Supplemental Appendix

Self-Reported Political Ideology

Eddy S. F. Yeung Kai Quek

TABLE OF CONTENTS

A Sample Characteristics **S2**

B Balance across Experimental Groups **S3**

C Preregistered Analyses Not Reported in the Main Text **S6**
C.1 Tests for Equality of Ideological Distributions S6
C.2 Tests for Equality of Correlations between Self-Reported Ideology and Fiscal Preferences S7
C.3 Tests for Equality of Correlations between Self-Reported Ideology and Preferences for Government Responsibility S8
C.4 Heterogeneity by Education S9
C.5 Word Counting for Open-Ended Responses S10

D Compliance with the Pre-Analysis Plan **S11**
D.1 Sample S11
D.2 Measurement S11
D.3 Analysis S11

E Ideological Understanding before Dropping the Free Trade Item **S13**

F Additional Analyses **S15**
F.1 Evaluating the Construct Validity of the Ideological Knowledge Measure S15
F.2 Evaluating How Self-Reported Ideology Changes across Partisanships S16
F.3 Replicating Figures and Tables with Speeders Included S18
F.4 Reanalyzing Ideological Knowledge with Alternative Operationalization S25
F.5 Evaluating the Extent to Which "Don't Know" Responses Bias the Results S27
F.6 Unpacking the Role of Political Knowledge in Generating Heterogeneity S28

G References **S29**

A Sample Characteristics

Table S1 shows that our sample characteristics are comparable to the national benchmarks.

Table S1. Sample Characteristics in Comparison with National Demographics

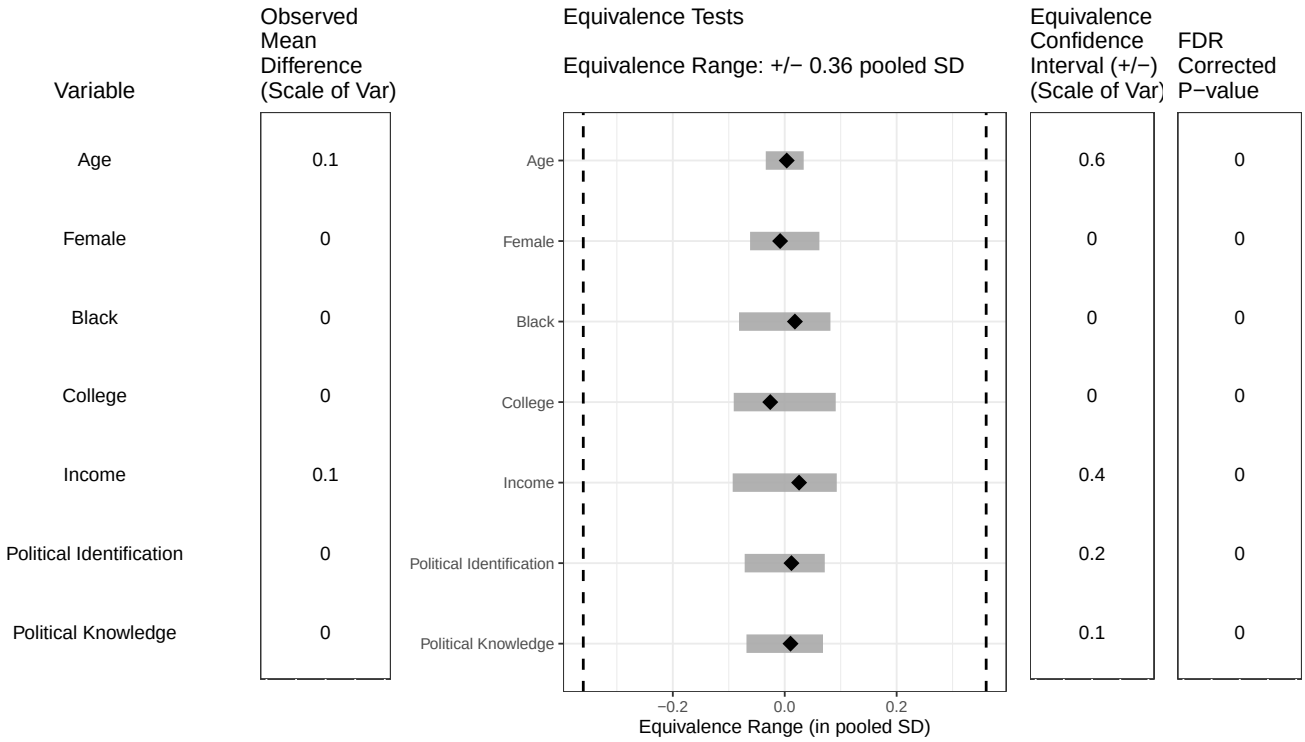
		Benchmark	Sample
GENDER	Male	48.7%	47.1%
	Female	51.3%	52.9%
AGE	18–24	11.9%	11.4%
	25–34	17.9%	15.2%
	35–44	16.4%	16.5%
	45–54	16.0%	15.0%
	55–64	16.6%	14.9%
	65+	21.2%	27.0%
RACE	White	64.1%	66.0%
	Black	12.0%	14.5%
	Native American	1.1%	2.0%
	Asian	6.1%	7.8%
	Pacific Islander	0.2%	0.4%
	Some other race	7.7%	5.0%
	Two or more races	8.8%	4.3%
HISPANIC	Yes	16.8%	17.9%
INCOME	< 15k	9.4%	11.4%
	15k–24k	8.7%	7.1%
	25k–49k	19.7%	21.2%
	50k–99k	28.6%	27.7%
	100k–149k	15.3%	17.4%
	$\geq$ 150k	18.2%	15.2%

Note: Percentages for sex, race and Hispanic origin are based on the adult population. Sex and age are calculated from Table S0101 of the 2019 American Community Survey (https://data.census.gov/cedsci/table?q=S0101&tid=ACSST1Y2019.S0101). Race figures are extracted from the Decennial Census (2020), PL 94-171, Table P3 (https://data.census.gov/cedsci/table?q=P3&tid=DECENNIALPL2020.P3). Hispanic origin figures are extracted from the Decennial Census (2020), PL 94-171, Table P4 (https://data.census.gov/cedsci/table?q=P4&tid=DECENNIALPL2020.P4). Household income is retrieved from CPS Annual Social and Economic Supplement (2021), Table HINC-01 (https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-hinc/hinc-01.html).

B Balance across Experimental Groups

We use equivalence testing recommended by Hartman and Hidalgo (2018) to test the covariate balance between experimental groups. In an equivalence test, difference between group means is the null hypothesis, and equivalence is the alternative hypothesis.[1] Figures S1–S3 show that, under the default equivalence ranges (i.e., $\pm 0.36\sigma$, with σ the pooled SD of the tested covariate), all important demographic and attitudinal variables are balanced across the three experimental groups.

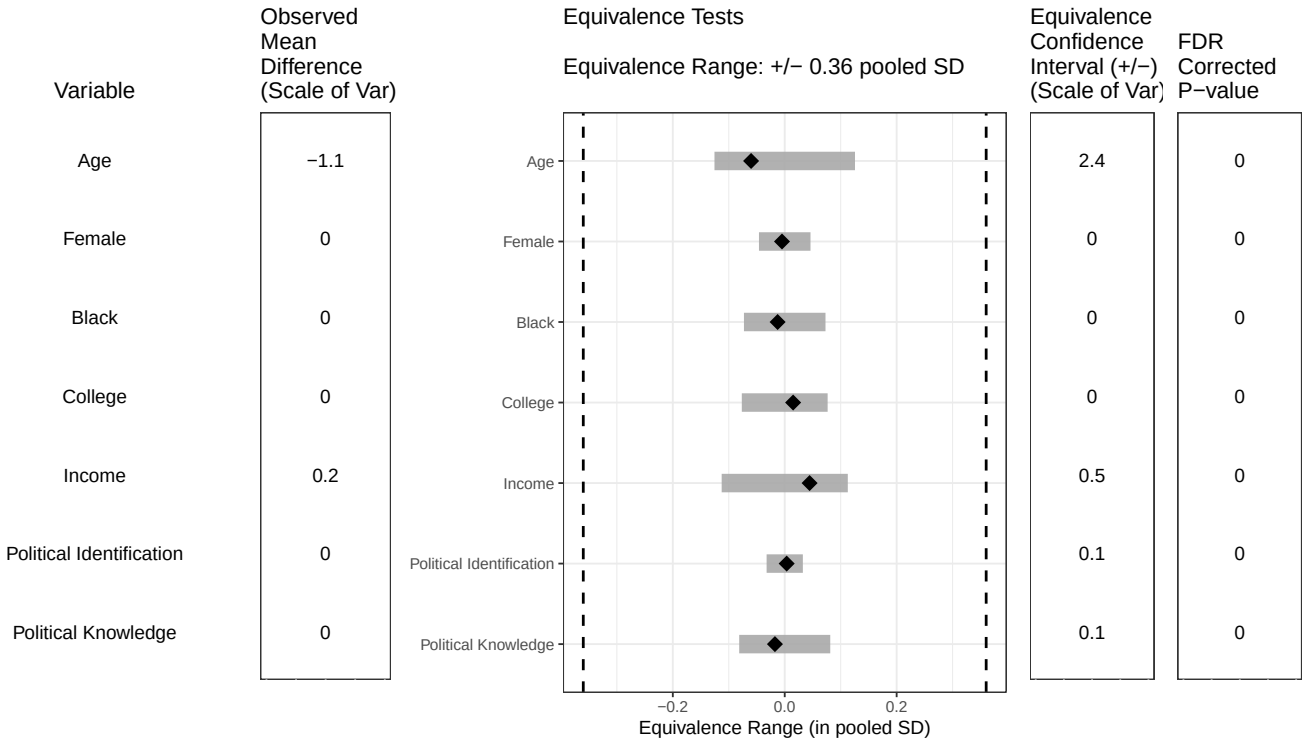
Figure S1. Covariate Balance between the Control Group and the Add Definitions Condition



Note: Gray bars are the inverted equivalence range in standardized differences. Black diamonds are the observed standardized difference of the covariate.

