Supplemental Material

1) Preliminary Survey Data Sets

2) Analysis of Participants' Demographic Information and Relevant Experience.

3) Predict conditions: Anchor Effect Analysis

4) Explain-Only conditions: Quantity and Quality of self-explanation

5) Screenshots of the main study interface

*Code and detailed analysis of the main study and the replication study are in the separated file (explaining_the_gap.R).

1. Preliminary Survey Data Sets

Fast Food Calorie Content

This dataset describes the calories in three menu items (fries, coke, shake) at McDonalds compared to the serving sizes. (Source: https://goo.gl/E3IfL5)
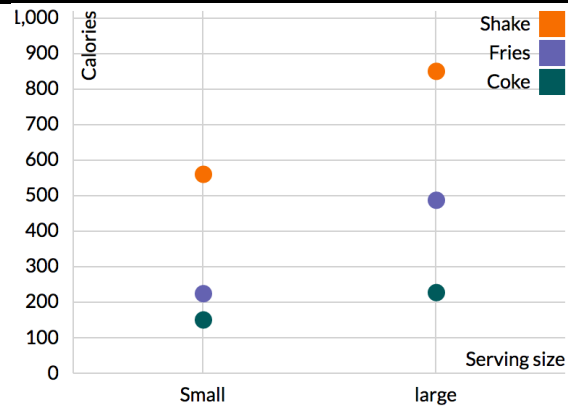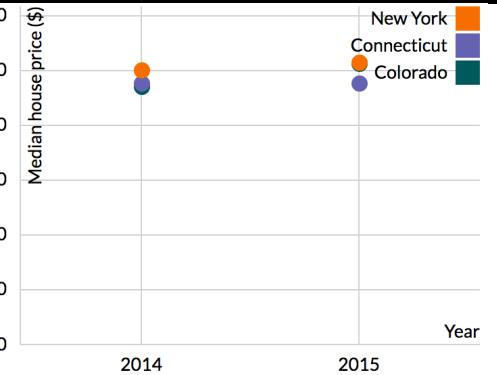
Smartphone Price

This dataset describes the average selling price of a single smartphone from three different smartphone companies (Apple, Android, Xiaomi) over time. (Source: https://goo.gl/QUBl91)
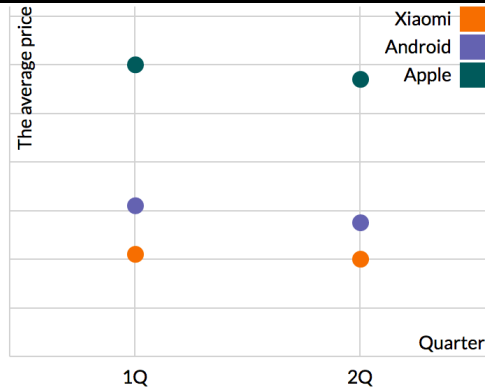
Median House Price

This dataset describes the median price of a house by square feet in three states (Colorado, Connecticut, New York) over time. (Source: https://goo.gl/QYGIeE)

Voting Result

This dataset describes the percentage of voters (by ethnicity: Hispanic, White, Black) who voted for the Republican candidate in the 2008 presidential election in a state, separated by income level. (Source: https://goo.gl/l69Vbj)
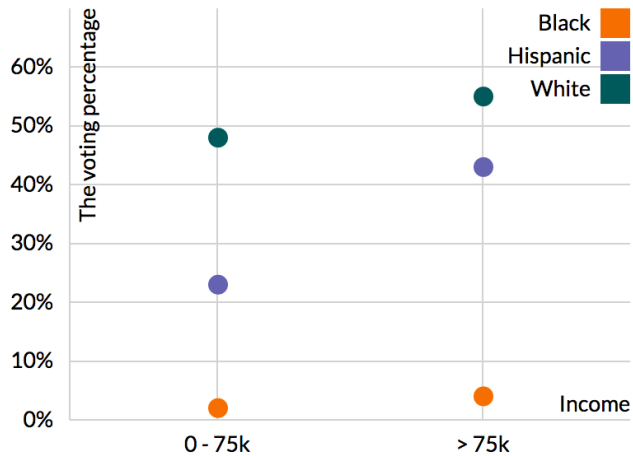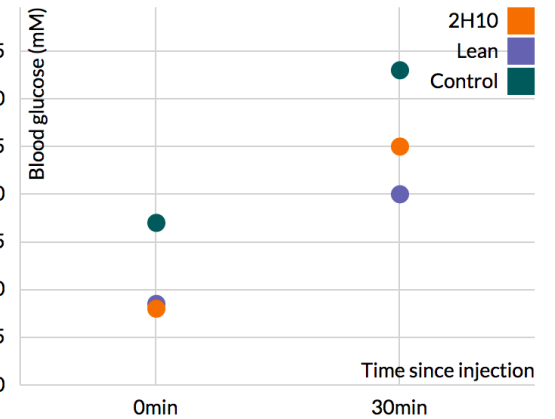
National Budget

This dataset describes the percentage of the U.S national budget by three categories (healthcare, defense, transportation) over time. (Source: https://goo.gl/9yzo5A)
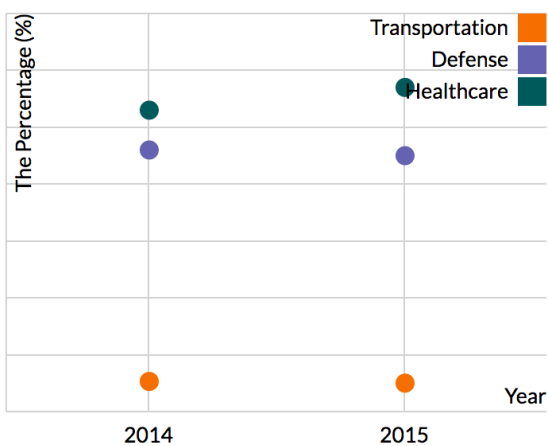
Pre-diabetic Experiment

This dataset describes the effect of three pre-diabetic conditions (Control,Lean,2H10) on the blood glucose level over time. (Source : https://goo.gl/gkmPph)

2. Participants' Demographic Information – Age, Gender

Condition	18-24	25-34	35-44	45-54	55+
None-Vis	7	22	8	4	1
Explain-Only-Vis	6	15	8	10	3
Predict-Explain-Vis	8	18	12	3	1
Predict-Feedback-Vis	5	28	2	6	1
Predict-Only-Vis	6	25	7	3	1
None-Text	7	19	10	5	1
Explain-Only-Text	8	14	8	6	3
Predict-Explain-Text	10	19	6	7	0
Predict-Only-Text	6	21	7	5	1

No difference across conditions (X^2=31.21, df=32, p=.506)

Condition	Female	Male
None-Vis	20	22
Explain-Only-Vis	24	17
Predict-Explain-Vis	18	24
Predict-Feedback-Vis	13	29
Predict-Only-Vis	17	25
None-Text	20	22
Explain-Only-Text	16	23
Predict-Explain-Text	16	25
Predict-Only-Text	21	19

No difference across conditions. (X^2=15.65, df=16, p=.478)

2. Participants' Demographic Information – Ethnicity, Education

Condition	Asian	Black	Hispanic	Native	White	other
None-Vis	3	5	2	0	32	0
Explain-Only-Vis	5	2	1	0	34	0
Predict-Explain-Vis	0	3	1	0	38	0
Predict-Feedback-Vis	6	4	6	0	26	0
Predict-Only-Vis	4	4	1	1	32	0
None-Text	4	4	4	0	30	0
Explain-Only-Text	1	1	4	0	31	2
Predict-Explain-Text	1	3	3	1	33	1
Predict-Only-Text	2	3	3	0	32	0

Condition	High	Bachelor	grad
None-Vis	10	26	6
Explain-Only-Vis	21	18	3
Predict-Explain-Vis	19	19	4
Predict-Feedback-Vis	18	21	3
Predict-Only-Vis	12	29	1
None-Text	13	26	3
Explain-Only-Text	17	15	7
Predict-Explain-Text	13	24	5
Predict-Only-Text	19	17	4

No difference across conditions (X^2=45.11, df=40, p=.267)

No difference across conditions (X^2= 22.31, df=16, p=.134)

2. Participants' Relevant Experience – Chart Experience

How often are you using charts?

