

# Women Also Know Stuff

## Challenging the Gender Gap in Political Sophistication\*

Patrick W. Kraft<sup>†</sup>

### Abstract

This paper proposes a simple but powerful framework to measure political sophistication based on open-ended survey responses. *Discursive sophistication* uses automated text analysis methods to capture the complexity of individual attitude expression. I validate the approach by comparing it to conventional political knowledge metrics using different batteries of open-ended items across five surveys spanning four languages (total  $N \simeq 35,000$ ). The new measure casts doubt on the oft-cited gender gap in political knowledge: Women might know fewer facts about institutions and elites, but they do not differ substantively in the sophistication of their expressed political attitudes.

*Keywords:* political sophistication, gender gap, open-ended responses, text analysis

---

\*This title is inspired by <https://womentalkpolitics.com/>, an organization that promotes the work of women in political science by providing a public database of relevant women experts for journalists, scholars, and others. Previous versions of this manuscript have been presented at Polmeth, MPSA, EPSA, ISPP, APSA, the Toronto Political Behavior Meeting, and Zurich Text as Data. I thank Alexa Bankert, Jason Barabas, Scott Clifford, Andy Delton, Milena Djourelova, Kathy Dolan, Peter DeScioli, Stanley Feldman, Roy Gava, Fabrizio Gilardi, Jennifer Jerit, Yanna Krupnikov, Emmy Lindstam, Hannah Nam, Michael Persess, Daniel Rubenson, Rüdiger Schmitt-Beck, and Arthur Spirling for helpful comments on previous versions of this manuscript. Special thanks to Céline Colombo, Scott Clifford, and Jennifer Jerit for sharing their data. The manuscript and code are available on GitHub: <https://github.com/pwkraft/knowledge>.

<sup>†</sup>Assistant Professor, University of Wisconsin-Milwaukee, [kraftp@uwm.edu](mailto:kraftp@uwm.edu).

Political sophistication is one of the most fundamental concepts in the study of political attitudes and behavior. While scholars have long emphasized people's alarmingly low levels of sophistication (Converse, 1964; Delli Carpini and Keeter, 1996), fundamental issues regarding its measurement continue to plague the discipline (Mondak, 2001; Sturgis, Allum, and Smith, 2008). Scholars usually rely on survey questions that assess individuals' ability to recall basic facts about political institutions and officeholders as a proxy for sophistication (Zaller, 1990; Carpini and Keeter, 1993). In principle, these factual knowledge questions should cover information that is necessary and/or sufficient for citizens to make competent decisions in a given context (Lupia, 2006), but determining such a set of items proves to be extremely difficult—especially since there are systematic differences in types of knowledge (Barabas et al., 2014). Even within a given policy area, people may disagree about which facts are crucial for political competence due to inherent value differences (Lupia, 2015). Furthermore, to the extent that items vary in difficulty across subgroups (i.e., differential item functioning), relying on standard knowledge batteries may introduce systematic measurement error (Petryka and MacIntosh, 2013).

One manifestation of such systematic measurement error is the oft-cited gender gap in sophistication. On the basis of conventional factual knowledge scores, women frequently appear to be less informed about politics than men (Verba, Burns, and Schlozman, 1997; Wolak and McDevitt, 2011). However, some scholars suggest that these differences might be an artifact of the measurement approach. For instance, Mondak and Anderson (2004) argue that at least part of the gender gap can be attributed to the fact that women are less likely to guess than men when facing a factual knowledge question for which they do not know an immediate answer. Others posit that the gap can be attenuated by focusing on gender-relevant political knowledge (Dolan, 2011) or by providing policy-specific information (Jerit and Barabas, 2017).

In this paper, I re-examine the gender gap by proposing a measure of *discursive sophistication* that is based on how people discuss their political preferences in open-ended

survey responses. Specifically, I develop a framework to assess whether beliefs and attitudes in a given political domain are expressed in a more elaborate manner—a question that is not directly discernible from off-the-shelf factual knowledge items. Measuring sophistication based on how people talk about politics provides two major advantages compared to off-the-shelf factual knowledge items: (1) it captures the extent to which a respondent’s political beliefs are based on elaborate reasoning, and (2) it can easily pinpoint competence in specific areas by incorporating targeted open-ended items. The resulting measurement is therefore conceptually closer to the underlying latent trait of interest: the degree of structure and constraint in political belief systems (Tetlock, 1983; Luskin, 1987).

I validate the measure across multiple data sets by comparing it to conventional factual knowledge scores as predictors of various indicators of civic competence and engagement. While discursive sophistication shares a considerable amount of variance with traditional metrics, they are far from equivalent. Indeed, discursive sophistication is a stronger predictor of turnout and other forms of political engagement. Contrary to previous research, however, I find no evidence for a gender gap in discursive sophistication. While women might score lower than men on factual knowledge about political institutions and elites, there are no differences in the complexity of expressed political attitudes. This divergence can be explained by the fact that open-ended responses allow women to focus on different issues than men. More generally, the results suggest that measuring political sophistication based on open-ended responses can provide new opportunities to examine political knowledge across time and contexts.

## Political Sophistication and Factual Knowledge

Public opinion researchers routinely incorporate political sophistication in their empirical analysis in one way or another—either directly as an outcome variable of interest, as a major explanatory factor, or as an important confounder to control for. In order to

measure the underlying latent trait, scholars commonly rely on short batteries of standard recall questions on basic facts about the political system. One canonical article proposing such a battery—[Carpini and Keeter \(1993\)](#)—has been cited more than 1000 times between its publication and the writing of this manuscript. Figure 1 shows the yearly citation count over time; the trend illustrates how political knowledge remains a concept of intense scholarly interest—and that it is frequently measured using standard off-the-shelf recall questions.

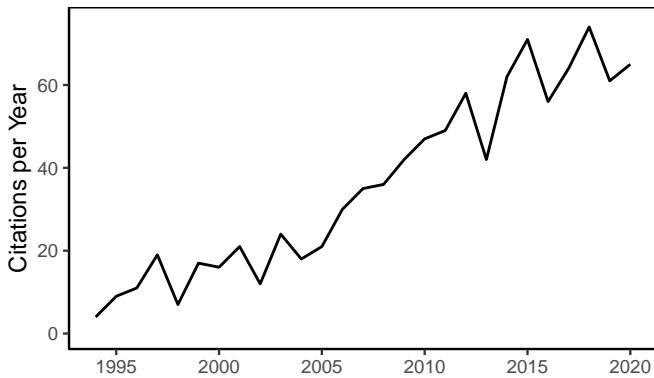


Figure 1: Yearly citation count of [Carpini and Keeter \(1993\)](#) based on Google Scholar.

As frequent as there are studies relying on factual knowledge measures are researchers' finding that people know too little.

[Delli Carpini and Keeter \(1996\)](#) show that large parts of the American electorate are not sufficiently informed about politics and that there are systematic differences in political attitudes and behavior between citizens who are well informed compared to those who are not. These findings are problematic from a normative perspective, since they indicate that differences in political information can result in unequal representation in the political system (see also [Althaus, 1998](#); [Kuklinski et al., 2000](#); [Gilens, 2001](#)). However, rather than relying on the degree to which individuals possess constrained belief systems, [Delli Carpini and Keeter \(1996\)](#) conceptualized knowledge as the awareness of key democratic values, using factual knowledge questions (see also [Carpini and Keeter, 1993](#)). Many studies focus on similar factual knowledge measures (e.g., [Zaller, 1991](#); [Gomez and](#)

Wilson, 2001). For example, Zaller (1992) measures political awareness using tests of neutral factual information about politics, since they “more directly than any of the alternative measures, capture what has actually gotten into peoples minds” (Zaller, 1992, 21). However, other research casts doubt on this assertion, both from methodological as well as theoretical perspectives.

From a methodological perspective, many studies raised issues related to the validity of factual knowledge questions as a measure of political sophistication. One problem are potential biases due to guessing (Mondak, 2000; Mondak and Davis, 2001; Miller and Orr, 2008). Responses to knowledge items that offer a “Don’t Know” option are simultaneously influenced individual information levels as well as the propensity to guess. Other scholars further criticized open-ended factual knowledge questions due to problematic coding rules, which do not capture partial knowledge (Krosnick et al., 2008; Gibson and Caldeira, 2009; DeBell, 2013).

Focusing solely on factual political knowledge has also been criticized on theoretical grounds. Lupia (2006) argues that the information asked for has no clear relevance to political participation. Instead, researchers should concentrate on heuristics that directly help citizens to make competent political decisions or focus only on knowledge relevant to a specific task (see also Lupia, 1994). Accordingly, there is no need for individuals to know all available facts, but only to possess the skills and resources to be able to find the information required in a specific context (Prior and Lupia, 2008). Furthermore, conventional items differ with regard to the dimension of political knowledge they measure (Barabas et al., 2014) and ignore important aspects such as visual cues (Prior, 2014).

## Please Mind the Gender Gap

Another common finding in public opinion research is that women have lower levels of observed political knowledge than men. For example, Verba, Burns, and Schlozman

(1997) report that women score lower on political information, interest, and efficacy, which decreases their respective levels of political participation. Since gender differences in political information and interest can only partly be explained by resource-related factors such as individual levels of education, the authors diagnose a “genuine difference in the taste for politics” between women and men, which they suspect is driven largely by socialization (see also Wolak and McDevitt, 2011). Indeed, Dow (2009, 117) describes the systematic gender differences in knowledge as “one of the most robust findings in the study of political behavior.” Other scholars, however, argue that the disparities between women and men may simply be an artifact of the measurement approach.

The discussion revolving around the apparent gender gap is therefore closely intertwined with the methodological debate about measuring political knowledge. For instance, Mondak and Anderson (2004) suggest that women are more likely to report that they do not know the answer to a recall question whereas men are more inclined to guess. Correcting for the systematic differences in the propensity to guess, however, mitigates the gender gap in knowledge but does not eliminate it completely (see also Lizotte and Sidman, 2009). Other aspects of the survey context have been shown to affect gender differences in political knowledge as well. McGlone, Aronson, and Kobrynowicz (2006), for example, present evidence that the gender gap is exacerbated in an environment that induces stereotype threat, such as if women are aware of the fact that the study focuses on gender differences or if they are interviewed by a male interviewer. However, gender differences are not only induced by *how* researchers ask their questions, but also by the question *content*, since focusing on gender-relevant political knowledge items such as information about women’s representation in the federal government has been shown to close the gap (Graber, 2001; Dolan, 2011; Fraile, 2014; Jerit and Barabas, 2017). Similarly, Stolle and Gidengil (2010) report that the gender gap disappears when people are asked about more practical issues related to the government (e.g., benefits and services).

Overall, the gender gap has been shown to be influenced by how we ask for political

information in surveys, as well as the kind of knowledge that is required for a correct response. Indeed, a comprehensive cross-national analysis of election studies in 47 countries between 1996 and 2011 suggests that question format and content account for large portions of the variance of gender disparities in political knowledge (Fortin-Rittberger, 2016, 2020). In short, conventional knowledge measures have problematic measurement properties that may exacerbate observed gender differences.

New cites: (Bos et al., 2021)

## Back to the Roots: The Structure of Belief Systems

In his seminal study on political attitudes, Converse (1964) examines whether citizens hold constrained belief systems about politics. Belief systems are defined as “a configuration of ideas and attitudes in which the elements are bound together by some form of constraint or functional interdependence” (Converse, 1964, 207). The majority of the electorate, it turns out, does not hold structured and constrained belief systems, understand abstract ideological concepts, or hold stable issue positions. This pessimistic view regarding the competence of the U.S. electorate has been supported in multiple subsequent analyses.

Conventional factual knowledge measures are but a proxy for political sophistication. From a theoretical perspective, political sophistication characterizes a particular *structure* of an individual’s system of beliefs about politics. For instance, Converse (1964) emphasizes the importance of the level of conceptualization as the main characteristic of sophistication rather than isolated pieces of factual information. Similarly, Tetlock (1983, 1993) uses the term *integrative complexity* to describe the degree to which considerations related to an issue are interconnected. Following, Luskin (1987)

First, let us examine response variation across survey items. Individuals who engage in elaborate processing should hold opinions about each political actor or policy they are

asked to discuss. Given a set of multiple open-ended probes focusing on different targets of evaluation, sophisticated respondents should be able to express their attitudes towards each question more or less equally.

**Proposition 1** (Size). *Political sophistication is characterized by a belief system that encompasses a large number of considerations.*

Next, the level of elaboration should also be related to the content and topics raised by each individual. Respondents who possess a large and wide-ranging belief system should be able to discuss their views from multiple perspectives rather than focusing only on a single consideration.

**Proposition 2** (Range). *Political sophistication is characterized by a belief system that includes a wide ranges of categories.*

Lastly, we can examine variation in the way each consideration itself is being talked about, particularly focusing on individual differences in word choice. Sophisticated respondents should use terms that are highly descriptive of a given topic—for example by mentioning interest rates or unemployment when talking about economic policies.

**Proposition 3** (Constraint). *Sophisticated respondents are aware of key concepts related to an issue and therefore use terms that are clearly associated with a given topic.*

also defines political sophistication based on the structure of individual belief systems, arguing that they can vary on three separate dimensions: (1) their *size* – i.e. the number of cognitions, (2) their *range* – i.e. the dispersion of cognition over categories, and (3) their *constraint* – i.e. the extent to which cognitions are interconnected in a meaningful way. Political sophistication, in turn, is seen as the conjunction of these dimensions: “A person is politically sophisticated to the extent to which his or her [political belief system] is large, wide-ranging, and highly constrained” ([Luskin, 1987](#), 860). Sophisticated political reasoning should reflect this notion of complex belief systems.<sup>1</sup>

---

<sup>1</sup>It should be no surprise that Converse and others examined open-ended responses in their early

# Measuring Discursive Sophistication

Rather than trying to develop a new item battery that addresses some of these issues, I propose an alternative approach. Instead of testing respondents on a specific set of predetermined facts, we can make inferences about political sophistication by analyzing how they discuss their attitudes and beliefs in their own words.

In the following, I propose a framework that leverages the content of open-ended responses in conjunction with the survey structure to evaluate the underlying reasoning that gives rise to individual statements. Consider for example a questionnaire where respondents are prompted to describe their attitudes toward major presidential candidates and their parties in multiple open-ended items. Each question asks for either positive or negative considerations related to one of the parties or candidates. In developing a measure of political sophistication, we are ultimately interested in people's level of elaboration when justifying their preferences in such a set of open-ended responses. So how would a politically sophisticated person who engages in in-depth processing discuss her views compared to a less informed individual?

Consider again the example of multiple open-ended questions regarding candidate and party preferences in a survey. In such a scenario, we can characterize elaborate reasoning by leveraging variation in response patterns on three levels—across questions, between topics mentioned, and within each topic. For each dimension, I make the following assumptions regarding the way sophistication manifests itself in verbatim responses.

In order to measure discursive sophistication according to the theoretical framework outlined above, we need to quantify individual-level variation in open-ended responses on each of the three dimensions. I begin by describing the measurement approach in more detail and explore a first set of results based on data from the 2012 and 2016 ANES. The same procedures are applied to the remaining datasets and additional information

---

studies—albeit from a slightly different perspective than the approach outlined here. Importantly, instead of relying on manual coding of open-ended responses, I develop an automated framework that is easily reproducible and can directly be applied to large surveys.

for these is provided in the appendix.

Sophisticated respondents evaluate political actors and issues on multiple dimensions and are therefore able to recall a large number of distinct considerations in response to each query presented in a survey.

The first component targets the *size* of the belief system by measuring the number of considerations discussed by each individual. I rely on the structural topic model framework to extract the number of topics mentioned by each respondent in a survey (Roberts et al., 2014).<sup>2</sup> I denote  $\mathcal{W}_i$  as the set of words contained in a response of individual  $i$ . Each word  $w \in \mathcal{W}_i$  is assigned to a topic  $t^* \in \{1, \dots, T\}$ , such that  $P(t^*|w, X_i) > P(t|w, X_i) \forall t \neq t^*$ .<sup>3</sup> In other words, each unique term in a response is assigned to the topic that has the highest likelihood of having generated that term, given the model. The set of topics that are mentioned by respondent  $i$  across all words in  $\mathcal{W}_i$  can then be described as  $\mathcal{T}_i^*$  and the number of considerations can be written as:

$$\text{size}_i = \frac{|\mathcal{T}_i^*|}{\max |\mathcal{T}_i^*|}. \quad (1)$$

I re-scale the measure to range from zero to one by dividing raw count of topics by the maximum number of topics observed across individuals.

Second, regarding response patterns across questions, high levels of elaboration have been argued to display a greater *range*, which is based on the degree to which individuals are able to respond to each query presented in a survey more or less equally. I quantify this diversity in relative lengths for each open-ended response as the Shannon entropy

---

<sup>2</sup>Please refer to the appendix for additional information. Specifically, see Appendix B for descriptive information on open-ended responses in each dataset and structural topic model results. Appendix C contains further details on pre-processing steps and modeling choices for the structural topic models as well as robustness checks, which include preText analyses proposed by Denny and Spirling (2018).

<sup>3</sup>Note that  $P(t|w, X_i) = \frac{P(w|t)P(t|X_i)}{P(w|X_i)}$ . In the context of structural topic models,  $X_i$  denotes the covariates used to predict individual topic prevalence (see Roberts et al., 2014, for details). I used measures for age, gender, education, party identification, as well as an interaction between education and party identification as covariates for topic prevalence. This variable selection—with the exception of including gender—is equivalent to the procedure described in Roberts et al. (2014).

across survey items:

$$\text{range}_i = \frac{-\sum_{j=1}^J p_{ij} \ln p_{ij}}{\ln J} \quad (2)$$

where  $p_{ij}$  is the proportion of words in the response of individual  $i$  to question  $j \in \{1, \dots, J\}$  relative to the overall size of the individuals' response. The variable ranges from 0 (only one question was answered) to 1 (all questions were answered with the same word length per answer).