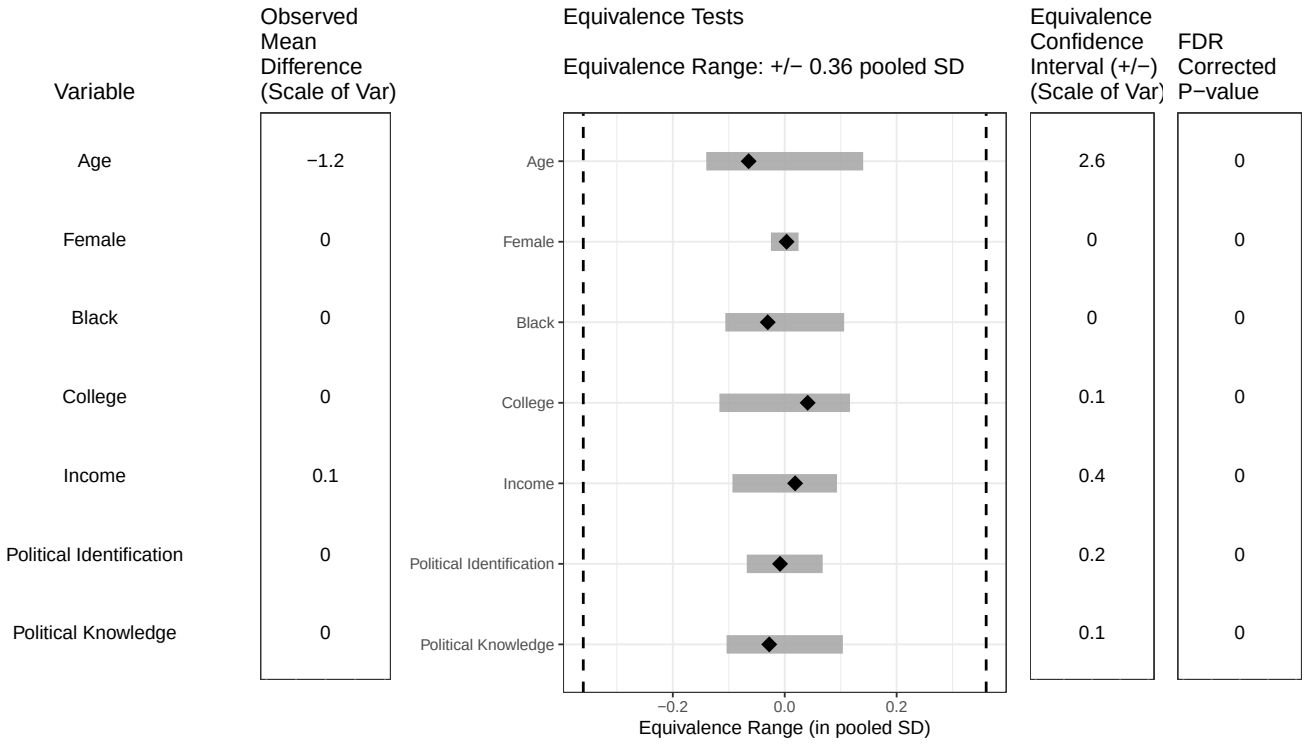
1. For readers interested in the mechanics of equivalence testing, we refer them to Hartman and Hidalgo (2018), Lakens (2017), and Rainey (2014).

Figure S2. Covariate Balance between the Control Group and the Subtract Labels Condition



Note: Gray bars are the inverted equivalence range in standardized differences. Black diamonds are the observed standardized difference of the covariate.

Figure S3. Covariate Balance between the Add Definitions Condition and the Subtract Labels Condition



Note: Gray bars are the inverted equivalence range in standardized differences. Black diamonds are the observed standardized difference of the covariate.

C Preregistered Analyses Not Reported in the Main Text

Due to space constraints, we are unable to report every preregistered analysis in the main text. This appendix conducts the unreported analyses by following our preregistered procedures. For completeness, we also conducted some additional analyses. Our pre-analysis plan can be found at https://osf.io/87bvs.

C.1 Tests for Equality of Ideological Distributions

In our pre-analysis plan, we preregistered on page 10 that we would perform a two-sample Kolmogorov-Smirnov test on (1) the distribution of self-reported ideology in Control Group and (2) the distribution of self-reported ideology in the *Add Definitions* condition or the *Subtract Labels* condition. We also preregistered on page 12 that we would do the same to compare the *Add Definitions* condition and the *Subtract Labels* condition.

Our two-sample Kolmogorov-Smirnov tests show the following:

- The control group vs. the *Add Definitions* condition: $D = 0.0606$, $p = 0.0212$. There is a statistically significant difference between the distributions of self-reported ideology in the control group and the *Add Definitions* condition. Defining the terms "liberal" and "conservative" for respondents changes the distribution of self-reported ideology in the full sample.
- The control group vs. the *Subtract Labels* condition: $D = 0.0432$, $p = 0.1917$. There is a statistically insignificant difference between the distributions of self-reported ideology in the control group and the *Subtract Labels* condition. Our main text shows that this masks important heterogeneity (e.g., Table 2, Figure 5, Figure 6).
- The *Add Definitions* condition vs. the *Subtract Labels* condition: $D = 0.0808$, $p = 0.0046$. There is a statistically significant difference between the distributions of self-reported ideology in the *Add Definitions* condition and the *Subtract Labels* condition. Taking away the ideological labels when defining them changes the distribution of self-reported ideology in the full sample.

C.2 Tests for Equality of Correlations between Self-Reported Ideology and Fiscal Preferences

In our pre-analysis plan, we preregistered on page 10 that we would adopt the bootstrap approach advocated by Rousselet, Pernet, and Wilcox (2021) to compare (1) the Pearson's r between fiscal preferences and self-reported ideology in the control group and (2) the Pearson's r between fiscal preferences and self-reported ideology in the *Add Definitions* condition or the *Subtract Labels* condition.

The question wording for fiscal preferences is taken directly from ANES: "Some people think the government should provide fewer services in order to reduce spending. Suppose these people are on one end of the scale, at point 1. Other people feel that it is important for the government to provide many more services even if it means an increase in spending. Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between. Where would you place yourself on this scale?" As preregistered, we reverse-code this variable such that it 1 represents "more services/spending" and 7 "cut services/spending."

We find the following:

- In the control group, the Pearson's r is 0.37 (95% CI = $[0.32, 0.41]$). In the *Add Definitions* condition, the Pearson's r is 0.41 (95% CI = $[0.35, 0.48]$). In the *Subtract Labels* condition, the Pearson's r is 0.31 (95% CI = $[0.23, 0.38]$).
- The difference in correlations between self-reported ideology and fiscal preferences in the control group and the *Add Definitions* condition is 0.049 ($p = 0.221$).
- The difference in correlations between self-reported ideology and fiscal preferences in the control group and the *Subtract Labels* condition is -0.058 ($p = 0.218$).
- The difference in correlations between self-reported ideology and fiscal preferences in the *Add Definitions* condition and the *Subtract Labels* condition is -0.107 ($p = 0.035$).[2]

2. While we did not preregister this analysis, we report it for completeness.

C.3 Tests for Equality of Correlations between Self-Reported Ideology and Preferences for Government Responsibility

In our pre-analysis plan, we preregistered on page 10 that we would adopt the bootstrap approach advocated by Rousselet, Pernet, and Wilcox (2021) to compare (1) the Pearson's r between preferences for government responsibility and self-reported ideology in the control group and (2) the Pearson's r between preferences for government responsibility and self-reported ideology in the *Add Definitions* condition or the *Subtract Labels* condition.

The question wording for preferences for government responsibility is taken directly from ANES: "Some people feel the government should see to it that every person has a job and a good standard of living. Suppose these people are at one end of a scale, at point 1. Others think the government should just let each person get ahead on their own. Suppose these people are at the other end, at point 7. And, of course, some other people have opinions somewhere in between, at points 2, 3, 4, 5, or 6. Where would you place yourself on this scale?" As preregistered, this variable ranges from 1 ("government should see to jobs and standard of living") to 7 ("government should let each person get ahead on own").

We find the following:

- In the control group, the Pearson's r is 0.30 (95% CI = $[0.25, 0.35]$). In the *Add Definitions* condition, the Pearson's r is 0.39 (95% CI = $[0.32, 0.45]$). In the *Subtract Labels* condition, the Pearson's r is 0.49 (95% CI = $[0.43, 0.55]$).
- The difference in correlations between self-reported ideology and preferences for government responsibility in the control group and the *Add Definitions* condition is 0.089 ($p = 0.035$).
- The difference in correlations between self-reported ideology and preferences for government responsibility in the control group and the *Subtract Labels* condition is 0.194 ($p < 0.001$).
- The difference in correlations between self-reported ideology and preferences for govern-

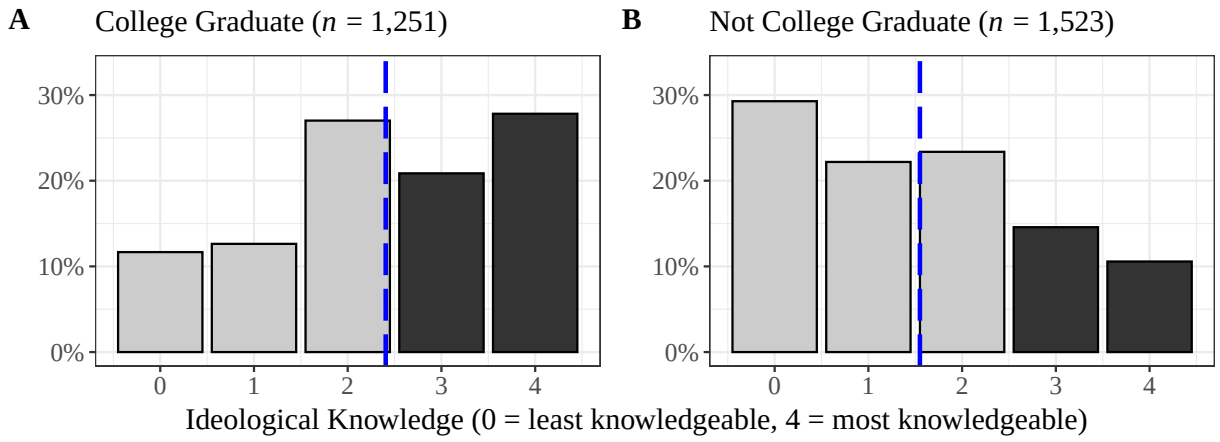
ment responsibility in the *Add Definitions* condition and the *Subtract Labels* condition is 0.105 ($p = 0.022$).[3]

Thus, our label-free measure of ideology (i.e., the measure adopted by the *Subtract Labels* condition) appears to perform much better in capturing respondents' preferences for government responsibility than the standard ANES measure does.

C.4 Heterogeneity by Education

In our pre-analysis plan, we preregistered on page 11 that we would conduct subgroup analysis by education. Figure S4 shows how ideological knowledge differs between college and non-college graduates. The average number of correct responses to the ideology questions was 2.41 among college graduates and was 1.55 among non-college graduates ($\beta = 0.86$, 95% CI = $[0.76, 0.95]$, $n = 2,774$). Thus, highly educated individuals appear to be more familiar and conversant with ideological terms.