Very often (about everyday)
Often (1 - 5 times per week)
Sometimes (1-5 times per month)
Rarely (less than once a month)
Never

Condition	Never	Rarely	Sometimes	Often	Very often
None-Vis	9	10	17	5	1
Explain-Only-Vis	3	15	19	2	3
Predict-Explain-Vis	2	23	10	6	1
Predict-Feedback-Vis	4	20	13	3	2
Predict-Only-Vis	6	14	18	3	1
None-Text	1	16	12	10	3
Explain-Only-Text	6	10	12	9	2
Predict-Explain-Text	5	14	11	9	3
Predict-Only-Text	6	12	12	7	3

No difference across conditions (X2= 41.57, df=32, p=.120)

3. Predict conditions: Anchor Effect Analysis
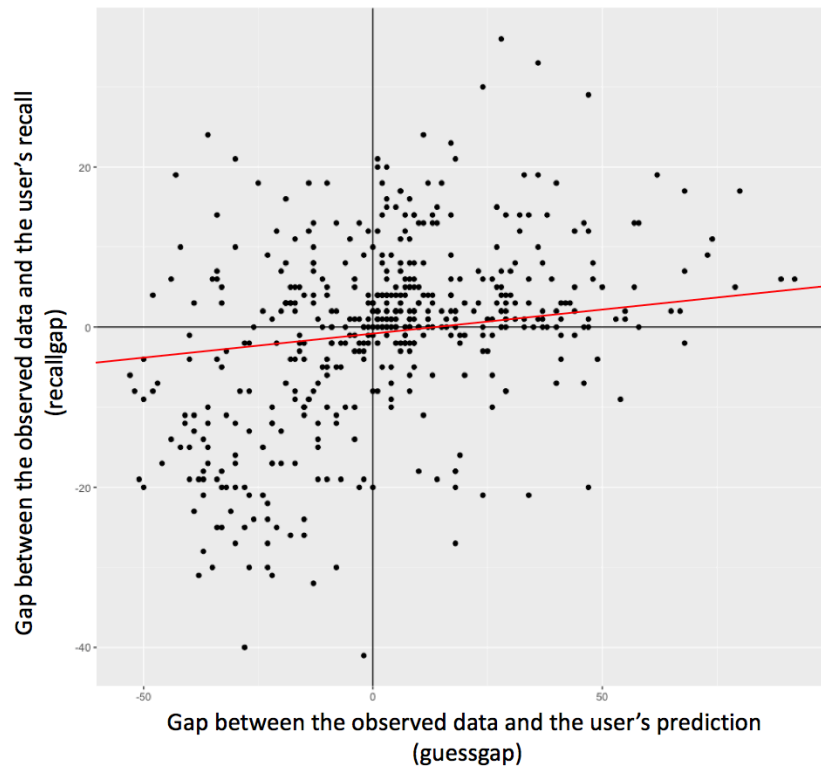
Model: lm(recallgap ~ guessgap)
Residuals:

Min	1Q	Median	3Q	Max
-39.350	-5.403	0.434	5.099	32.145

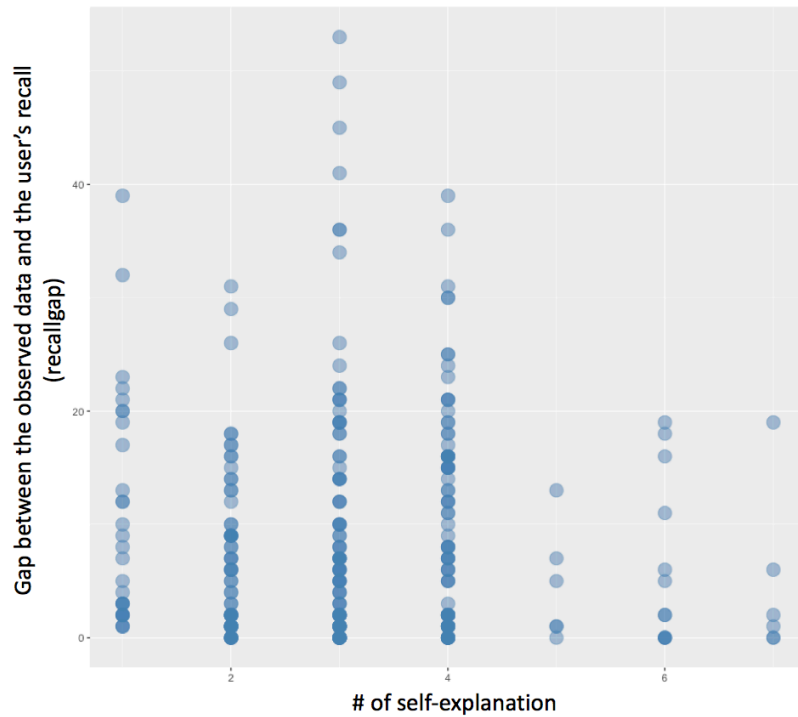**Coefficients:**   **Estimate Std. Error t value Pr(>|t|)**
(Intercept) -1.28275   0.46571 -2.754 0.00609 **
guessgap 0.18349   0.01773 10.349 < 2e-16 ***

Multiple R-squared: 0.1758, Adjusted R-squared: 0.1742
F-statistic: 107.1 on 1 and 502 DF, p-value: < 2.2e-16

4. Explain–Only conditions: Quantity of self–explanation

Model: lm(recallGap ~ #ofSelf-Explanation)
Residuals:
  Min   1Q Median   3Q   Max

-6.830 -3.669 -1.165 1.456 16.903

Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept)   10.4651   2.4906   4.202 0.000144 ***
# of self-explanation -0.2327   0.7601 -0.306 0.761031

Multiple R-squared: 0.002339,   Adjusted R-squared: -0.0226
F-statistic: 0.09377 on 1 and 40 DF, p-value: 0.761

4. Explain-Only conditions: Quality of self-explanation

Example self-explanations:

Inference without prior knowledge + High detail

Generally, people with incomes over 75k were less likely to vote for John McCain in 2008. Blacks who made over 75k were slightly more likely to vote for him, but it was a very small increase to 3% meaning not many blacks voted for McCain in any income category.

Inference without prior knowledge + Low detail

Majority of people no matter ethnicity voted for the Democrats and not Republicans.

Inference with prior knowledge + High detail

There was a slight increase in the White voting population with the higher income bracket, I could assume that this is due to McCain's policies which benefit the wealthier. This could also account for the 20% increase in the Hispanic voting numbers, I would assume that McCain favored policies which affected the more affluent voters. Since the overall numbers of Black voters was low and did not vary significantly between income brackets, I would assume that the majority of Black voters were not Republican.