The last component of discursive sophistication addresses the way individuals talk about each consideration. More specifically, I argue that when discussing a specific topic, sophisticated respondents should use words that are clearly associated with it since they are aware of relevant key concepts. We can again rely on the structural topic model result to summarize the extent to which *word choice* consists of highly descriptive terms for each topic, which implies that a specific term has a high likelihood to be mentioned if a given topic is discussed. Thus, highly descriptive word choice is conceptualized as the sum of term likelihoods  $P(w|t^*)$  given topic assignments over the entire set of words in  $\mathcal{W}_i$ :

$$\text{constraint}_i = \frac{\text{CONJ}_i + \text{EXCL}_i}{\max [\text{CONJ}_i + \text{EXCL}_i]} \quad (3)$$

As before, I re-scale the measure to range from zero to one by dividing all values by the empirical maximum observed across all individuals in the data.

In a questionnaire where respondents are asked to discuss their political attitudes in multiple open-ended items, I therefore define discursive sophistication as the conjunction of the three dimensions described here. Overall, a highly sophisticated individual can be expected to give a more elaborate response to every question that focuses on multiple considerations using terms that are highly descriptive of each topic. Note that this simple framework makes no strong assumptions about the type of information necessary to make competent decisions in a given context; neither does it make assumptions about whether a given argument is correct or not. Crucially, since it is solely based on

*how* individuals discuss their preferences, it can be directly applied in various settings to target specific political tasks such as choosing between candidates, parties, or policy propositions. Rather than having to devise a new set of questions that attempt to capture information necessary to make competent decisions, we can simply analyze how respondents describe and justify their political preferences in verbatim.

Of course, this is not the first time a framework is developed to assess the complexity of written (or spoken) word. In fact, this task has been the subject of longstanding research in linguistics and educational sciences, resulting in a multitude of alternative metrics. Recently, these measures have been employed by political scientists who study different forms of elite communication. [Spirling \(2016\)](#), for example, uses a standard readability score based on the ratio of words per sentence and syllables per word to study the linguistic complexity of speeches in the British House of Commons over time. More recently, [Benoit, Munger, and Spirling \(2019\)](#) expanded on previous metrics to develop a measure of comprehensibility that is more applicable in the realm of politics.

These approaches—and especially the development of metrics specifically suited for political text—are particularly useful when studying elite communication. Yet, in contrast to the framework outlined above, they focus on the *comprehensibility* as a measure of complexity; sophistication is evaluated based on a recipient's ease to understand the message, which is largely driven by linguistic and syntactic difficulty rather than actual political content. Again, while this is certainly a reasonable approach when studying the effects of elite communication, the inference of interest outlined in this paper is markedly different. My focus is to examine verbatim attitude expression to assess the underlying degree of elaborate political reasoning. Pure linguistic style is therefore not of central concern so long as it is unrelated to the actual political content.<sup>4</sup> After all, being hard to comprehend does not necessarily imply that someone put a lot of thought into a state-

---

<sup>4</sup>In fact, one might be concerned that pure linguistic complexity is ultimately driven by other factors such as a person's general verbosity or linguistic prowess and therefore not a valid measure of political sophistication.

ment.

## A First Look at Discursive Sophistication

Applying this approach to both waves of the ANES shows that all three components are highly correlated.<sup>5</sup> Furthermore, an exploratory factor analysis confirms that all three components load on a single factor with all loadings exceeding 0.5. Detailed results are presented in Table 1.

Variable	2018 CES	2020 ANES	2016 ANES	2012 ANES
Size	0.840	0.997	0.997	0.997
Range	0.526	0.536	0.548	0.576
Constraint	0.513	0.684	0.623	0.709

Table 1: Factor Loadings of Discursive Sophistication Components

Together, the three measures can therefore be combined in a composite metric of sophistication in political attitude expression by averaging them for each respondent. Like each individual component, the resulting *discursive sophistication* score ranges from 0 to 1:

$$\text{discursive sophistication}_i = \frac{1}{3}(\text{opinionation}_i + \text{considerations}_i + \text{word choice}_i). \quad (4)$$

We can now compare discursive sophistication to alternative metrics of political knowledge. As discussed in the beginning, the standard approach to measuring political knowledge in surveys is to ask a set of factual questions about political institutions. The ANES surveys include such a basic item battery, inquiring for example about the number of times an individual can be elected President of the United States, or how the current U.S. federal budget deficit compares to the deficit in the 1990s. I combine responses on these items to form an additive index of *factual knowledge* about politics. As an additional benchmark, I consider *interviewer assessments* of each respondent's political sophistication

---

<sup>5</sup>See Appendix B for correlation matrices between individual components for the ANES as well as the remaining two datasets.

(see Bartels 2005 for an example of a study that relies on interviewer assessments; but cf. Ryan 2011).<sup>6</sup>

---

<sup>6</sup>Interviewer assessments were only recorded in the face-to-face sample of the ANES.

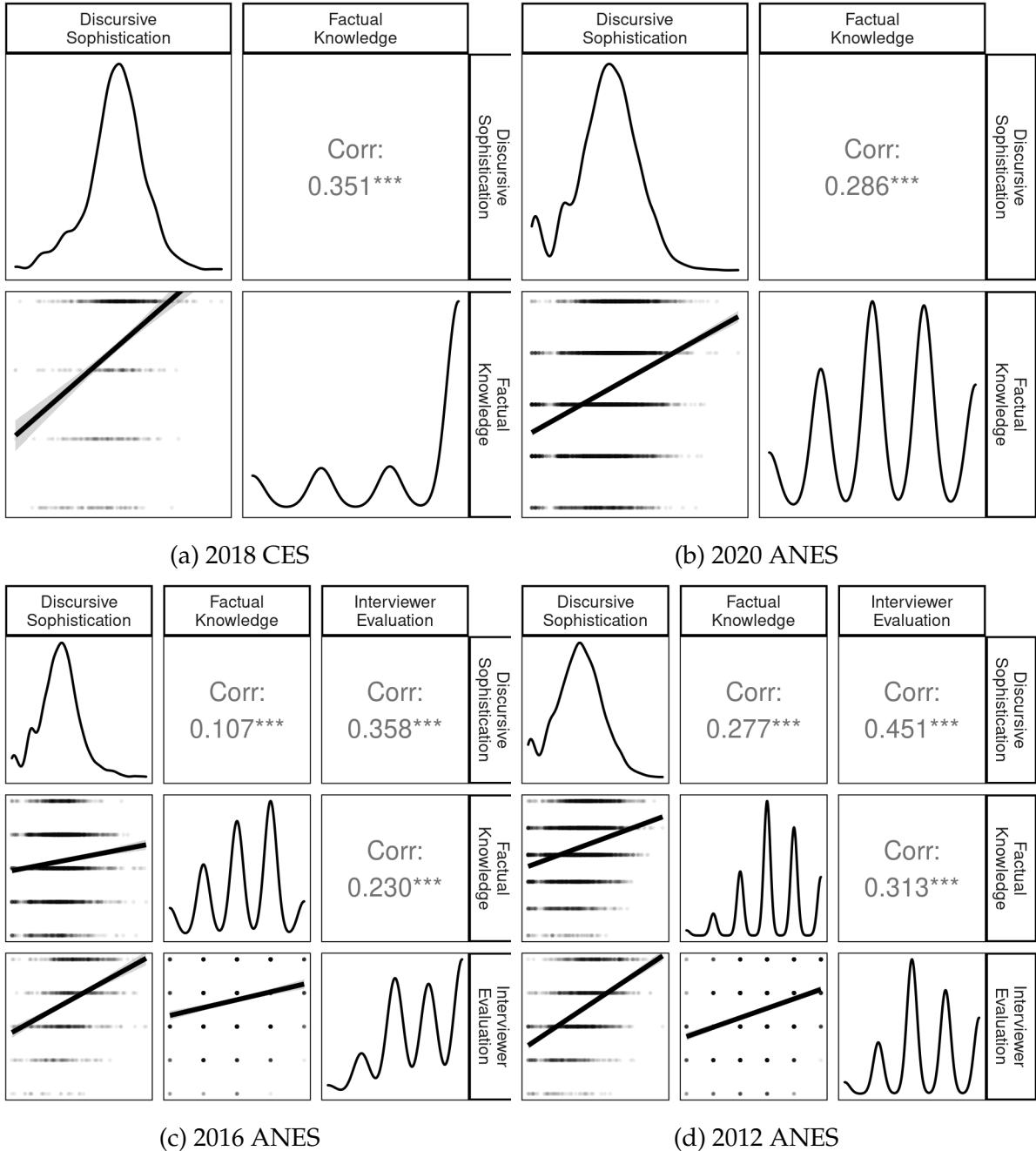


Figure 2: Correlation matrix of discursive sophistication and conventional political knowledge metrics. The plots on the diagonal display univariate densities for each variable. The panels in the lower triangular display the scatter plot of two measures as well as a linear fit. The upper triangular displays the correlation coefficient. All correlations reported are statistically significant with  $p < .05$ .

Figure 2 compares discursive sophistication to the conventional knowledge metrics for both waves of the ANES. Each figure presents scatterplots between individual measures (lower triangular), univariate densities (diagonal), and correlation coefficients (upper triangular). The measure of discursive sophistication is positively correlated with both conventional metrics while capturing some additional variation. Interestingly, there is a stronger correlation between discursive sophistication and interviewer evaluations than between factual knowledge and interviewer evaluations ( $r = .45$  vs.  $r = .31$  in 2012, and  $r = .36$  vs.  $r = .23$  in 2016), which indicates that the open-ended measure captures characteristics that influence subjective assessments of sophistication. Interviewers certainly form their impressions throughout the entire survey, but a respondent's verbatim answers seem to be more influential for subsequent knowledge assessments than a respondent's performance on the factual knowledge questions.

Overall, while discursive sophistication and the alternative measures are clearly correlated, the relationship between each metric is far from perfect. To provide some intuition as to whether the variation in discursive sophistication is theoretically meaningful, I present an example of open-ended responses of two individuals in the 2016 ANES who identified as Republicans and *scored equally on the factual knowledge score* (3 out of 4 correct responses), but varied highly in discursive sophistication. The results are presented in Table 2.

Each row in the table represents one of the open-ended responses (like/dislike for each candidate/party). Column A displays the responses of an individual who scored low on discursive sophistication and column B displays the responses of a high scoring individual. Cells are empty if a respondent refused to provide a response. Even though both individuals have the same factual knowledge score, there are systematic differences in their response behavior that can be attributed to their political sophistication. Overall, respondent A provided a less elaborate response, only focused on a narrow range of issues, and only reported attitudes on two items. Irrespective of whether one agrees

	A: Low Sophistication Response	B: High Sophistication Response
Clinton (+)		Politician.
Clinton (-)	The fact that she has links to Al-Qaeda.	Caught in lies.
Trump (+)		Says what he thinks.
Trump (-)	He is going to start a civil war. I feel like he is racist.	Reality TV star, poor businessman
Democrats (+)		Middle class minded.
Democrats (-)		Too many handouts.
Republicans (+)		Economic growth conscious.
Republicans (-)		For the big business.
Disc. Soph.	0.162	0.461

Table 2: Example of open-ended responses for low and high scores on discursive sophistication with equal factual knowledge scores (3 out of 4 correct responses). Column A displays the verbatim responses of an individual who scored low on discursive sophistication and column B displays the verbatim responses of an individual who scored high on the open-ended measure. Each row represents one of the likes/dislikes items included in the analysis. Note that the responses in this table were slightly edited for readability (spelling errors removed, etc.).

with the specific statements or whether they are factually accurate (e.g., Clinton’s connection to Al-Qaeda), A’s response pattern is suggestive of a less sophisticated political belief system and a lower level of motivation to engage in in-depth reasoning about both parties and candidates. Overall, this initial result suggests that the variation in discursive sophistication captures meaningful differences in response behavior that overlaps with traditional knowledge metrics while displaying some unique variation. The following sections will show that this variation is also politically consequential.

## Validating the Measure

A crucial step in validating any measure of political sophistication is to examine the extent to which it is positively related to citizen competence in modern democracies (Lupia, 2006, 2015). Accordingly, I consider the potential role of discursive sophistication in promoting (1) engagement and participation in politics, (2) the ability to incorporate new information, and (3) well-justified policy preferences. Each point is addressed using one

of the three data sets described above.

## Engagement and Participation in Politics

Political sophistication is commonly expected to be strongly associated with individual engagement and participation in politics. In fact, factual knowledge items have been validated in the past based on their strong relationship with outcomes such as turnout and other forms of participation (Lupia, 2015, 230–233). Figure 3 compares the effects of discursive sophistication and factual knowledge in the 2012 and 2016 ANES on four dependent variables related to political engagement: turnout, non-conventional participation, internal efficacy, and external efficacy. The model predicting turnout is estimated via logistic regression while the estimates for the three remaining dependent variables are based on OLS. Each model equation includes both sophistication measures while controlling for gender, education, income, age, race, church attendance, survey mode (face-to-face vs. online), as well as the Wordsum vocabulary score measuring verbal intelligence.<sup>7</sup>

Each panel displays the expected difference in the respective dependent variable for a two standard deviation increase in each sophistication measure, while holding all other variables constant at their means. Overall, discursive sophistication is a stronger predictor of turnout, non-conventional participation, as well as (to a lesser extent) internal and external efficacy. In the 2012 ANES, the positive effect of factual knowledge on participation is statistically indistinguishable from zero when controlling for discursive sophistication. Furthermore, there is a negative effect of factual knowledge on external efficacy in the 2016 ANES. In contrast, the positive effect of discursive sophistication on external efficacy is more consistent with previous research. Considering these initial results, a potential concern may be that discursive sophistication is confounded by personality characteristics that influence verbatim response patterns as well as engagement. Appendix E

---

<sup>7</sup>Appendix D provides additional information on these as well as remaining variables included in subsequent analyses.

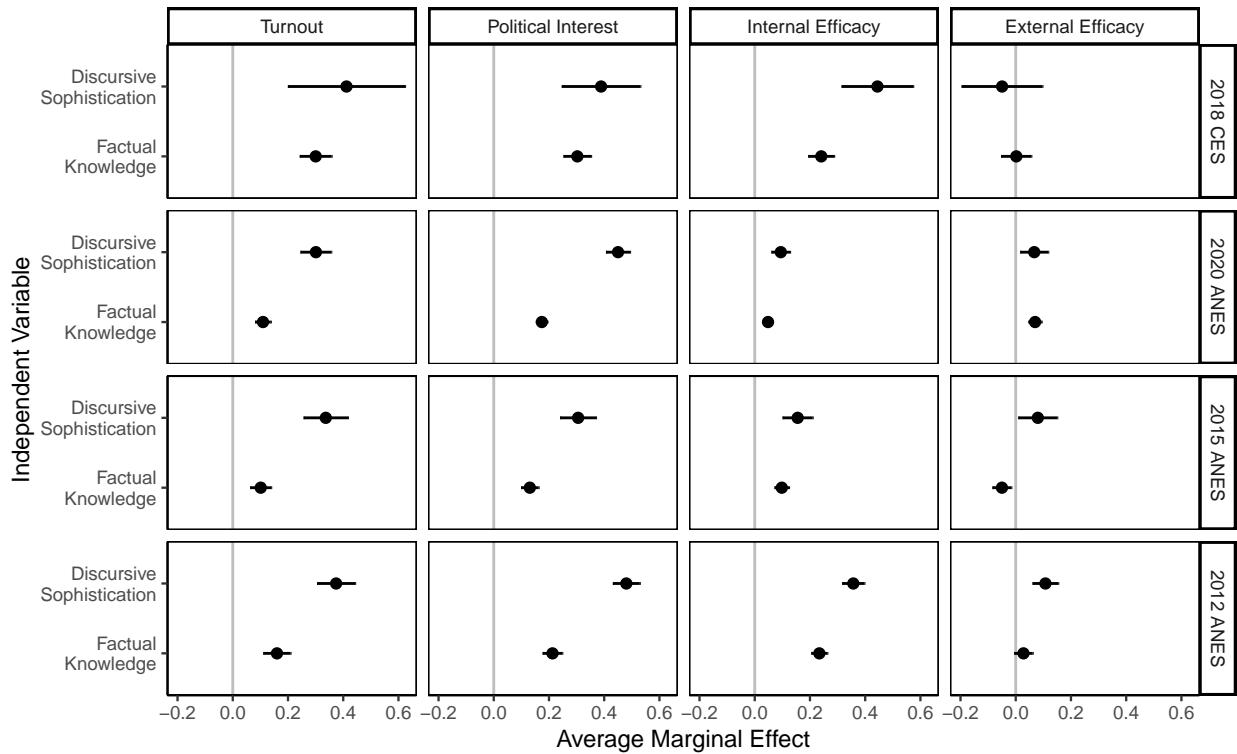


Figure 3: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy in the 2012 and 2016 ANES. For each dependent variable, the figure displays the change in expected values after increasing each sophistication measure from  $-1$  to  $+1$  standard deviation from its mean (including 95% confidence intervals). Model estimates are based on logistic regression (turnout) or OLS (non-conventional participation, internal efficacy, external efficacy). Both sophistication measures are included simultaneously while controlling for gender, education, income, age, race, church attendance, survey mode, and Wordsum vocabulary scores. Full model results are displayed in the appendix, Tables E.1 and E.2.

provides additional analyses controlling for such factors that might drive verbosity (extraversion and being reserved) as well as individual response length itself. The substantive conclusions remain unchanged.

## Incorporation of New Information

Competent citizens should not only engage in politics but are also expected to be sufficiently informed about the issues of the day. As such, they have to be attentive to their media environments and incorporate potentially relevant new information about parties,

office-holders, and policies. Indeed, Zaller (1990, 1992) and others argue that tests of factual information about politics are the best available proxy for awareness. In this analysis I draw on the 2015 YouGov study to explore whether discursive sophistication or factual knowledge serves as a better predictor of people's ability to incorporate new information from media sources. As part of the survey, respondents were asked to read a newspaper article about a fictional infectious disease and were subsequently asked to answer questions about information provided in the article (e.g. regarding symptoms, modes of contraction etc.). I compute an additive index counting the pieces of information that were correctly recalled (*information retrieval*) as a measure of the ability to retrieve information from a news article on a non-partisan issue that is related to public health policies.