Figure S4. Distribution of Ideological Knowledge among College and Non-College Graduates



Note: The dashed lines indicate subgroup means.

We explore the existence of heterogeneous treatment effects by education. To do so, we estimate the following equation: Self-Reported Ideology$_i = \alpha + \beta_1$Add Definitions$_i + \beta_2$Subtract Labels$_i +$ γ_1(Add Definitions$_i \times$ College Graduate$_i$) + γ_2(Subtract Labels$_i \times$ College Graduate$_i$) + λCollege

3. While we did not preregister this analysis, we report it for completeness.

Graduate$_i + \varepsilon_i$, where Self-Reported Ideology$_i$ is respondent i's self-reported ideology in the corresponding treatment condition, Add Definitions$_i$ and Subtract Labels$_i$ indicate i's treatment conditions, and College Graduate$_i$ is a dummy variable indicating whether i already graduated from college. Estimates of γ_1 and γ_2 are 0.16 ($SE = 0.14$) and –0.05 ($SE = 0.14$) respectively, suggesting a lack of treatment effect heterogeneity.

C.5 Word Counting for Open-Ended Responses

In our pre-analysis plan, we preregistered on page 12 that we would count how many self-reported liberals, in an open-ended question,[4] mention any of the following terms: "Democrat," "Democrats," and "Democratic Party." We also preregistered that we would count how many self-reported conservatives, in another open-ended question,[5] mention any of the following terms: "Republican," "Republicans," and "Republican Party." We find 17 respondents for the former and 22 for the latter. To provide a reference point, we also counted how many respondents mentioned any of the following terms: "activ*", "chang*", "tradition*", "interven*" (where the asterisks indicate any possible combinations of letters). Although this dictionary for word counting is not preregistered, they are the key terms that appear in Lowi et al.'s (2019) definitions of "liberal" and "conservative." We find that 36 self-reported liberals and 40 self-reported conservatives mentioned these terms.

4. The question wording is "In a sentence or two, please tell us what defines a LIBERAL?"
5. The question wording is "In a sentence or two, please tell us what defines a CONSERVATIVE?"

D Compliance with the Pre-Analysis Plan

This appendix documents and lists any deviations of our paper from the pre-analysis plan. The pre-analysis plan is available at https://osf.io/87bvs.

D.1 Sample

- *Recruitment:* As preregistered, we recruited our respondents via Lucid.
- *Sampling:* Consistent with our pre-analysis plan, we set demographic benchmarks for Lucid based on sex, age, race, and household income. Our benchmarks in sex and age are drawn from the 2019 American Community Survey (Table S0101). For race, we draw from the 2020 Decennial Census (PL 94–171, Tables P3–P4). For household income, we draw from the 2021 CPS Annual Social and Economic Supplement (Table HINC-01). Table S1 shows that our sample is close to these preregistered benchmarks.
- *Size:* As stated in the pre-analysis plan, the targeted sample size was 2,500 American adults. An increase in research budget allowed us to recruit more respondents. This increased our study's power, especially in terms of detecting treatment effect heterogeneity.
- *Inattentiveness:* As preregistered, we excluded respondents whose survey completion time is less than 5 minutes. This procedure dropped 529 respondents, accounting for 12.14% of the original sample.

D.2 Measurement

Only the measurement of ideological knowledge deviates from our preregistration. We explain that in detail in Appendix E. All other variables reported in this study, including the demographics and outcome variables, are measured and coded in accordance with the pre-analysis plan.

D.3 Analysis

Our reported analyses in the main text closely adhere to our pre-analysis plan. For transparency, we record which reported analyses are preregistered, and which are not, in Table S2.

Table S2. Compliance of Reported Analyses with the Pre-Analysis Plan

Analysis Reported in the Main Text	**Preregistered?**	**Reference from the PAP**
Table 1: Percentage of correct responses to each ideology question	Yes	Page 9
Figure 1: Percentage of correct responses to each ideology question by self-reported ideology	No	N/A
Figure 2: Average number of correct responses by question type and self-reported ideology	No	N/A
Figure 3: Distribution of ideological knowledge by PID, race, sex, and political knowledge	Yes	Page 11
Figure 4: Distribution of self-reported ideology across Experimental Conditions	Yes	Page 9
Table 2: Whether and how self-reported ideology is changed by the question wording by PID, race, sex, and political knowledge	Yes	Page 11
Figure 5: Within-subjects differences in ideology by PID in the *Subtract Labels* condition	No	N/A
Figure 6: Ideological differences between Democrats and Republicans based on two different self-reported measures in the *Subtract Labels* condition	No	N/A
Figure 7: Correlation between self-reported ideology and PID across different measures of ideology	Yes	Page 12

E Ideological Understanding before Dropping the Free Trade Item

In our survey, we asked a fifth question for the ideological knowledge test that concerns free trade. The question item was: "In general, free trade is good for our country." We defined the correct answer as "Neither." The percentage of correct responses was only 19.91%. This unusually low accuracy raises concerns about measurement error. To test its impact on the internal consistency of our measure of ideological knowledge, we find that the Cronbach's alpha of the variable *ideological knowledge* drops from $\alpha = 0.65$ to $\alpha = 0.51$ if we include the free trade item. Thus, we deviate from our pre-analysis plan by analyzing this outcome variable with the free trade item dropped.

For completeness, however, we also present a version of analysis of ideological understanding with the free trade item included. Keeping the free trade item, our *ideological knowledge* variable now takes a 6-point scale and ranges from 0 ("Least knowledgeable") to 5 ("Most knowledgeable"). Our reanalysis of ideological understanding, as reported below, suggests that the substantive results remain unchanged.

Figure S5 shows the distribution of the new *ideological knowledge* variable. Figure S6 shows the distribution of the new *ideological knowledge* variable across demographic subgroups.

Figure S5. Distribution of Ideological Knowledge before Dropping the Free Trade Item

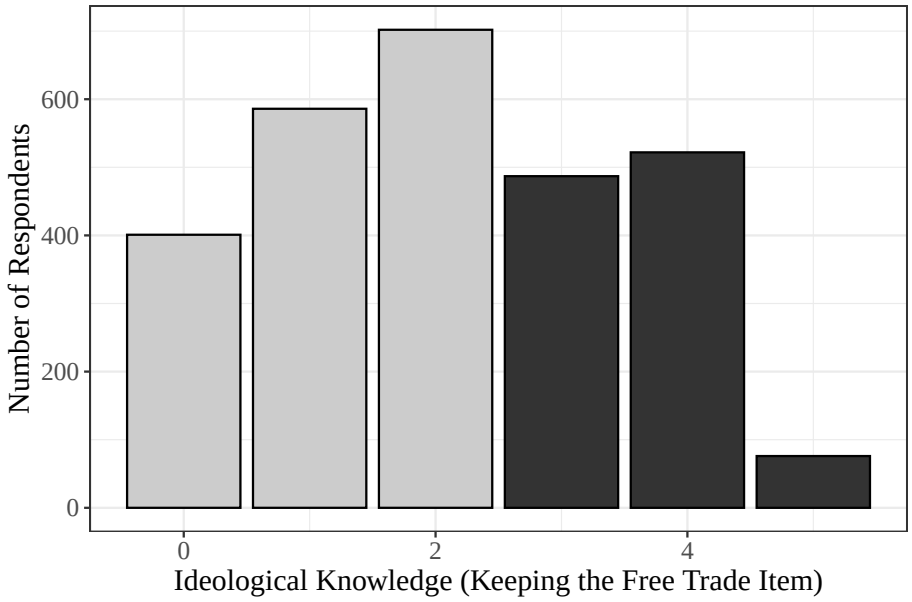
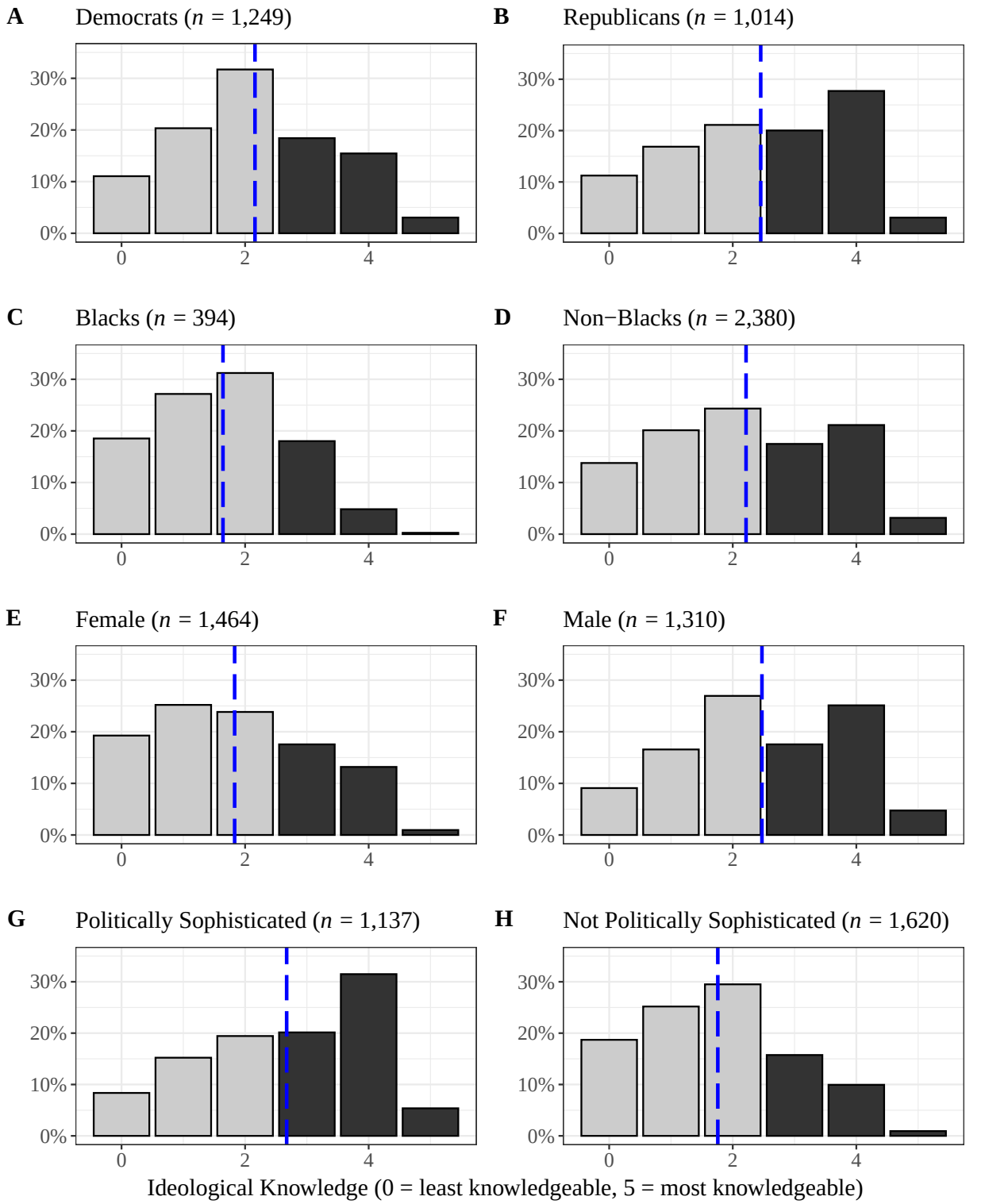


Figure S6. Distribution of Ideological Knowledge across Demographic Subgroups, before Dropping the Free Trade Item



Note: The dashed lines indicate subgroup means.