Inference with prior knowledge + Low detail

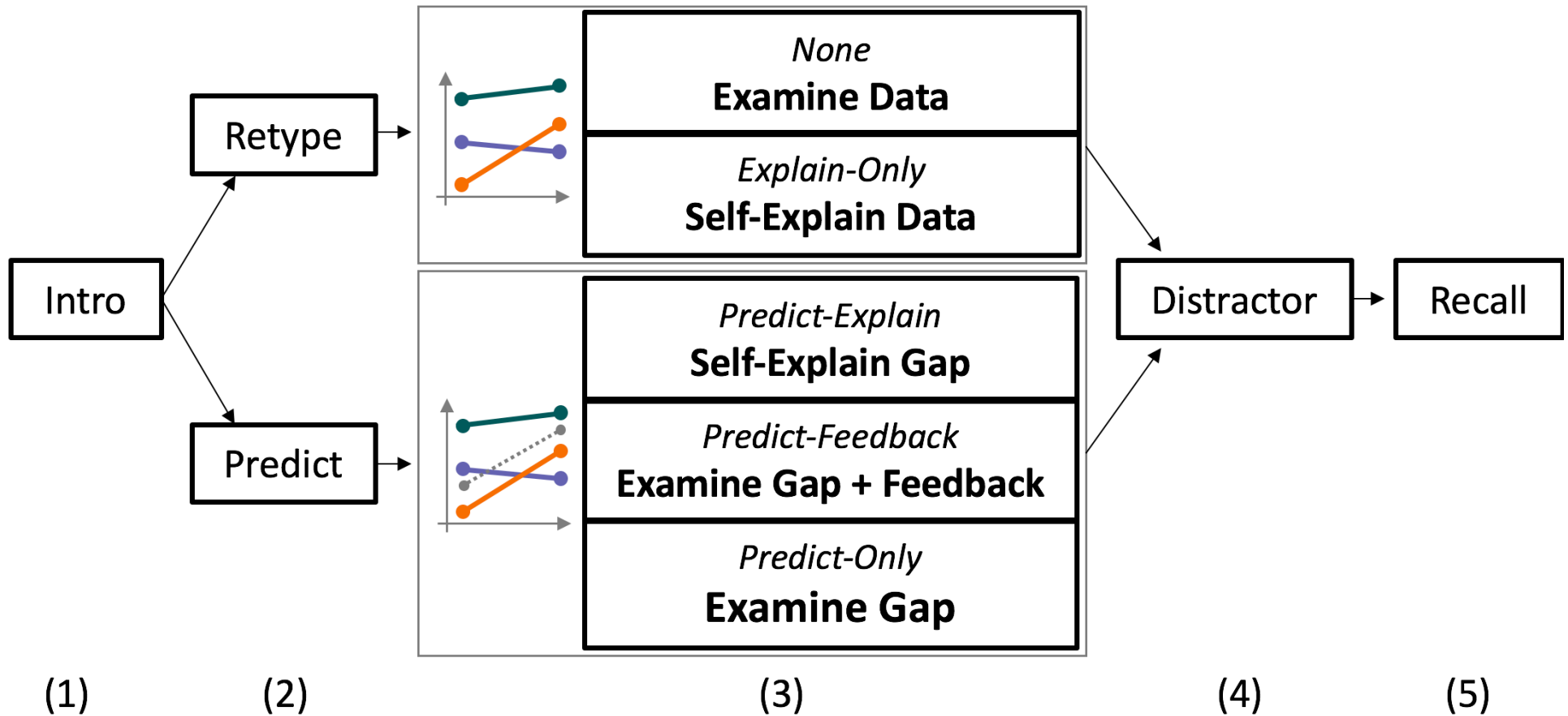People are more conservative in Colorado.

No inference

Each colored line is a different race. Each point is a different income bracket.

Model: aov(Average participants' error ~ inference * detail)

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Inference	1	11.1	11.09	0.359	0.553
Detail	1	12.8	12.85	0.415	0.523
Inference * Detail	1	70.1	70.13	2.268	0.140

5. Screenshots of the main study interface – Overview

Intro

Retype

Predict

None
Examine Data

Explain-Only
Self-Explain Data

Predict-Explain
Self-Explain Gap

Predict-Feedback
Examine Gap + Feedback

Predict-Only
Examine Gap

Distractor

Recall

(1)   (2)   (3)   (4)   (5)

5. Screenshots of the main study interface – (1) Intro

"Data understanding task"

In this task, you will examine small data sets of the voting behavior of different races in the 2008 presidential election.

After reading about the experiment and viewing the data, you will complete a paper folding task.

This HIT should take no more than 30 minutes to complete.

This HIT is part of a research project. We are interested in how well you do on the task without the help of other resources. Please do NOT use any outside resources (e.g., web searches). Doing so will NOT increase your reward.

Reward : You will receive $1.5 for your work.

Submit

5. Screenshots of the main study interface – (2) Retype

"2008 Presidential election voting results"

On the next page you will see a dataset describes the percentage of voters of different races who voted for the Republican candidate John McCain in the 2008 presidental election.

First, read this paragraph from Wikipedia on elections:

The election of the President and Vice President of the United States is an indirect vote in which citizens of the United States registered to vote in one of the fifty states orWashington, D.C cast ballots for a set of members of the U.S. Electoral College. These electors then cast direct votes for the President and Vice President. If both votes result in anabsolute majority, the election is over.

Please re-type the above paragraph before continuing to the next page.

Continue

5. Screenshots of the main study interface – (2) Predict-Vis

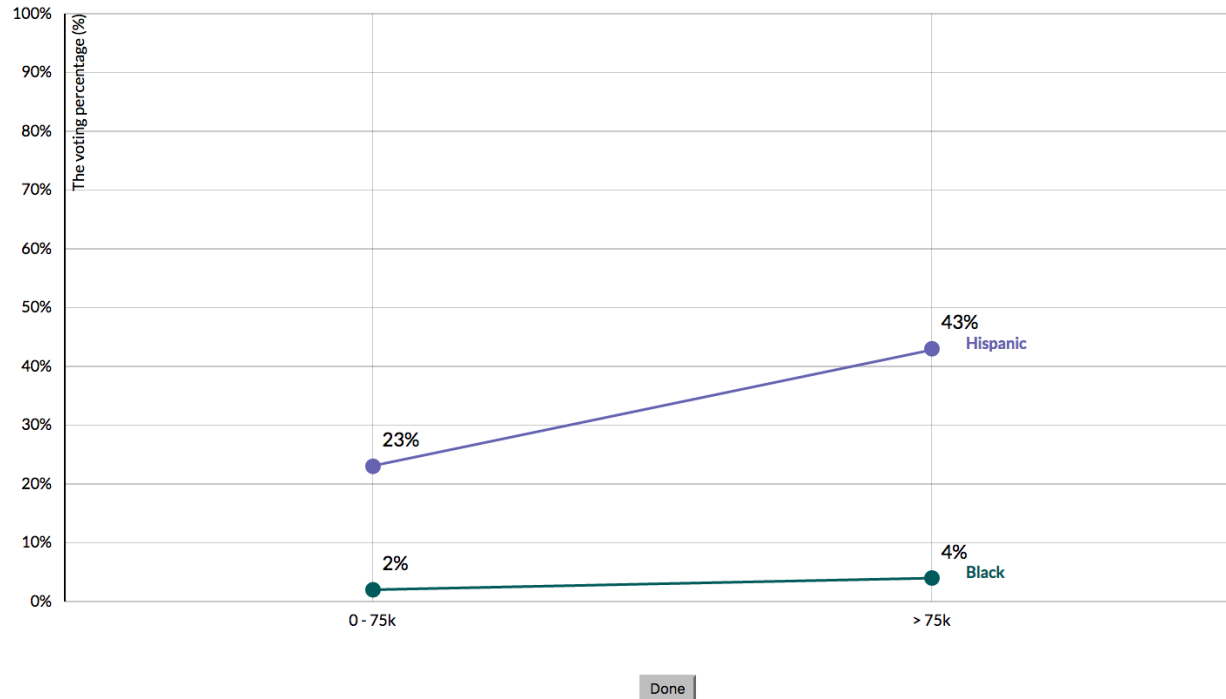**"2008 Presidential election voting results"**

Please read the description of the dataset below and follow the instructions.

An analyst has collected data on the percentage of eligible voters of different races who voted for the Republican candidate John McCain in the 2008 presidential election.

Each line represents a different race. The two points in each line represent two income brackets (up to $75,000 and more than $75,000) in **Colorado**.

