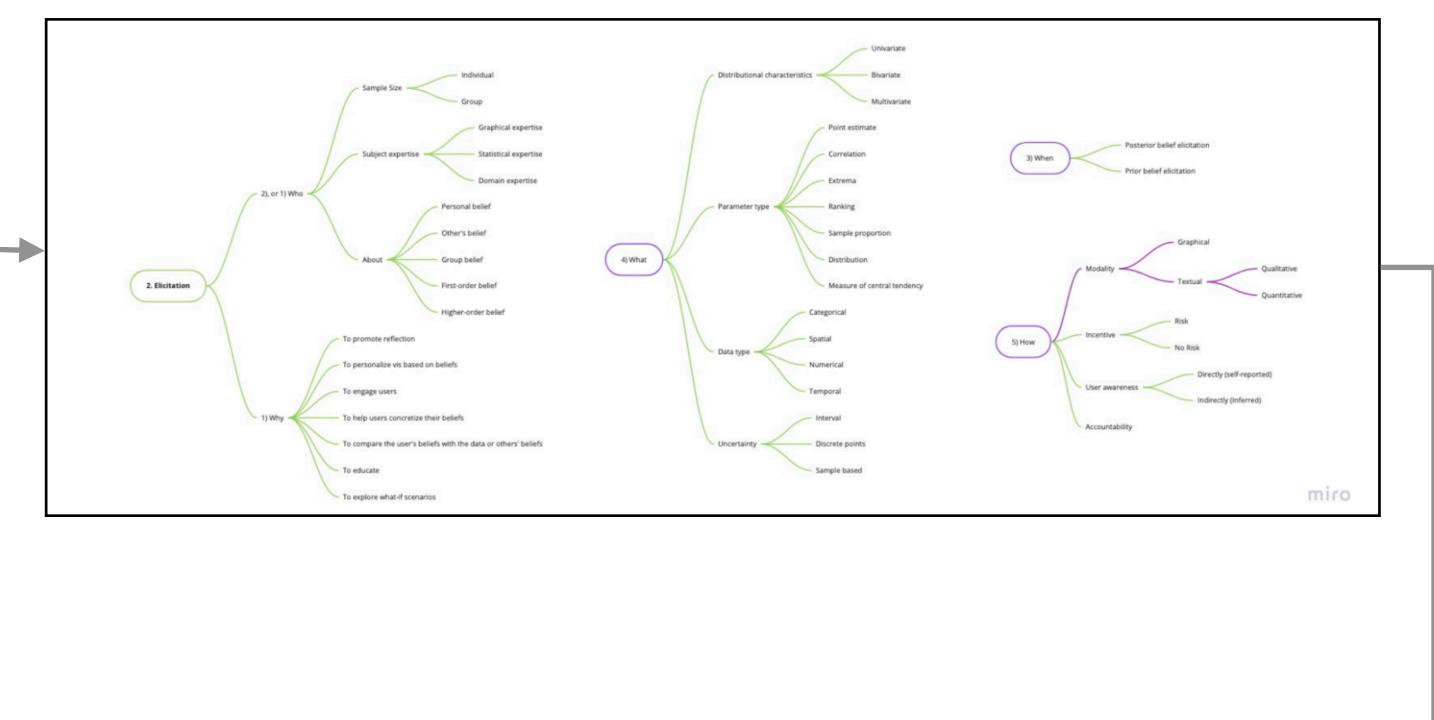
Iteration 1
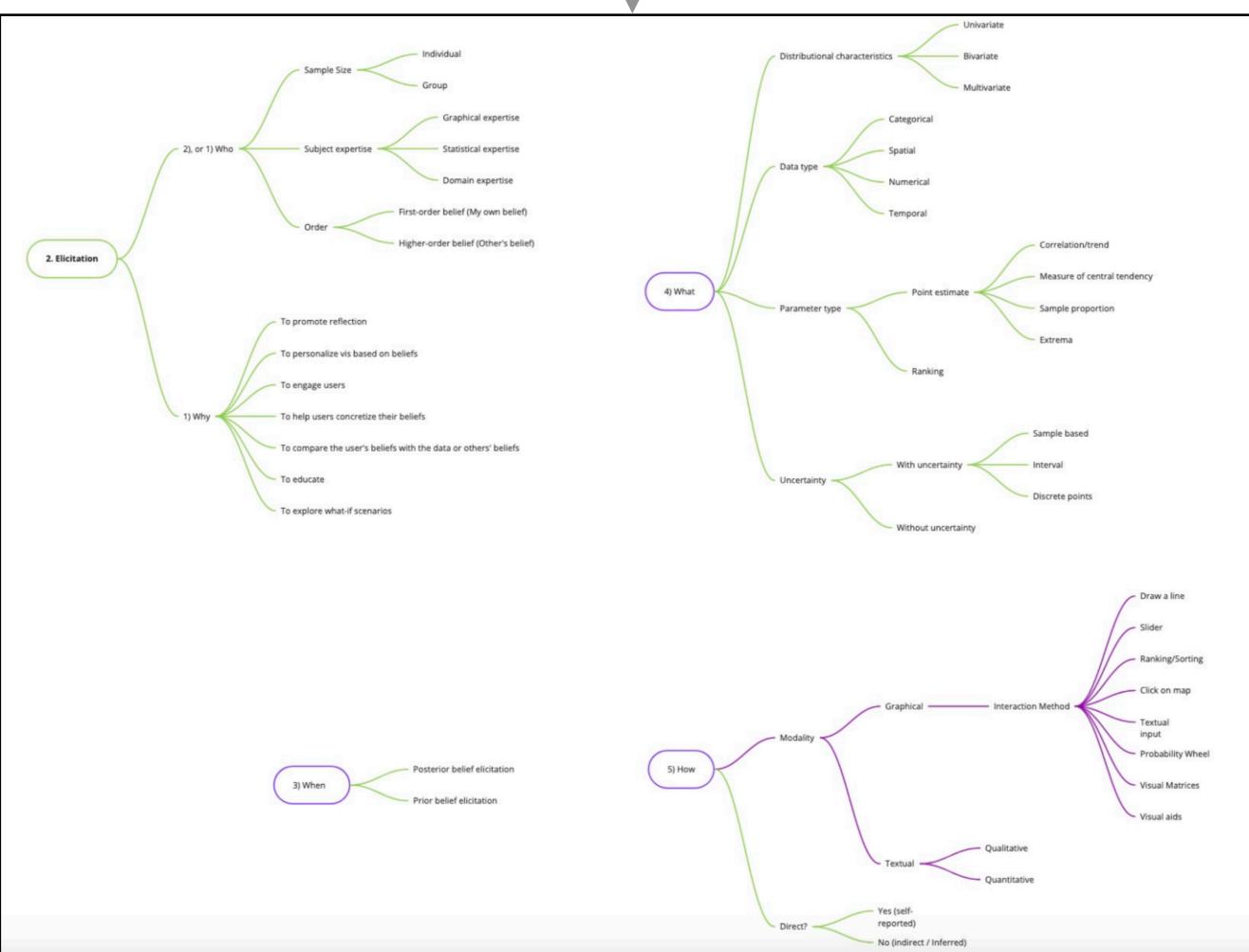
Iteration 2
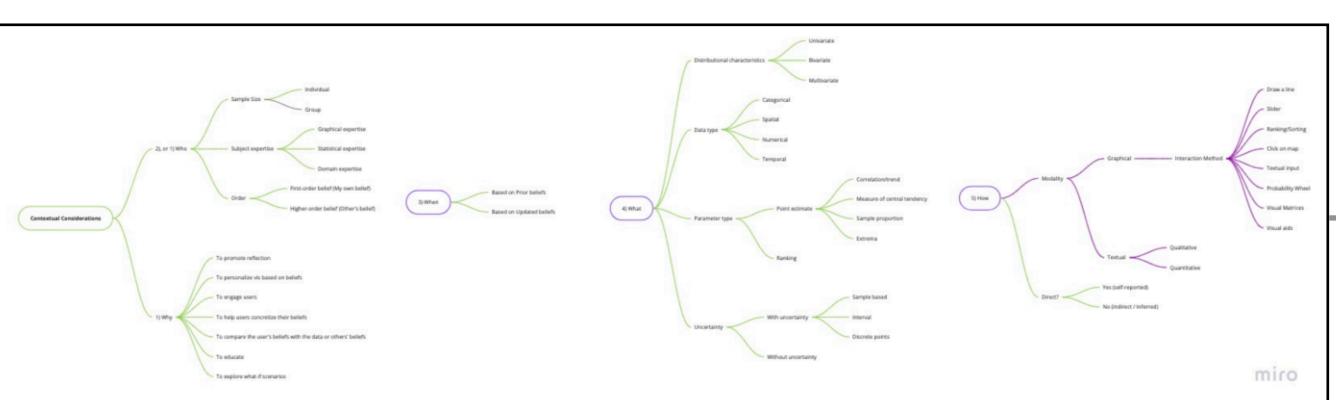
Iteration 3
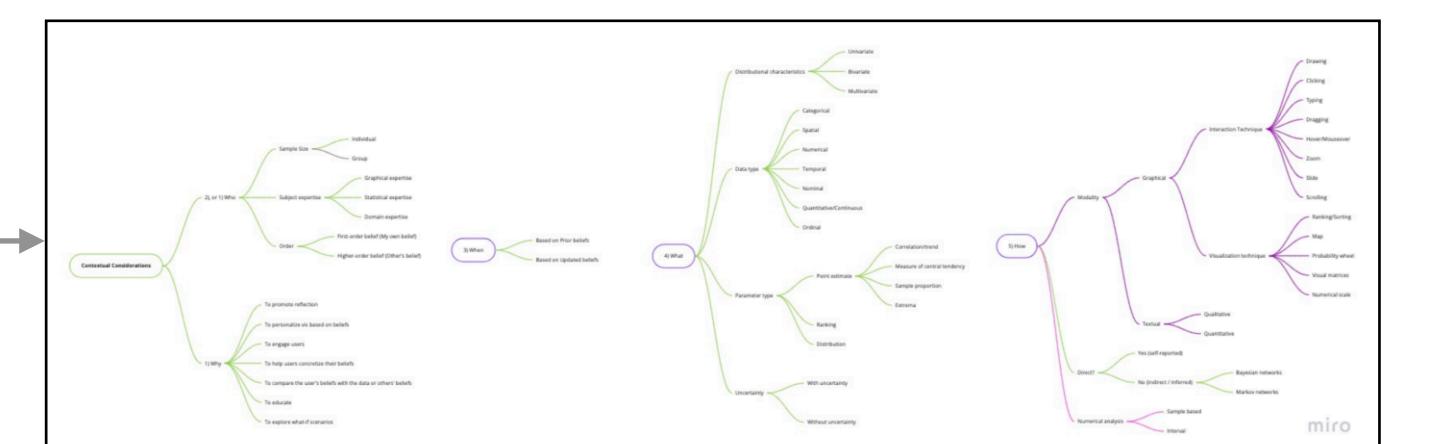
Iteration 4
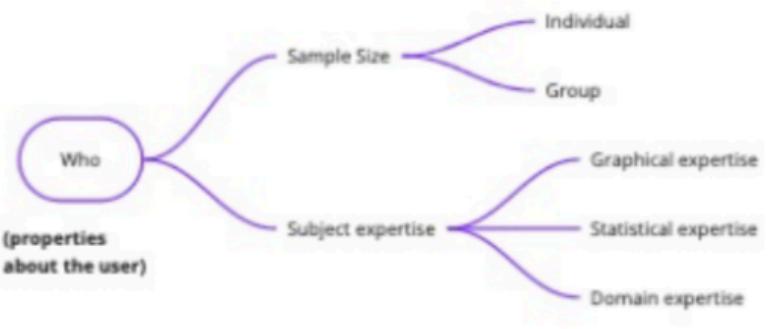
Iteration 5