F Additional Analyses

This appendix presents six sets of additional analysis. First, we evaluate the construct validity of our measure of ideological knowledge. Second, we evaluate how self-reported ideology changes across experimental conditions among Democrats, Republicans, and Independents, respectively. Third, we replicate all the main analyses by using the sample with speeders included. Fourth, we reanalyze the ideological knowledge variable using an alternative operationalization that treats "Neither" as a correct answer on top of "Conservative" for the second item of the variable. Fifth, we evaluate the extent to which "Don't know" responses bias the results reported in our main text. Sixth, we unpack the role of political knowledge in driving the heterogeneous treatment effects detected in the *Subtract Labels* condition.

F.1 Evaluating the Construct Validity of the Ideological Knowledge Measure

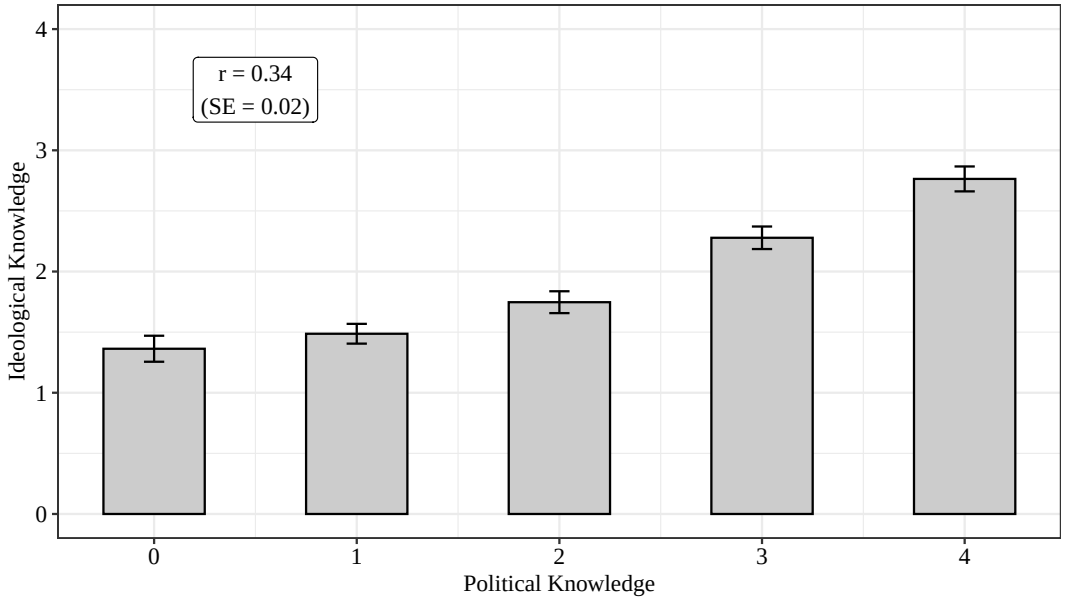
We assess the construct validity of our measure of ideological knowledge by testing how it correlates with political knowledge. To measure political knowledge, we use the following items:

1. To your knowledge, John Roberts's current office is: [Chief Justice / Secretary of Homeland Security / Secretary of Defense / Federal Reserve Board Chairman]
2. To your knowledge, Janet Yellen's current office is: [Secretary of State / Attorney General / Senator / Secretary of the Treasury]
3. To your knowledge, Lee Kuan Yew was formerly the leader of: [South Korea / Singapore / Taiwan / North Korea]
4. The U.S. Bureau of Labor Statistics counts a person as unemployed if they are not employed at any job and are looking for work. By this definition, what percentage of Americans was unemployed in January of 2022? [Around 2 percent / Around 4 percent / Around 6 percent / Around 8 percent]

A long line of literature has argued that ideological labels are more familiar to individuals with higher political knowledge (e.g., Claassen, Tucker, and Smith 2014; Converse 1964; Feldman and

Johnston 2014; Kalmoe 2020; Kinder and Kalmoe 2017). Therefore, if our measure of ideological knowledge has convergent validity, our test should reveal that respondents with higher levels of political knowledge, by answering more questions correctly, scored higher in ideological knowledge. This is what we find. Figure S7 shows that, among our respondents, political knowledge increases with ideological knowledge. The Pearson's r between the two variables is 0.34 ($SE = 0.02$).

Figure S7. Political Knowledge Increases with Ideological Knowledge



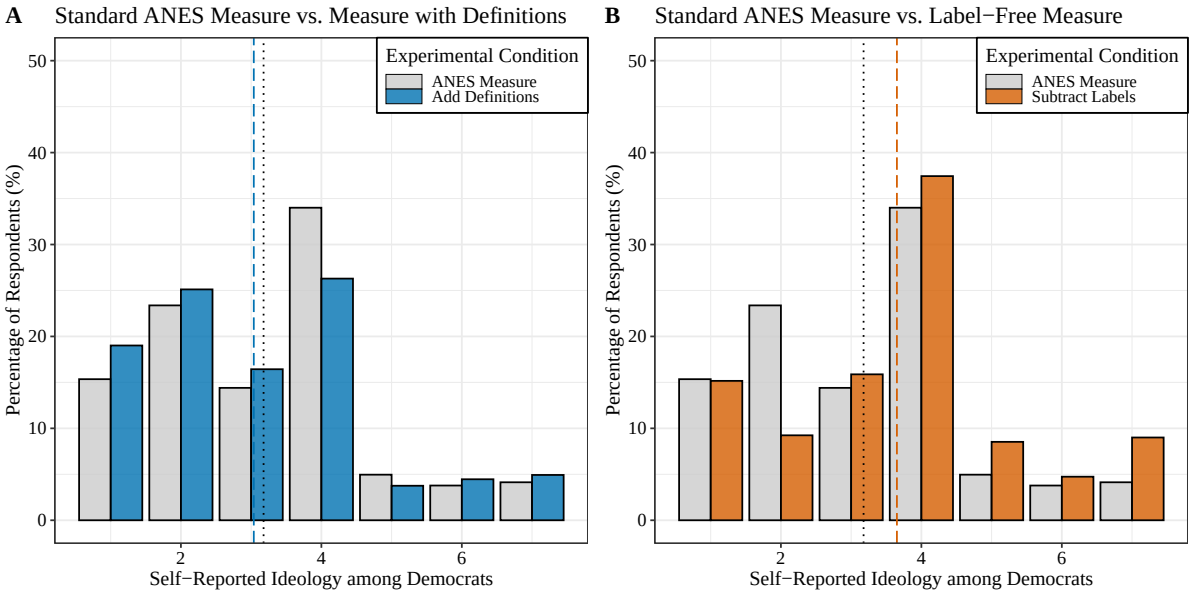
Note: Error bars represent 95% confidence intervals.

F.2 Evaluating How Self-Reported Ideology Changes across Partisanships

We evaluate more comprehensively how self-reported ideology changes across experimental conditions. We do so by visualizing the distributions of self-reported ideology among different subgroups: Democrats, Republicans, and Independents. Figure S8B shows that the label-free measure in the *Subtract Labels* condition pushed Democrats to report a more moderate and conservative ideology.

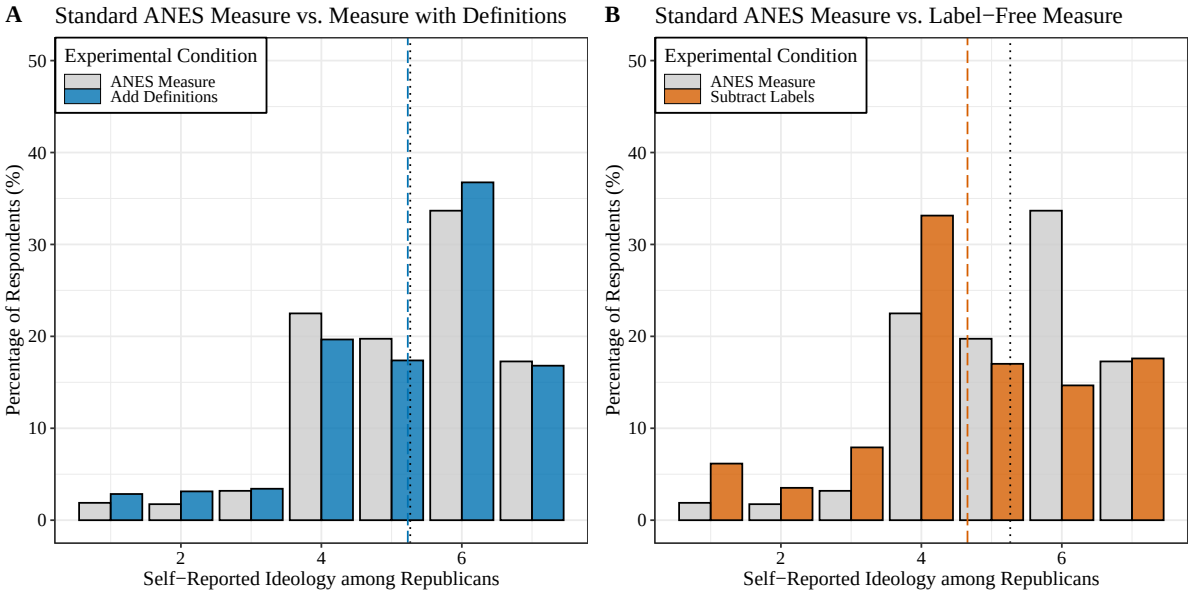
Similarly, Figure S9B shows that the label-free measure in the *Subtract Labels* condition pushed Republicans to report a more moderate and liberal ideology. Independents, however, appeared less likely to report as moderates in both treatment groups, where definitions of ideology were provided (Figure S10).