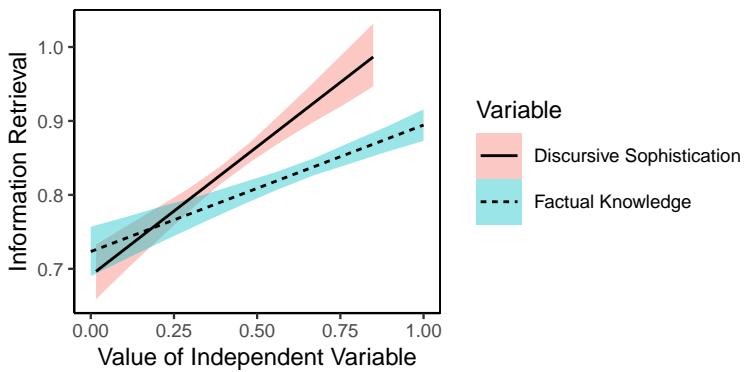


Figure 4: Expected information retrieval in the 2015 YouGov Study as a function of political sophistication (including 95% confidence intervals). Estimates are based on a linear regression model controlling for education, income, age, church attendance, gender, and race. Full model results are displayed in the appendix, Table E.3.

Figure 4 displays the relationship between political sophistication and disease information retrieval in the 2015 YouGov study. Estimates are based on a linear regression model controlling for education, income, age, church attendance, gender, and race. As a benchmark for discursive sophistication, I again consider the effect of factual knowledge based on a battery of eight items similar to the knowledge questions in the ANES. Both discursive sophistication as well as factual knowledge increase the amount of information individuals are able to recall from a news article discussing a fictional disease.

Similar to the previous results, the effects are stronger for discursive sophistication than for factual knowledge scores. The degree to which citizens discuss their own political beliefs in a more elaborate manner is not only a stronger predictor of political engagement but also serves as a better proxy for the ability to incorporate new information about a non-partisan issue.

## Well-Justified Policy Preferences

Beyond the ability to incorporate new information, competent citizens should be knowledgeable about the underlying policies themselves and be able to justify their own preferences. Here, I explore the extent to which high levels of discursive sophistication correspond to well-justified policy preferences in open-ended responses. As mentioned above, the Swiss surveys included items that asked respondents to explain why they voted in favor or against a given proposition in multiple policy referenda. To corroborate the face validity of discursive sophistication, I examine whether the measure is related to Colombo's (2016) manual coding of the respondents' *level of justification*, which assessed the content, elaboration, and complexity of open-ended responses.

The results are presented in Figure 5, which displays the distribution of discursive sophistication for each level of justification coded by Colombo (2016) as well as the correlation coefficients for both respective variables. Across all three language groups, discursive sophistication is systematically higher among respondents with the highest level of justification and both measures are positively correlated ( $r = 0.26, 0.30$ , and  $0.35$ , respectively). The proposed measure of discursive sophistication therefore shows a high degree of correspondence with individual levels of justification assessed by independent manual coders.

To summarize, the results presented thus far indicate that discursive sophistication shares common characteristics with factual political knowledge measures. Compared to conventional metrics, the proposed measure performs at least as well as a predictor

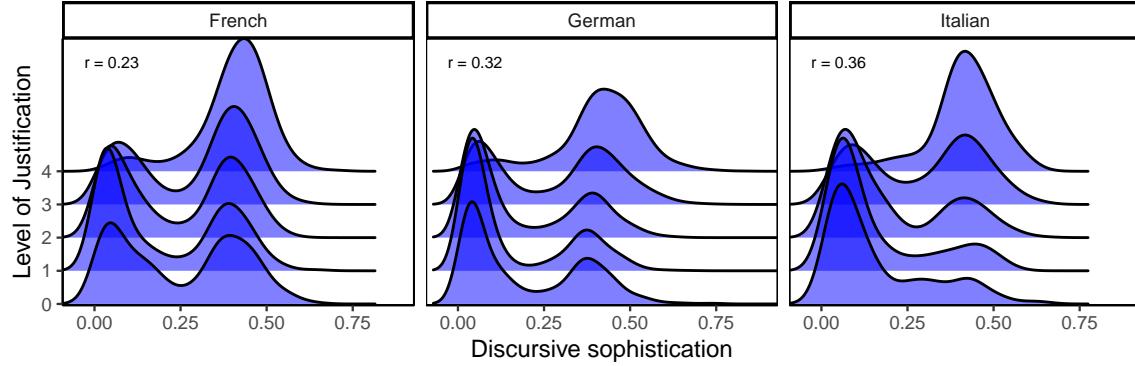


Figure 5: Discursive sophistication and manually coded level of justification (Colombo, 2016) in Swiss post-referendum surveys. The plot compares kernel densities of discursive sophistication for each manually coded level of justification.

of essential competences that allow citizens to engage successfully in politics. In fact, discursive sophistication is a stronger predictor of certain outcomes (such as different forms of political participation) than conventional knowledge scores. In the following, I turn to an application to illustrate how discursive sophistication can help refine important previous insights from the literature on political knowledge.

## Reassessing the Gender Gap

How do women and men compare on the different metrics of political sophistication in the surveys analyzed in the present study? The top row of Figure 6 displays the average levels of discursive sophistication as well as conventional metrics comparing both genders. While we observe a sizable and statistically significant gender gap for factual knowledge in both ANES surveys, this difference disappears for discursive sophistication. These results are replicated in the 2015 YouGov survey. Even though women do not perform as well as men on political quizzes, they do not differ substantially in complexity and sophistication when describing their political preferences.

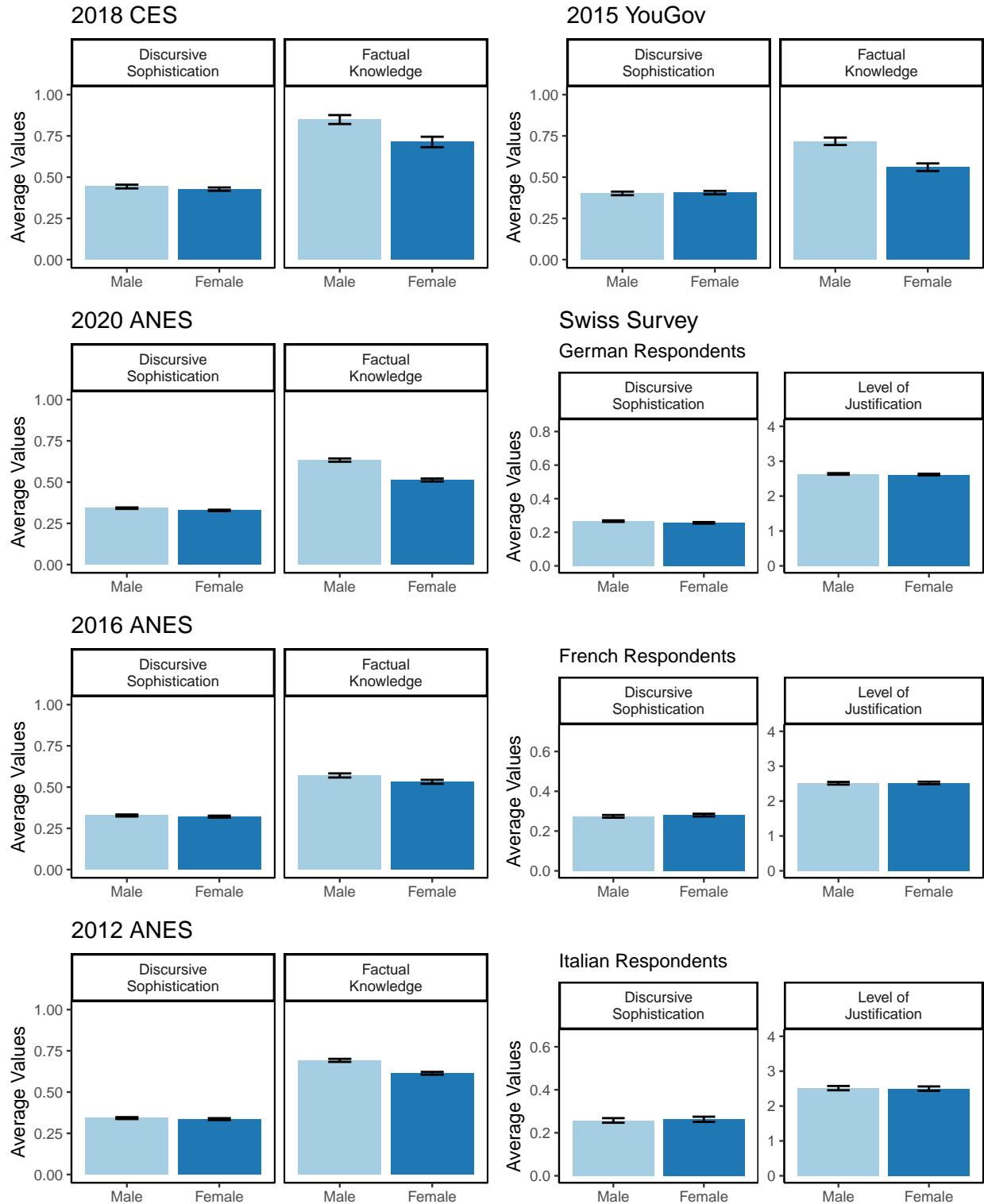


Figure 6: The gender gap in political sophistication. The figures display mean levels of sophistication for each measure comparing women and men (including 95% confidence intervals). Gender differences in factual knowledge in the 2012/2016 ANES and 2015 YouGov survey (top row) are statistically significant with  $p < .05$ . Gender differences in discursive sophistication and manually coded levels of justification (Colombo, 2016) are not statistically significant.

Of course, we need to make sure that this absence of a gender gap in discursive sophistication is not idiosyncratic to the particular measurement approach proposed here. One way to investigate this question is to examine gender differences in discursive sophistication using data from Colombo (2016) and comparing them to her manually coded measure. That way, we can not only determine whether the lack of a gender gap in discursive sophistication replicates in the third survey, but also check whether there is an equivalent lack of gender differences in Colombo's alternative measure of citizen competence in direct democracies. If discursive sophistication captures a person's motivation to undertake in-depth reasoning and form quality opinions (and assuming these characteristics do not differ by gender), there should be no difference between women and men on either metric (discursive sophistication and Colombo's measure). As shown in the bottom row of Figure 6 there are indeed no significant gender differences on *both* metrics across all three languages in the Swiss referendum surveys. The absence of a gender gap is consistent whether open-ended responses are coded manually or using the proposed measure of discursive sophistication.

Next, we have to consider whether the apparent gender gap in factual knowledge is a manifestation of real differences between women and men. Prior research attributes at least part of the gap to actual discrepancies in individual resources and engagement. Accordingly, we need to control for these determinants of political knowledge to provide a more comprehensive examination of the veracity of observed gender differences. Figure 7 displays estimated effects of various potential common determinants of factual knowledge and discursive sophistication on both measures. Previous studies consistently showed that political information levels are positively related to high media exposure, frequent political discussions, education, and income. Furthermore, I include age, race, church attendance, and survey mode (face-to-face vs. online) as additional control variables.

After controlling for common determinants, discursive sophistication again reveals

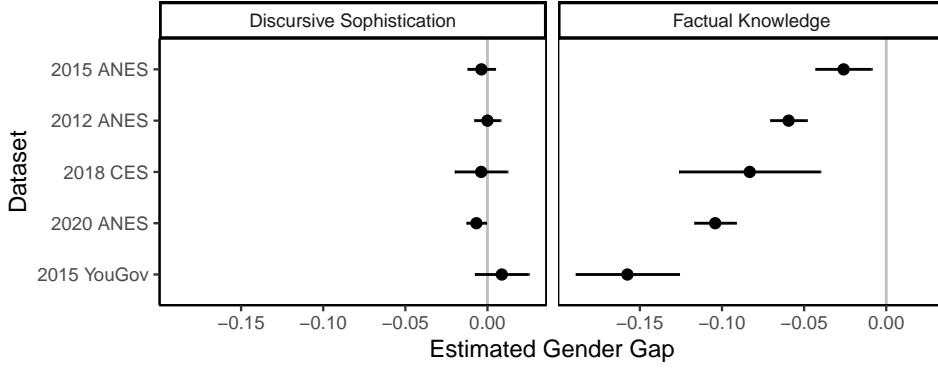


Figure 7: Common determinants of political sophistication. Estimates are OLS regression coefficients with 95% confidence intervals. Dependent variables are discursive sophistication as well as factual political knowledge. Full model results are displayed in the appendix, Tables E.4 and E.5.

no significant differences between women and men in both ANES surveys as well as the 2015 YouGov study. The gender gap in factual political knowledge, however, persists and is substantively as well as statistically significant. Thus, a considerable portion of the observed differences in factual knowledge between women and men cannot be attributed to underlying disparities in resource-related factors or engagement. It is worth pointing out in this context, that the effects of the control variables are quite similar across both measures and different surveys. Knowledge and discursive sophistication are significantly higher among respondents who are more exposed to political news media, discuss politics frequently, are more educated, and have higher income.<sup>8</sup> Overall, the finding that determinants of political sophistication are consistent across models lends additional validity to the open-ended measure.

One could argue (see R3) that this lack of gender differences is due to higher measurement error in discursive sophistication. This is unlikely to be the case since other common predictors of discursive sophistication have a similar size / magnitude as the

<sup>8</sup>An interesting deviation, however, is the effect of survey mode in the 2012 and 2016 ANES. Respondents in online surveys score significantly higher on factual knowledge than in face-to-face interviews. This difference can be attributed to the fact that individuals are able to look up answers for factual knowledge questions while taking an online survey (cf. Clifford and Jerit, 2016). For discursive sophistication, on the other hand, individuals perform better in the face-to-face survey. Open-ended answers in online surveys may be less elaborate because respondents have to manually type their responses.

factual knowledge model. See table in Appendix for more information.

## Explaining the (Lack of a) Gender Gap

In summary, the analyses only reveal a significant gender gap based on conventional recall-based measures of political knowledge, a result that previous research attributed—at least partly—to the format (e.g., availability of “Don’t Know” options) and content (e.g., focusing on issues that are less relevant to women) of the question batteries. This section explores whether these arguments provide a sufficient explanation for the conflicting results for discursive sophistication—namely the complete lack of systematic differences between women and men. In other words, is it the *existence* of a gender gap in factual knowledge or the *absence* of a gap in discursive sophistication, that is more likely to be an artifact of the respective measurement approach?

The first set of arguments about why conventional metrics may overstate potential gender differences is based on the finding that women are less likely to guess than men ([Mondak and Anderson, 2004](#)). Arguably, respondents’ differential willingness to admit not knowing the answer to a question is certainly less of an issue when they are simply asked to voice their opinions rather than being quizzed on political facts. Following best practices, however, both waves of the ANES as well as the YouGov study omitted “Don’t Know” options in their recall questions. Differential propensity to guess can therefore not be viewed as a valid explanation for the gender gap in factual knowledge observed here. At the same time, the lack of significant differences between women and men in discursive sophistication may itself be the product of selection biases in women’s willingness to answer open-ended question in the first place. Following this argument, it could be the case that only women who are highly sophisticated provide a response, thereby misleadingly closing the gender gap in the discursive measure. There are two reasons why that is unlikely to be the case. First, as the analyses have shown, this proposed selec-

tion mechanism does not diminish gender differences in factual knowledge. Second, and more importantly, there are no significant differences between men's and women's willingness to answer open-ended questions (ANES 2012:  $z = 1.630$ ,  $p = 0.103$ ; ANES 2016:  $z = 0.464$ ,  $p = 0.643$ ). In fact, adjusting for potential selection effects when examining determinants of sophistication does not change the substantive conclusions presented in the discussion of Figure 7.

The second major explanation for the gender gap in political knowledge focuses on the question content. At the outset of this paper, I discussed how Cramer and Toff (2017), along with other scholars, criticized traditional knowledge batteries for focusing erroneously on what people do not know. By choosing a specific set of recall questions as a general metric for political knowledge, researchers are making strong assumptions about the information deemed necessary for competent decision-making. As it turns out, these item batteries usually focus on male-dominated topics in politics (Dolan, 2011). Open-ended questions, on the other hand, make it possible to directly study the information that is in fact available to citizens and—importantly—to examine how they apply their knowledge when discussing their political preferences.

Accordingly, if it is the case that the gender gap in discursive sophistication is nonexistent simply because open-ended questions allow women to raise political considerations particularly salient to them, then we should be able to observe systematic variation in types of issues discussed by women and men, respectively. Luckily, we can directly examine such gender differences in topic prevalence within the structural topic model framework used to measure discursive sophistication. More specifically, gender is included in the model as one of the covariates that influences how often each topic is discussed by a respondent (see also Roberts et al., 2014, for details).

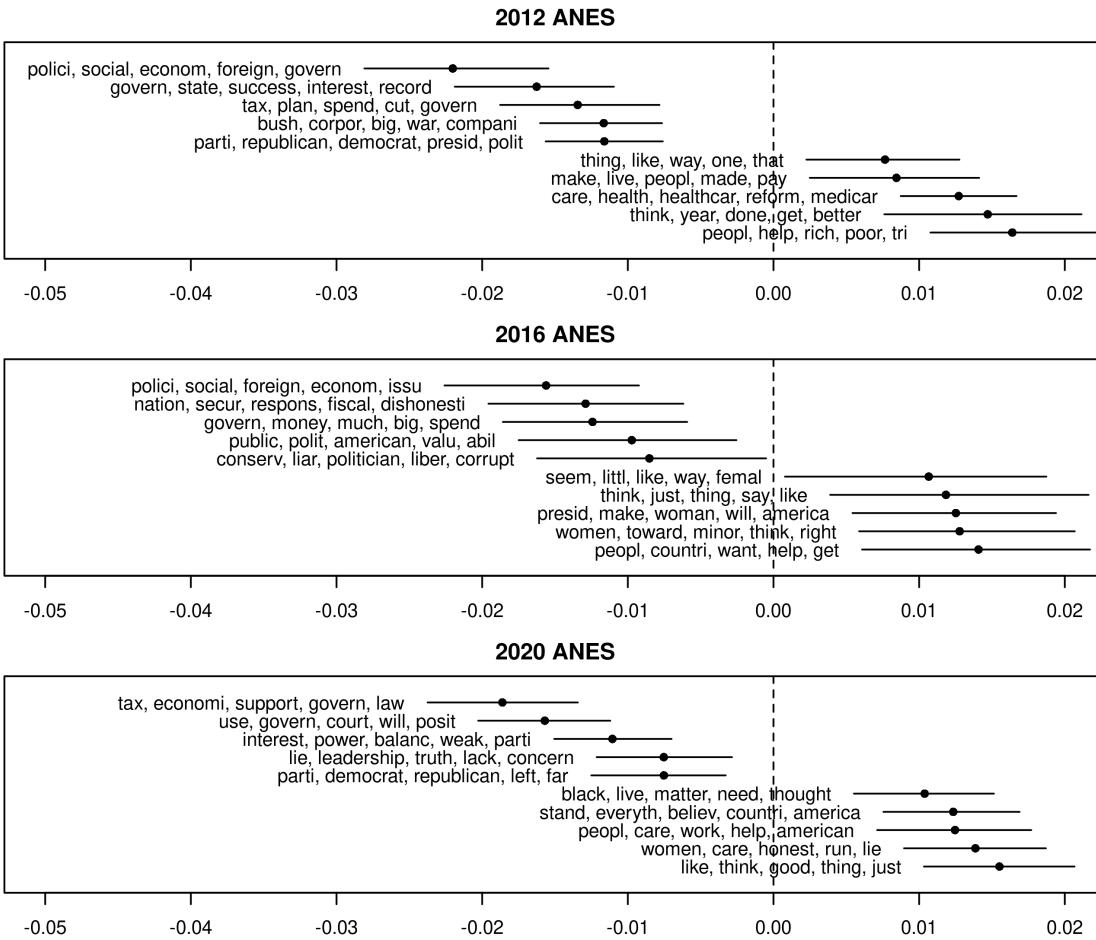


Figure 8: Gender differences in topic proportions in open-ended responses based on the structural topic model used to compute discursive sophistication (including 95% confidence intervals). Coefficients indicate the difference in predicted topic prevalence among women and men; positive values indicate higher prevalence among women. Labels are based on the five highest probability terms related to the topic.