Iteration 6

Contextual considerations

Who *(properties about the user)*
- Sample Size
  - Individual
  - Group
- Subject expertise
  - Graphical expertise
  - Statistical expertise
  - Domain expertise

Why *(some example motivations for eliciting beliefs)*
- To promote reflection
- To personalize vis based on beliefs
- To engage users
- To help users concretize their beliefs
- To compare the user's beliefs with the data or others' beliefs
- To educate
- To explore what-if scenarios

VIBE 1.0

Preliminary design space

Iteration 7

Elicitation design considerations

When *(the stage of the belief process that we elicit)*
- Based on Prior beliefs
- Based on Updated beliefs

What *(the thing we are eliciting)*
- Parameter
  - Point estimate
    - Correlation/trend
    - Measure of central tendency
    - Sample proportion
    - Extrema
    - Binary
  - Ranking
  - Distribution
- Uncertainty (opt.)
  - With uncertainty
  - Without uncertainty

(other properties about the belief we are eliciting)

Distributional characteristics
- Univariate
- Bivariate
- Multivariate

Data type
- Categorical
- Spatial
- Numerical
- Temporal
- Nominal
- Quantitative/Continuous
- Ordinal

Order
- First-order belief (My own belief)
- Higher-order belief (Other's belief)

How *(interface methods for eliciting the belief)*
- Interaction Technique
  - Drawing
  - Clicking
  - Typing
  - Dragging
  - Hover/Mouseover
  - Zoom
  - Slide
  - Scrolling
- Visualization technique
  - Ranking/Sorting
  - Map
  - Probability wheel
  - Visual matrices
  - Numerical scale
- Inference Method
  - Direct / Self-reported
  - Indirect / Inferred
    - Bayesian networks
    - Markov networks
- Numerical analysis
  - Sample based
  - Interval

Textual
- Qualitative
- Quantitative

VIBE 1.0
Contextual considerations

Who
(properties about the user)

- Sample Size
  - Individual
  - Group
- Subject expertise
  - Graphical expertise
  - Statistical expertise
  - Domain expertise

Why
(some example motivations for eliciting beliefs)

- To promote reflection
- To personalize vis based on beliefs
- To engage users
- To help users concretize their beliefs
- To compare the user's beliefs with the data or others' beliefs
- To educate
- To explore what-if scenarios