Figure S8. Distribution of Democrats' Self-Reported Ideology across Experimental Conditions



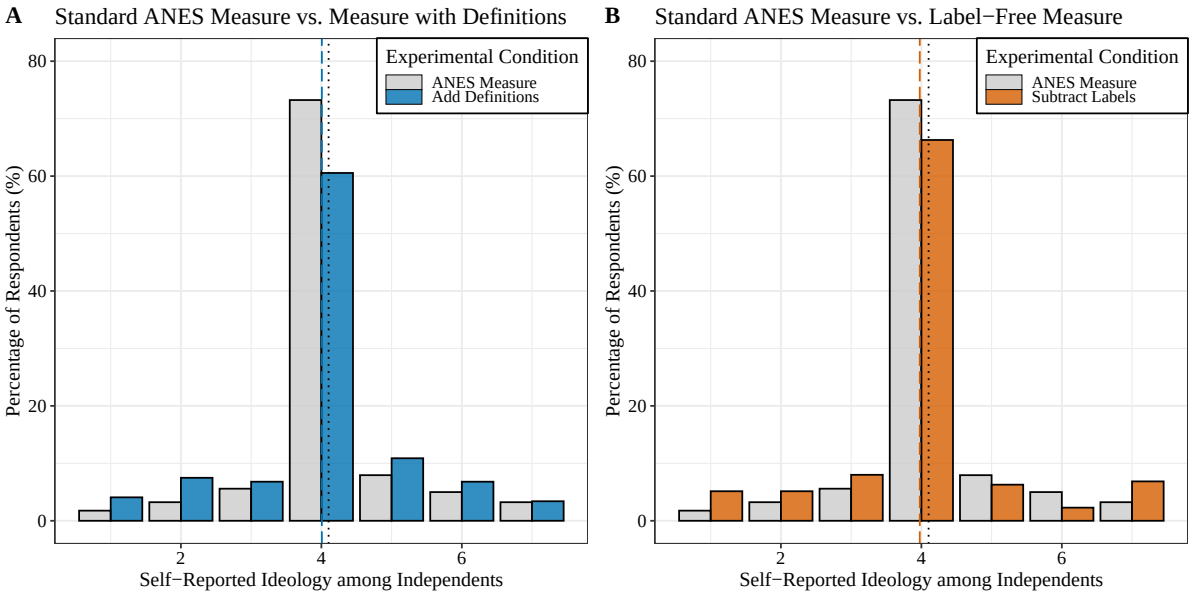
Note: The dashed lines indicate subgroup means.

Figure S9. Distribution of Republicans' Self-Reported Ideology across Experimental Conditions



Note: The dashed lines indicate subgroup means.

Figure S10. Distribution of Independents' Self-Reported Ideology across Experimental Conditions



Note: The dashed lines indicate subgroup means.

F.3 Replicating Figures and Tables with Speeders Included

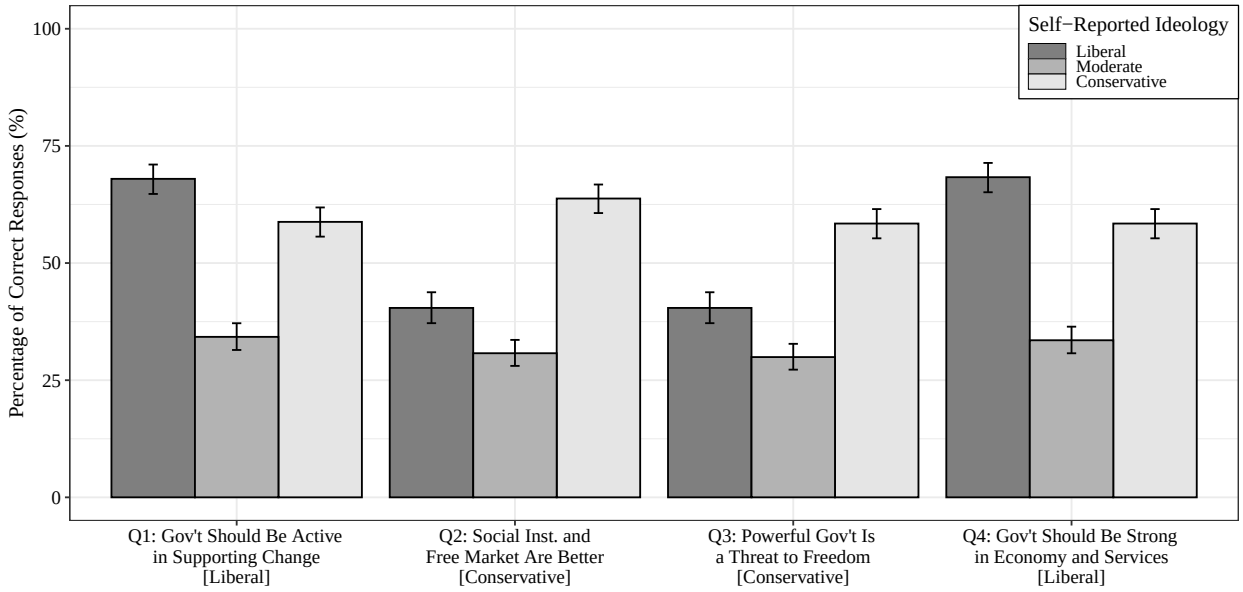
We evaluate whether the analyses reported in the main text are sensitive to the inclusion of speeders. Using the entire sample to replicate all figures and tables shown in the main text, we find that the results remain robust (Figures S11–S17; Tables S3–S4).

Table S3. Percentage of Correct Responses and Don't Knows for Each Ideology Question (Speeders Included)

Ideology Question	Correct Answer	Percentage of Correct Responses	Percentage of Don't Knows
Q1. Government should be active in supporting social and political change	LIBERAL	52.39% [50.56%, 54.22%]	11.81% [10.67%, 13.04%]
Q2. Social institutions and the free market solve problems better than governments	CONSERVATIVE	44.62% [42.81%, 46.45%]	14.81% [13.56%, 16.16%]
Q3. A powerful government is a threat to citizens' freedom	CONSERVATIVE	42.54% [40.74%, 44.36%]	13.35% [12.15%, 14.65%]
Q4. Gov't should play a strong role in the economy and provision of social services	LIBERAL	52.43% [50.60%, 54.25%]	11.50% [10.38%, 12.72%]

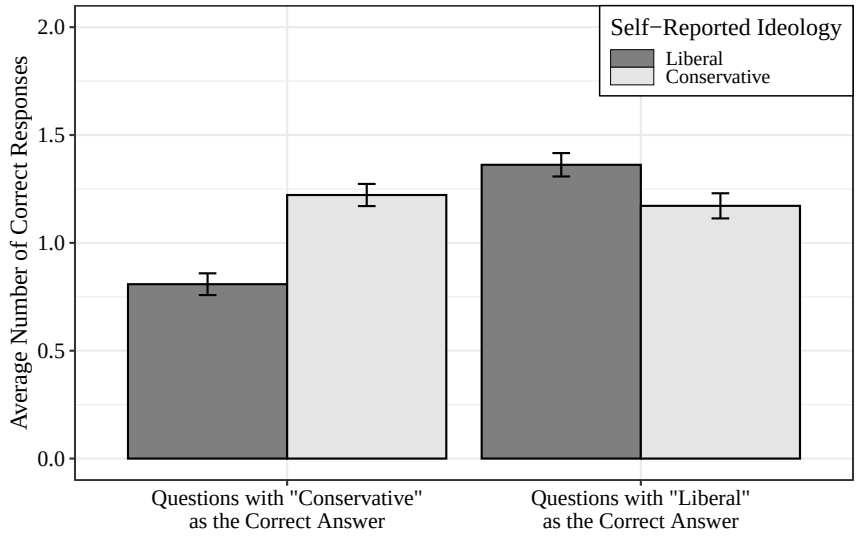
Note: 95% confidence intervals are shown in brackets.

Figure S11. Percentage of Correct Responses to Each Ideology Question by Self-Reported Ideology (Speeders Included)



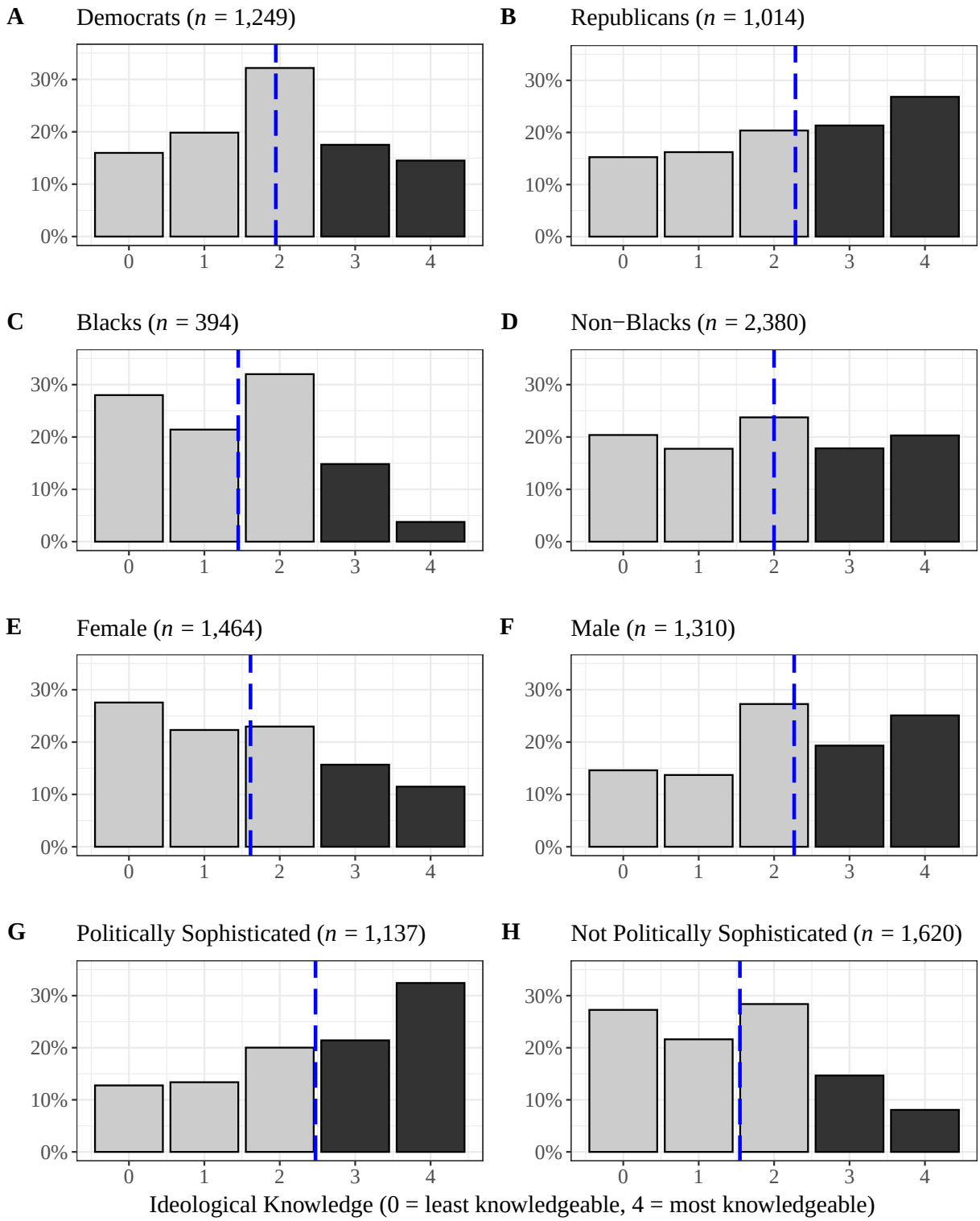
Note: Conservatives are respondents who selected "slightly conservative," "conservative," or "extremely conservative" under ANES's 7-point ideology question. Liberals are those who selected "slightly liberal," "liberal," or "extremely liberal." Moderates are those who selected "moderate; middle of the road." Error bars represent 95% confidence intervals.