Please draw your prediction for voting percentage among **White voters** for each income bracket in **Colorado**. **(Click to add two points. Adding the second point will add a line between them. You can adjust either points by dragging the circles).**

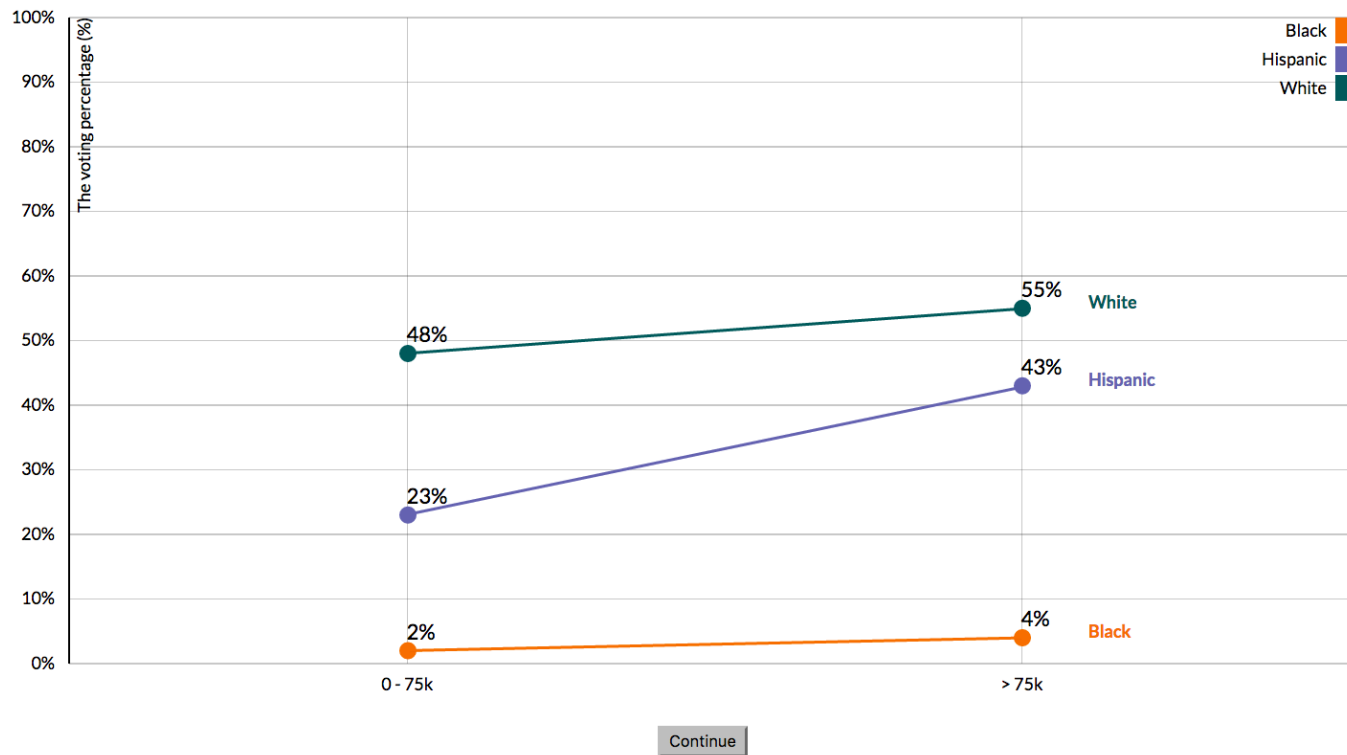**When you are finished, press Done.**

"2008 Presidential election voting results"

Please read the description of the dataset below and follow the instructions.

An analyst has collected data on the percentage of eligible voters of different races who voted for the Republican candidate John McCain in the 2008 presidential election.

The dataset provides information of the voting result by races and by two income brackets (up to $75,000 and more than $75,000) in **Colorado**.

Dataset:

23% of Hispanic voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 43% of Hispanic voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

2% of Black voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 4% of Black voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

Given what the dataset you just saw please guess for **the percentage of White voters** who voted for the Republican candidate John McCain in **each income brackets.**

```


```

Continue

5. Screenshots of the main study interface – (3) None–Vis

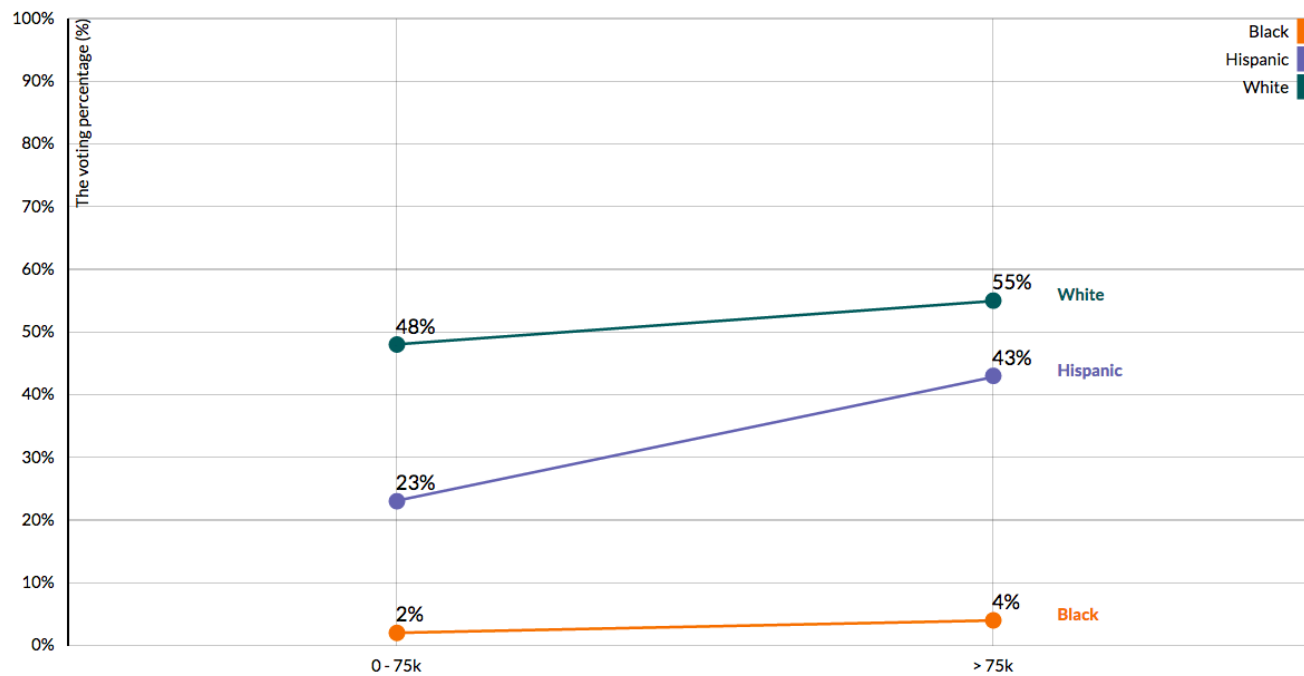## "2008 Presidential election voting results"

Please read the description of the dataset below and follow the instructions.

An analyst has collected data on the percentage of eligible voters of different races who voted for the Republican candidate John McCain in the 2008 presidential election.

Each line represents a different race. The two points in each line represent two income brackets (up to $75,000 and more than $75,000) in **Colorado**.

Please examine the chart (read all labels) **twice** before continuing to the next page.
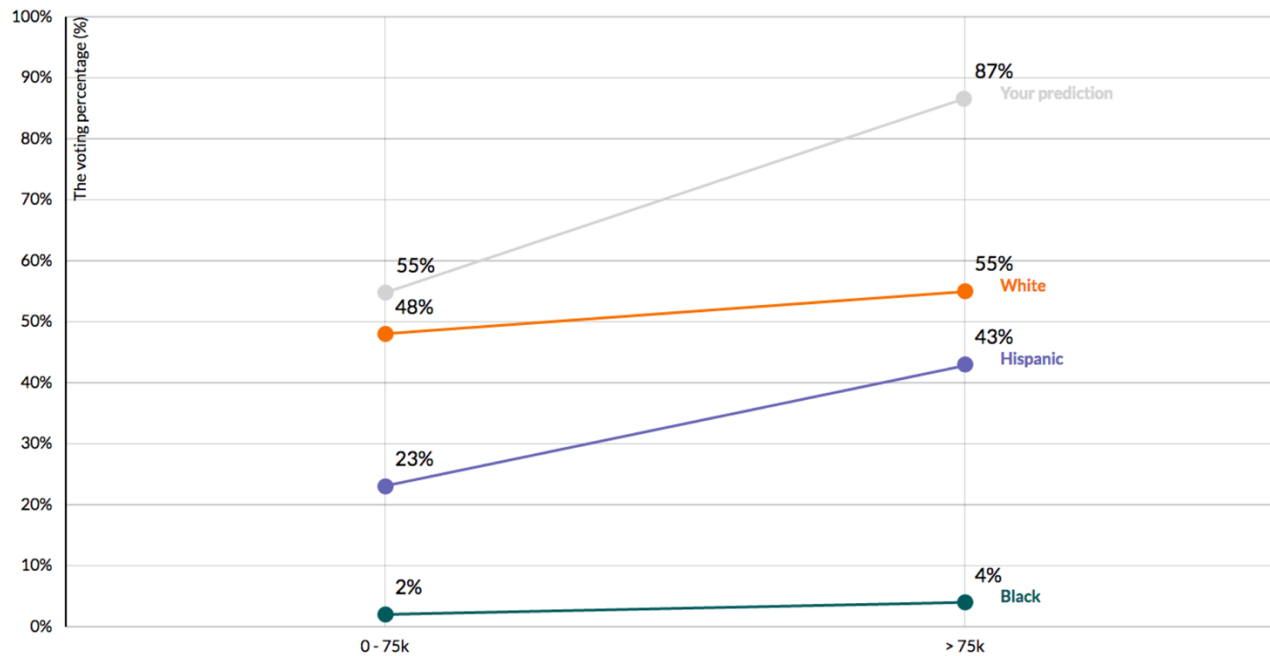

Continue

5. Screenshots of the main study interface – (3) Explain-Only-Vis

"2008 Presidential election voting results"

Please read the description of the dataset below and follow the instructions.