VIBE 1.0
Elicitation design considerations
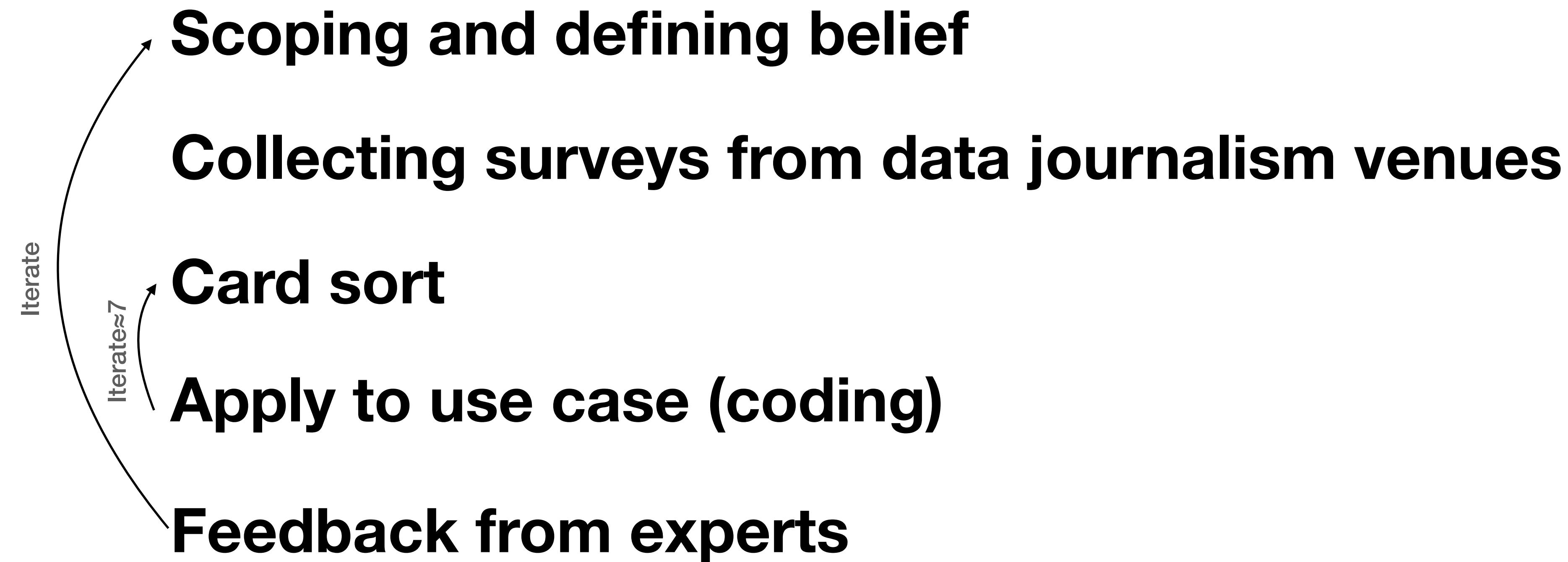
When
- Based on Prior beliefs
- Based on Updated beliefs

(the stage of the belief process that we elicit)

Scoping and defining belief

Collecting surveys from data journalism venues

Card sort

Apply to use case (coding)

Feedback from experts

Iterate≈7

Scoping and defining belief

Collecting surveys from data journalism venues

Card sort

Apply to use case (coding)

Feedback from experts

Iterate

Iterate≈7

Formative Study

Feedback from experts

5 participants

2 data visualization experts

2 data journalists

1 user experience designer

Goals

1. Assess the utility of current version of design space

2. Collect feedback to revise design space

⏱ TASK 1

Design *without* VIBE 1.0

Create a system that helps people learn about their food choices impacting climate change

or

Create a system to learn if people know America's geography

⏱ **TASK 2**

Design *with* VIBE 1.0

How do Americans feel about guns?

or

How do socio-economic factors affect higher education for Americans?

Feedback

"It is a **comprehensive framework**. I would have missed out on several design considerations in my usual design process."

"The **contextual considerations** were a **good anchor** to ask questions about the problem space."

Feedback

1. Doesn't help with scoping the problem space

✓ *Added key assumption of starting with a dataset*

2. Visually overwhelming, and not self-explanatory

✓ *Showed design considerations within the stages of belief*

Insights from expert feedback

Former scope

Belief formation

Belief elicitation

Seeing new data

Belief updating

Insights from expert feedback
Updated scope

Belief formation

Belief elicitation

Seeing new data

Belief updating

Final design space

VIBE 2.0

1. Belief formation

2. Belief elicitation

Contextual Considerations

⊕ **Who** properties about the user

Whose beliefs
- Individual
- Group

Subject expertise
- Graphical
- Statistical
- Domain

⊕ **Why** some example motivations
- To promote reflection
- To personalize vis based on beliefs
- To help users concretize their beliefs
- To compare with data or others' beliefs
- To educate
- To explore what-if scenarios

⊕ **When** which stage of belief
- Existing beliefs
- Updated beliefs

⊕ **Dataset source**
- Social
- Ground Truth

Iterate between steps

Interface Considerations

⊕ **What** the thing we are eliciting

Parameter
- Point estimate
  - Correlation/trend
  - Extrema
  - Measure of central tendency
  - Binary
  - Sample proportion
- Distribution
- Ranking

Uncertainty
- With
- Without

☆ How to elicit truthful beliefs:
- Make the belief more salient in the user's memory e.g. show a picture
- Make the mechanism match the mental model for how people represent their beliefs in their head
- Consider incentivizing to improve accuracy

⊕ **How** some example interface methods

Interaction technique
- Drawing
- Clicking
- Typing
- Dragging
- Hover/mouseover
- Zoom
- Slide
- Scrolling

Visualization technique
- Bar chart
- Map
- Pie chart
- Image
- Scatterplot
- Likert scale
- Line chart

Inference method
- Direct/self-reported
- Indirect/inferred

Uncertainty elicitation method
- Sample based
- Interval

3. Seeing new data

⊕ **Feedback**
- Veracity
- Deviation
- Explanation

4. Belief updating

Contextual considerations

⊕ Who properties about the user

— Whose beliefs
  — Individual   Group

— Subject expertise
  — Graphical   Statistical   Domain

⊕ When which stage of belief

— Existing beliefs   Updated beliefs

⊕ Why some example motivations

— To promote reflection

— To personalize vis based on beliefs

— To compare with
  Data   Others' beliefs   Both

— To educate

⊕ Data set

should have problem and data set fixed

Here's a simple question we'd like you to answer before you read further: **Which of the nine maps below best describes what you think of as the heartland?**

