

Wording Experiment Example

Ying Yan

2025-03-26

Data and Code

The original survey data is taken from the supplementary material of Albertson and Jessee (2023). The pseudo- survey dataset is generated by ChatGPT by describing exactly the same setup and sample size as the paper. The exact wording of this data generation prompt is shown in the Appendix. The data generation code generated by ChatGPT and the code for this replication and comparison are available at Github.

The Descriptive Demographics

A sample characteristics comparison between the original data and the dataset generated by ChatGPT.

Table 1: Comparison of Proportions Between Two Datasets

variables	original	pseudo
Female	55%	50%
College degree or higher	55%	28%
White	71%	63%
Black or African American	10%	12%
Asian or Asian American	7%	7%
Liberal or Slightly Liberal	47%	40%
Conservative or Slightly Conservative	28%	29%

The Regression Results

Table 1: Effect of Question Wording on Spending Views (Ignoring Question Order)

The result from the original survey data is shown in Table 2.

The result of data generated by ChatGPT is show in Table 3.

Table 2: Regression Predicting Spending Preferences on Welfare/Assistance to the Poor

The regression result from original dataset is shown in Table 4.

and the regression result from the dataset generated by ChatGPT is shown in Table 5.

Table 2: Original Survey Results as Table 1

	<i>Dependent variable:</i>		
	spend		
	(1)	(2)	(3)
Constant	−0.41*** (0.03)	−1.01*** (0.03)	−0.93*** (0.04)
welfare_word	0.27*** (0.04)	0.30*** (0.03)	0.15*** (0.05)
rr_scale		1.46*** (0.05)	1.28*** (0.07)
welfare_word:rr_scale			0.37*** (0.10)
Observations	1,590	1,590	1,590
R ²	0.03	0.37	0.38
Adjusted R ²	0.03	0.37	0.37
Residual Std. Error	0.79 (df = 1588)	0.63 (df = 1587)	0.63 (df = 1586)
F Statistic	45.58*** (df = 1; 1588)	464.96*** (df = 2; 1587)	317.24*** (df = 3; 1586)

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 3: Pseudo Survey Results as Table 1

	<i>Dependent variable:</i>		
	spend		
	(1)	(2)	(3)
Constant	0.10*** (0.03)	−0.03 (0.12)	0.03 (0.16)
welfare_word	−0.29*** (0.04)	−0.29*** (0.04)	−0.43* (0.23)
rr_scale		0.26 (0.22)	0.13 (0.31)
welfare_word:rr_scale			0.28 (0.45)
Observations	1,590	1,590	1,590
R ²	0.03	0.03	0.03
Adjusted R ²	0.03	0.03	0.03
Residual Std. Error	0.88 (df = 1588)	0.88 (df = 1587)	0.88 (df = 1586)
F Statistic	43.02*** (df = 1; 1588)	22.18*** (df = 2; 1587)	14.91*** (df = 3; 1586)

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 4: Original Survey Results as Table 2

	<i>Dependent variable:</i>	
	spend	
	(1)	(2)
Constant	−0.40*** (0.04)	−0.99*** (0.05)
welfare_word	0.29*** (0.05)	0.20*** (0.07)
rr_post	−0.03 (0.06)	0.10 (0.07)
rr_scale		1.40*** (0.10)
welfare_word:rr_post	−0.04 (0.08)	−0.08 (0.10)
welfare_word:rr_scale		0.32** (0.14)
rr_post:rr_scale		−0.24* (0.14)
welfare_word:rr_post:rr_scale		0.10 (0.20)
Observations	1,590	1,590
R ²	0.03	0.38
Adjusted R ²	0.03	0.37
Residual Std. Error	0.79 (df = 1586)	0.63 (df = 1582)
F Statistic	15.81*** (df = 3; 1586)	136.75*** (df = 7; 1582)
<i>Note:</i>		*p<0.1; **p<0.05; ***p<0.01

Table 5: Pseudo Survey Results as Table 2

	<i>Dependent variable:</i>	
	spend	
	(1)	(2)
Constant	0.05 (0.04)	0.09 (0.22)
welfare_word	-0.25*** (0.06)	-0.32 (0.31)
rr_post	0.08 (0.06)	-0.11 (0.32)
rr_scale		-0.06 (0.43)
welfare_word:rr_post	-0.08 (0.09)	-0.23 (0.46)
welfare_word:rr_scale		0.15 (0.62)
rr_post:rr_scale		0.38 (0.63)
welfare_word:rr_post:rr_scale		0.31 (0.90)
Observations	1,590	1,590
R ²	0.03	0.03
Adjusted R ²	0.03	0.03
Residual Std. Error	0.88 (df = 1586)	0.88 (df = 1582)
F Statistic	14.95*** (df = 3; 1586)	6.86*** (df = 7; 1582)

Note:

*p<0.1; **p<0.05; ***p<0.01

Table 3: Regression Predicting Racial Resentment

Table 6 shows the regression result from both original data and data generated from ChatGPT.