In this last analysis, I therefore explore how women and men differ in topical prevalence across open-ended responses in the 2012 and 2016 ANES. Figure 8 displays the subset of topics that shows the largest absolute gender difference in topic prevalence in both waves. Positive coefficients indicate that women are more likely than men to mention a given topic, and vice versa. As such, the top six topics are more prevalent among men and the bottom six are more likely to be mentioned by women. The label for each coefficient consists of the five highest probability terms related to the topic in order to illustrate its content.

Taking the 2012 ANES as an example, the topic consisting of terms such as *right*, *women*, and *equal* is significantly more likely to be mentioned by women. On the other hand, men are more likely to mention the topic revolving around terms like *budget*, *balance*, and *privat*. Overall, across both waves of the ANES, women were less likely than men to discuss foreign affairs, economic issues, or the Supreme Court. Instead, they focused on issues related to women's rights, equality, and health care. The considerations raised by women when discussing their political preferences are therefore clearly different from men's and—crucially—the issues discussed by men happen to be more aligned with the type of questions usually covered in standard political knowledge batteries (i.e., pertaining to the economy, institutions, elites, etc.). For example, men are more likely to mention considerations related to the federal budget in their open-ended responses. At the same time, two of the five knowledge questions included in the 2012 ANES pertain to government spending: one asking respondents to compare the federal deficit to levels in 1990, the other requiring a comparison of federal spending on different programs such as foreign aid, medicare, and national defense.

Overall, the results indicate that gender differences in conventional knowledge metrics can at least partly be explained by the fact that the issues women care about are not represented in standard item batteries. When using the alternative measure—discursive sophistication—any evidence for systematic differences between women and men disap-

pears since open-ended questions about political preferences allow respondents to focus on different issues.

## Discussion and Conclusion

Political scientists should worry less about pure levels of *factual knowledge* and instead focus on the necessary conditions for individuals to make *competent* decisions. Competence in elections, for example, requires citizens to hold informed attitudes about their representatives. Factual knowledge about political institutions might be a useful proxy in certain scenarios, but it cannot address directly whether individuals hold well-considered opinions about political actors they try to hold accountable. In comparison, the measure of discursive sophistication proposed here is agnostic about the specific contents of people's beliefs, but directly targets the complexity of expressed attitudes related to the political task. It can therefore be easily applied to assess sophistication in any decision-making context (such as policy referenda or local elections) by fielding targeted open-ended questions related to the relevant underlying beliefs and preferences.

The findings presented in this paper show that conventional knowledge indices and the open-ended measure share a substantial amount of variance. However, they are far from being identical and capture different aspects of sophistication. In fact, discursive sophistication is a stronger predictor of political engagement and efficacy than traditional metrics. It is also strongly related to people's ability to incorporate new information from news sources and shows a high degree of overlap with manually coded levels of justification. Most importantly, using the discursive measure, any evidence for the gender gap commonly reported using factual knowledge scales disappears. Women might know fewer facts about political institutions, but they do not differ substantively in the complexity of their expressed political beliefs. The fact that women perform just as well as men on discursive sophistication across various surveys can be attributed to the fact that

they focus on different considerations when evaluating political parties and candidates.

From a normative perspective, there is no reason to assume that a particular set of issues should be more important for citizens' preference formation or competence in elections. Whether one cares more about the federal budget or reproductive rights, the most important question is whether citizens think deeply about the issues they care about and incorporate them accordingly when making their vote choices. As Druckman (2014) argues, citizen competence (for example in elections) should not be evaluated based on their ability to recall unrelated facts about political institutions, but rather focus people's motivation to form quality opinions—which implies that they focus on the issues most important to them. As it turns out, while the types of issues raised women and men differ systematically, there is no reason to assume that women are therefore less sophisticated or competent in the realm of politics.

This issue has been recognized in the literature before (e.g., Graber, 2001; Dolan, 2011), but it cannot be properly addressed while relying exclusively on off-the-shelf recall questions to measure political knowledge. What is more, there is thus far no principled approach to develop new sets of items that focus less on male-dominated issues. Beyond proposing an alternative measurement approach, the framework presented in this paper can help provide such a first step towards devising balanced recall items. More specifically, examining the types of issues women and men emphasize when discussing their political preferences can serve as a guide to select new sets of knowledge questions. Granted, more research is necessary to explore whether factual knowledge questions selected based on open-ended responses are indeed more balanced with regard to gender differences. To the extent that this indeed proves to be a useful heuristic for item selection, researchers planning a survey could rely on pilot studies fielding open-ended questions in order to devise balanced factual knowledge items in the main survey.

That being said, relying on open-ended responses to assess political sophistication also comes with potential drawbacks. First and foremost, elaboration in verbatim atti-

tude expression may be more prone to biases due to differential levels of motivation to answer survey questions. It should be noted, however, that conventional knowledge metrics are not free from survey effort effects either—as indicated for example by the fact that scores can be improved by providing monetary incentives for correct responses (Prior and Lupia, 2008)—and future studies should investigate the extent to which discursive sophistication is subject to similar deviations. A related potential confounding factor that is unique to open-ended responses is the respondents' general linguistic skills or verbal verbosity, which may again influence elaboration in open-ended responses but is orthogonal to political sophistication.

One reason why these potential drawbacks may be less worrisome is that the proportion of respondents who refuse to answer any open-ended question in the first place is very low, which indicates that people are sufficiently motivated to engage with the survey. Furthermore, controlling for pure response length did not change the substantive conclusions regarding the effects of discursive sophistication on, for example, political participation or efficacy. The results were also robust for the inclusion of measures of linguistic skills or personality characteristics like extraversion. In a similar vein, the gender gap finding did not appear to be driven by selection effects, which again suggests that survey effort—albeit an important confounding factor to consider—is unlikely to jeopardize the substantive conclusions presented in this paper.

Nevertheless, it is important to keep in mind the differential role of survey mode when comparing factual knowledge and discursive sophistication. Open-ended responses in face-to-face or phone interviews are relatively effortless since they are not unlike voicing your opinion in regular conversations and do not require respondents to transform their thoughts into fixed response categories (e.g., Sudman, Bradburn, and Schwarz, 1996). Unsurprisingly though, respondents tend to provide less elaborate responses in online surveys, resulting in systematically lower discursive sophistication scores (see Figure 7). Knowledge quizzes conducted online, on the other hand, are prone to bias in the opposite

direction due to respondents' tendency to cheat by looking up correct answers (Clifford and Jerit, 2016). Ultimately, more work is needed to explore how survey mode affects discursive sophistication and factual knowledge scores, especially focusing on ways to reduce the effort in answering open-ended questions in online surveys.

Lastly, even if one supports the general notion that open-ended responses can provide useful insights, a skeptic may still argue that manual coding is preferable to the automated framework presented here. However, manual coding of open-ended responses is not always feasible in the context of large-scale surveys, since it can be labor-intensive and requires extensive contextual knowledge such as high levels of language proficiency. The Swiss surveys in Colombo's (2016) study, for example, were conducted in three different languages (German, French, and Italian) and ranged across numerous policy referenda. More importantly, knowledge assessments can be biased by the level of political agreement between individuals (e.g., Ryan, 2011). The measurement approach presented here, on the other hand, is easily replicable and reproducible, is not affected by subjective judgments, and can be directly applied to large-scale surveys in multiple contexts across different languages.

In the past, scholars have argued that testing for factual information, despite its shortcomings, still provides the best available measure of political awareness as it captures "what has actually gotten into people's minds, which, in turn, is critical for intellectual engagement with politics" Zaller (1992, 21). The results presented in this paper suggest that a direct examination of open-ended responses provides a viable supplemental approach that promises new insights into how people make up their mind about politics.

## References

- Althaus, Scott L. 1998. "Information effects in collective preferences." *American Political Science Review* 92 (3): 545-558.

- Barabas, Jason, Jennifer Jerit, William Pollock, and Carlisle Rainey. 2014. "The Question(s) of Political Knowledge." *American Political Science Review* 108 (04): 840–855.
- Bartels, Larry M. 2005. "Homer gets a tax cut: Inequality and public policy in the American mind." *Perspectives on Politics* 3 (1): 15–31.
- Benoit, Kenneth, Kevin Munger, and Arthur Spirling. 2019. "Measuring and explaining political sophistication through textual complexity." *American Journal of Political Science* 63 (2): 491–508.
- Bos, Angela L, Jill S Greenlee, Mirya R Holman, Zoe M Oxley, and J Celeste Lay. 2021. "This One's for the Boys: How Gendered Political Socialization Limits Girls' Political Ambition and Interest." *American Political Science Review*: 1–18.
- Carpini, Michael X Delli, and Scott Keeter. 1993. "Measuring political knowledge: Putting first things first." *American Journal of Political Science* 37 (4): 1179–1206.
- Clifford, Scott, and Jennifer Jerit. 2016. "Cheating on Political Knowledge Questions in Online Surveys: An Assessment of the Problem and Solutions." *Public Opinion Quarterly*.
- Colombo, Céline. 2016. "Justifications and Citizen Competence in Direct Democracy: A Multilevel Analysis." *British Journal of Political Science*: 1–20.
- Converse, Philip E. 1964. "The Nature of Belief Systems in Mass Publics." In *Ideology and Discontent*, ed. David E. Apter. New York: Free Press.
- Cramer, Katherine J, and Benjamin Toff. 2017. "The Fact of Experience: Rethinking Political Knowledge and Civic Competence." *Perspectives on Politics* 15 (3): 754–770.
- DeBell, Matthew. 2013. "Harder than it looks: Coding political knowledge on the ANES." *Political Analysis* 21 (4): 393–406.

Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know about Politics and Why It Matters*. New Haven, London: Yale University Press.

Denny, Matthew J., and Arthur Spirling. 2018. "Text Preprocessing For Unsupervised Learning: Why It Matters, When It Misleads, And What To Do About It." *Political Analysis* 26 (2): 168–189.

Dolan, Kathleen. 2011. "Do women and men know different things? Measuring gender differences in political knowledge." *The Journal of Politics* 73 (01): 97–107.

Dow, Jay K. 2009. "Gender differences in political knowledge: Distinguishing characteristics-based and returns-based differences." *Political Behavior* 31 (1): 117–136.

Druckman, James N. 2014. "Pathologies of studying public opinion, political communication, and democratic responsiveness." *Political Communication* 31 (3): 467–492.

Fortin-Rittberger, Jessica. 2016. "Cross-National Gender Gaps in Political Knowledge How Much Is Due to Context?" *Political Research Quarterly* 69 (3): 391–402.

Fortin-Rittberger, Jessica. 2020. "Political knowledge: Assessing the stability of gender gaps cross-nationally." *International Journal of Public Opinion Research* 32 (1): 46–65.

Fraile, Marta. 2014. "Does deliberation contribute to decreasing the gender gap in knowledge?" *European Union Politics* 15 (3): 372–388.

Gibson, James L, and Gregory A Caldeira. 2009. "Knowing the Supreme Court? A reconsideration of public ignorance of the high court." *The Journal of Politics* 71 (02): 429–441.

Gilens, Martin. 2001. "Political ignorance and collective policy preferences." *American Political Science Review* 95 (02): 379–396.

Gomez, B.T., and J.M. Wilson. 2001. "Political sophistication and economic voting in the American electorate: A theory of heterogeneous attribution." *American Journal of Political Science* 45 (4): 899–914.

Graber, Doris A. 2001. *Processing politics: Learning from television in the Internet age*. University of Chicago Press.

Jerit, Jennifer, and Jason Barabas. 2017. "Revisiting the Gender Gap in Political Knowledge." *Political Behavior* 39 (4): 817–838.

Krosnick, Jon A, Arthur Lupia, Matthew DeBell, and Darrell Donakowski. 2008. Problems with ANES questions measuring political knowledge. Technical report Ann Arbor, MI: American National Election Studies Report.

Kuklinski, James H, Paul J Quirk, Jennifer Jerit, David Schwieder, and Robert F Rich. 2000. "Misinformation and the currency of democratic citizenship." *Journal of Politics* 62 (3): 790–816.

Lee, Moontae, and David Mimno. 2014. "Low-dimensional embeddings for interpretable anchor-based topic inference." In *Proceedings of Empirical Methods in Natural Language Processing*. Citeseer.

Lizotte, Mary-Kate, and Andrew H Sidman. 2009. "Explaining the gender gap in political knowledge." *Politics & Gender* 5 (02): 127–151.

Lupia, Arthur. 1994. "Shortcuts Versus Encyclopedias: Information and Voting Behavior in California Insurance Reform Elections." *American Political Science Review* 88 (1): 63–76.

Lupia, Arthur. 2006. "How elitism undermines the study of voter competence." *Critical Review* 18 (1-3): 217–232.

Lupia, Arthur. 2015. *Uninformed: Why people seem to know so little about politics and what we can do about it*. Oxford University Press.

Luskin, Robert C. 1987. "Measuring political sophistication." *American Journal of Political Science* 31 (4): 856–899.

Manning, Christopher D., Prabhakar Raghavan, Hinrich Schütze et al. 2008. *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.

McGlone, Matthew S, Joshua Aronson, and Diane Kobrynowicz. 2006. "Stereotype threat and the gender gap in political knowledge." *Psychology of Women Quarterly* 30 (4): 392–398.

Miller, Melissa K, and Shannon K Orr. 2008. "Experimenting with a "Third Way" in political knowledge estimation." *Public Opinion Quarterly* 72 (4): 768–780.

Mondak, Jeffery J. 2000. "Reconsidering the measurement of political knowledge." *Political Analysis* 8 (1): 57–82.

Mondak, Jeffery J, and Belinda Creel Davis. 2001. "Asked and answered: Knowledge levels when we will not take 'don't know' for an answer." *Political Behavior* 23 (3): 199–224.

Mondak, Jeffery J, and Mary R Anderson. 2004. "The knowledge gap: A reexamination of gender-based differences in political knowledge." *Journal of Politics* 66 (2): 492–512.

Mondak, Jeffrey J. 2001. "Developing valid knowledge scales." *American Journal of Political Science* 45 (1): 224–238.

Pietryka, Matthew T, and Randall C MacIntosh. 2013. "An Analysis of ANES Items and Their Use in the Construction of Political Knowledge Scales." *Political Analysis* 21 (4): 407–429.

Prior, Markus. 2014. "Visual political knowledge: A different road to competence?" *Journal of Politics* 76 (1): 41–57.

Prior, Markus, and Arthur Lupia. 2008. "Money, time, and political knowledge: Distinguishing quick recall and political learning skills." *American Journal of Political Science* 52 (1): 169–183.

Roberts, Margaret E., Brandon M. Stewart, and Dustin Tingley. 2014. "stm: R package for structural topic models." *Journal of Statistical Software* 1: 1–49.

Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Kushner Gadarian, Bethany Albertson, and David G. Rand. 2014. "Structural Topic Models for Open-Ended Survey Responses." *American Journal of Political Science* 58 (4): 1064–1082.

Ryan, John Barry. 2011. "Accuracy and bias in perceptions of political knowledge." *Political Behavior* 33 (2): 335–356.

Spirling, Arthur. 2016. "Democratization and Linguistic Complexity The Effect of Franchise Extension on Parliamentary Discourse, 1832–1915." *The Journal of Politics* 78 (1): 120–136.

Stolle, Dietlind, and Elisabeth Gidengil. 2010. "What do women really know? A gendered analysis of varieties of political knowledge." *Perspectives on Politics* 8 (01): 93–109.

Sturgis, Patrick, Nick Allum, and Patten Smith. 2008. "An experiment on the measurement of political knowledge in surveys." *Public Opinion Quarterly* 72 (1): 90–102.

Sudman, Seymour, Norman M Bradburn, and Norbert Schwarz. 1996. *Thinking about answers: The application of cognitive processes to survey methodology*. Jossey-Bass.

Tetlock, Philip E. 1983. "Cognitive style and political ideology." *Journal of Personality and Social Psychology* 45 (1): 118.

Tetlock, Philip E. 1993. "Cognitive structural analysis of political rhetoric: Methodological and theoretical issues." In *Explorations in political psychology*. Duke University Press Durham.

Verba, Sidney, Nancy Burns, and Kay Lehman Schlozman. 1997. "Knowing and caring

- about politics: Gender and political engagement." *The Journal of Politics* 59 (04): 1051–1072.
- Wolak, Jennifer, and Michael McDevitt. 2011. "The roots of the gender gap in political knowledge in adolescence." *Political Behavior* 33 (3): 505–533.
- Zaller, John. 1990. "Political awareness, elite opinion leadership, and the mass survey response." *Social Cognition* 8 (1): 125.
- Zaller, John. 1991. "Information, values, and opinion." *American Political Science Review* 85 (04): 1215–1237.
- Zaller, John. 1992. *The nature and origins of mass opinion*. Cambridge University Press.