To help you get started: Do big cities belong in the heartland? (If not, choose a map with "holes" in it.) Does the heartland rigidly follow state lines? Does it venture south into Texas, or east into Pennsylvania? Choose one of the maps below.
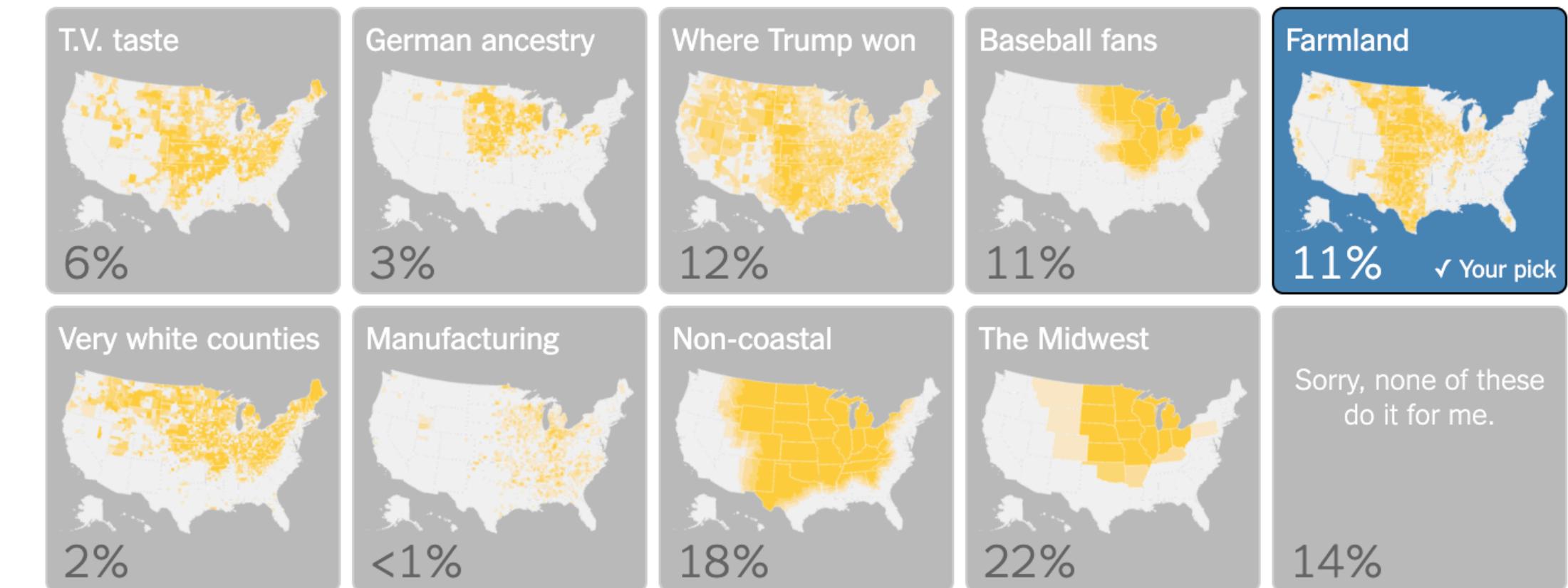
Sorry, none of these do it for me.
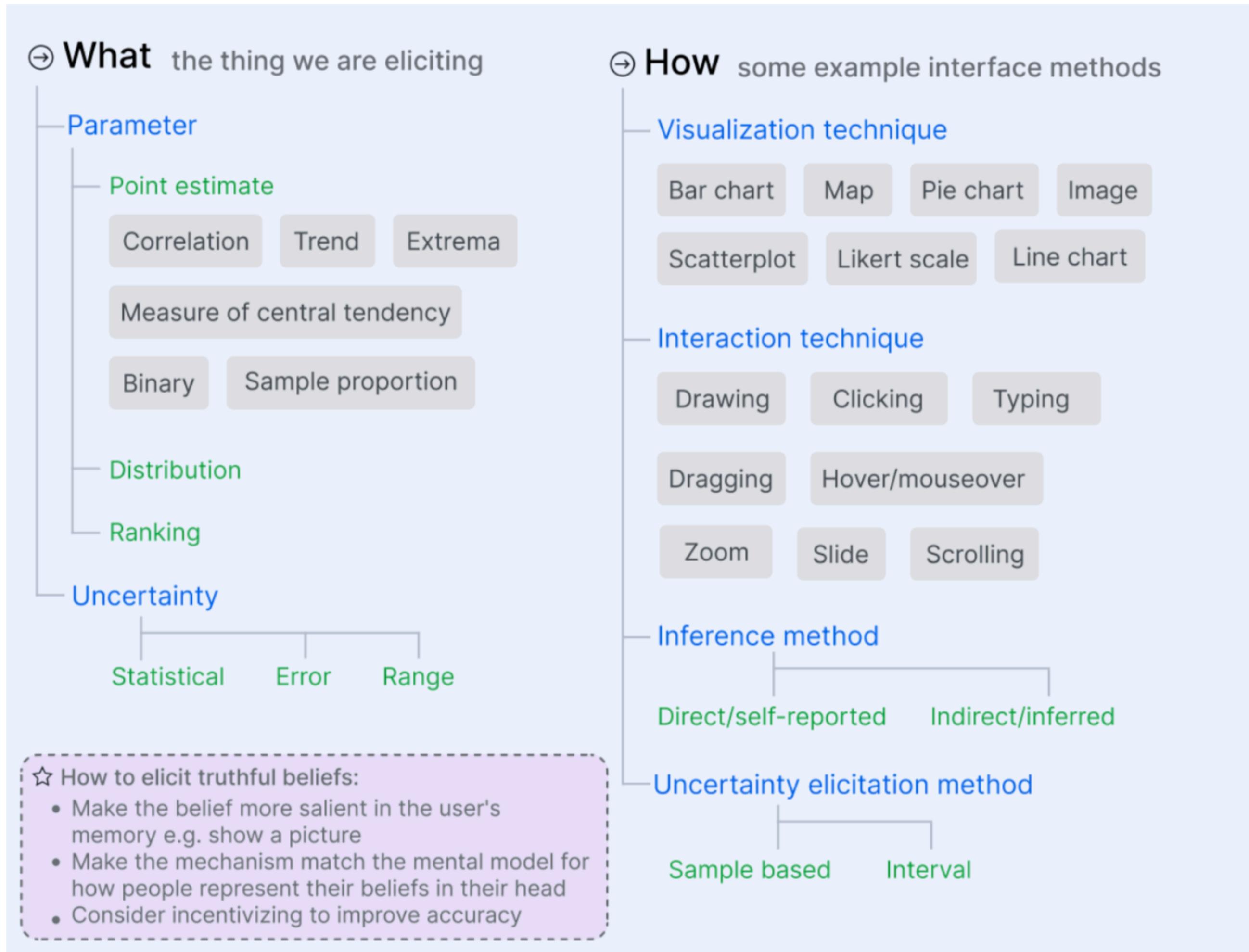
Count my vote!

Feedback: social data

What the maps actually show, and what other Times readers thought

T.V. taste	German ancestry	Where Trump won	Baseball fans	Farmland
6%	3%	12%	11%	11% ✓ Your pick
Very white counties	Manufacturing	Non-coastal	The Midwest	Sorry, none of these do it for me.
2%	<1%	18%	22%	14%

Interface considerations

⊕ What the thing we are eliciting

Parameter
— Point estimate
  - Correlation
  - Trend
  - Extrema
  - Measure of central tendency
  - Binary
  - Sample proportion
— Distribution
— Ranking
Uncertainty
  - Statistical
  - Error
  - Range

☆ How to elicit truthful beliefs:
- Make the belief more salient in the user's memory e.g. show a picture
- Make the mechanism match the mental model for how people represent their beliefs in their head
- Consider incentivizing to improve accuracy

⊕ How some example interface methods

Visualization technique
  - Bar chart
  - Map
  - Pie chart
  - Image
  - Scatterplot
  - Likert scale
  - Line chart

Interaction technique
  - Drawing
  - Clicking
  - Typing
  - Dragging
  - Hover/mouseover
  - Zoom
  - Slide
  - Scrolling

Inference method
  - Direct/self-reported
  - Indirect/inferred

Uncertainty elicitation method
  - Sample based
  - Interval

What

Measure of central tendency	Median value of home pricing
Extrema	Most expensive homes in a district
Discrete distribution	Number of semesters taken to graduate from different majors
Continuous distribution	Age of graduating students in a major
Ranking	Rank order of graduating students by GPA

How

Drag and drop these choices to rank the hardest graduate CS programs to get into.

1	Harvard University
2	Stanford University
3	Princeton University
4	Columbia University

What are the acceptance rates at these universities for their graduate CS program?

Harvard University		%
Stanford University		%
Princeton University		%
Columbia University		%