Figure S12. Average Number of Correct Responses by Question Type and Self-Reported Ideology (Speeders Included)



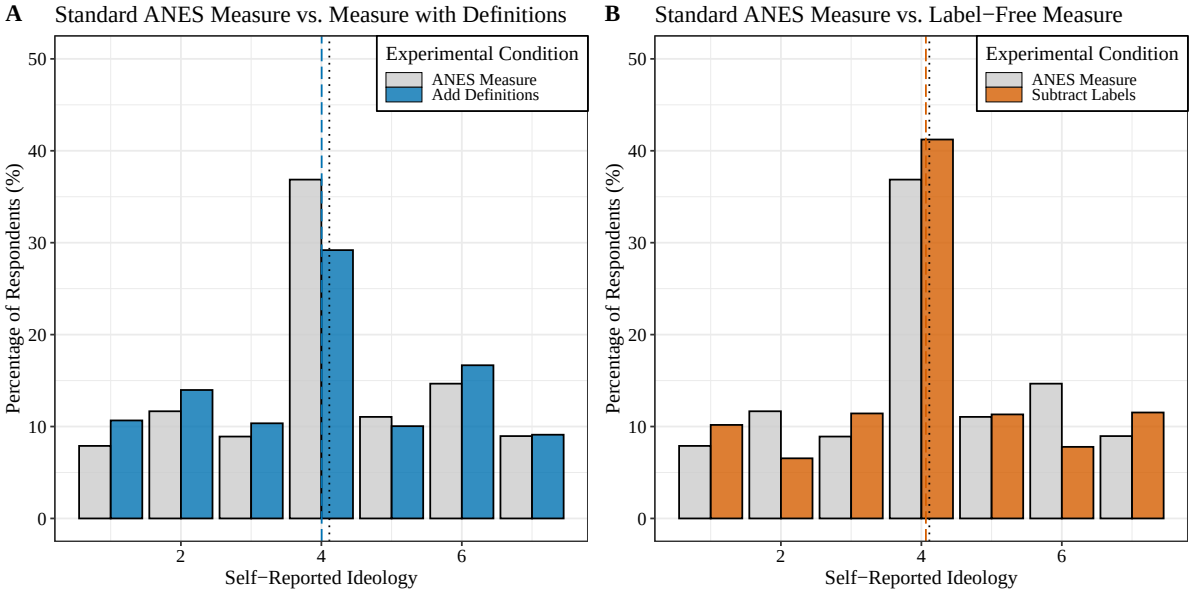
Note: Conservatives are respondents who selected “slightly conservative,” “conservative,” or “extremely conservative” under ANES’s 7-point ideology question. Liberals are those who selected “slightly liberal,” “liberal,” or “extremely liberal.” Error bars represent 95% confidence intervals.

Figure S13. Distribution of Ideological Knowledge across Demographic Subgroups (Speeders Included)



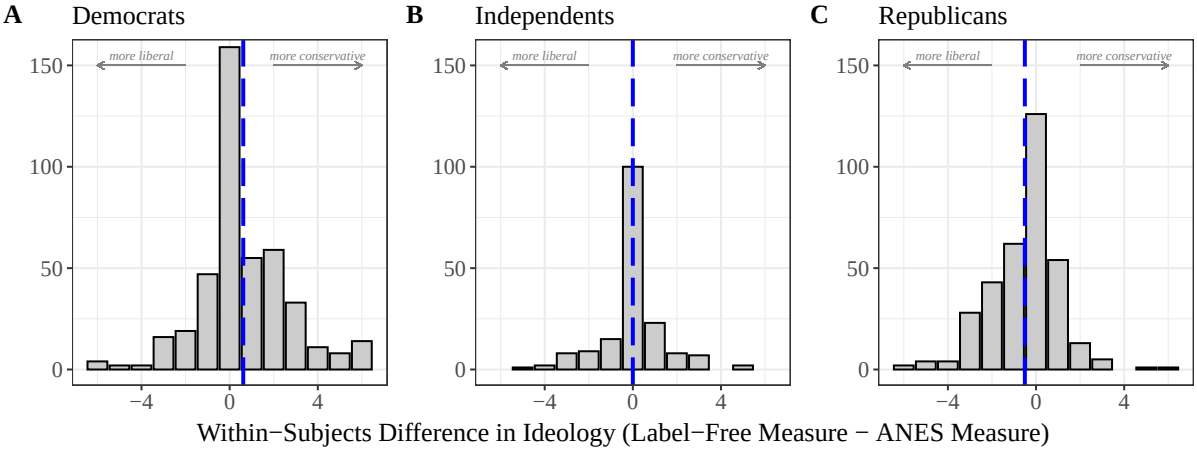
Note: The dashed lines indicate subgroup means.

Figure S14. Distribution of Self-Reported Ideology across Experimental Conditions (Speeders Included)



Note: The dashed lines indicate subgroup means.

Figure S15. Within-Subjects Differences in Ideology by Partisanship in the Subtract Labels Condition (Speeders Included)



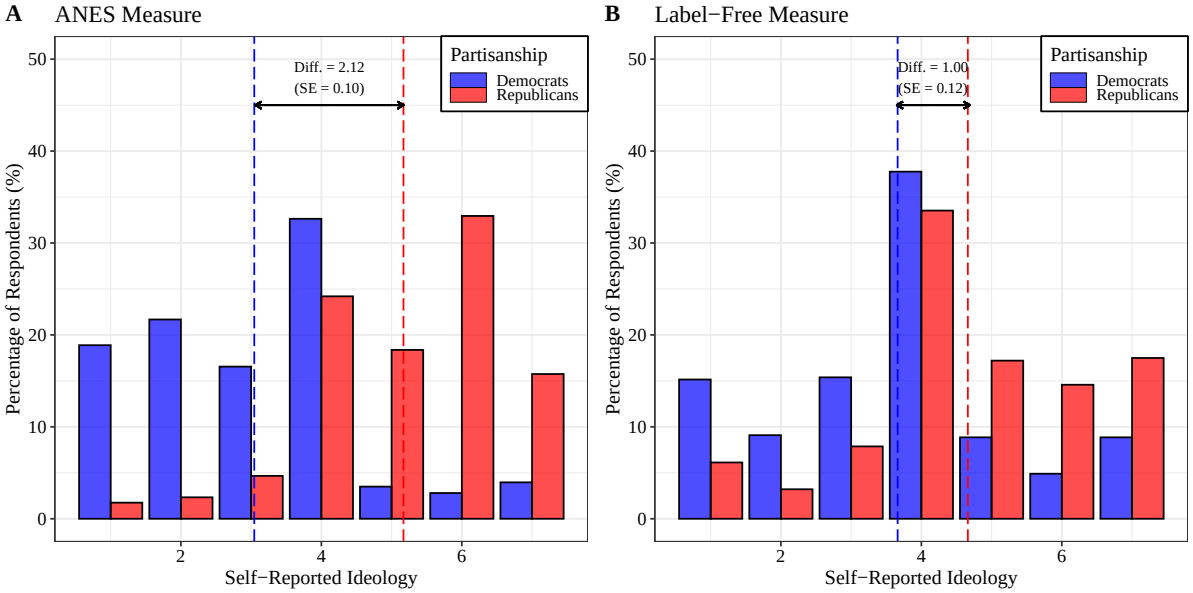
Note: Positive (negative) values indicate that the respondent reported a more conservative (liberal) ideology under the label-free measure of ideology. The dashed lines indicate subgroup means.