# Appendices

*Online Appendix:*

## Let's Talk Politics – A Naive Approach for Measuring Political Sophistication

<b>A Overview of Data Sources and Open Ended Items</b>	<b>2</b>
I 2012 & 2016 American National Election Study . . . . .	2
II 2015 YouGov Survey . . . . .	2
III Swiss Referendum Survey . . . . .	3
<b>B Information on Discursive Sophistication Components</b>	<b>3</b>
I Distribution of Word Counts in Open-Ended Responses . . . . .	3
II Overview of Topic Proportions . . . . .	5
III Discursive Sophistication Components . . . . .	8
<b>C Pre-Processing and Topic Model Specification</b>	<b>9</b>
I PreText Analysis . . . . .	9
II Robustness Checks for Varying Model Specifications . . . . .	10
<b>D Additional Information on Remaining Variables</b>	<b>13</b>
I Item Selection and Recoding . . . . .	13
II Variable Distributions – 2012 ANES . . . . .	15
III Variable Distributions – 2016 ANES . . . . .	17
IV Variable Distributions – 2015 YouGov . . . . .	19
<b>E Robustness Checks</b>	<b>20</b>
I Controlling for Personality Characteristics . . . . .	20
II Controlling for Individual Response Length . . . . .	21
<b>F Tables of Model Estimates</b>	<b>22</b>
I Main Analyses . . . . .	22
II Robustness Checks . . . . .	27

# **Appendix A Overview of Data Sources and Open Ended Items**

Throughout the remainder of this paper, I apply the theoretical framework across multiple surveys fielded in two countries that employed different sets of open-ended questions. Each survey focuses on sophistication in the context of distinct political tasks, namely the evaluation of (1) candidates running for public office, (2) broad issue areas such as health care and gun legislation, and (3) specific legislative policy referenda. The data sets are briefly described below.

## **I 2012 & 2016 American National Election Study**

The main analyses are based on the 2012 and 2016 wave of the American National Election Study (ANES), each of which consists of a representative survey of about 5000 adults in the months before the US Presidential election in each year. About 2000 respondents in each wave participated in face-to-face interviews while the remaining respondents filled out the survey online. For the purpose of the present analyses, I rely on pooled datasets while controlling for differences in survey mode. Discursive sophistication is evaluated using a set of open-ended questions in which respondents were asked to list anything in particular that they like/dislike about the Democratic/Republican party as well as anything that might make them vote/not vote for either of the Presidential candidates. They were probed by the interviewer asking “anything else?” until the respondent answered “no.” Overall, there are a total number of 8 open-ended responses where individuals described their beliefs and attitudes towards political actors. Individuals who did not respond to all of the open-ended items (420 in 2012; 204 in 2016), or who responded in Spanish (228 in 2012; 43 in 2016), are excluded from the analysis since discursive sophistication for them is not directly comparable to the remaining respondents.

## **II 2015 YouGov Survey**

In order to replicate and extend the main analyses, I rely on a separate nationally representative survey employing an alternative set of open-ended responses. The data was collected by YouGov in December 2015 and contains responses of 1000 U.S. citizens. As part of this study, respondents were asked to describe their attitudes towards two prominent political issues that were discussed frequently in the media. First, they were asked in a closed format whether they favor or oppose stricter gun laws. Subsequently, they were asked to respond to the following two questions:

- Still thinking about the question you just answered, what thoughts came to mind while you were answering that question? Please try to list everything that came to mind.
- Thinking about the mass shootings that have occurred in the U.S. in the last few years, what factors do you think are responsible for the shootings?

Second, the respondents reported on their attitudes towards the Affordable Care Act in a closed format and were then asked to elaborate in their own words by answering the following questions:

- Still thinking about the question you just answered, what thoughts came to mind while you were answering that question? Please try to list everything that came to mind.
- For decades, experts have observed that the United States spends far more per person on health care than any other country. However, the U.S. falls behind on most measures of health care outcomes, such as life expectancy. What factors do you think are responsible for the state of our health care system?

Compared to the open-ended likes/dislikes items included in the 2012 and 2016 ANES, the questions directly address considerations related to specific policy issues that were prominent in the political discourse at the time of the survey. Respondents who did not provide an answer to all of the open-ended questions were removed from the analysis (48).

### **III Swiss Referendum Survey**

Lastly, I examine survey data on Swiss citizens justifying their vote choices on multiple referenda used in a recent analysis by [Colombo \(2016\)](#). The author compiled a data set of cross-sectional surveys administered in Switzerland after national popular votes on multiple policy propositions. The original surveys were conducted as representative samples after each of thirty-four national policy votes that were held between 2008 and 2012 resulting in a total of about 27,000 observations. However, respondents were only asked to justify their decision for or against a given proposition in verbatim if they participated in the vote in the first place. As such, about 5,000 individuals in the data set did not provide an open-ended response. The remaining respondents were asked to describe the main reason as well as additional justifications for their decision in two separate items. As before, discursive sophistication is evaluated based on verbatim responses to both questions. Since the Swiss survey was conducted in three different languages (German, French, and Italian), I computed the measure of discursive sophistication separately for each group of respondents.

## **Appendix B Detailed Information on Open-Ended Responses and Discursive Sophistication Components**

### **I Distribution of Word Counts in Open-Ended Responses**

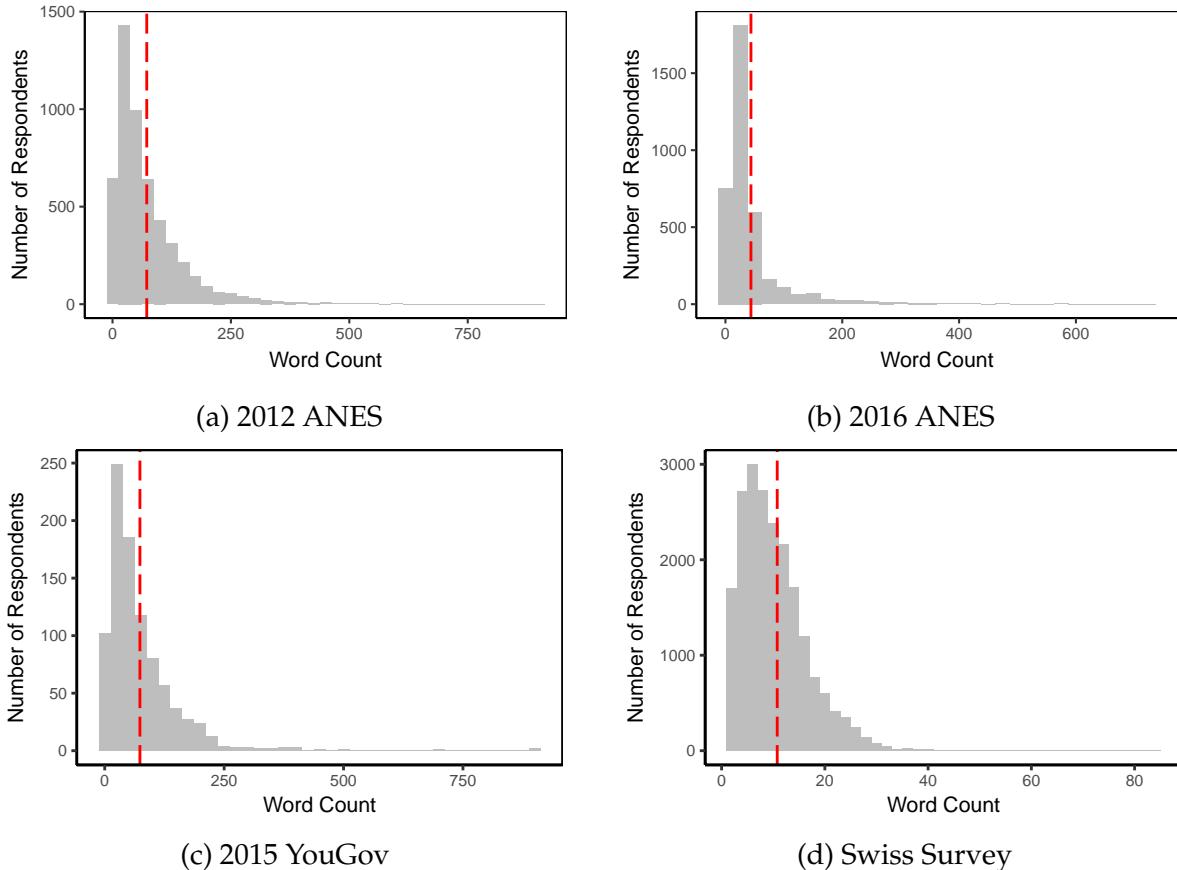


Figure B.1: Histograms of total word count in the collection of open-ended responses for each individual. The dashed red lines indicate the average response lengths in each survey.

## II Overview of Topic Proportions

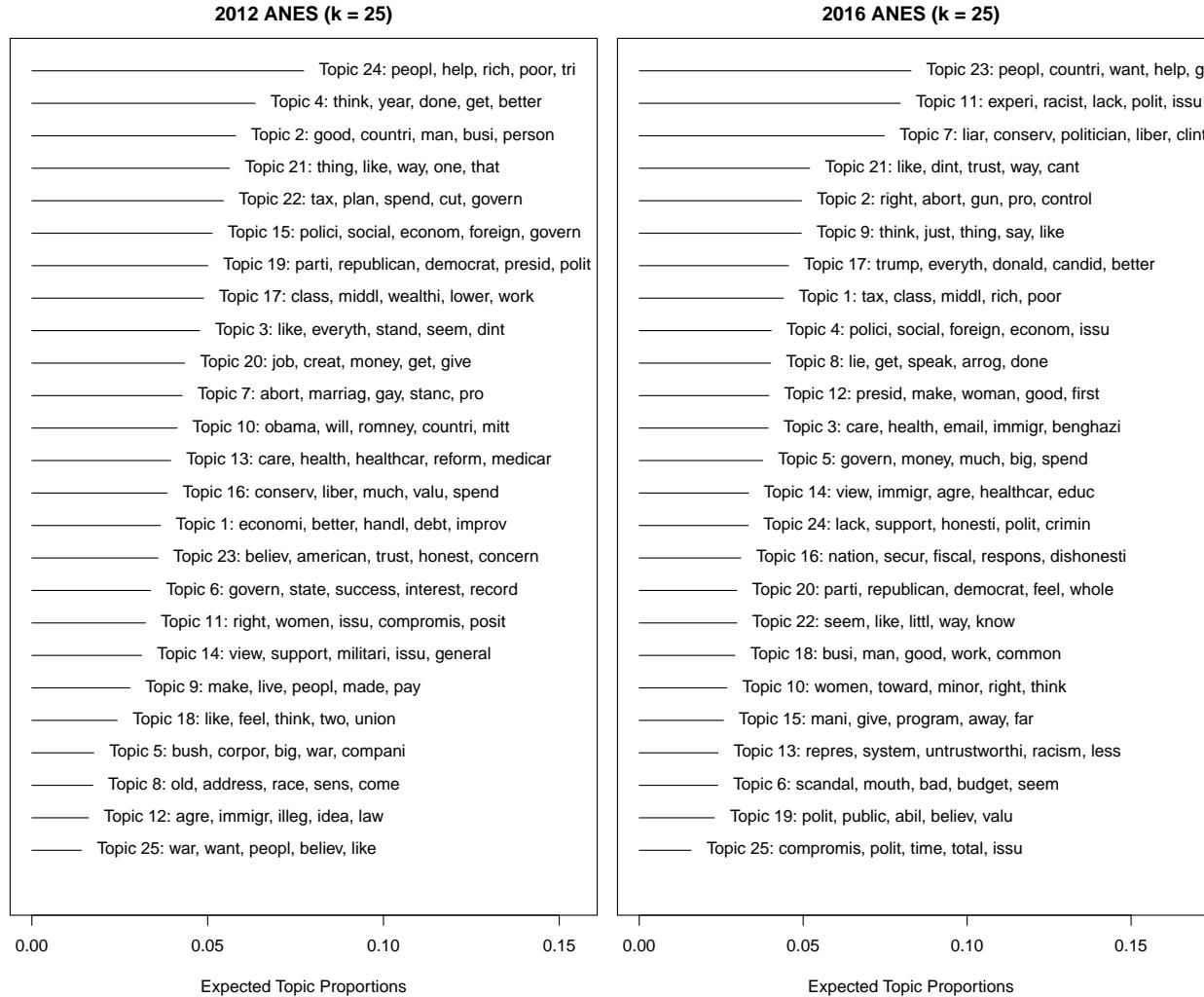


Figure B.2: Estimated topic proportions in the 2012 and 2016 ANES based on the structural topic model. See Appendix C for details on the model specification.

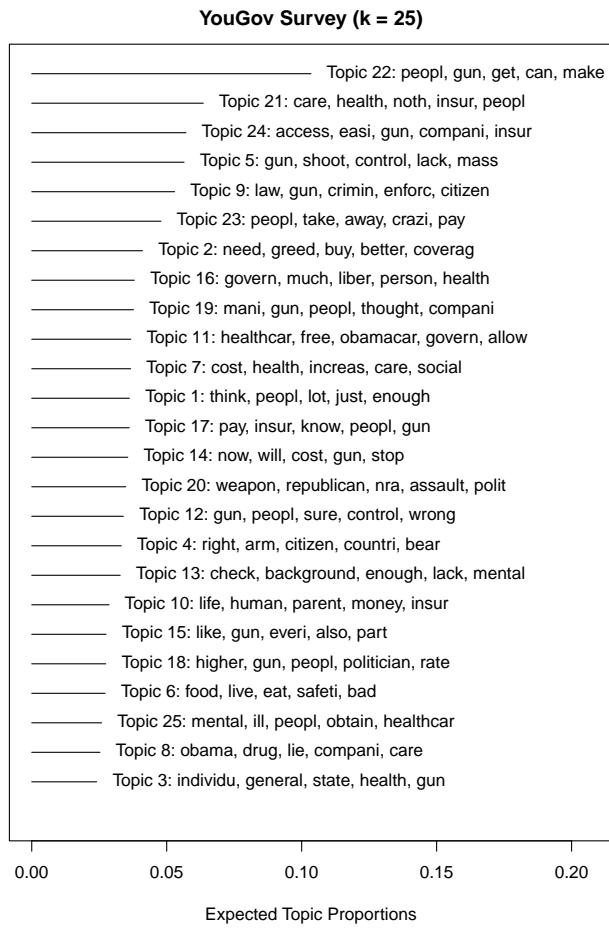


Figure B.3: Estimated topic proportions in the 2015 YouGov survey based on the structural topic model. See Appendix C for details on the model specification.

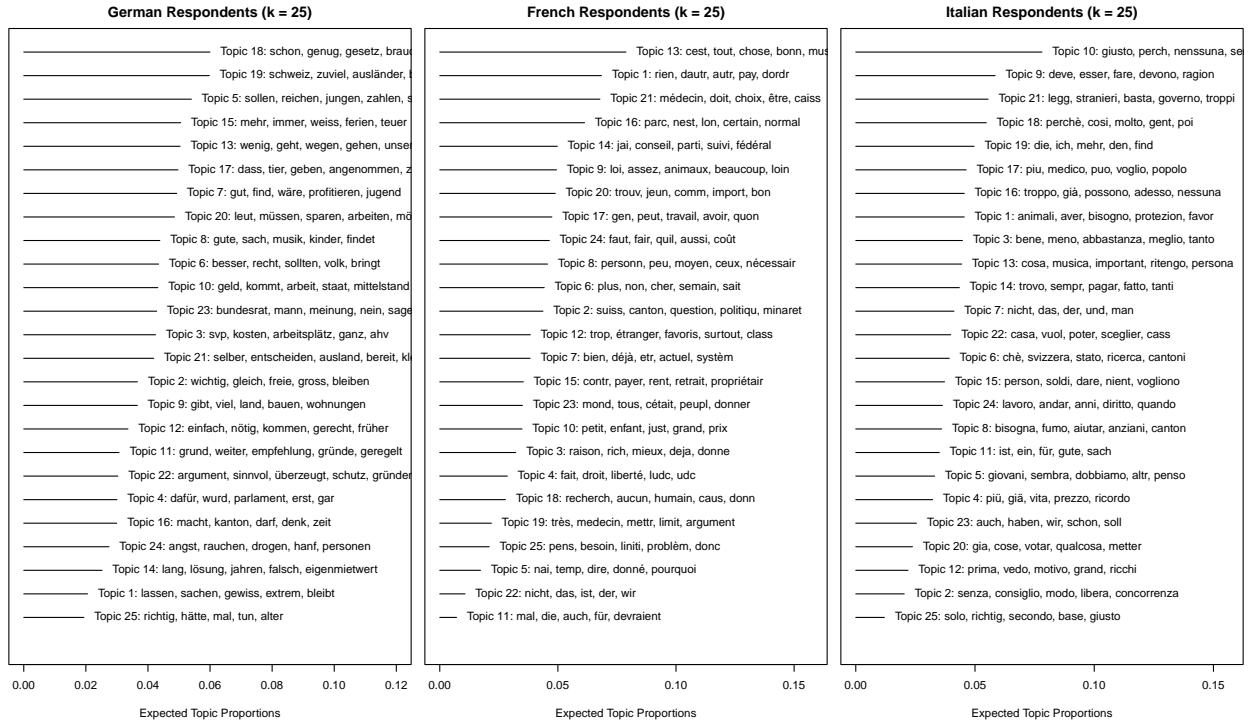


Figure B.4: Estimated topic proportions in the Swiss survey based on the structural topic model. See Appendix C for details on the model specification.