An analyst has collected data on the percentage of eligible voters of different races who voted for the Republican candidate John McCain in the 2008 presidential election.

Each line represents a different race. The two points in each line represent two income brackets (up to $75,000 and more than $75,000) in **Colorado**.

You will now write a few sentences of explanation to help you understand. Think about the voting precentage separately for each income bracket. and consider the trends in how the three ethnic groups compare.

You must provide explanations, but it will not affect your score.

Continue

5. Screenshots of the main study interface – (3) Predict–Explain–Vis

"2008 Presidential election voting results"

Please read the description of the dataset below and follow the instructions.

An analyst has collected data on the percentage of eligible voters of different races who voted for the Republican candidate John McCain in the 2008 presidential election.

Each line represents a different race. The two points in each line represent two income brackets (up to $75,000 and more than $75,000) in **Colorado**.

You will now write a few sentences of explanation to help you understand where your predictions were off. Think about how you did in predicting the voting precentage separately for each income bracket.

For example, consider **how much your predictions were off** from the true trends, and **why you might have been wrong**.

You must provide explanations, but it will not affect your score.
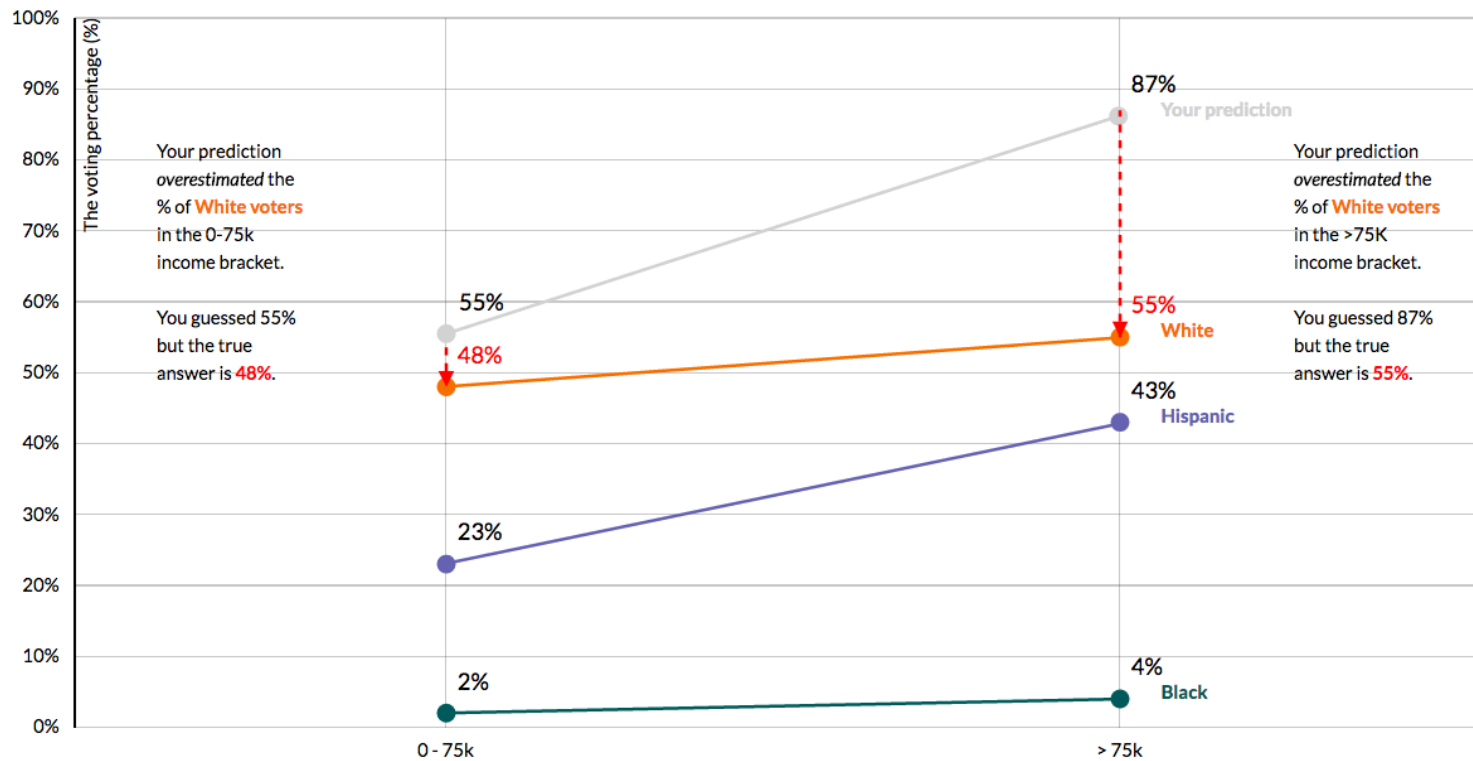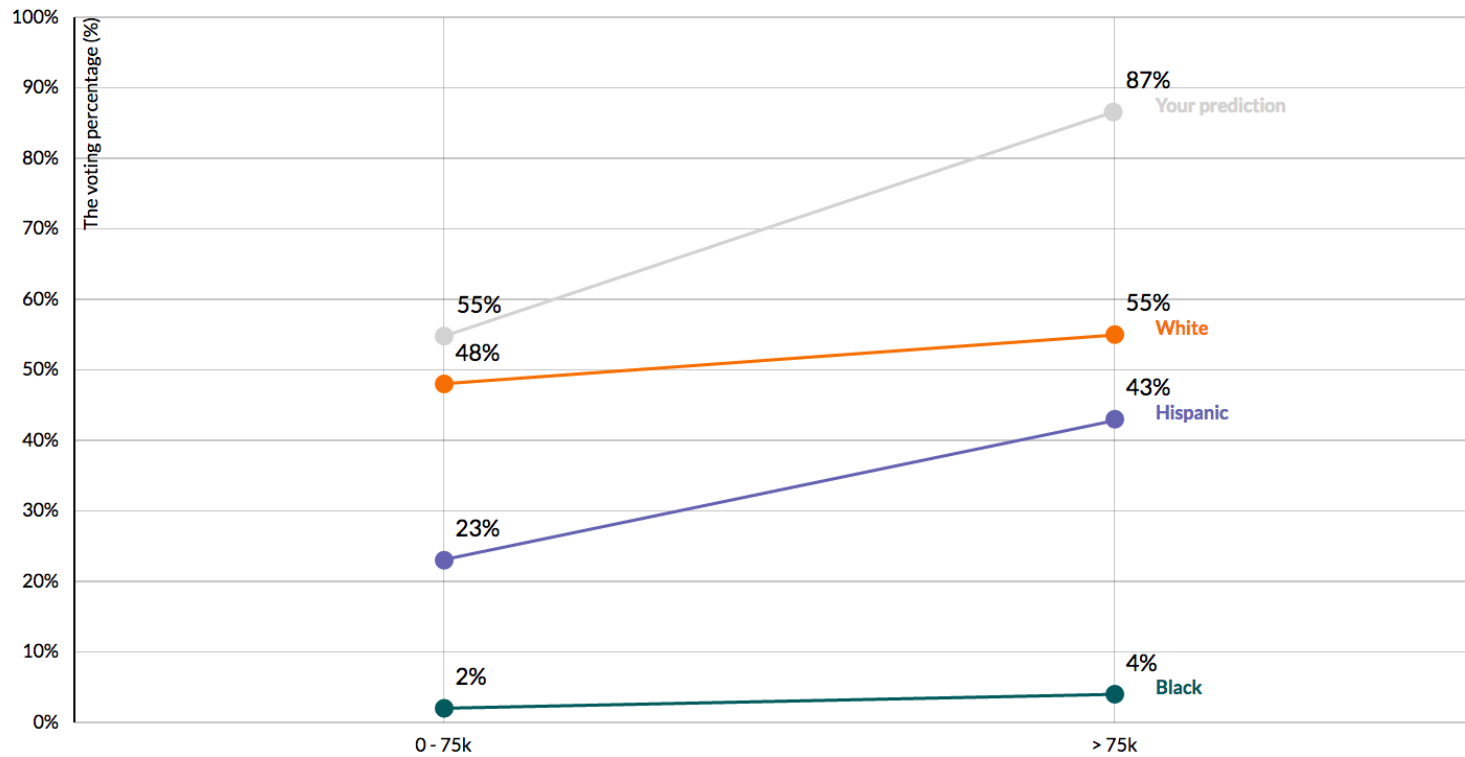
[Continue]