Table S4. Self-Reported Ideologies in Different Social Groups Are Changed by Question Wording (Speeders Included)

	Dependent Variable: Self-Reported Ideology						
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Constant	4.11	4.14	4.16	4.15	4.10	4.11	4.09
	(0.04)	(0.06)	(0.07)	(0.06)	(0.07)	(0.07)	(0.08)
Add Definitions	−0.11	−0.11	−0.11	−0.13	0.01	−0.16	−0.04
	(0.07)	(0.06)	(0.11)	(0.06)	(0.09)	(0.08)	(0.16)
Subtract Labels	−0.05	−0.02	−0.13	−0.07	0.01	0.12	0.07
	(0.07)	(0.06)	(0.11)	(0.07)	(0.10)	(0.08)	(0.16)
Republican		1.03	1.13	1.03	1.03	1.03	1.12
		(0.06)	(0.07)	(0.06)	(0.06)	(0.06)	(0.08)
Democrat		−0.76	−0.89	−0.76	−0.76	−0.75	−0.89
		(0.06)	(0.07)	(0.06)	(0.06)	(0.06)	(0.08)
Black		0.03	0.03	−0.06	0.03	0.03	0.01
		(0.08)	(0.08)	(0.10)	(0.08)	(0.08)	(0.10)
Female		−0.17	−0.16	−0.17	−0.10	−0.17	−0.07
		(0.05)	(0.05)	(0.05)	(0.07)	(0.05)	(0.07)
Politically Sophisticated (PS)		0.06	0.07	0.06	0.06	0.12	0.13
		(0.05)	(0.05)	(0.05)	(0.05)	(0.07)	(0.07)
Add Definitions × Republican			0.05				0.04
			(0.15)				(0.15)
Subtract Labels × Republican			−0.44				−0.42
			(0.15)				(0.15)
Add Definitions × Democrat			−0.05				−0.05
			(0.15)				(0.15)
Subtract Labels × Democrat			0.58				0.59
			(0.15)				(0.15)
Add Definitions × Black				0.08			0.12
				(0.18)			(0.19)
Subtract Labels × Black				0.29			−0.03
				(0.19)			(0.19)
Add Definitions × Female					−0.23		−0.20
					(0.12)		(0.13)
Subtract Labels × Female					−0.06		−0.15
					(0.13)		(0.13)
Add Definitions × PS						0.12	0.07
						(0.12)	(0.12)
Subtract Labels × PS						−0.36	−0.31
						(0.12)	(0.13)
Adjusted R^2	0.00	0.23	0.25	0.23	0.23	0.23	0.25
Number of Observations	3893	3790	3790	3790	3790	3790	3790
RMSE	1.69	1.48	1.47	1.48	1.48	1.48	1.47

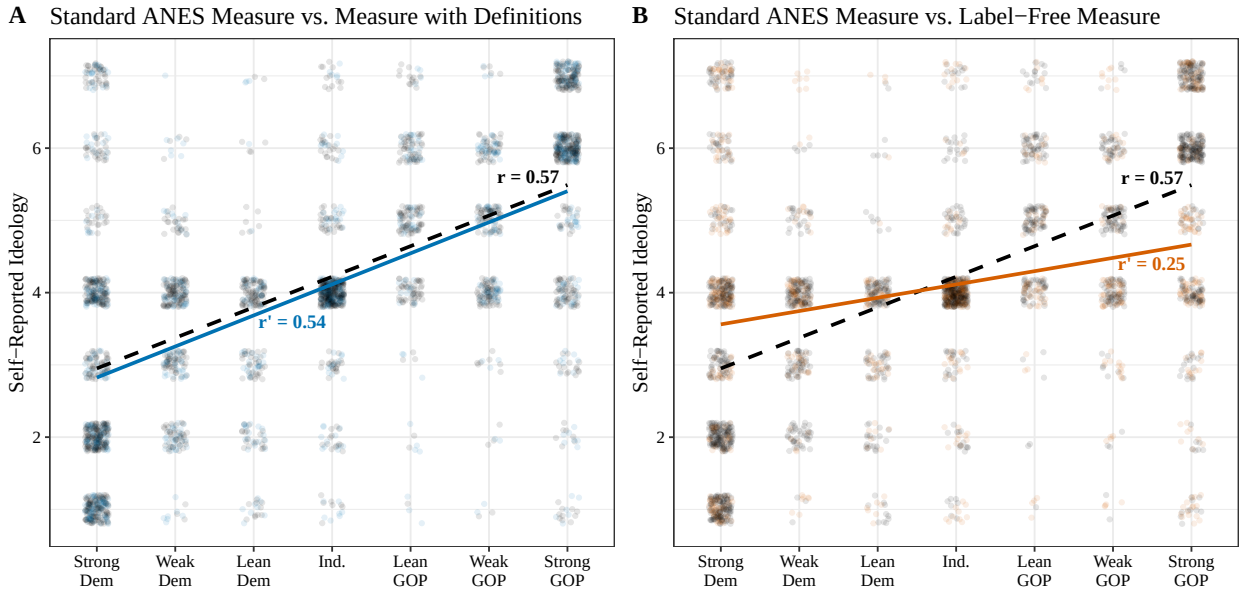
Note: Entries are OLS estimates with robust standard errors in parentheses.

Figure S16. Ideological Differences between Democrats and Republicans Based on Two Different Self-Reported Measures in the Subtract Labels Condition (Speeders Included)



Note: The dashed lines indicate subgroup means.

Figure S17. Correlation between Self-Reported Ideology and Partisanship under Different Measures of Ideology (Speeders Included)

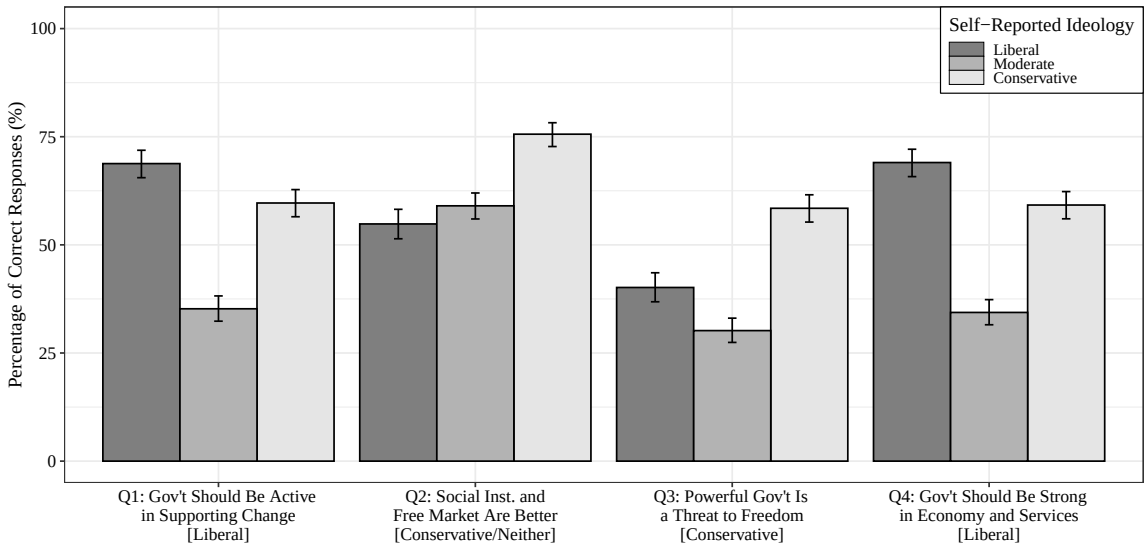


Note: The solid lines in Panel A and Panel B are best-fit lines for self-reported ideology and partisanship in the *Add Definitions* condition and the *Subtract Labels* condition, respectively. The dashed lines are best-fit lines under the standard ANES measure.

F.4 Reanalyzing Ideological Knowledge with Alternative Operationalization

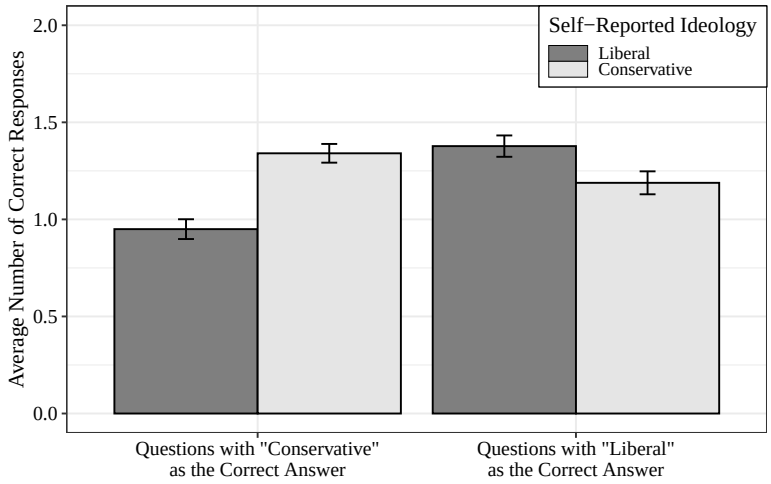
We present a version of analysis that treats "Neither" as a correct answer on top of "Conservative" for Q2 of the ideological knowledge variable. Figures S18–S20 show the substantive findings remain the same.

Figure S18. Percentage of Correct Responses to Each Ideology Question by Self-Reported Ideology (Alternative Operationalization)



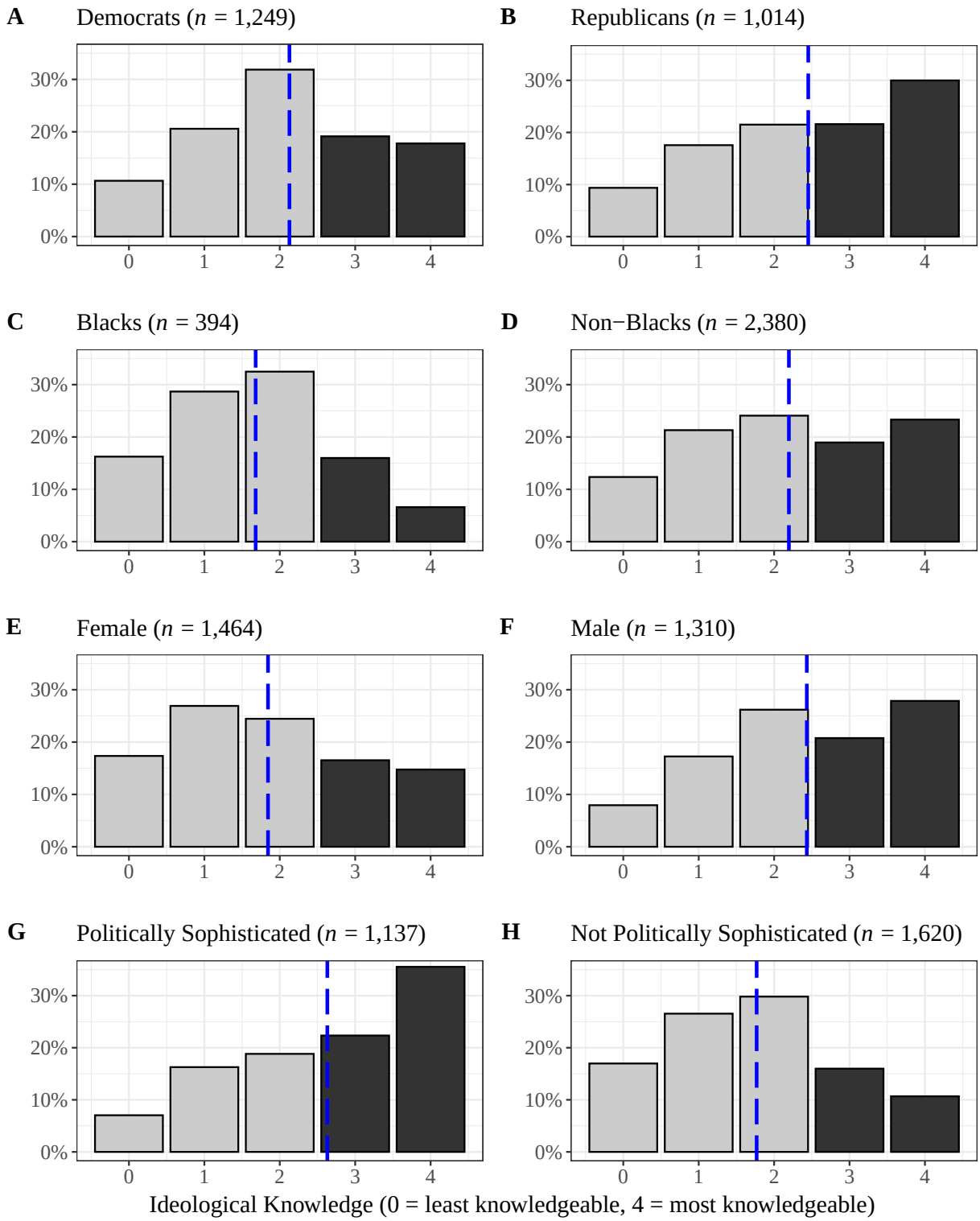
Note: Error bars represent 95% confidence intervals.