✅

❌

Guides for eliciting truthful beliefs

Make the mechanism match their mental model
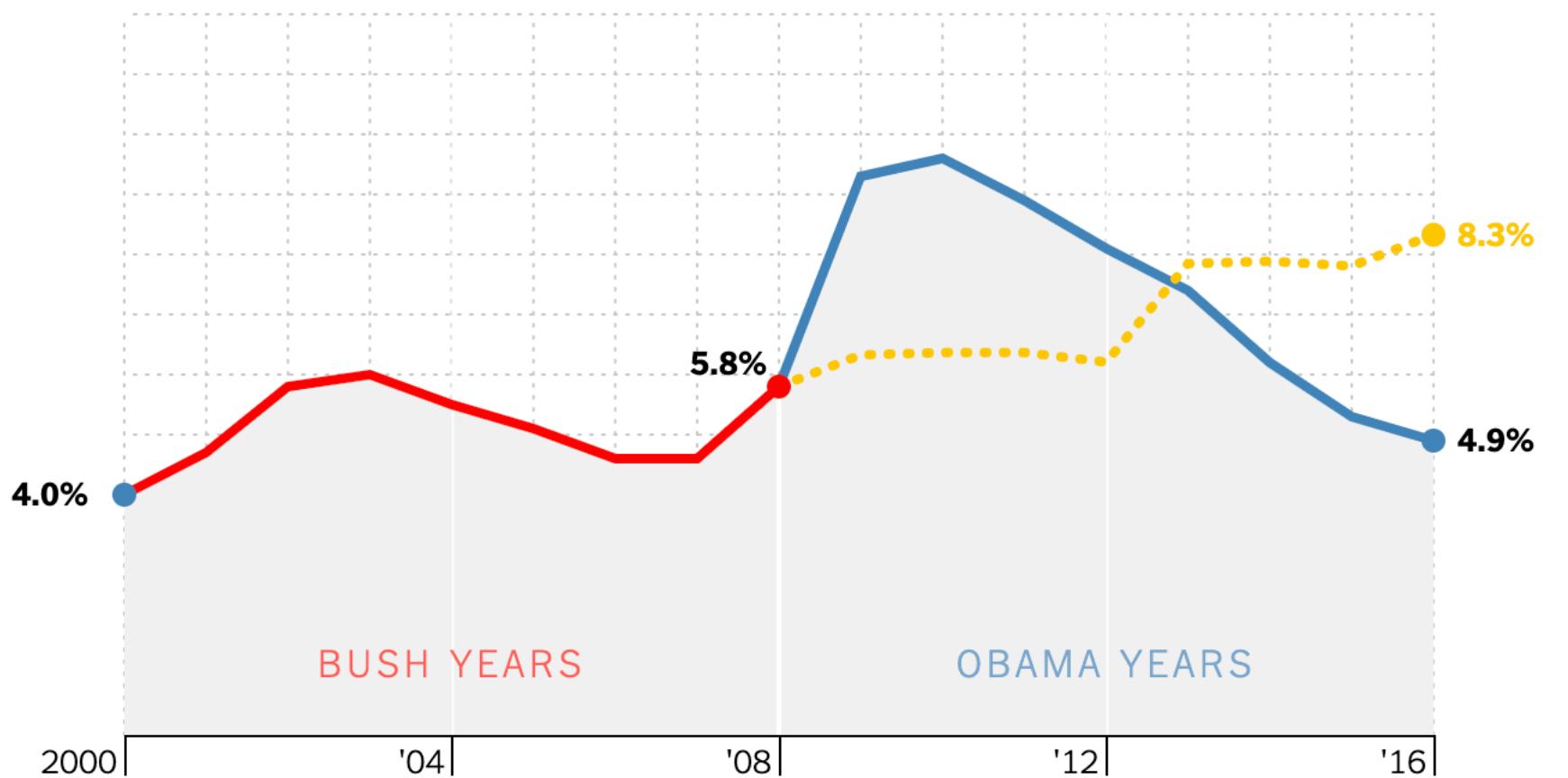
Make the beliefs more salient in their memory

Consider incentivizing

Seeing new data

Feedback

Type			Source	
Veracity	Deviation	Explanation	Social	Ground truth

Under President Obama, the **unemployment rate** …

4.0%

5.8%

8.3%

4.9%

BUSH YEARS

OBAMA YEARS

2000   '04   '08   '12   '16
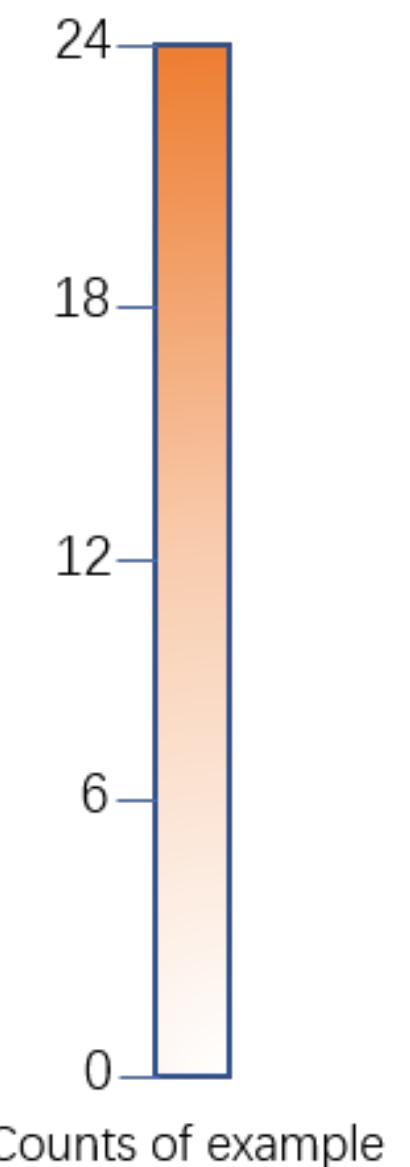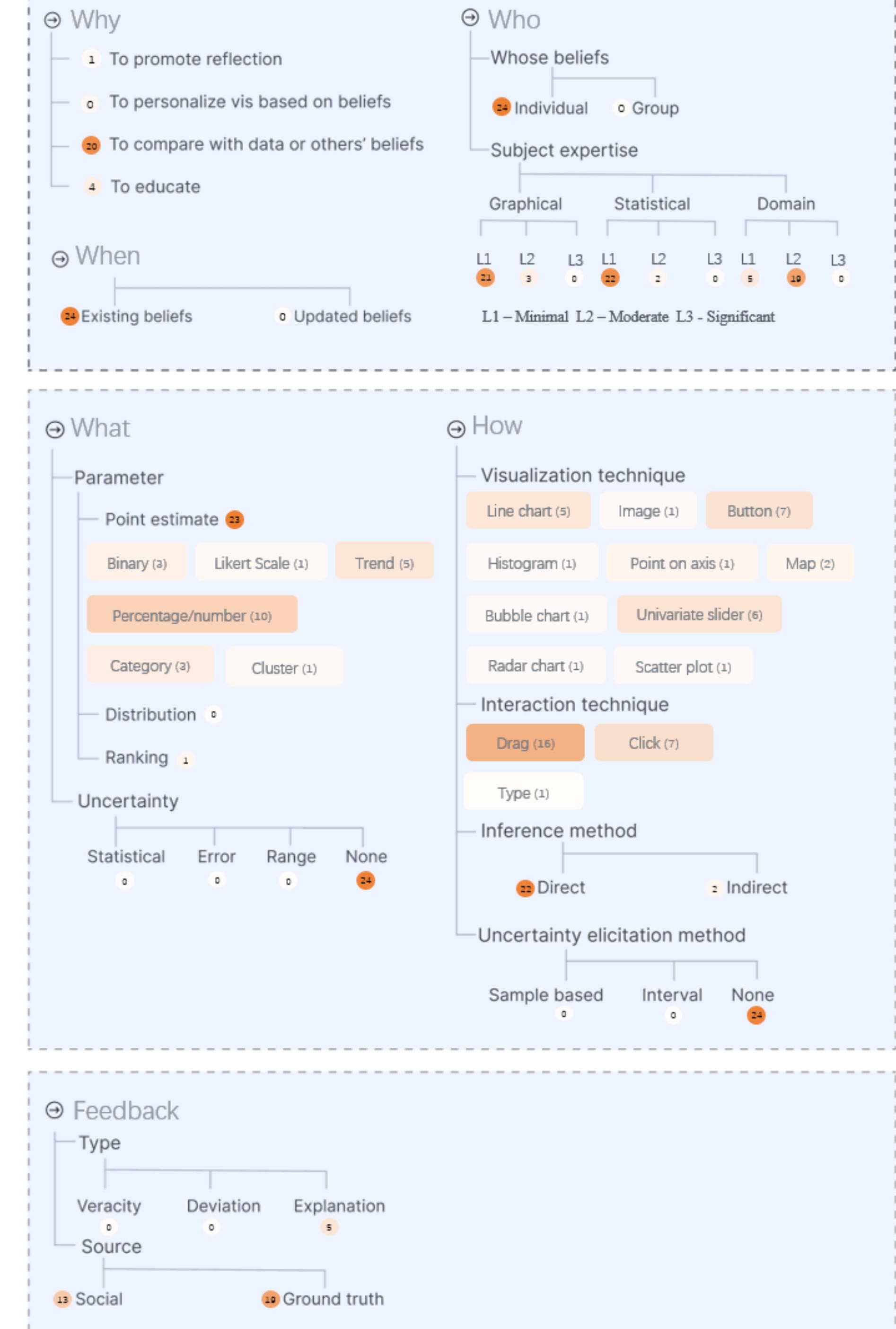
… **reached its lowest level since 2007**. The current rate is a stunning decline from the 9.3 percent in 2009, the year Mr. Obama took office.