"2008 Presidential election voting results"

Please read the description of the dataset below and follow the instructions.

An analyst has collected data on the percentage of eligible voters of different races who voted for the Republican candidate John McCain in the 2008 presidential election.

Each line represents a different race. The two points in each line represent two income brackets (up to $75,000 and more than $75,000) in **Colorado**.

Continue

5. Screenshots of the main study interface – (3) Predict-Only-Vis

"2008 Presidential election voting results"

Please read the description of the dataset below and follow the instructions.

An analyst has collected data on the percentage of eligible voters of different races who voted for the Republican candidate John McCain in the 2008 presidential election.

Each line represents a different race. The two points in each line represent two income brackets (up to $75,000 and more than $75,000) in **Colorado**.

[Continue]

5. Screenshots of the main study interface – (3) None–Text

"2008 Presidential election voting results"

Please read the description of the dataset below and follow the instructions.

An analyst has collected data on the percentage of eligible voters of different races who voted for the Republican candidate John McCain in the 2008 presidential election.

Each line represents a different race. The two points in each line represent two income brackets (up to $75,000 and more than $75,000) in **Colorado**.

Please read through the dataset **twice** before continuing to the next page.

Dataset:

48% of White voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 55% of White voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

23% of Hispanic voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 43% of Hispanic voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

2% of Black voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 4% of Black voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

Continue

5. Screenshots of the main study interface – (3) Explain-Only-Text

"2008 Presidential election voting results"

You will now write a few sentences of explanation to help you understand. Think about the voting precentage separately for each income bracket. and consider the trends in how the three ethnic groups compare.

You must provide explanations, but it will not affect your score.

Dataset:

48% of White voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 55% of White voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

23% of Hispanic voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 43% of Hispanic voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

2% of Black voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 4% of Black voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

Your explanations:

Continue

5. Screenshots of the main study interface – (3) Predict–Explain–Text

"2008 Presidential election voting results"

Full dataset is now shown along with your prediction.

Dataset:

48% of White voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 55% of White voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

23% of Hispanic voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 43% of Hispanic voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

2% of Black voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 4% of Black voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

Your prediction:

17% of White voters in up to $75,000 income bracket and 29% in over $75,000 bracket.

You will now write a few sentences of explanation to help you understand where your predictions were off. Think about how you did in predicting the voting precentage separately for each income bracket.

For example, consider **how much your predictions were off** from the true trends, and **why you might have been wrong.**

You must provide explanations, but it will not affect your score.

Continue

5. Screenshots of the main study interface – (3) Predict–Only–Text

"2008 Presidential election voting results"

Full dataset is now shown along with your prediction.

Dataset:

48% of White voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 55% of White voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

23% of Hispanic voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 43% of Hispanic voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

2% of Black voters in up to $75,000 income bracket in Colorado voted for John McCain in 2008. 4% of Black voters in over $75,000 income bracket in Colorado voted for John McCain in 2008.

Your prediction:

17% of White voters in up to $75,000 income bracket and 29% in over $75,000 bracket.

Continue

"The paper folding task"

You will now complete a paper folding task. Please read the instructions and do your best.

You have 3 minutes to complete 10 questions. **Your answer will be automatically submited after 3 minutes.**

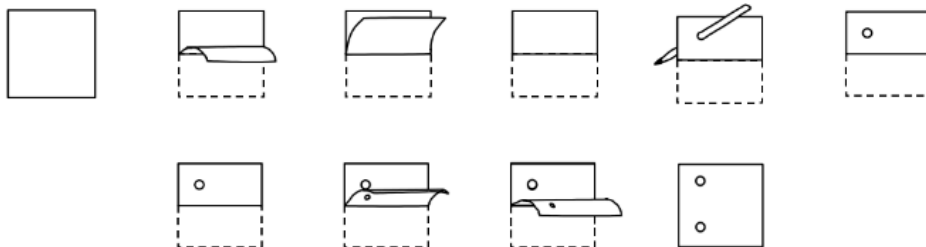The score will not affect your payment or HIT approval, so just do your best.

We are interested in your score, however, for our analysis. To score the test, we will take the number of questions you answer correctly minus a fraction of the questions you answered incorrectly. Therefore, **it will not be to your advantage to guess. Instead, leave the default response of 'I don't know.'**

In this test you are to imagine the folding and unfolding of pieces of paper. In each problem in the test there are some figures drawn at the left of a vertical line and there are others drawn at the right of the line. The figures at the left represent a square piece of paper being folded, and the last of these figures has one or two small circles drawn on it to show where the paper has been punched. Each hole is punched through all the thicknesses of paper at that point. One of the five figures on the right of the vertical line shows where the holes will be when the paper is completely unfolded. You are to decide which one of these figures is correct and draw an X through that figure. Now try the sample problem below. (In this problem only one hole was punched in the folded paper).
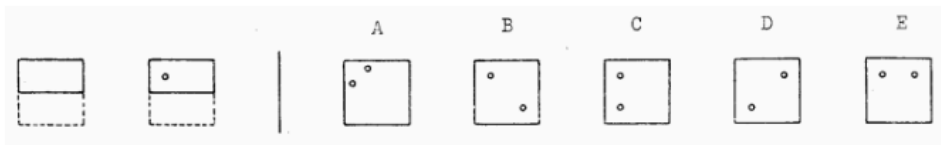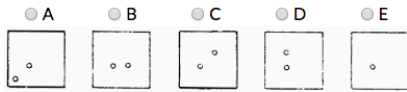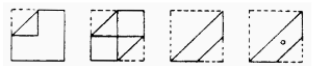

The correct answer to the sample problem above is C and so it should have been marked with an X. The figures below show how the paper was folded and why C is the correct answer.
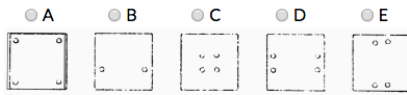
In these problems all of the folds that are made are shown in the figures at the left of the line, and the paper is not turned or moved in any way except to make the folds shown in the figures. Remember, the correct answer is the figure that shows the positions of the holes when the paper is completely unfolded.
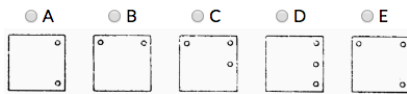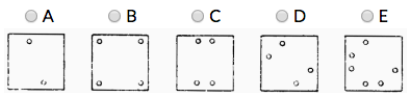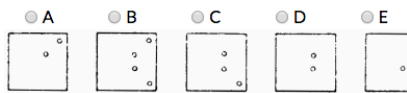
A   B   C   D   E   I don't know.

A   B   C   D   E   I don't know.

A   B   C   D   E   I don't know.

A   B   C   D   E   I don't know.

A   B   C   D   E   I don't know.

A   B   C   D   E   I don't know.

A   B   C   D   E   I don't know.

A   B   C   D   E   I don't know.

A   B   C   D   E   I don't know.

A   B   C   D   E   I don't know.

Continue

"Recall test"

Please answer questions below. You can't go back to the previous page. Please do NOT use any outside resources.