### III Discursive Sophistication Components

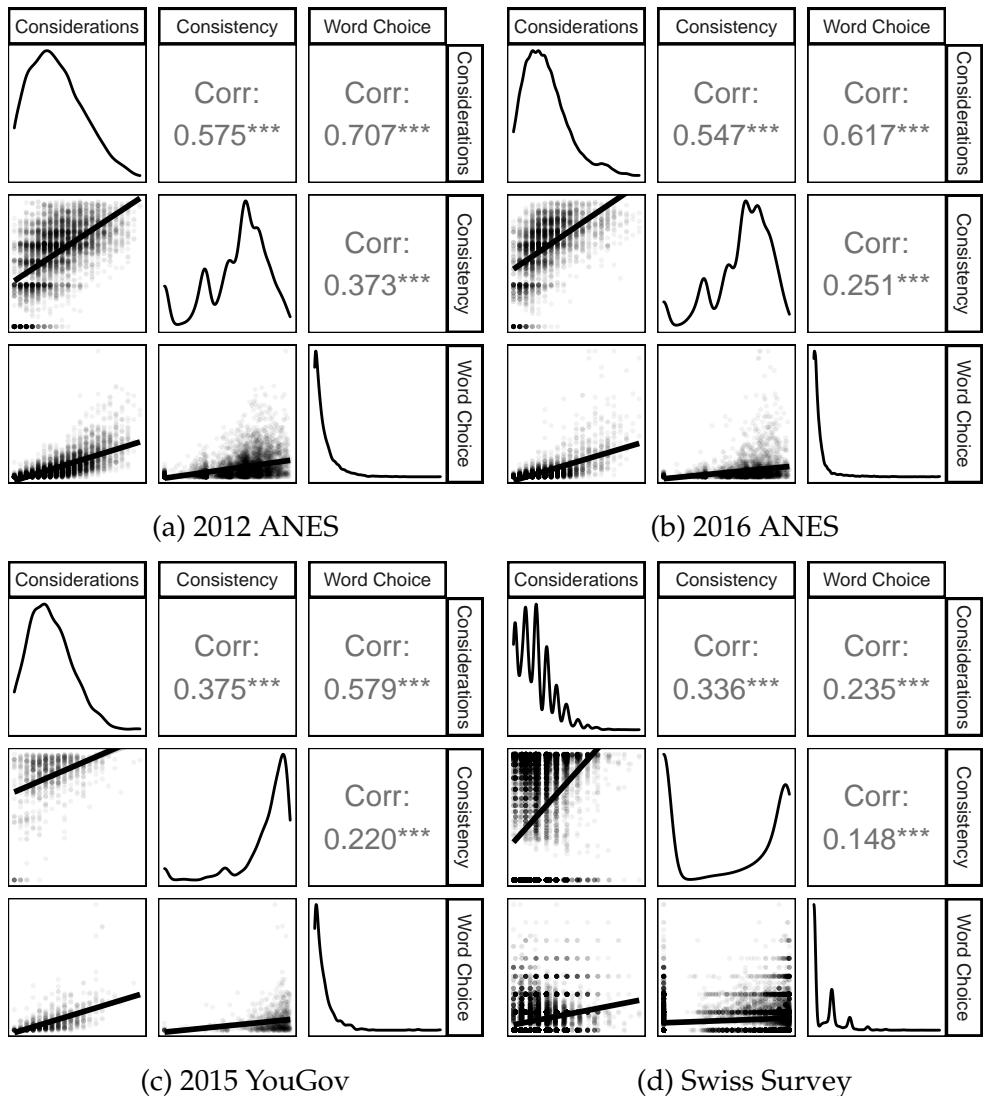


Figure B.5: Correlation matrix of individual components of discursive sophistication. The plots on the diagonal display univariate densities for each component. The panels in the lower triangular display the scatter plot of two measures as well as a linear fit.

# Appendix C Pre-Processing and Topic Model Specification

## I PreText Analysis

Two components of discursive sophistication (*considerations* and *word choice*) rely on quantities extracted from structural topic models (Roberts et al., 2014). As with any other text-as-data approach, a necessary first step before estimating the topic model is to pre-process the raw text and convert it into a document term matrix (DTM, see for example Manning et al., 2008). Common pre-processing procedures include stemming and lower-casing, as well as the removal of numbers, punctuation, stopwords, and infrequent terms. However, topic models and other unsupervised learning techniques can be sensitive to these pre-processing choices (c.f., Denny and Spirling, 2018). To address this issue, Denny and Spirling (2018) recommend that researchers compare DTMs under all possible pre-processing regimes. The authors propose *preText scores* as a measure to quantify the extent to which varying pre-processing regimes may yield unusual results compared to a baseline without any pre-processing.

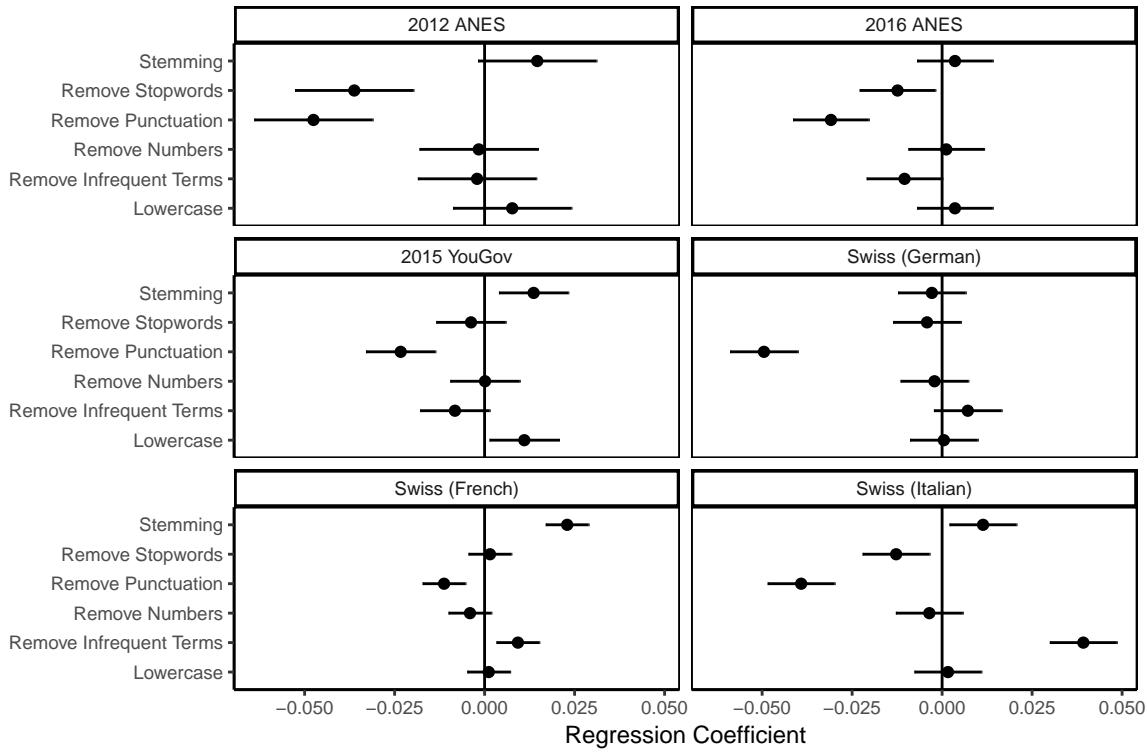


Figure C.1: PreText analysis of pre-processing decisions of open-ended responses across all datasets. Regression coefficients display the effects of each of the six pre-processing choices on the resulting preText score.

## II Robustness Checks for Varying Model Specifications

Following the procedure outlined in [Denny and Spirling \(2018\)](#), Figure C.1 displays the results of a linear model regressing preText scores resulting from all possible pre-processing regimes on each individual step for a random subset of 500 open-ended responses in each of the studies included in the analyses. Significant coefficients indicate that the topic model results may be sensitive to the respective pre-processing step. As such, removing stopwords and punctuation, as well as removing infrequent terms (at least in the 2016 ANES) might be problematic. [Denny and Spirling \(2018\)](#), however, emphasize that the most important consideration in choosing pre-processing steps are theoretical. Given that the purpose of the topic model is to extract considerations related to political preferences, there are strong theoretical reasons to remove stopwords and punctuation from open-ended responses as they do not contain any relevant content. Furthermore, I apply lowercasing and stemming of terms to reduce resulting document term matrix to a computationally more manageable size and since these pre-processing steps should not be influential according to the preText analysis.

It is less obvious from a theoretical perspective whether to remove infrequent terms from open-ended responses, although it is preferred in order to make the estimation of the discursive sophistication components computationally efficient. Since the preText analysis for the 2016 ANES suggests that this pre-processing step might be influential, I compare discursive sophistication for both alternative regimes below (c.f., [Denny and Spirling, 2018](#)). Before turning to this sensitivity check, however, I consider another crucial modeling choice when working with topic models: determining the total number of topics  $k$  to be estimated. For all analyses reported below, the number of topics was selected using the algorithm proposed by [Lee and Mimno \(2014\)](#) and implemented in the `stm` package in R ([Roberts, Stewart, and Tingley, 2014](#)).

Figure C.2 examines whether the proposed measure of discursive sophistication is sensitive to the removal of infrequent terms as well as the chosen number of topics  $k$ . The y-axis depicts the preferred pre-processing regime including all steps discussed above while the x-axis plots results for alternative specifications. The panels on the left compare the preferred specification to discursive sophistication based on a reduced number of topics ( $k = 20$ ). The middle panels additionally include infrequent terms instead of removing them.<sup>9</sup> The panels on the right do not perform stemming as part of the pre-processing step. Across all panels, discursive sophistication scores are highly correlated and therefore insensitive to pre-processing choices and varying numbers of topics.

In summary, open-ended responses in the analyses reported in the main text are pre-processed by stemming and lowercasing, as well as the removing numbers, punctuation, stopwords, and infrequent terms (i.e., terms that appear in fewer than 10 responses).<sup>10</sup> While the results discussed in the manuscript are based on this preferred specification, the substantive results are robust for alternative pre-processing regimes or varying numbers

<sup>9</sup>Calculating discursive sophistication with large numbers of topics while including infrequent terms is computationally prohibitive.

<sup>10</sup>Prior to applying these pre-processing steps, open-ended responses in the 2012 & 2016 ANES as well as the 2015 YouGov survey are cleaned by correcting spelling errors using an implementation of the Aspell spell-checking algorithm ([www.aspell.net](http://www.aspell.net)).

of topics.

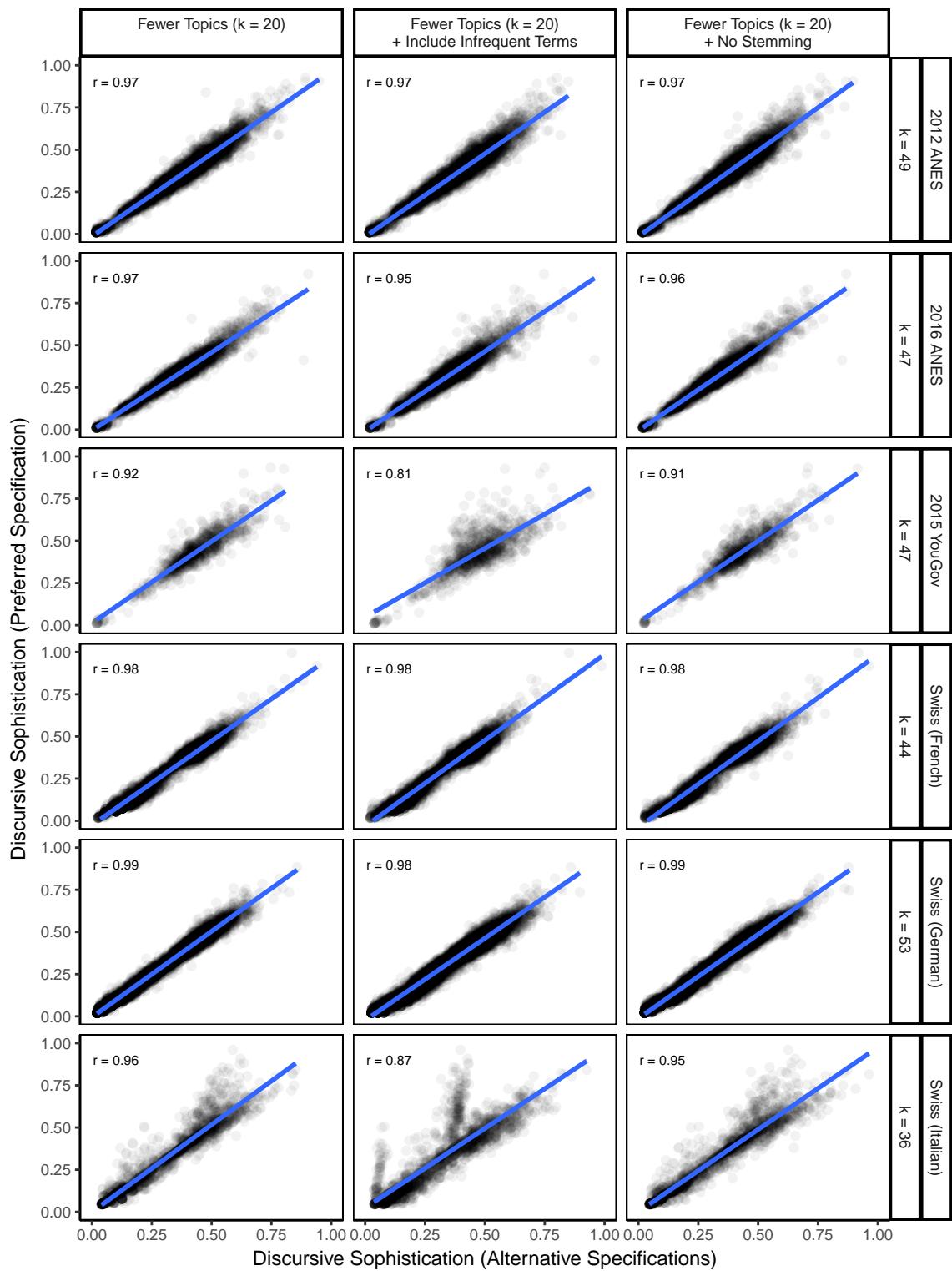


Figure C.2: Robustness of discursive sophistication measure for different pre-processing choices and topic model specifications.

## Appendix D Additional Information on Remaining Variables

### I Item Selection and Recoding

#### Conventional measures of political knowledge:

- 2012 ANES: Additive index of correct responses to 5 knowledge items included in the pre-election wave (number of Presidential terms, size of budget deficit, length of Senate term, meaning of Medicare, federal government spending). ‘Don’t know’ responses are considered incorrect. Interviewer evaluations are based on the assessment of the respondent’s general level of information about politics recorded at the end of the pre-election wave.
- 2016 ANES: Additive index of correct responses to 4 knowledge items included in the pre-election wave (length of Senate term, federal government spending, majority in House, majority in Senate). ‘Don’t know’ responses are considered incorrect. Interviewer evaluations are based on the assessment of the respondent’s general level of information about politics recorded at the end of the pre-election wave.
- 2015 YouGov Survey: Additive index of correct responses to 8 knowledge items (Speaker of the House, meaning of TPP, Chair of Federal Reserve Board, current unemployment rate, Presidential veto override, meaning of Common Core, leading source of electricity in US, majority in Senate). ‘Don’t know’ responses are considered incorrect.

#### Dependent variables:

- Turnout (2012 & 2016 ANES): Dichotomous indicator, based on post-election wave.
- Non-conventional participation (2012 & 2016 ANES): Additive index of different forms of political engagement (participated in protest, signed petition, wore campaign button, wrote letter to public official).
- Internal efficacy (2012 & 2016 ANES): Sum of two agree/disagree items (politics too complicated, good understanding of political issues [reversed]).
- External efficacy (2012 & 2016 ANES): Sum of two agree/disagree items (public officials don’t care, people have no say about what the government does).
- Information retrieval (2015 YouGov Survey): Additive index of correct answers to 9 questions about the fictional disease described in the news article (symptoms: fatigue, headaches, diarrhea, joint pain, boils, warts, fever; virus spread; cure for the virus)
- Candidate policy positions (2012 & 2016 ANES): Placement of Republican and Democratic Presidential candidates on 7-point scale (ideology, government spending, defense spending, insurance policy, job guarantee, aid to Blacks, environment vs jobs).

### **Control variables:**

- *Gender* (2012 & 2016 ANES, 2015 YouGov Survey): Dichotomous indicator for female respondents.
- *Wordsum vocabulary scores* (2012 & 2016 ANES): Modified version of the GSS wordsum vocabulary test consisting of 10 terms.
- *Media exposure* (2012 & 2016 ANES): Additive index of the frequency of weekly exposure to various political information sources such as newspapers or TV news (2012 ANES). In the 2016 ANES, it only consists of a single item measuring the number of days in the past week the respondent has spent watching/reading/listening news on any media.
- *Political discussion frequency* (2012 & 2016 ANES): Self-reported count of days in the past week spent discussing politics with family or friends.
- *College education* (2012 & 2016 ANES, 2015 YouGov Survey): Dichotomous indicator for Bachelor's degree or higher.
- *Family/Household income* (2012 & 2016 ANES, 2015 YouGov Survey): Self-reported household income categories.
- *Age* (2012 & 2016 ANES, 2015 YouGov Survey): Logged age in years.
- *Race* (2012 & 2016 ANES, 2015 YouGov Survey): Dichotomous indicator for black non-Hispanic vs. others.
- *Church attendance* (2012 & 2016 ANES, 2015 YouGov Survey): Six-category indicator of the frequency of church attendance.
- *Survey Mode* (2012 & 2016 ANES): Dichotomous indicator for face-to-face vs. online samples of the ANES surveys.
- *Personality characteristics* (2012 & 2016 ANES): Measures of extraversion and being reserved, part of the Ten Item Personality Inventory (TIPI) measuring the "Big Five" personality traits.
- *Response length* (2012 & 2016 ANES): Logged number of words in the collection of open-ended responses by each individual.