How do existing articles fit into VIBE?

Coding the design space

Why
- 1 To promote reflection
- 0 To personalize vis based on beliefs
- 20 To compare with data or others' beliefs
- 4 To educate

When
- 24 Existing beliefs   0 Updated beliefs

Who
Whose beliefs
- 24 Individual   0 Group

Subject expertise

Graphical			Statistical			Domain		
L1	L2	L3	L1	L2	L3	L1	L2	L3
21	3	0	22	2	0	5	19	0

L1 – Minimal   L2 – Moderate   L3 - Significant

Counts of example

What
Parameter
- Point estimate 23

| Binary (3) | Likert Scale (1) | Trend (5) |

Percentage/number (10)

| Category (3) | Cluster (1) |

- Distribution 0
- Ranking 1

Uncertainty

| Statistical | Error | Range | None |
| 0 | 0 | 0 | 24 |

How
Visualization technique

Line chart (5)	Image (1)	Button (7)
Histogram (1)	Point on axis (1)	Map (2)
Bubble chart (1)	Univariate slider (6)	
Radar chart (1)	Scatter plot (1)	

Interaction technique

| Drag (16) | Click (7) |

Type (1)

Inference method
- 22 Direct   2 Indirect

Uncertainty elicitation method

| Sample based | Interval | None |
| 0 | 0 | 24 |

Feedback
Type

| Veracity | Deviation | Explanation |
| 0 | 0 | 5 |

Source
- 13 Social   19 Ground truth

0.833 Cohen's Kappa
(high inter-rater reliability)

Insights from coding
Opportunities

Who
Group and social beliefs

What
Uncertainty

How
Inferred beliefs

How to apply VIBE

You Draw It: How Family Income Predicts Children's College Chances

By **GREGOR AISCH**, **AMANDA COX** and **KEVIN QUEALY**   MAY 28, 2015

How likely is it that children who grow up in very poor families go to college? How about children who grow up in very rich families?

We'd like you to **draw your guess** for every income level on the chart below.

If you think the chances of enrolling in college (or vocational school) are about the same for everyone, you should draw something like this: ——— . If you think the odds are especially harsh for children from the poorest families, but higher for middle- and higher-income children, your drawing would instead look like this: ⟋ . Or here is one for a situation in which chances level off after a certain income threshold: ⟋ . Or for one that spikes ⌐ or dips ⌐ for the very richest.

When you've finished drawing, we'll compare your line to the reality for children born in the early 1980s, based on research by a team of economists. We've started you off with one free point: 58 percent of children who were born in the early 1980s and raised in median-income families enrolled in higher education by the time they were 21. One way or another, your chart should go through that point.

Time to draw!

Coding VIBE
1. Contextual considerations

Who

Minimum to moderate graphical and statistical expertise

Why

To educate

When

Prior beliefs

Dataset

Parents' income percentile vs children's educational attainment

Coding VIBE

2. Interface considerations
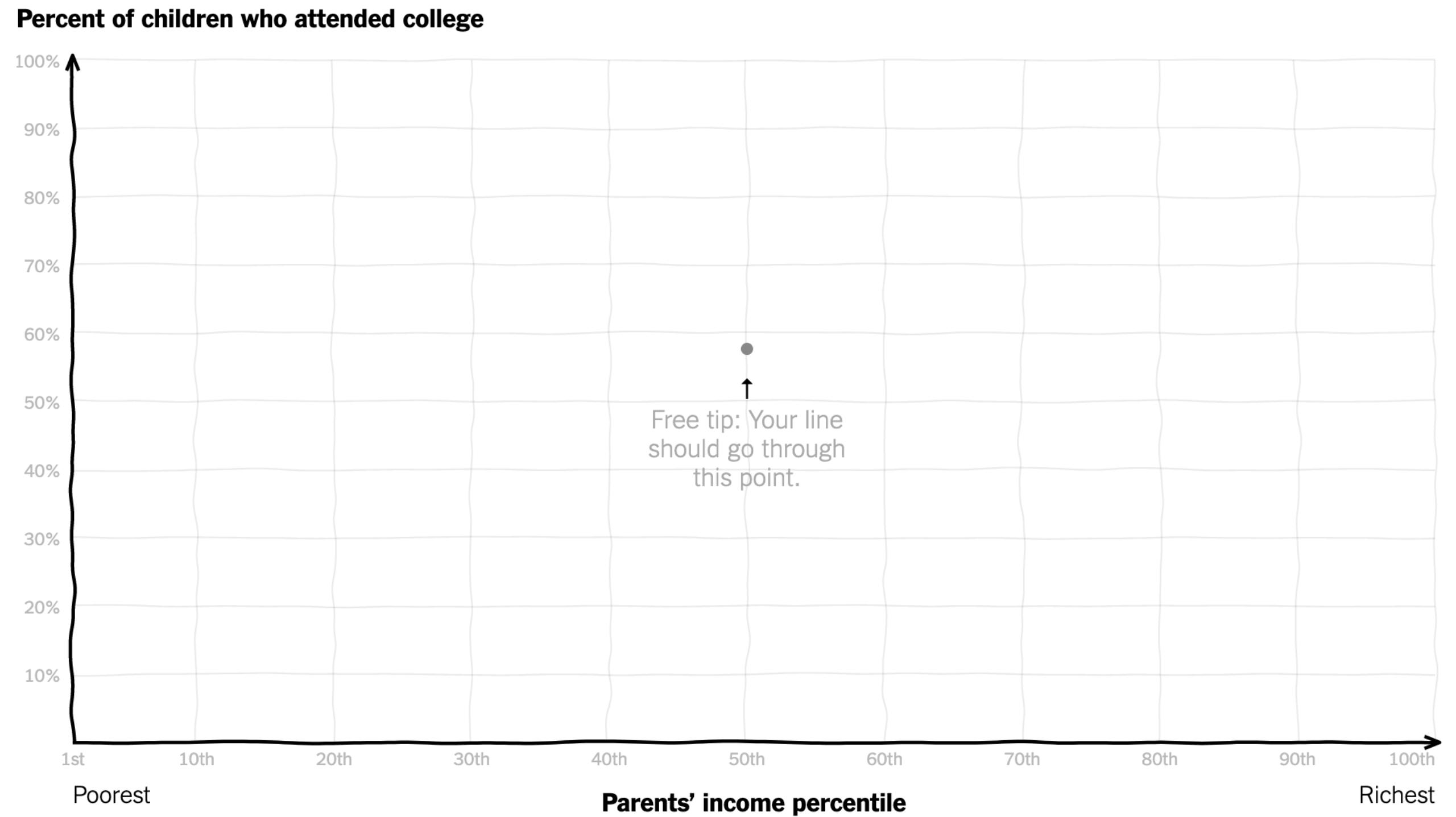
What

Parameter: Correlation

How

1. **Vis technique: Line chart**
2. **Interaction technique: Drawing**

Draw your line on the chart below

Percent of children who attended college

100%

90%

80%

70%

60%

50%

Free tip: Your line should go through this point.