We are interested in how well you do without any help.
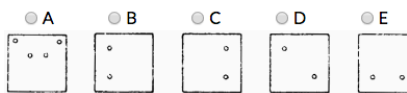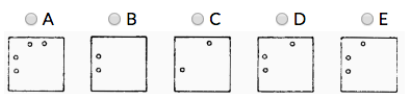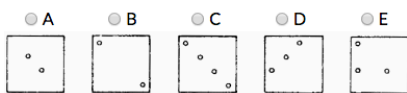
Provide the percentage of vote for **White** for John McCain in **Colorado** by clicking in the chart for that income bracket.

Even if you are not sure of the answer, please make your best guess.

White

Continue

"Recall test"

Please answer questions below. You can't go back to the previous page. Please do NOT use any outside resources.

We are interested in how well you do without any help.

Provide the percentage of vote for **Hispanic** for John McCain in **Colorado** by clicking in the chart for that income bracket.

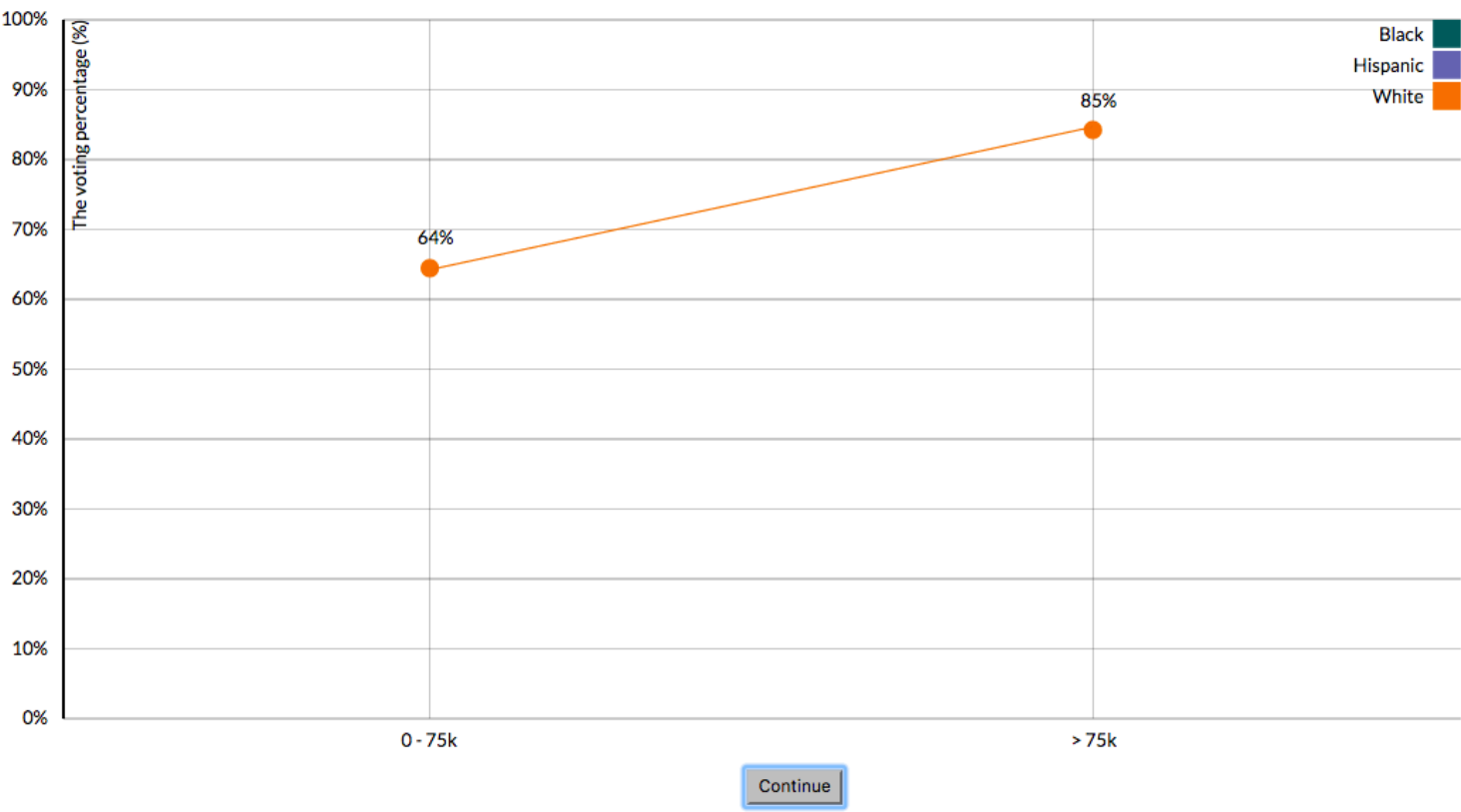Even if you are not sure of the answer, please make your best guess.

Hispanic

Continue

"Recall test"

Please answer questions below. You can't go back to the previous page. Please do NOT use any outside resources.

We are interested in how well you do without any help.

Provide the percentage of vote for **Black** for John McCain in **Colorado** by clicking in the chart for that income bracket.

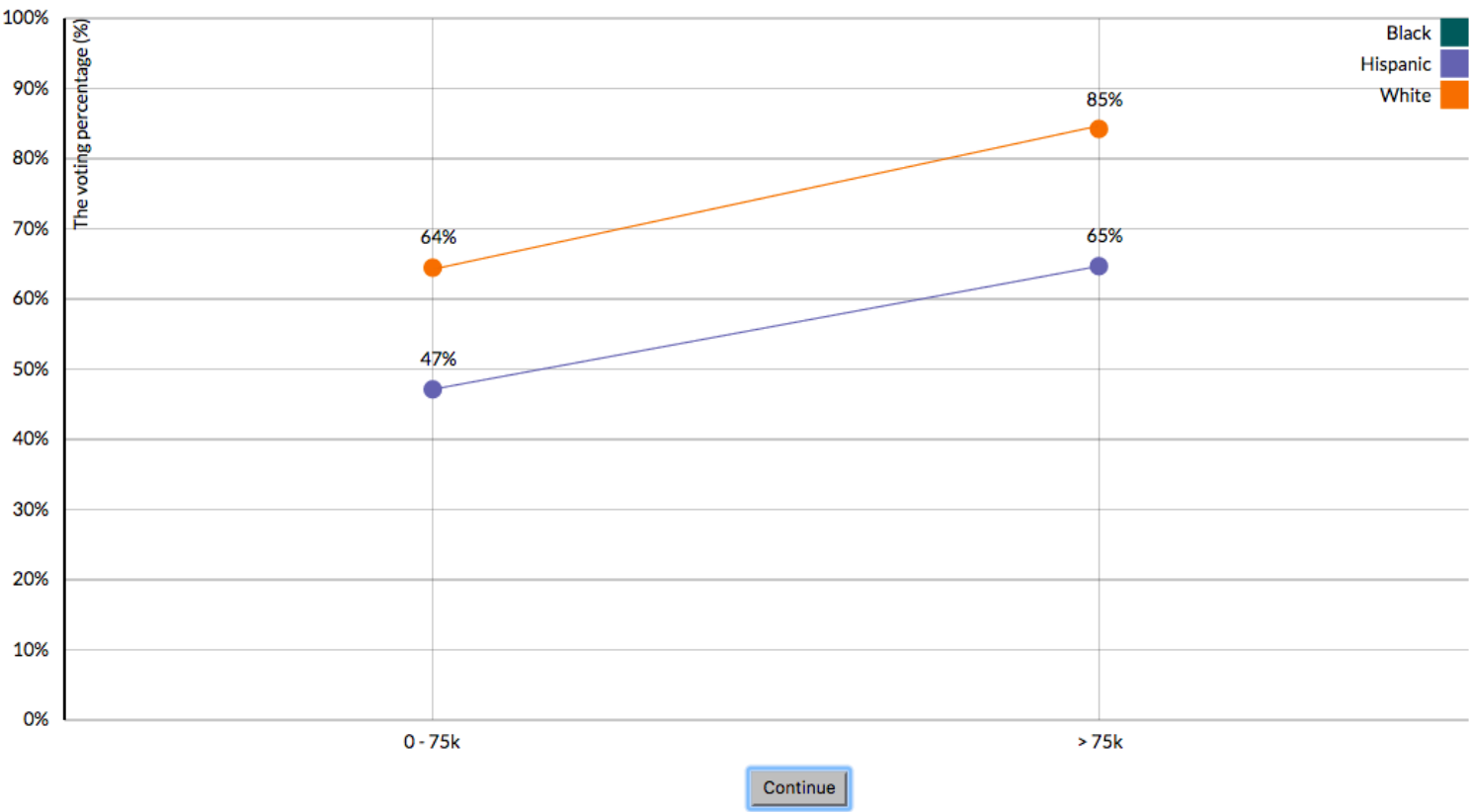Even if you are not sure of the answer, please make your best guess.