## II Variable Distributions – 2012 ANES

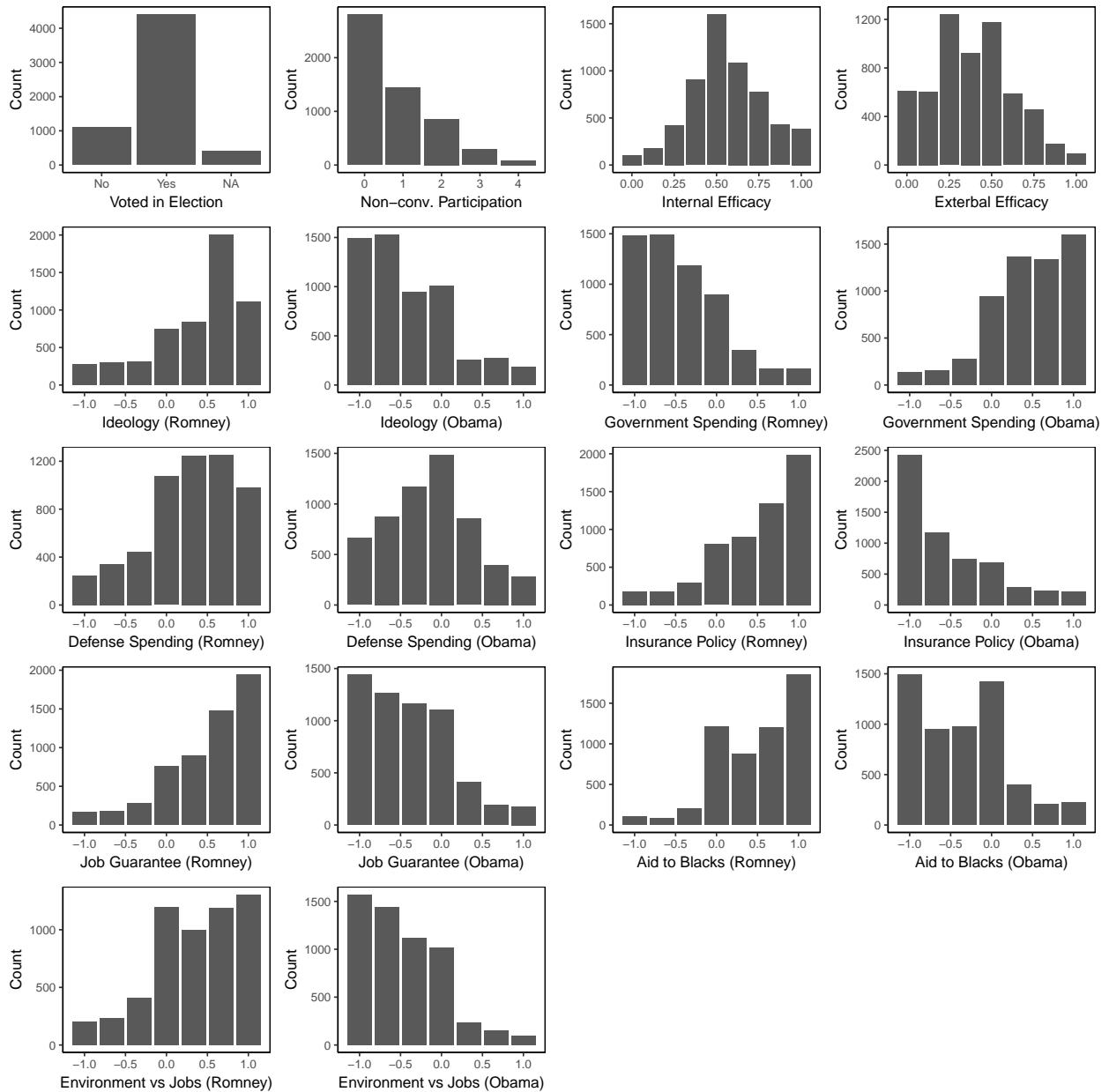


Figure D.1: Histograms of dependent variables included in 2012 ANES.

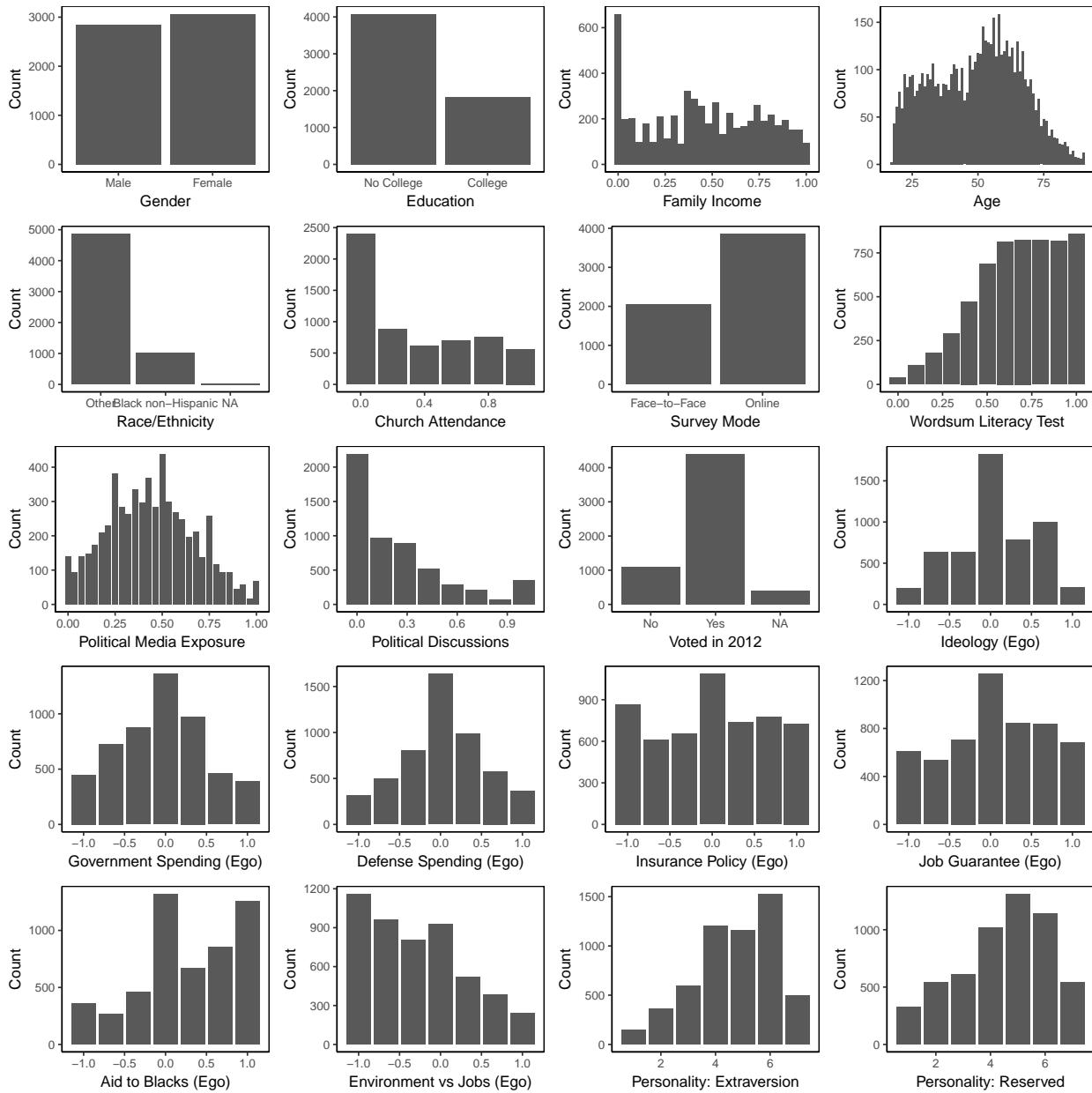


Figure D.2: Histograms of independent variables included in 2012 ANES.

### III Variable Distributions – 2016 ANES

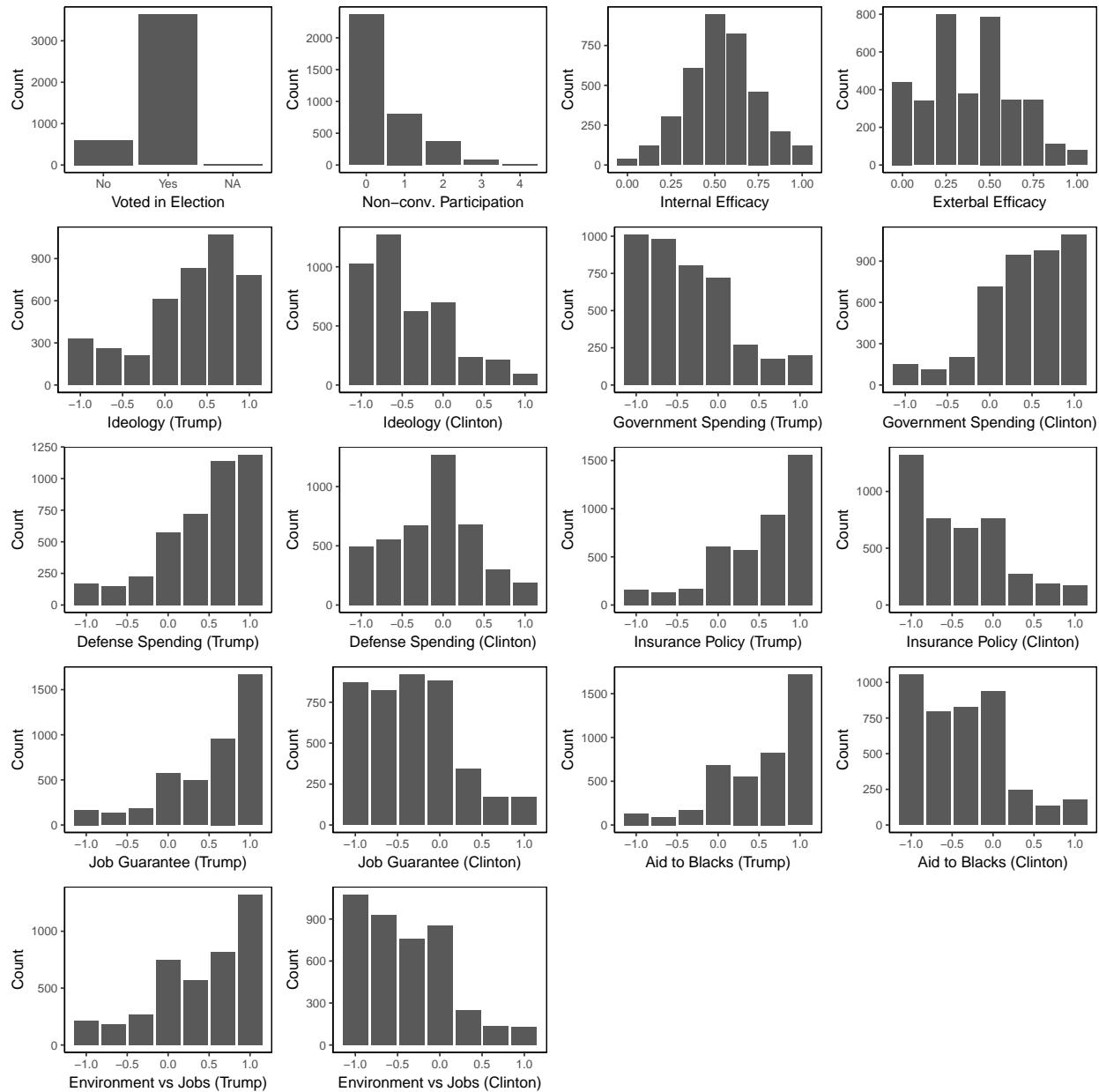


Figure D.3: Histograms of dependent variables included in 2016 ANES.

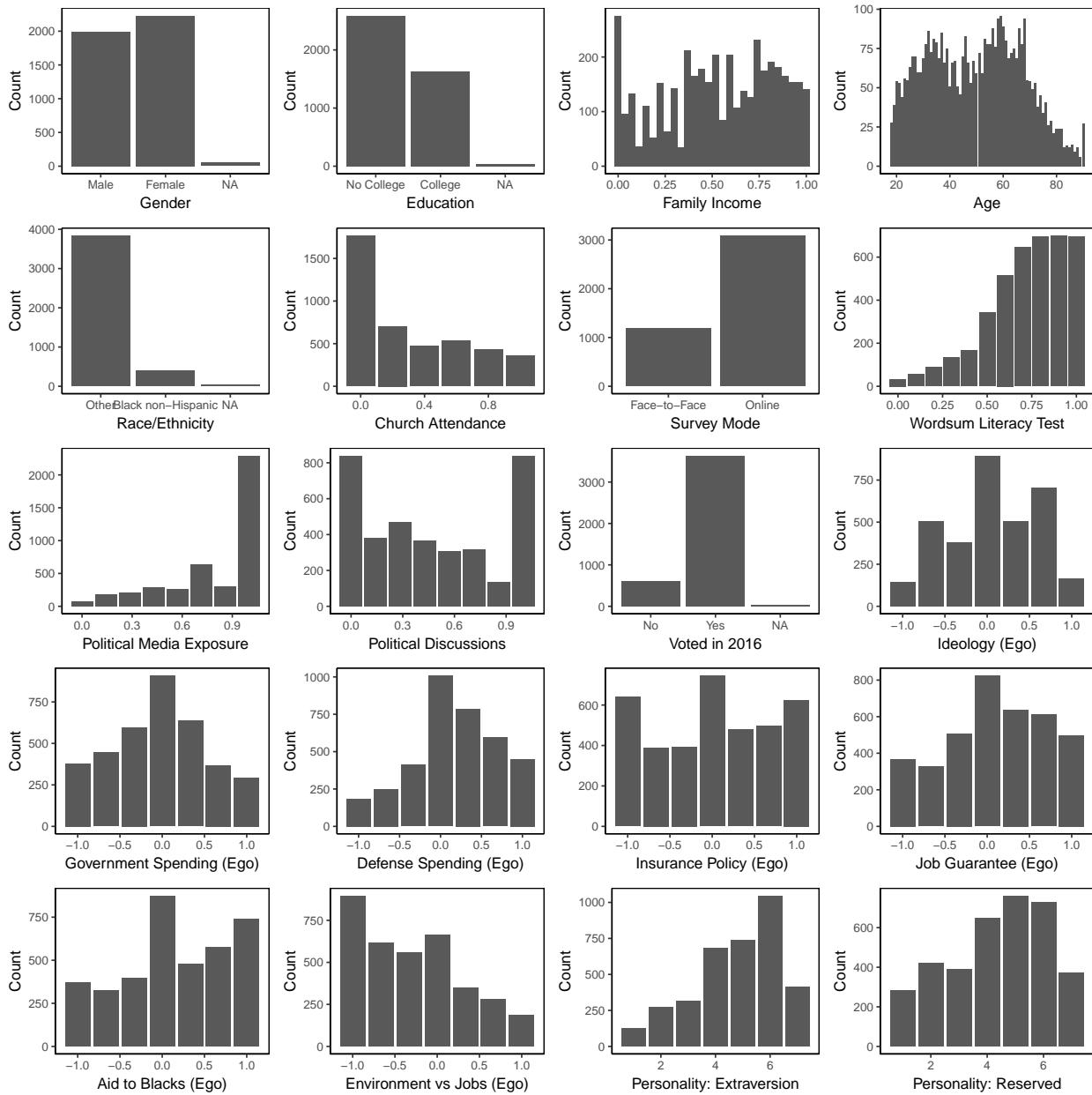


Figure D.4: Histograms of independent variables included in 2016 ANES.

## IV Variable Distributions – 2015 YouGov

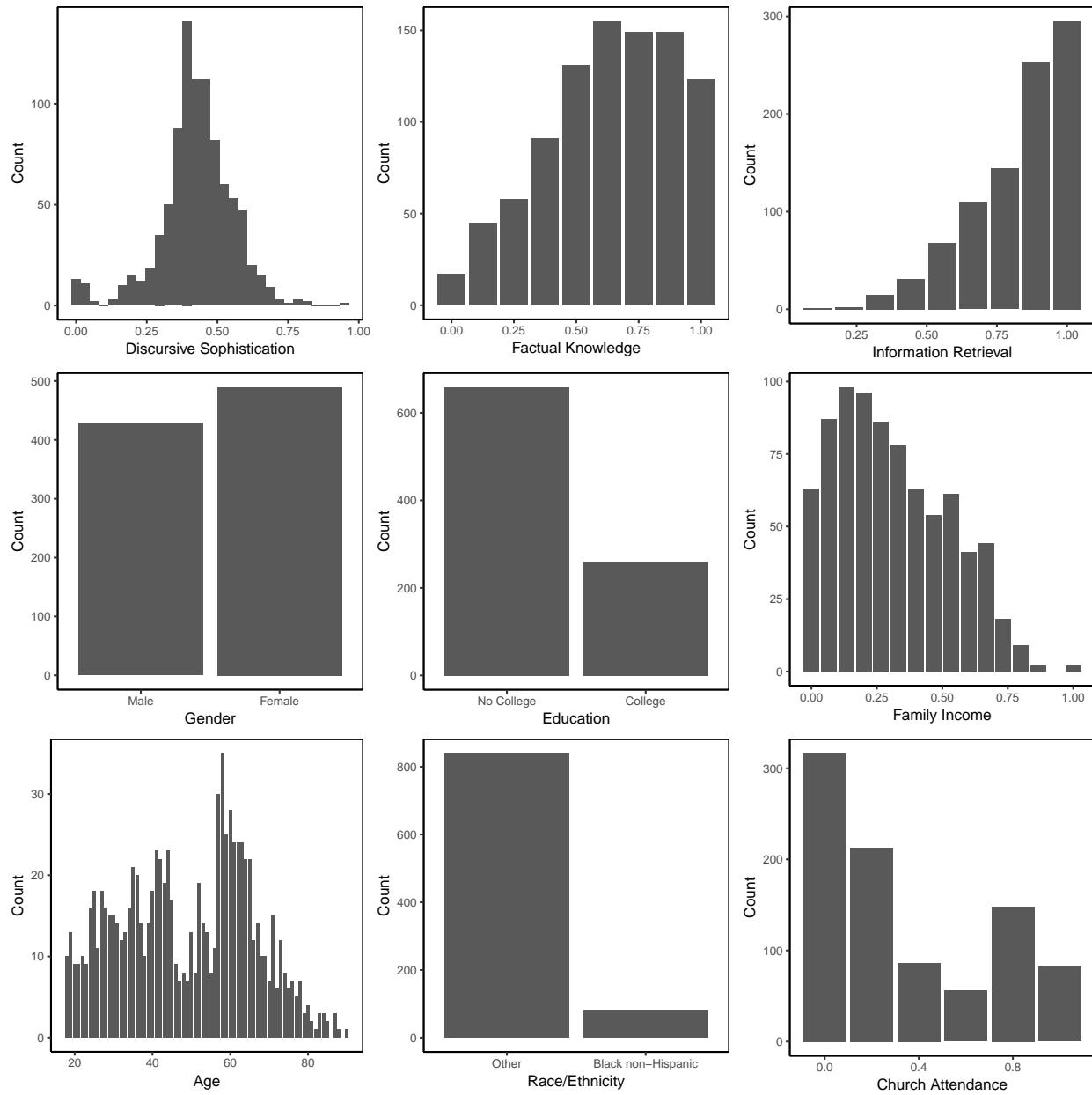


Figure D.5: Histogram of variables included in 2015 YouGov survey.