40%

30%

20%

10%

| 1st | 10th | 20th | 30th | 40th | 50th | 60th | 70th | 80th | 90th | 100th |

Poorest

Parents' income percentile

Richest
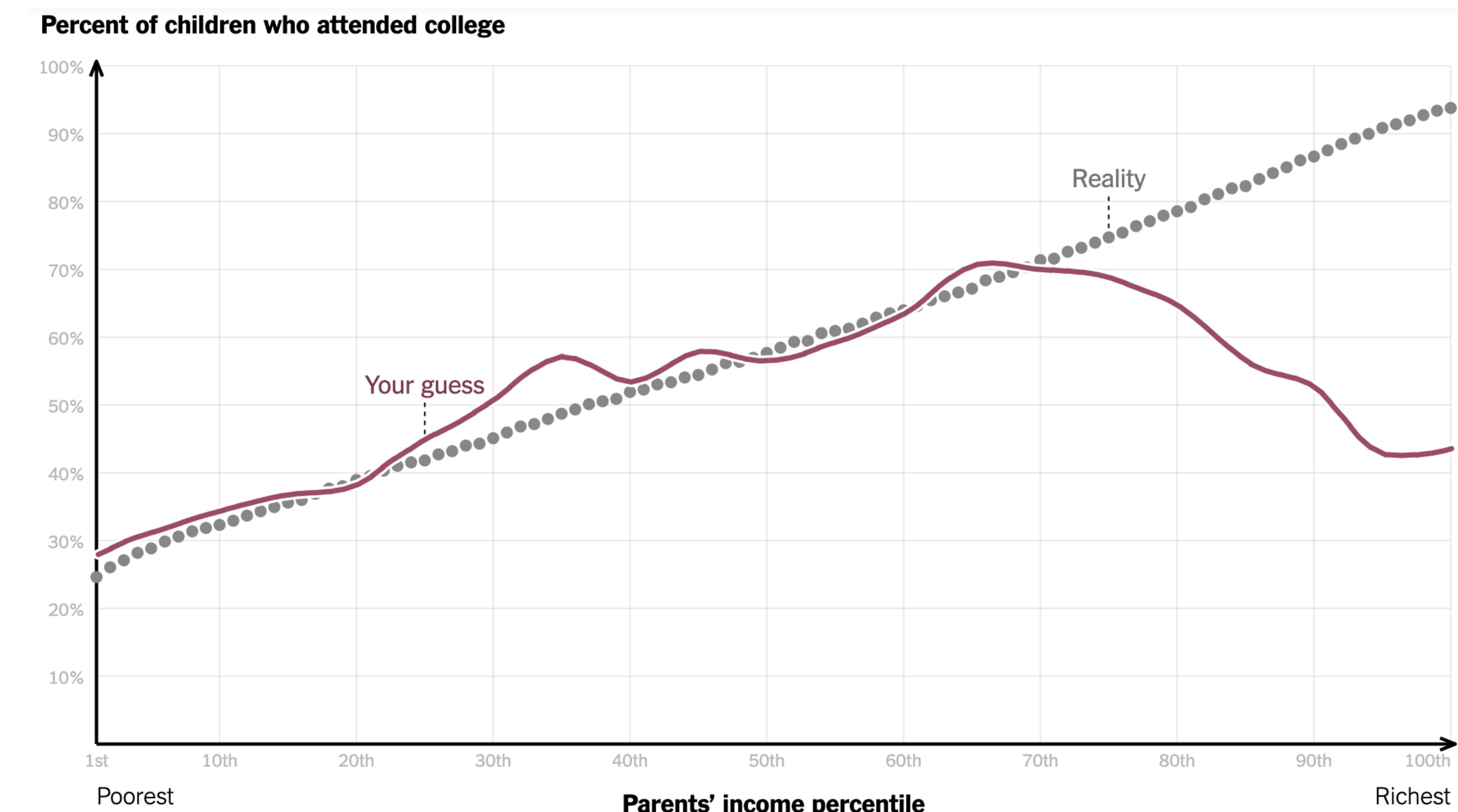
I'm done   Start over

Coding VIBE
3. Seeing new data

Feedback

1. Type: Deviation
2. Source: Ground truth

Percent of children who attended college

Thanks for drawing. Here's how you did:

- About 75 percent of people drew a more accurate picture of reality than you did.

- You correctly guessed that children from the very poorest families face tough odds in going to college – only about one in four do.

- You underestimated the chances of college enrollment for the very richest children. In reality, about 94 percent of children from America's richest families go to college. (You guessed around 44 percent.)
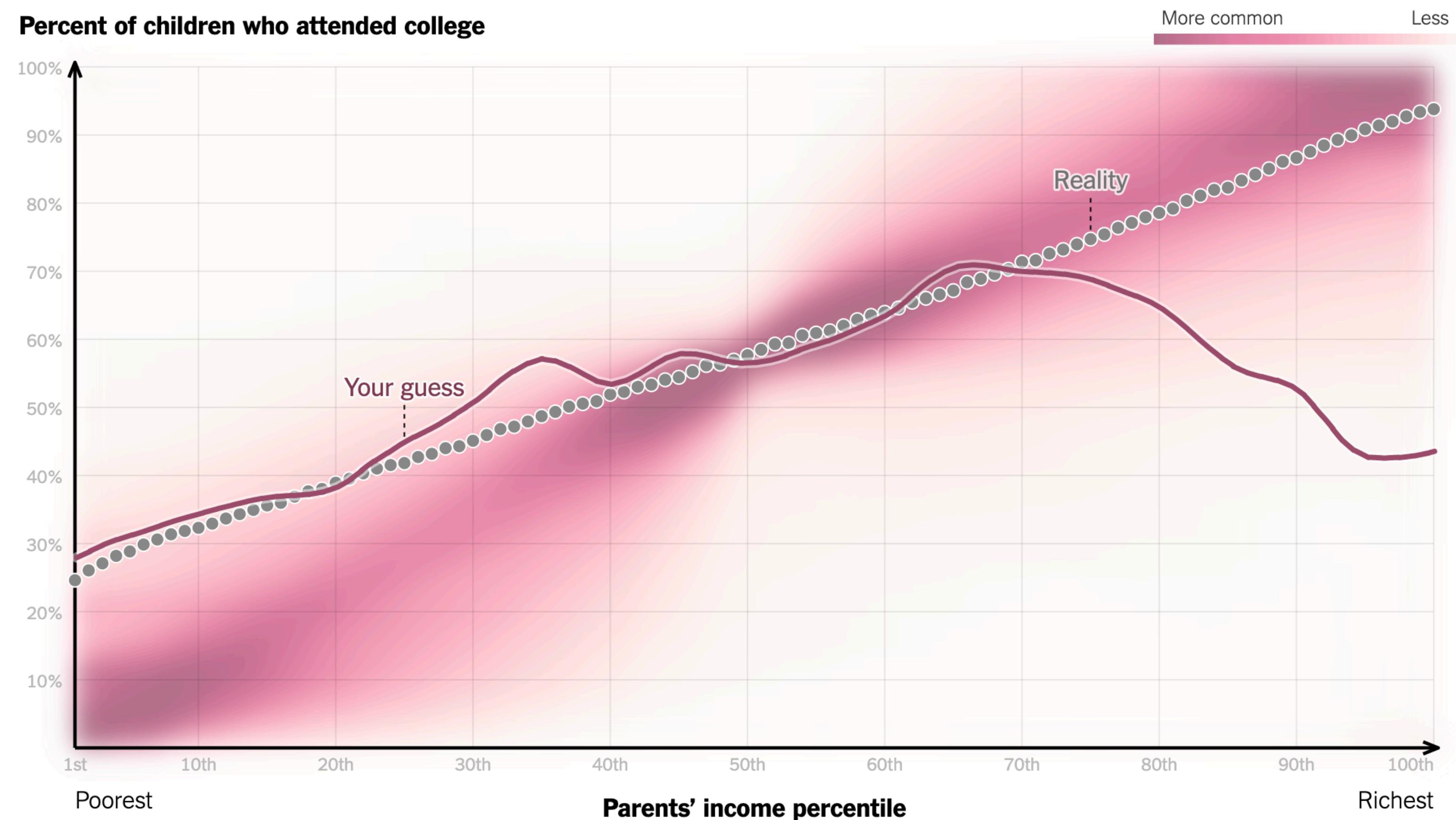
Coding VIBE

3. Seeing new data

Feedback

1. Type: Deviation
2. Source: Social Data

In case you were wondering, here's a chart showing the aggregate choices of 78,022 other New York Times readers. (Remember, we gave you a free point at the 50th percentile, which explains why so many people were so accurate there.)