Black

"Recall test"

Please answer questions below. You can't go back to the previous page. Please do NOT use any outside resources.

We are interested in how well you do without any help.

Please adjust your answer if you would like by dragging the circles and click Continue to proceed when you are satisfied with your answer.
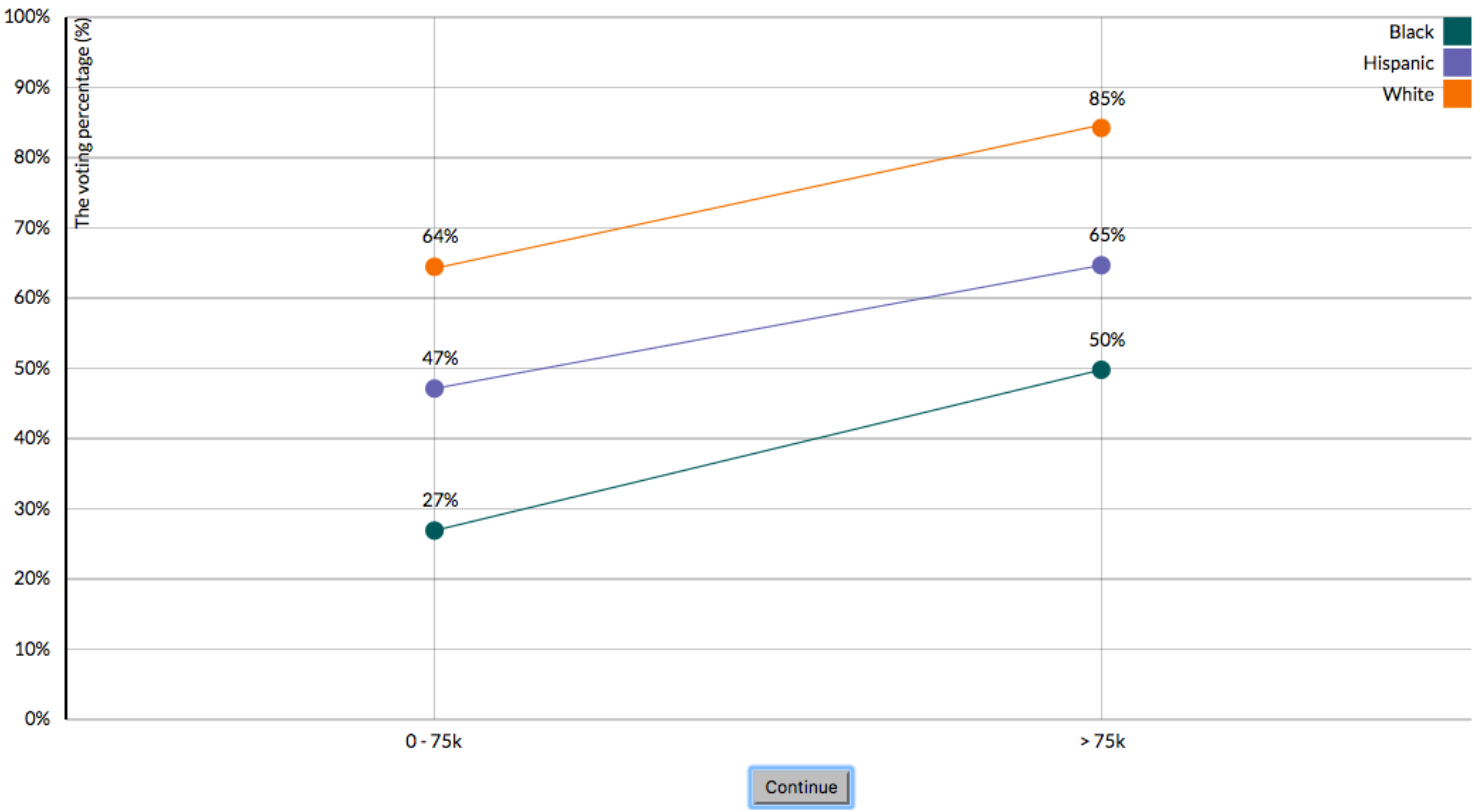

White

Hispanic

Black

Continue

"Recall test"

Please answer questions below. You can't go back to the previous page. Please do NOT use any outside resources.

We are interested in how well you do without any help.

Provide the voting percentage, in percent, for by entering a number in the box.

Even if you are not sure of the answer, please make your best guess.

Q: **What percentage of Black voters** in up to $75,000 income bracket voted for John McCain in **Colorado?**

[   ] percentage

Q: **What percentage of Black voters** in over $75,000 income bracket voted for John McCain in **Colorado?**

[   ] percentage

[Continue]

"Recall test"

Please answer questions below. You can't go back to the previous page. Please do NOT use any outside resources.

We are interested in how well you do without any help.

Provide the voting percentage, in percent, for by entering a number in the box.

Even if you are not sure of the answer, please make your best guess.

Q: **What percentage of Black voters** in up to $75,000 income bracket voted for John McCain in **Colorado?**

| 2 | percentage

Q: **What percentage of Black voters** in over $75,000 income bracket voted for John McCain in **Colorado?**

| 4 | percentage

Q: **What percentage of White voters** in up to $75,000 income bracket voted for John McCain in **Colorado?**

| | percentage

Q: **What percentage of White voters** in over $75,000 income bracket voted for John McCain in **Colorado?**

| | percentage

[Continue]

"Recall test"

Please answer questions below. You can't go back to the previous page. Please do NOT use any outside resources.

We are interested in how well you do without any help.

Provide the voting percentage, in percent, for by entering a number in the box.

Even if you are not sure of the answer, please make your best guess.

Q: **What percentage of Black voters** in up to $75,000 income bracket voted for John McCain in **Colorado?**

| 2 | percentage

Q: **What percentage of Black voters** in over $75,000 income bracket voted for John McCain in **Colorado?**

| 4 | percentage

Q: **What percentage of White voters** in up to $75,000 income bracket voted for John McCain in **Colorado?**

| 70 | percentage

Q: **What percentage of White voters** in over $75,000 income bracket voted for John McCain in **Colorado?**

| 80 | percentage

Q: **What percentage of Hispanic voters** in up to $75,000 income bracket voted for John McCain in **Colorado?**

| | percentage

Q: **What percentage of Hispanic voters** in over $75,000 income bracket voted for John McCain in **Colorado?**

| | percentage

Continue

"Recall test"

Please answer questions below. You can't go back to the previous page. Please do NOT use any outside resources.

We are interested in how well you do without any help.

Provide the voting percentage, in percent, for by entering a number in the box.

Even if you are not sure of the answer, please make your best guess.

Please adjust your answer if you would like and click Continue to proceed when you are satisfied with your answer.

Q: **What percentage of Black voters** in up to $75,000 income bracket voted for John McCain in **Colorado?**

| 2 | percentage

Q: **What percentage of Black voters** in over $75,000 income bracket voted for John McCain in **Colorado?**

| 4 | percentage

Q: **What percentage of White voters** in up to $75,000 income bracket voted for John McCain in **Colorado?**

| 70 | percentage

Q: **What percentage of White voters** in over $75,000 income bracket voted for John McCain in **Colorado?**

| 80 | percentage

Q: **What percentage of Hispanic voters** in up to $75,000 income bracket voted for John McCain in **Colorado?**

| 45 | percentage

Q: **What percentage of Hispanic voters** in over $75,000 income bracket voted for John McCain in **Colorado?**

| 48 | percentage

Continue