# Appendix E Robustness Checks

## I Controlling for Personality Characteristics

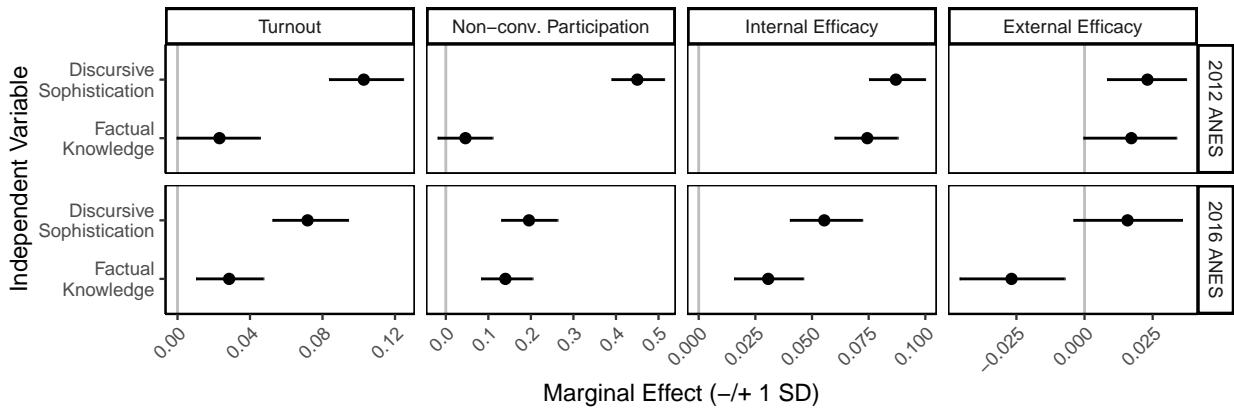


Figure E.1: Effects of sophistication on internal efficacy, external efficacy, non-conventional participation, and turnout in the 2012 and 2016 ANES. For each dependent variable, the figure displays the change in expected values after increasing each sophistication measure from -1 to +1 standard deviation from its mean (including 95% confidence intervals). Model estimates are based on logistic regression (turnout) or OLS (internal efficacy, external efficacy, non-conventional participation). Both sophistication measure are included simultaneously while controlling for gender, education, income, age, race, church attendance, survey mode, Wordsum vocabulary scores, as well as personality characteristics (extraversion and being reserved). Full model results are presented in the appendix, Tables E.6 and E.7.

## II Controlling for Individual Response Length

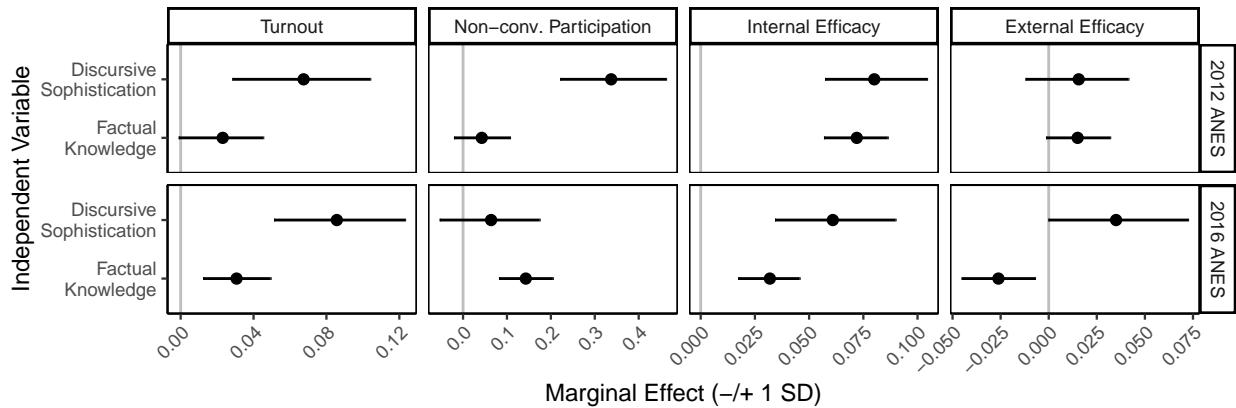


Figure E.2: Effects of sophistication on internal efficacy, external efficacy, non-conventional participation, and turnout in the 2012 and 2016 ANES. For each dependent variable, the figure displays the change in expected values after increasing each sophistication measure from -1 to +1 standard deviation from its mean (including 95% confidence intervals). Model estimates are based on logistic regression (turnout) or OLS (internal efficacy, external efficacy, non-conventional participation). Both sophistication measure are included simultaneously while controlling for gender, education, income, age, race, church attendance, survey mode, Wordsum vocabulary scores, as well as the logged word count across open-ended responses. Full model results are presented in the appendix, Tables E.8 and E.9.

## Appendix F Tables of Model Estimates

### I Main Analyses

Table E.1: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy in the 2012 ANES. Standard errors in parentheses. Estimates are used for Figure 3 in the main text.

	<i>Dependent variable:</i>			
	Turnout	Participation	Internal Efficacy	External Efficacy
Discursive Soph.	2.921 (0.299)	1.440 (0.098)	0.278 (0.020)	0.084 (0.024)
Factual Knowledge	0.432 (0.218)	0.099 (0.075)	0.158 (0.016)	0.032 (0.018)
Female	0.086 (0.085)	-0.067 (0.028)	-0.053 (0.006)	0.016 (0.007)
College Degree	0.350 (0.112)	0.159 (0.034)	0.022 (0.007)	0.035 (0.008)
Family Income	0.947 (0.156)	0.022 (0.052)	0.010 (0.011)	0.016 (0.013)
Age (log)	0.988 (0.105)	0.102 (0.038)	-0.006 (0.008)	-0.014 (0.009)
African American	0.910 (0.123)	0.096 (0.038)	0.066 (0.008)	0.082 (0.009)
Church Attendance	0.752 (0.129)	0.112 (0.040)	0.010 (0.008)	0.048 (0.010)
Mode: Online	0.530 (0.094)	0.227 (0.033)	0.017 (0.007)	-0.039 (0.008)
Wordsum Score	0.638 (0.219)	0.403 (0.076)	0.092 (0.016)	0.013 (0.019)
Constant	-5.019 (0.401)	-0.598 (0.145)	0.326 (0.030)	0.352 (0.035)
Observations	4,716	4,692	4,996	4,985
R <sup>2</sup>		0.124	0.161	0.043
Log Likelihood	-1,868.199			

Table E.2: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy in the 2016 ANES. Standard errors in parentheses. Estimates are used for Figure 3 in the main text.

	<i>Dependent variable:</i>			
	Turnout	Participation	Internal Efficacy	External Efficacy
Discursive Soph.	3.891 (0.487)	0.792 (0.131)	0.229 (0.031)	0.076 (0.041)
Factual Knowledge	0.730 (0.219)	0.269 (0.058)	0.058 (0.014)	-0.051 (0.018)
Female	0.177 (0.114)	0.061 (0.029)	-0.059 (0.007)	-0.003 (0.009)
College Degree	0.554 (0.141)	0.092 (0.033)	0.058 (0.008)	0.057 (0.010)
Family Income	0.487 (0.207)	-0.077 (0.055)	0.020 (0.013)	0.062 (0.017)
Age (log)	0.840 (0.139)	-0.110 (0.038)	0.020 (0.009)	-0.005 (0.012)
African American	1.126 (0.237)	0.097 (0.051)	0.057 (0.012)	-0.020 (0.016)
Church Attendance	1.067 (0.191)	-0.175 (0.043)	-0.006 (0.010)	0.080 (0.014)
Mode: Online	0.192 (0.136)	0.110 (0.036)	0.069 (0.009)	-0.016 (0.011)
Wordsum Score	0.931 (0.275)	0.402 (0.079)	0.103 (0.019)	0.017 (0.025)
Constant	-4.276 (0.535)	0.213 (0.151)	0.231 (0.036)	0.338 (0.048)
Observations	3,486	3,039	3,037	3,038
R <sup>2</sup>		0.062	0.147	0.043
Log Likelihood	-1,085.646			

Table E.3: Effects of sophistication on information retrieval in the 2015 YouGov study. Standard errors in parentheses. Estimates are used for Figure 4 in the main text.

	<i>Dependent variable:</i> Information Retrieval
Discursive Soph.	0.247 (0.045)
Factual Knowledge	0.133 (0.027)
Female	0.044 (0.012)
College Degree	0.015 (0.014)
Family Income	−0.062 (0.030)
Age (log)	0.032 (0.016)
African American	−0.026 (0.021)
Church Attendance	−0.059 (0.017)
Constant	0.534 (0.061)
Observations	792
R <sup>2</sup>	0.136

Table E.4: Effects of gender on discursive sophistication and factual knowledge in the 2012 ANES and 2016 ANES. Standard errors in parentheses. Estimates are used for Figure 7 in the main text.

	<i>Dependent variable:</i>			
	Discursive Sophistication		Factual Knowledge	
	2012 ANES	2016 ANES	2012 ANES	2016 ANES
Female	0.001 (0.004)	-0.002 (0.004)	-0.050 (0.006)	-0.029 (0.009)
Media Exposure	0.042 (0.010)	0.034 (0.008)	0.067 (0.014)	0.056 (0.019)
Discussion Frequency	0.121 (0.008)	0.058 (0.006)	0.062 (0.010)	0.078 (0.013)
College Degree	0.070 (0.005)	0.045 (0.004)	0.091 (0.007)	0.045 (0.010)
Family Income	0.054 (0.008)	0.043 (0.007)	0.116 (0.010)	0.089 (0.017)
Age (log)	0.017 (0.006)	0.003 (0.005)	0.081 (0.008)	0.092 (0.013)
African American	-0.010 (0.006)	-0.026 (0.007)	-0.085 (0.008)	-0.004 (0.016)
Church Attendance	0.004 (0.006)	0.005 (0.006)	-0.003 (0.008)	-0.006 (0.013)
Mode: Online	-0.044 (0.005)	-0.110 (0.004)	0.092 (0.006)	0.085 (0.010)
Constant	0.199 (0.022)	0.284 (0.020)	0.195 (0.030)	0.009 (0.046)
Observations	4,690	3,119	4,690	3,108
R <sup>2</sup>	0.178	0.265	0.282	0.110

Table E.5: Effects of gender on discursive sophistication and factual knowledge in the 2015 YouGov Study. Standard errors in parentheses. Estimates are used for Figure 7 in the main text.

	<i>Dependent variable:</i>	
	Discursive Sophistication	Factual Knowledge
Female	0.008 (0.009)	-0.157 (0.016)
College Degree	0.027 (0.011)	0.108 (0.019)
Family Income	0.003 (0.024)	0.233 (0.041)
Age (log)	0.040 (0.013)	0.135 (0.021)
African American	-0.064 (0.017)	-0.086 (0.029)
Church Attendance	-0.008 (0.013)	-0.051 (0.023)
Constant	0.272 (0.049)	0.125 (0.083)
Observations	792	792
R <sup>2</sup>	0.043	0.272

## II Robustness Checks

Table E.6: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for personality characteristics in the 2012 ANES. Standard errors in parentheses. Estimates are used for Figure E.1 in the appendix.

	<i>Dependent variable:</i>			
	Turnout	Participation	Internal Efficacy	External Efficacy
Discursive Soph.	2.912 (0.300)	1.419 (0.098)	0.272 (0.021)	0.085 (0.025)
Factual Knowledge	0.434 (0.218)	0.105 (0.075)	0.165 (0.016)	0.036 (0.019)
Female	0.088 (0.085)	-0.078 (0.028)	-0.059 (0.006)	0.009 (0.007)
College Degree	0.352 (0.112)	0.160 (0.034)	0.020 (0.007)	0.033 (0.008)
Family Income	0.941 (0.157)	-0.002 (0.052)	0.006 (0.011)	0.010 (0.013)
Age (log)	0.981 (0.106)	0.102 (0.039)	-0.003 (0.008)	-0.017 (0.010)
African American	0.902 (0.124)	0.101 (0.038)	0.067 (0.008)	0.078 (0.010)
Church Attendance	0.732 (0.130)	0.104 (0.040)	0.006 (0.008)	0.044 (0.010)
Mode: Online	0.541 (0.096)	0.249 (0.033)	0.022 (0.007)	-0.033 (0.008)
Wordsum Score	0.649 (0.220)	0.409 (0.076)	0.092 (0.016)	0.021 (0.019)
Extraversion	0.029 (0.030)	0.032 (0.010)	0.009 (0.002)	0.020 (0.002)
Reserved	0.009 (0.026)	-0.023 (0.009)	-0.005 (0.002)	0.002 (0.002)
Constant	-5.181 (0.449)	-0.645 (0.158)	0.296 (0.034)	0.255 (0.039)
Observations	4,691	4,685	4,691	4,679
R <sup>2</sup>		0.129	0.171	0.058
Log Likelihood	-1,861.790			

Table E.7: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for personality characteristics in the 2016 ANES. Standard errors in parentheses. Estimates are used for Figure E.1 in the appendix.

	<i>Dependent variable:</i>			
	Turnout	Participation	Internal Efficacy	External Efficacy
Discursive Soph.	3.936 (0.543)	0.760 (0.132)	0.213 (0.031)	0.070 (0.042)
Factual Knowledge	0.717 (0.241)	0.260 (0.059)	0.056 (0.014)	-0.052 (0.019)
Female	0.172 (0.125)	0.046 (0.029)	-0.064 (0.007)	-0.004 (0.009)
College Degree	0.616 (0.155)	0.085 (0.033)	0.056 (0.008)	0.055 (0.010)
Family Income	0.471 (0.230)	-0.085 (0.055)	0.017 (0.013)	0.060 (0.017)
Age (log)	0.894 (0.153)	-0.119 (0.039)	0.018 (0.009)	-0.010 (0.012)
African American	1.228 (0.267)	0.102 (0.052)	0.057 (0.012)	-0.026 (0.016)
Church Attendance	0.930 (0.205)	-0.172 (0.043)	-0.007 (0.010)	0.078 (0.014)
Mode: Online	0.199 (0.148)	0.111 (0.036)	0.069 (0.009)	-0.015 (0.011)
Wordsum Score	0.999 (0.311)	0.436 (0.080)	0.114 (0.019)	0.022 (0.025)
Extraversion	0.090 (0.041)	0.023 (0.010)	0.011 (0.002)	0.012 (0.003)
Reserved	0.005 (0.037)	-0.013 (0.009)	-0.002 (0.002)	0.005 (0.003)
Constant	-4.948 (0.646)	0.196 (0.163)	0.199 (0.039)	0.277 (0.051)
Observations	2,999	3,011	3,012	3,012
R <sup>2</sup>		0.065	0.155	0.047
Log Likelihood	-907.430			

Table E.8: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for individual response length in the 2012 ANES. Standard errors in parentheses. Estimates are used for Figure E.2 in the appendix.

	<i>Dependent variable:</i>			
	Turnout	Participation	Internal Efficacy	External Efficacy
Discursive Soph.	2.053 (0.543)	0.967 (0.186)	0.215 (0.039)	0.082 (0.045)
Factual Knowledge	0.426 (0.218)	0.096 (0.075)	0.158 (0.016)	0.032 (0.018)
Female	0.067 (0.085)	-0.076 (0.028)	-0.054 (0.006)	0.016 (0.007)
College Degree	0.370 (0.113)	0.167 (0.034)	0.023 (0.007)	0.035 (0.008)
Family Income	0.948 (0.156)	0.026 (0.052)	0.011 (0.011)	0.016 (0.013)
Age (log)	0.976 (0.106)	0.094 (0.039)	-0.007 (0.008)	-0.014 (0.009)
African American	0.896 (0.123)	0.088 (0.038)	0.064 (0.008)	0.082 (0.009)
Church Attendance	0.758 (0.129)	0.113 (0.040)	0.010 (0.008)	0.048 (0.010)
Mode: Online	0.574 (0.097)	0.246 (0.034)	0.020 (0.007)	-0.039 (0.008)
Wordsum Score	0.626 (0.219)	0.396 (0.076)	0.091 (0.016)	0.013 (0.019)
Word Count (log)	0.970 (0.507)	0.539 (0.180)	0.071 (0.037)	0.003 (0.044)
Constant	-5.237 (0.418)	-0.710 (0.150)	0.311 (0.031)	0.352 (0.036)
Observations	4,716	4,692	4,996	4,985
R <sup>2</sup>		0.126	0.162	0.043
Log Likelihood	-1,866.366			

Table E.9: Effects of sophistication on turnout, non-conventional participation, internal efficacy, and external efficacy controlling for individual response length in the 2016 ANES. Standard errors in parentheses. Estimates are used for Figure E.2 in the appendix.

	<i>Dependent variable:</i>			
	Turnout	Participation	Internal Efficacy	External Efficacy
Discursive Soph.	4.762 (0.843)	0.224 (0.231)	0.222 (0.055)	0.153 (0.073)
Factual Knowledge	0.727 (0.219)	0.269 (0.058)	0.058 (0.014)	-0.051 (0.018)
Female	0.189 (0.114)	0.054 (0.029)	-0.059 (0.007)	-0.002 (0.009)
College Degree	0.547 (0.141)	0.096 (0.033)	0.058 (0.008)	0.056 (0.010)
Family Income	0.487 (0.207)	-0.074 (0.055)	0.020 (0.013)	0.062 (0.017)
Age (log)	0.842 (0.139)	-0.111 (0.038)	0.020 (0.009)	-0.005 (0.012)
African American	1.137 (0.237)	0.090 (0.051)	0.057 (0.012)	-0.019 (0.016)
Church Attendance	1.057 (0.191)	-0.171 (0.043)	-0.006 (0.010)	0.079 (0.014)
Mode: Online	0.081 (0.162)	0.180 (0.043)	0.069 (0.010)	-0.026 (0.014)
Wordsum Score	0.931 (0.275)	0.401 (0.079)	0.103 (0.019)	0.017 (0.025)
Word Count (log)	-1.016 (0.802)	0.659 (0.220)	0.009 (0.053)	-0.089 (0.070)
Constant	-3.972 (0.586)	0.017 (0.165)	0.228 (0.039)	0.365 (0.052)
Observations	3,486	3,039	3,037	3,038
R <sup>2</sup>		0.065	0.147	0.043
Log Likelihood	-1,084.839			