Percent of children who attended college

More common — Less

Reality

Your guess

Parents' income percentile

Poorest

Richest

Redesign using VIBE

Iteration 1

Belief Elicitation

How Does Family Income Predict Children's College Chances?

| Almost No relationship | Weakly Influences | Moderately Influences | Strongly Influences | Almost Completely Influences |

Feedback

| Almost No relationship | Weakly Influences | Moderately Influences | Strongly Influences | Almost Completely Influences |

Your guess | Reality!

What

Parameter: Correlation

How

1. **Visualization technique: Likert scale**
2. **Interaction technique: Clicking**

Redesign using VIBE
Iteration 2
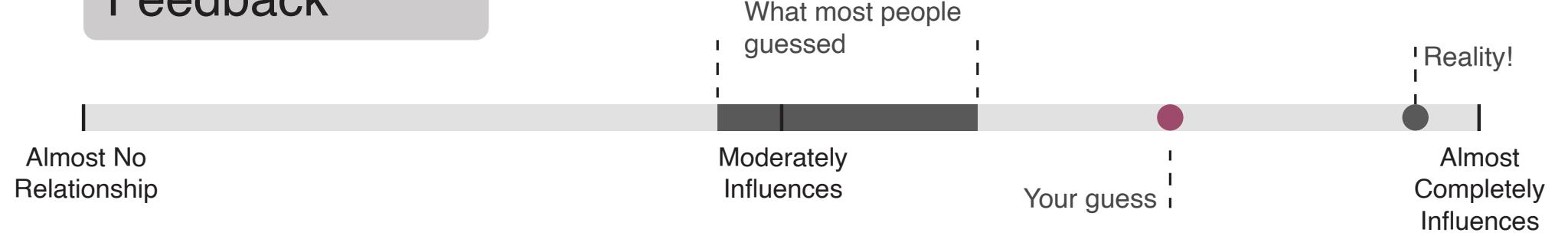
Belief Elicitation

How Does Family Income Predict Children's College Chances?

Almost No Relationship | Moderately Influences | Almost Completely Influences

Feedback

What most people guessed

Reality!

Almost No Relationship | Moderately Influences | Almost Completely Influences

Your guess

What

Parameter: Correlation

How
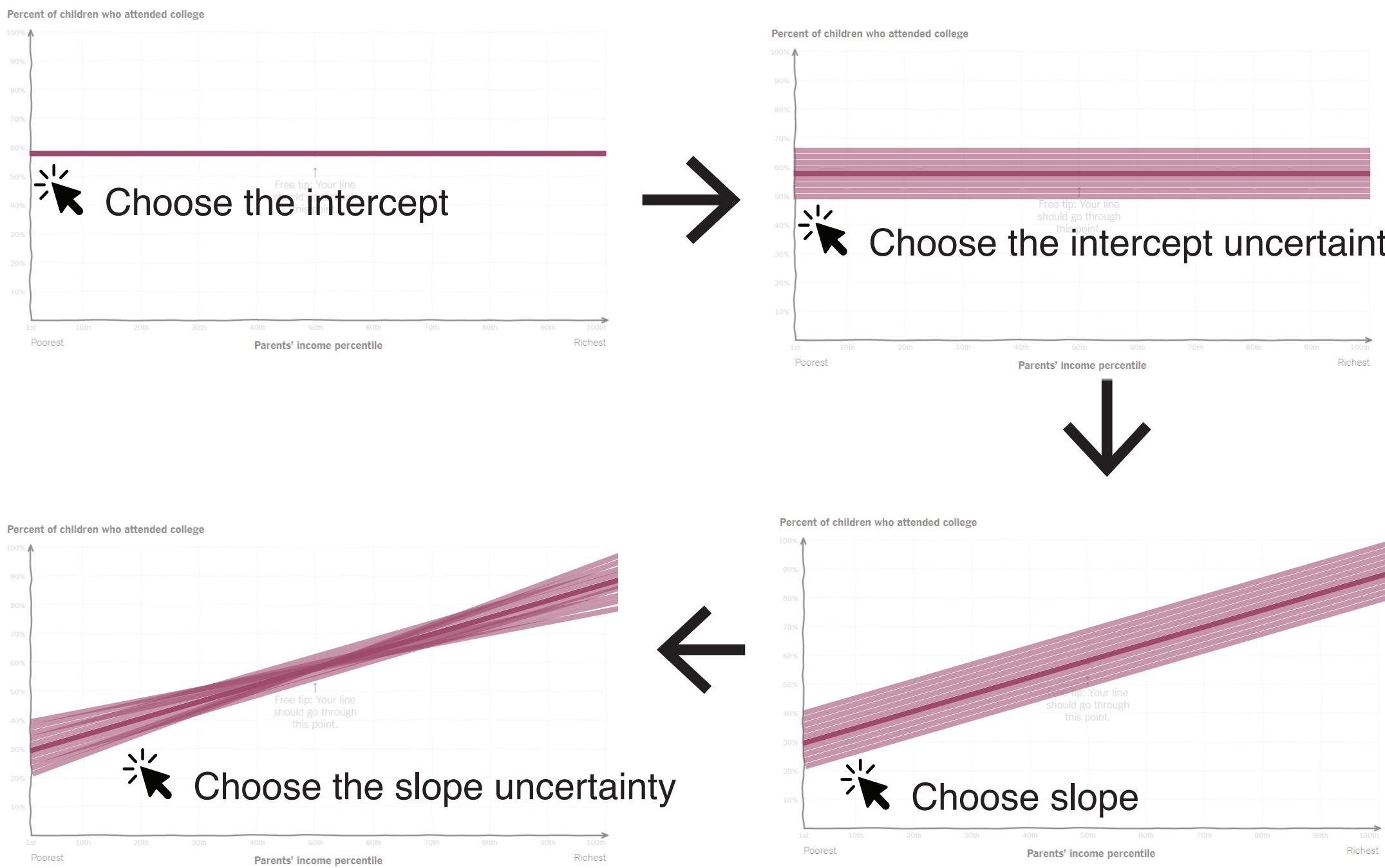
1. **Visualization technique: Continuous scale**
2. **Interaction technique: Dragging**

Redesign using VIBE
Iteration 3

[Karduni et. al, TVCG 2020]

Who

High graphical and statistical expertise

What

**Parameter: Correlation
Uncertainty**

How

1. **Visualization technique: Line chart**
2. **Interaction technique: Hover + Click**

Conclusion

Limitations & Future Work

Scope currently limited to data journalism

Can explore non-graphical modalities

Future work required to be more prescriptive (i.e., to investigate the effectiveness of alternative elicitation techniques)

Summary

Described formative stages of VIBE, a design space for belief elicitation in data journalism

Demonstrated how 24 data journalism examples fit into the design space

Utilized VIBE to redesign an example article

Thank you