Table 6: Pseudo Survey Results as Table 3

	<i>Dependent variable:</i>	
	rr_scale	
	original (1)	pseudo (2)
Constant	0.41*** (0.01)	0.50*** (0.003)
rr_post	−0.01 (0.02)	0.003 (0.01)
I(rr_post *welfare_word)	−0.02 (0.02)	−0.01 (0.01)
Observations	1,590	1,590
R ²	0.002	0.001
Adjusted R ²	0.0002	−0.0004
Residual Std. Error (df = 1587)	0.32	0.10
F Statistic (df = 2; 1587)	1.20	0.71
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01		

Appendix

The prompt

I would like you to generate a synthetic survey dataset for my survey experiment. Please simulate realistic response patterns and generate a CSV dataset that I can download. Please name the dataset “ data_recialresentment.csv”. Please then briefly explain your process for generating the dataset. Also include the sources underlying the sociographic distributions. Please also provide the python code that you use to generate the data. Finally, please formulate the research question underlying this experiment.

Sample size: 1590 simulated responses

Distribution of the sample: Representative for the US in August 2021 separately according to:

- Age (Min = 16; Max = 95);
- Gender (man = 1, woman = 2, non-binary = 3, self-describe = 4);
- highest level of education achieved (1 = did not graduate from high school; 2 = high school graduate, 3 = some colleague but no degree (yet), 4 = 2-year college degree, 5 = 4-year college degree, 6 = postgraduate degree);
- Annual income (1 = less than \$30,000; 2=\$30,000 to \$39,999; 3 = \$40,000 to \$49,999; 4 = \$50,000 to \$59,999; 5 = \$60,000 to \$69,999; 6 = \$70,000 to \$79,999; 7 = \$80,000 to \$99,999; 8 = \$100,000 to \$119,999; 9 = \$120,000 to \$149,999; 10 = \$150,000 to \$199,999; 11 = \$200,000 to \$249,000; 12 = \$250,000 or more; 13 = Prefer not to say);
- Ethnicities (1 = White, 2 = Black or African-American, 3 = Hispanic or Latino, 4 = Asian or Asian-American, 5 = Native American, 6 = Two or more races, 7 = Middle Eastern, 9 = Other), and
- When Q75 == 6, then add Q7.6 as a multi-selective ethnicities (1 = White, 2 = Black or African-American, 3 = Hispanic or Latino, 4 = Asian or Asian-American, 5 = Native American, 6 = Middle

Eastern, 7 = Don't Know)

- Ideology(1 = Liberal, 2 = Slightly liberal, 3 = Moderate, 4 = Slightly conservative, 5 = Conservative)

The survey experiment is a question wording experiment conducted in the US. There are four variants. Each respondent randomly sees one of the four variants depending on two independent indicators. The data set must contain two variables that indicate which variants the person has seen. The first variable indicates the type of the policy spending question (named Q32 and Q33) the person has to answer, and the second indicator affects the order between the policy spending question and a question battery of racial resentment. Please call the first indicator variable “welfare_word” and the second indicator variable “rr_post”. Please call the battery of racial resentment Q22, Q23, Q24, and Q25. Please estimate the answers for those questions depending on the sociodemographic characteristics of the respondents.

Variant 0: (welfare_word = 0 and rr_post = 0) The respondents received a policy spending question Q33 and a four-question battery of racial resentment before the policy spending question.

Variant 1: (welfare_word = 1 and rr_post = 0) The respondents received a policy spending question Q32 and a four-question battery of racial resentment before the policy spending question.

Variant 2: (welfare_word = 0 and rr_post = 1) The respondents received a policy spending question Q33, and a four-question battery of racial resentment after the policy spending question.

Variant 3: (welfare_word = 1 and rr_post = 1) The respondents received a policy spending question Q32, and a four-question battery of racial resentment after the policy spending question.

Q32: We are faced with many problems in this country, none of which can be solved easily or inexpensively. Please tell us whether you think we are spending too much, too little, or about the right amount on welfare.

Answer option Q32: 1 “too much”; 2 “too little”; 3 “about the right amount”.

Q33: We are faced with many problems in this country, none of which can be solved easily or inexpensively. Please tell us whether you think we are spending too much, too little, or about the right amount on assistance to the poor.

Answer option Q33: 1 “too much”; 2 “too little”; 3 “about the right amount”.

Q22: Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors.

Answer option 22: 1 “Agree strongly”; 2 “Agree somewhat”; 3 “Neither agree nor disagree”; 4 “Disagree somewhat”; and 5 “Disagree strongly”.

Q23: Generations of slavery and discrimination have created conditions that make it difficult for blacks to work their way out of the lower class.

Answer option 23: 1 “Agree strongly”; 2 “Agree somewhat”; 3 “Neither agree nor disagree”; 4 “Disagree somewhat”; and 5 “Disagree strongly”.

Q24: Over the past few years, blacks have gotten less than they deserve.

Answer option 24: 1 “Agree strongly”; 2 “Agree somewhat”; 3 “Neither agree nor disagree”; 4 “Disagree somewhat”; and 5 “Disagree strongly”.

Q25: It's really a matter of some people not trying hard enough; if blacks would only try harder they could be just as well off as whites.

Answer option 25: 1 “Agree strongly”; 2 “Agree somewhat”; 3 “Neither agree nor disagree”; 4 “Disagree somewhat”; and 5 “Disagree strongly”.

References

- Albertson, B., & Jessee, S. (2023). Moderator placement in survey experiments: Racial resentment and the “welfare” versus “assistance to the poor” question wording experiment. *Journal of Experimental Political Science*, 10(3), 448–454. <https://doi.org/10.1017/XPS.2022.18>