Figure S19. Average Number of Correct Responses by Question Type and Self-Reported Ideology (Alternative Operationalization)



Note: The category on the left includes Q2, which now treats "Neither" also as a correct answer.

Figure S20. Distribution of Ideological Knowledge across Demographic Subgroups with Alternative Operationalization

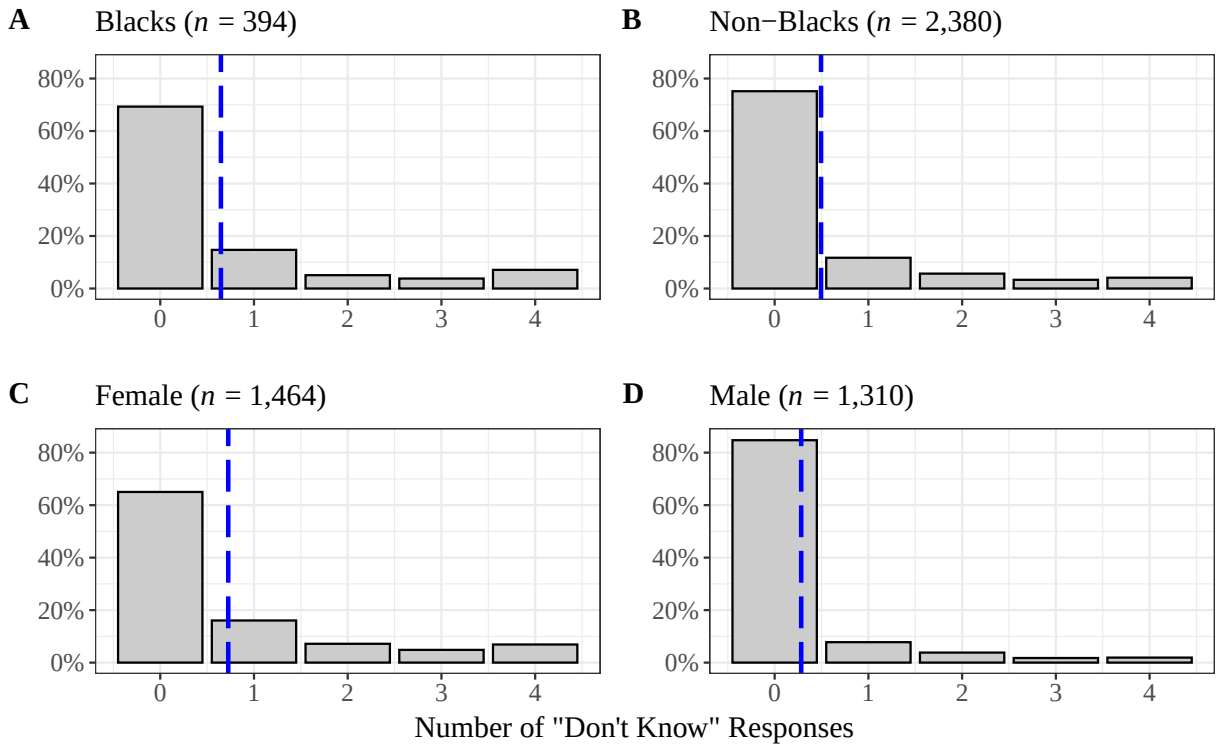


Note: The dashed lines indicate subgroup means.

F.5 Evaluating the Extent to Which "Don't Know" Responses Bias the Results

We find that black and female respondents were more likely to report "Don't know" than their counterparts (Figure S21). On average, the number of "Don't Know" responses to the ideological knowledge questions among Blacks was 0.15 higher than that among non-Blacks ($p = 0.02$, $n = 2,774$), and the number among female respondents was 0.44 higher than the number among male respondents ($p < 0.001$, $n = 2,774$).

Figure S21. The Number of "Don't know" Responses across Racial and Gender Groups



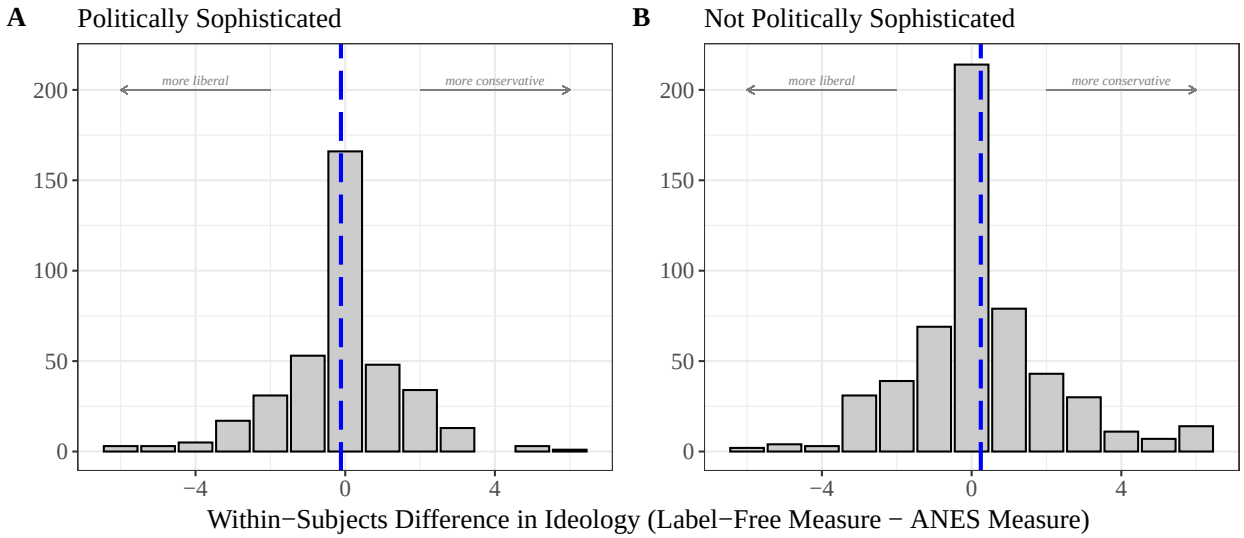
Note: The dashed lines indicate subgroup means.

What if we assume to the extreme that all "Don't know" responses would have been correct? We find that this assumption would reduce part of the subgroup differences we documented in the main text. The racial gap in ideological knowledge would close from $\beta = -0.58$ (95% CI = $[-0.71, -0.46]$, $n = 2,774$) to $\beta = -0.43$ (95% CI = $[-0.56, -0.30]$, $n = 2,774$). More substantially, the gender gap would close from $\beta = -0.66$ (95% CI = $[-0.76, -0.56]$, $n = 2,774$) to $\beta = -0.22$ (95% CI = $[-0.31, -0.12]$, $n = 2,774$). The gaps, however, still exist even with this extreme assumption.

F.6 Unpacking the Role of Political Knowledge in Generating Heterogeneity

In the main text, we report a heterogeneous treatment effect by political knowledge in the *Subtract Labels* condition. But is the heterogeneous effect mainly driven by politically sophisticated or less sophisticated respondents? To shed light on this question, we zoom in on the *Subtract Labels* condition, which allows us to make within-subjects comparisons. Figure S22 shows the results. We find that while the politically sophisticated reported slightly lower conservatism (mean difference = –0.11, $n = 377$, $p = 0.175$), less sophisticated respondents reported much greater conservatism (mean difference = 0.25, $n = 546$, $p = 0.003$). This finding suggests that the treatment effect heterogeneity is mostly driven by the less politically sophisticated respondents. Under our label-free measure which defines "liberal" and "conservative" for them without displaying ideological labels, they reported greater conservatism.

Figure S22. Within-Subjects Differences in Self-Reported Ideology by Political Knowledge in the Subtract Labels Condition



Note: The dashed lines indicate subgroup means.

G References

Claassen, Christopher, Patrick Tucker, and Steven S. Smith. 2015. "Ideological Labels in America." *Political Behavior* 37 (2): 253–78.

Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David E. Apter, 206–61. Gelncoe: The Free Press.

Feldman, Stanley, and Christopher Johnston. 2014. "Understanding the Determinants of Political Ideology: Implications of Structural Complexity." *Political Psychology* 35 (3): 337–58.

Hartman, Erin, and F. Daniel Hidalgo. 2018. "An Equivalence Approach to Balance and Placebo Tests." *American Journal of Political Science* 62 (4): 1000–13.

Kalmoe, Nathan P. 2020. "Uses and Abuses of Ideology in Political Psychology." *Political Psychology* 41 (4): 771–93.

Kinder, Donald R., and Nathan P. Kalmoe. 2017. *Neither Liberal nor Conservative: Ideological Innocence in the American Public*. University of Chicago Press.

Lakens, Daniël. 2017. "Equivalence Tests: A Practical Primer for *t* Tests, Correlations, and Meta-Analyses." *Social Psychological and Personality Science* 8 (4): 355–62.

Lowi, Theodore J., Benjamin Ginsberg, Kenneth A. Shepsle, and Stephen Ansolabehere. 2019. *American Government: Power and Purpose*, 15th edition. New York: W. W. Norton.

Rainey, Carlisle. 2014. "Arguing for a Negligible Effect." *American Journal of Political Science* 58 (4): 1083–91.

Rousselet, Guillaume A., Cyril R. Pernet, and Rand R. Wilcox. 2021. "The Percentile Bootstrap: A Primer with Step-by-Step Instructions in R." *Advances in Methods and Practices in Psychological Science* 4 (1): 1–10.